FREE STEIN DISCREPANCY AS A REGULARITY CONDITION

IAN CHARLESWORTH\textsuperscript{a} AND BRENT NELSON\textsuperscript{b}

Abstract. We introduce a free probabilistic quantity called free Stein information, which is defined in terms of free Stein discrepancies. It turns out that this quantity exactly measures the von Neumann dimension of the closure of the domain of the adjoint of the non-commutative Jacobian associated to Voiculescu’s free difference quotients. One consequence of this is that the free Stein information is a $*$-algebra invariant. We also relate this quantity to the free Fisher information, the non-microstates free entropy, and the non-microstates free entropy dimension.

Introduction.
In free probability, given an $n$-tuple of self-adjoint operators $X := (x_1, \ldots, x_n)$ in a tracial von Neumann algebra $(M, \tau)$, a regularity condition is some quantitative behavior of the joint distribution of $X$ that implies some qualitative behavior of the individual operators $x_1, \ldots, x_n$ or the algebras (von Neumann or otherwise) that they generate. All of the well-studied regularity conditions are expressed in terms of free probabilistic quantities that fall into two categories: microstates or non-microstates. Examples of the former include Voiculescu’s microstates free entropy $\chi(X)$, microstates free entropy dimension $\delta(X)$ (see [Voi94]), modified microstates free entropy dimension $\delta_0(X)$ (see [Voi96]), upper free orbit dimension $\Delta_2(X)$ (see [HS07]), and 1-bounded entropy $h(W^*(X))$ (see [Hay18]). Examples of the latter include non-microstates free entropy $\chi^*(X)$, free Fisher information $\Phi^*(X)$ (see [Voi98]), non-microstates free entropy dimensions $\delta^*(X)$ and $\delta'^*(X)$, and $\Delta(X)$ (see [CS05]).

Roughly speaking, microstates quantities examine the joint distribution of $X$ in terms of how well it is approximated by finite dimensional matrix algebras, whereas non-microstates quantities consider the behavior of certain derivations on the polynomial algebra generated by $x_1, \ldots, x_n$. We recall a few of the regularity conditions corresponding to the aforementioned free probabilistic quantities:

- If $\Phi^*(x) < \infty$, then the spectral measure of $x$ is Lebesgue absolutely continuous with density in $L^1(\mathbb{R}, m)$ [Voi93].
- If $\delta(x) = 1$, then $x$ is diffuse (i.e. its spectral measures has no atoms) [Voi94].
- If $\delta_0(X) > 1$, then $W^*(X)$ has no Cartan subalgebras and does not have property $\Gamma$ [Voi96].
- If $\delta_0(X) > 1$, then $W^*(X)$ is prime [Ge98].
- If $\Phi^*(X) < \infty$, then $W^*(X)$ does not have property $\Gamma$ [Dab10].
- If $\delta^*(X) = n > 1$, then $W^*(X)$ is a factor [Dab10].
- If $\delta^*(X) = n$, then every non-constant, self-adjoint $p \in \mathbb{C} \langle x_1, \ldots, x_n \rangle$ is diffuse [CS16, MSW17].
- If $\Phi^*(X) < \infty$, then $\chi^*(p) > -\infty$ for every non-constant, self-adjoint $p \in \mathbb{C} \langle x_1, \ldots, x_n \rangle$ [BM18].

In the present paper, we propose new quantities that fall into the non-microstates category: free Stein information and $R$-bounded free Stein information (see Definition 2.1). Motivated by work of the second author in [FN17], these quantities are defined via the free analogues of Stein information and Stein discrepancy (see [LNP15] and its references). Given an $n$-tuple $(\xi_1, \ldots, \xi_n) \in L^2(M)^n$, the free Stein discrepancy of $X$ relative to this $n$-tuple (see Subsection 1.3) is a non-negative quantity that measures how close $\xi_1, \ldots, \xi_n$ are to being the conjugate variables to $x_1, \ldots, x_n$. In particular, the free Stein discrepancy is zero if and only if $\xi_1, \ldots, \xi_n$ are the conjugate variables, in which case $\Phi^*(X) < \infty$ and so the above results tell us that $W^*(X)$ does not have property $\Gamma$ and $\chi^*(p) > -\infty$ for every non-constant, self-adjoint $p \in \mathbb{C} \langle x_1, \ldots, x_n \rangle$. Of course, determining that the free Stein discrepancy was zero required preexisting knowledge of the $n$-tuple $(\xi_1, \ldots, \xi_n)$ — or a very lucky guess.

\textsuperscript{a}DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY
\textsuperscript{b}DEPARTMENT OF MATHEMATICS, VANDERBILT UNIVERSITY

ilc@math.berkeley.edu
brent.nelson@vanderbilt.edu
In this paper, we explore what can be said if instead one merely supposes that the free Stein discrepancy can be made arbitrarily small by varying the n-tuple \((\xi_1, \ldots, \xi_n) \in L^2(M)^n\). We are therefore naturally driven to consider the infimum of free Stein discrepancies, which we define as the free Stein information, and the situation of interest is simply the regularity condition of having zero free Stein information. One immediately has that this is a weaker regularity condition than \(\Phi^*(X) < \infty\), but it turns out to be a stronger condition than \(\delta^*(X) = n\) (see Corollary 2.13). Interestingly, in the one variable case \(\delta^*(X) = 1\) is equivalent to having zero free Stein information. This is because for \(X = (x_1)\), the square of the free Stein information can be computed explicitly and is given by the sum of the squares of masses of any atoms in the spectral measure of \(x_1\) (see Proposition 2.14). In the general case, the free Stein information is (somewhat surprisingly) given by a formula involving the von Neumann dimension of the domain of an unbounded operator (see Theorem 2.5): namely, the adjoint of the non-commutative Jacobian associated to Voiculescu’s free difference quotients (see Subsection 1.1). From this characterization it follows that the free Stein information is a \(*\)-algebra invariant (see Theorem 2.9).

The structure of the paper is as follows. In Section 1 we establish some notation and recall the definitions of free Stein kernels and free Stein discrepancy. In Section 2 (the bulk of the paper) we define free Stein information and \(R\)-bounded free Stein information, provide alternate characterizations, establish properties, give explicit computations in certain examples, and finally relate the free Stein information being zero to other regularity conditions. We conclude with a few appendices detailing interesting examples and computations.

Acknowledgments. The authors would like to thank Dimi Shlyakhtenko for his useful comments and suggestions. They would also like to thank Michael Hartglass, Benjamin Hayes, and David Jekel for helpful discussions related to this paper. This work was initiated while the authors were attending the Park City Mathematics Institute (PCMI) Summer Session on Random Matrices. Part of this research was performed while the authors were visiting the Institute for Pure and Applied Mathematics (IPAM), which is supported by the National Science Foundation. The first and second authors were supported by NSF grants DMS-1803557 and DMS-1502822, respectively.

1. Preliminaries.

1.1 Notation. Fix a tracial \(W^\ast\)-probability space \((M, \tau)\). We denote by \(L^2(M)\) the GNS Hilbert space corresponding to \(\tau\) and identify \(M\) with its representation on this space. We let \(M^\circ = \{x^\circ : x \in M\}\) denote the opposite von Neumann algebra, represented on \(L^2(M^\circ)\) which can be identified with the dual Hilbert space to \(L^2(M)\). We let \(M \circ \otimes M^\circ\) denote the von Neumann algebra tensor product, which is equipped with the tensor product trace \(\tau \otimes \tau^\circ\).

Fix \(X := (x_1, \ldots, x_n) \in M_{\times, n}\) such that \(W^\ast(x_1, \ldots, x_n) = M\), and define \(\mathbb{C}\langle X \rangle := \mathbb{C}\langle x_1, \ldots, x_n \rangle\). Assume that \(x_1, \ldots, x_n\) are algebraically free (i.e. they do not satisfy any non-trivial polynomial relations). Recall, that \(\text{Voiculescu’s free difference quotients}\) for \(X\) are maps \(\partial_j : \mathbb{C}\langle X \rangle \to \mathbb{C}\langle X \rangle \otimes \mathbb{C}\langle X \rangle, j = 1, \ldots, n\), defined by

\[
\partial_j(x_{i_1} \cdots x_{i_d}) = \sum_{k=1}^d \delta_{i_k = j} x_{i_1} \cdots \hat{x}_{i_k} \cdots x_{i_d} \otimes x_{i_k} \cdots x_{i_d}.
\]

(Note that these maps are well-defined by the assumed algebraic freeness.) By mapping \(a \otimes b \mapsto a \otimes b^\circ\), we can think of \(\partial_j\) as valued in \(L^2(M \circ \otimes M^\circ)\). We define \(\partial : \mathbb{C}\langle X \rangle \to L^2(M \circ \otimes M^\circ)^\ast\) by

\[
\partial(p) = (\partial_1p, \ldots, \partial_np).
\]

We let \(\mathcal{J} : \mathbb{C}\langle X \rangle^n \to M_n(L^2(M \circ \otimes M^\circ))\) denote the \(\text{non-commutative Jacobian}\), which is defined by

\[
\mathcal{J}(p_1, \ldots, p_n) = \begin{bmatrix}
\partial_1 p_1 & \cdots & \partial_n p_1 \\
\partial_1 p_2 & \cdots & \partial_n p_2 \\
\vdots & \ddots & \vdots \\
\partial_1 p_n & \cdots & \partial_n p_n
\end{bmatrix} = \begin{bmatrix}
\partial p_1 \\
\partial p_2 \\
\vdots \\
\partial p_n
\end{bmatrix}.
\]
Denote
\[
I := \begin{bmatrix} 1 \otimes 1^\circ & 0 \\
0 & 1 \otimes 1^\circ \end{bmatrix},
\]
so that \( \mathcal{J}(X) = I \).

For \( E = (\xi_1, \ldots, \xi_n) \), \( H = (\eta_1, \ldots, \eta_n) \) in either \( L^2(M)^n \) or \( L^2(M \otimes M^\circ)^n \) denote
\[
\langle \Xi, H \rangle_2 := \sum_{j=1}^{n} \langle \xi_j, \eta_j \rangle_2.
\]

For \( A, B \in M_n(L^2(M \otimes M^\circ)) \) denote
\[
\langle A, B \rangle_{HS} := \sum_{j,k=1}^{n} \langle [A]_{jk}, [B]_{jk} \rangle_2.
\]

We denote by \( \# \) the usual product in \( M \otimes M^\circ \) \( ((a \otimes b^\circ) \# (c \otimes d^\circ)) = (ac) \otimes (db)^\circ \), the usual product in \( M_n(M \otimes M^\circ) \), the action of \( M \otimes M^\circ \) on \( L^2(M) \) \( ((a \otimes b)^\circ c = abc) \), and the action of \( M_n(M \otimes M^\circ) \) on \( L^2(M)^n \).

1.2 Non-commutative power series. Let \( t_1, \ldots, t_n \) be formal, non-commuting, self-adjoint variables. After [CS16], for \( R > 0 \) we denote by \( C \langle t_1, \ldots, t_n : R \rangle \) the completion of \( C \langle t_1, \ldots, t_n \rangle \) in the norm
\[
\left\| \sum_{d=1}^{N} \sum_{j=(j_1,\ldots,j_d) \in [n]^d} c_J t_{j_1} \cdots t_{j_d} \right\|_R := \sum_{d=1}^{N} \sum_{j=(j_1,\ldots,j_d) \in [n]^d} |c_J|R^d.
\]

We also denote
\[
A[t_1, \ldots, t_n; R] := \bigcup_{R' > R} C \langle t_1, \ldots, t_n; R' \rangle.
\]

This space should be regarded as non-commutative power series with radius of convergence strictly greater than \( R \). Moreover, observe that if \( R \geq \max_i \|x_i\| \), then there is a unique homomorphism from \( A[t_1, \ldots, t_n; R] \) to \( M \) that sends \( t_i \) to \( x_i \). We let \( p(x_1, \ldots, x_n) \) denote the image of \( p \in A[t_1, \ldots, t_n; R] \) under this map, and define
\[
C \langle X \rangle_R := \{ p(x_1, \ldots, x_n) : p \in A[t_1, \ldots, t_n; R] \}.
\]

Finally, we remark that for each \( j = 1, \ldots, n, \) \( \partial_j \) extends to a derivation
\[
\partial_j : C \langle X \rangle_R \rightarrow M \otimes M^\circ.
\]

More precisely, \( \partial_j \) is valued in the image of the projective tensor product \( A[t_1, \ldots, t_n; R] \otimes A[t_1, \ldots, t_n; R] \) in \( M \otimes M^\circ \) under the unique homomorphism that sends \( p_1 \otimes p_2 \) to \( p_1(x_1, \ldots, x_n) \otimes p_2(x_1, \ldots, x_n) \). Consequently, \( \partial \) and \( \mathcal{J} \) each extend to \( C \langle X \rangle_R \) and \( C \langle X \rangle_R^n \), respectively.

1.3 Free Stein kernels and free Stein discrepancy. Given \( \Xi \in L^2(M)^n \), we say (after [FN17]) that
\[
A \in M_n(L^2(M \otimes M^\circ))
\]
is a free Stein kernel of \( X \) relative to \( \Xi \) if
\[
\langle \Xi, P \rangle_2 = \langle A, P \rangle_{HS} \quad \forall P \in C \langle X \rangle^n.
\]

In this case we say (after [Shl04]) that \( \Xi \) is a partial conjugate variable to \( X \) corresponding to \( A \).

The free Stein discrepancy of \( X \) relative to \( \Xi \) is the quantity
\[
\Sigma^*(X \mid \Xi) := \inf_A \| A - I \|_{HS},
\]
where the infimum is over all free Stein kernels of \( X \) relative to \( \Xi \). Equivalently, \( \Sigma^*(X \mid \Xi) = \| \Pi(A) - I \|_{HS} \) where \( A \) is any free Stein kernel of \( X \) relative to \( \Xi \) and \( \Pi \) is the orthogonal projection onto the closure of the range of \( \mathcal{J} \).
A priori the free Stein discrepancy could be infinite, since a free Stein kernel for $X$ need not exist. Indeed, if $\Xi$ is not orthogonal to $C_n \subset L^2(M)^n$ then for some $Z \in C^n$ we have
\[
\langle \Xi, Z \rangle_2 \neq 0 = (A, \mathcal{J} Z)_{HS} \quad \forall A \in M_n(L^2(M \otimes M^o)).
\]
However, if one merely assumes $\Xi \perp C^n$ then by [CFM18, Theorem 2.1] the free Stein discrepancy will necessarily be finite. We state a slightly more general version here for the convenience of the reader, but remark that this holds by exactly the same proof as the original.

**Proposition 1.1.** For $\Xi = (\xi_1, \ldots, \xi_n) \in L^2(M)^n \otimes C^n$,
\[
B_{\Xi} := \left[\frac{1}{2}(\xi_i \otimes 1 - 1 \otimes \xi_i) \#(x_j \otimes 1 - 1 \otimes x_j)\right]_{i,j=1}^n \in M_n(L^2(M \otimes M^o))
\]
is a free Stein kernel for $X$ relative to $\Xi$. Consequently, $\Sigma^*(X \mid \Xi) < \infty$ always.

One might hope that $B_{\Xi}$ is the free Stein kernel which attains the free Stein discrepancy of $X$, but unfortunately this holds if and only if $\Xi = 0$ (see Appendix A). However, we do obtain the following corollary:

**Corollary 1.2.** The map
\[
L^2(M)^n \otimes C^n \ni \Xi \mapsto \Sigma^*(X \mid \Xi)
\]
is continuous.

**Proof.** For $\Xi, \Xi' \in L^2(M)^n \otimes C^n$ let $B_{\Xi}$ and $B_{\Xi'}$ be as in Proposition 1.1. Then
\[
|\Sigma^*(X \mid \Xi) - \Sigma^*(X \mid \Xi')| = \|\Pi(B_{\Xi}) - \Pi(B_{\Xi'})\|_{HS} \leq C\|\Xi - \Xi'\|_2,
\]
where $C > 0$ is a constant depending only on $n$ and $X$.

**Remark 1.3.** If $\Sigma^*(X \mid \Xi) = 0$, then $\mathbb{1}$ is a free Stein kernel for $X$ and hence
\[
\langle \Xi, \mathbb{1} \rangle_2 = (\mathbb{1}, \mathcal{J} \mathbb{1})_{HS} \quad \forall \mathbb{1} \in \mathcal{C}(X)^n.
\]
That is, $\Xi$ is the usual conjugate variable to $X$. In fact, this is precisely why the free Stein discrepancy is defined to measure the distance between a free Stein kernel $A$ and $\mathbb{1}$. We remind the reader that the free Fisher information of $X$ is defined as the quantity
\[
\Phi^*(X) := \|\Xi\|_2^2
\]
if the conjugate variable to $X$ exists and is $\Xi$, and is otherwise defined to be $+\infty$ (cf. [Voi98, Definition 6.1]).

**Remark 1.4.** $\Sigma^*(X \mid X) = 0$ if and only if $X$ is the conjugate variable to $X$ if and only if $X$ is a free semicircular family.

**Remark 1.5.** Equation (1) is equivalent to saying $A \in \text{dom}(\mathcal{J}^*)$ with $\mathcal{J}^*(A) = \Xi$, where we think of $\mathcal{J} : L^2(M)^n \rightarrow M_n(L^2(M \otimes M^o))$ as a densely defined (unbounded) operator.

## 2. Free Stein Information.

We begin by introducing two new quantities determined by the joint distribution of $X$:

**Definition 2.1.** The **free Stein information** of $X$ is the quantity:
\[
\Sigma^*(X) := \inf\{\Sigma^*(X \mid \Xi) : \Xi \in L^2(M)^n\}.
\]
For $R > 0$, the **$R$-bounded free Stein information** of $X$ is the quantity
\[
\Sigma_R^*(X) := \inf\{\Sigma^*(X \mid \Xi) : \Xi \in L^2(M)^n \text{ with } \|\Xi\|_2 \leq R\}.
\]
Remark 2.2. Since \( \Sigma_R^*(X) \) is a decreasing function in \( R \) it is easy to see that
\[
\Sigma^*(X) = \inf_{R > 0} \Sigma_R^*(X) = \lim_{R \to \infty} \Sigma_R^*(X).
\]

Also, by Corollary 1.2 it follows that the values \( \Sigma^*(X) \) and \( \Sigma_R^*(X) \) are unchanged if we replace \( L^2(M) \) in their definitions with any dense subset. In particular, it suffices to consider \( \Xi \in \mathbb{C}(X)^n \). Furthermore, by Remark 1.5 we see that \( \Sigma^*(X) \) is precisely the distance from \( \mathbb{I} \) to \( \text{dom}(\mathcal{J}^*) \) in \( M_n(L^2(M \otimes M^o)) \).

2.1 Alternate characterizations. We first prove some alternate characterizations.

Theorem 2.3. For \( R > 0 \), \( \Sigma_R^*(X) = 0 \) if and only if \( \Phi^* (X) \leq R^2 \).

Proof. Suppose \( \Sigma_R^*(X) = 0 \). Then there exists a sequence \( (\Xi^{(k)})_{k \in \mathbb{N}} \subset L^2(M) \) such that \( \|\Xi^{(k)}\|_2 \leq R \) and \( \Sigma^*(X | \Xi^{(k)}) < \frac{1}{k} \) for all \( k \in \mathbb{N} \). Consequently, letting \( B_{\Xi^{(k)}} \) be as in Proposition 1.1 and setting \( A_{\Xi^{(k)}} = \Pi(B_{\Xi^{(k)}}) \), we have
\[
\|A_{\Xi^{(k)}} - \mathbb{I}\|_{\text{HS}} = \Sigma^*(X | \Xi^{(k)}) \to 0.
\]

Hence for every \( P \in \mathbb{C}(X)^n \) we have
\[
\lim_{k \to \infty} \langle \Xi^{(k)}, P \rangle_2 = \lim_{k \to \infty} \langle A_{\Xi^{(k)}}, \mathcal{J} P \rangle_{\text{HS}} = \langle \mathbb{I}, \mathcal{J} P \rangle_{\text{HS}}.
\]

The density of \( \mathbb{C}(X)^n \) in \( L^2(M) \) implies the sequence \( (\Xi^{(k)})_{k \in \mathbb{N}} \) (since it is uniformly bounded) converges weakly to some \( \Xi \in L^2(M)^n \). Moreover, the above limit implies \( \Xi \) is the conjugate variable to \( X \) and
\[
\Phi^*(X) = \|\Xi\|_2^2 \leq \lim_{k \to \infty} \|\Xi^{(k)}\|_2^2 \leq R^2.
\]

The converse is immediate. \( \square \)

In order to provide an alternate characterization for \( \Sigma^*(X) \), we will study the domains of \( \partial^* \) and \( \mathcal{J}^* \).

We first show their closures are left \( M \otimes M^o \)-modules. The following is the multivariate analogue of [Vol98, Proposition 4.1] and follows by an identical proof:

Lemma 2.4. For \( \eta \in \text{dom}(\partial^*) \in L^2(M \otimes M^o)^n \) and \( p, q \in \mathbb{C}(X) \), \( (p \otimes q \otimes I_n) \# \eta \in \text{dom}(\partial^*) \) with
\[
\partial^*((p \otimes q \otimes I_n) \# \eta) = (p \otimes q) \# \partial^*(\eta) - m \circ (1 \otimes \tau \otimes 1) \circ (1 \otimes \partial + \partial \otimes 1)(p \otimes q),
\]
where \( m(a \circ b \circ c) = abc \).

From this lemma we see that \( \text{dom}(\partial^*) \) is invariant under the action of \( \mathbb{C}(X) \otimes \mathbb{C}(X)^o \). Consequently, the Kaplansky density theorem implies that \( \text{dom}(\partial^*) \) is a closed, left \( M \otimes M^o \)-module. Observe that for \( A \in \text{dom}(\mathcal{J}^*) \), if \( A_i = (A_{i1}, \ldots, A_{in}) \) (i.e the \( i \)-th row of \( A \)) for \( i = 1, \ldots, n \), then \( A_{i1}, \ldots, A_{in} \in \text{dom}(\partial^*) \). It then follows that \( \text{dom}(\mathcal{J}^*) \) is also a closed, left \( M \otimes M^o \)-module satisfying \( \text{dom}(\mathcal{J}^*) \cong \text{dom}(\partial^*) \). This identification immediately gives the second equality in the following theorem.

Theorem 2.5. For \( X = (x_1, \ldots, x_n) \in M_{s.a.}^n \), generating \( M \) and algebraically free,
\[
n - \Sigma^*(X)^2 = \dim_{M \otimes M^o} \text{dom}(\partial^*) = \frac{1}{n} \dim_{M \otimes M^o} (\text{dom}(\mathcal{J}^*)�).
\]

Consequently, \( \Sigma^*(X) = 0 \) if and only if \( \mathcal{J} \) is closable, and if and only if \( \partial \) is closable.

Proof. Let \( e \in M_n(L^2(M \otimes M^o)) \) be the projection of \( \mathbb{I} \) onto \( \text{dom}(\mathcal{J}^*) \). Then \( \Sigma^*(X) = \|e - \mathbb{I}\|_{\text{HS}} \). Hence
\[
n - \Sigma^*(X)^2 = n - \|e\|_2^2 + 2\Re (e, \mathbb{I})_{\text{HS}} - \|\mathbb{I}\|_2^2 = \|e\|_{\text{HS}}.
\]

Now, identify \( M \otimes M^o \) with its diagonal representation \( M \otimes M^o \otimes I_n \) on \( L^2(M \otimes M^o)^n \). Then \( N := (M \otimes M^o) \cap \mathcal{B}(L^2(M \otimes M^o)) \) is identified with \( M_n(N) \). Let \( f \) be the projection of \( L^2(M \otimes M^o)^n \) onto \( \text{dom}(\partial^*) \), so that \( f \in M_n(N) \). Observe that
\[
\text{dom}(\mathcal{J}^*) \cong \text{dom}(\partial^*) \cong \left\{(Tv_1, \ldots, Tv_n) \in M_n(L^2(M \otimes M^o)): T \in M_n(N), TL^2(M \otimes M^o)^n \subset \text{dom}(\partial^*)\right\},
\]
where \( v_j \in L^2(M \otimes M^o)^n \) is the vector with \( 1 \otimes 1^o \) in the \( j \)-th entry and zeros elsewhere. Actually, \( (Tv_1, \ldots, Tv_n) \) in the right-most space is sent to its transpose in the left-most space. Hence \( f^T \in \text{dom}(\mathcal{J}^*) \),
Proof. For we have Proposition 2.7.

\[ \langle f^{T}, A \rangle_{HS} = \langle 1, f^{T}A \rangle_{HS} = \sum_{i,j=1}^{n} \langle 1 \otimes 1^{o}, [f^{T}]_{ij}[A]_{ij} \rangle_{HS} \]

\[ = \sum_{i,j=1}^{n} \langle 1 \otimes 1^{o}, [f]_{ij} [A]_{ij} \rangle_{HS} = \sum_{i,j,k=1}^{n} \langle \delta_{i=k} 1 \otimes 1^{o}, [f]_{kj}(A_{i})_{j} \rangle_{HS} \]

\[ = \sum_{i,k=1}^{n} \langle \delta_{i=k} 1 \otimes 1^{o}, (fA_{i})_{k} \rangle_{HS} = \sum_{i,k=1}^{n} \langle \delta_{i=k} 1 \otimes 1^{o}, (A_{i})_{k} \rangle_{HS} \]

\[ = \sum_{i=1}^{n} \langle 1 \otimes 1^{o}, [A]_{ii} \rangle_{HS} = \langle 1, A \rangle_{HS} = \langle e, A \rangle_{HS} . \]

Thus \( f^{T} = e \) and

\[ \dim_{M \otimes M^{o}}(\text{dom}(\vartheta^{*})) = \| f \|^{2}_{HS} = \| f^{T} \|^{2}_{HS} = \| e \|^{2}_{HS} . \]

So the result follows by our previous computation. The final statement follows from that fact that an operator is closable if and only if the domain of its adjoint is dense.

\[ \square \]

2.2 Properties. We derive some useful properties of free Stein information.

Proposition 2.6. The function \( R \mapsto \Sigma_{R}^{*}(X) \) is convex.

Proof. Let \( 0 < R_{1} < R_{2} \). Let \( A_{1}, A_{2} \in \text{dom}(\vartheta^{*}) \) with \( \| \vartheta^{*}(A_{i}) \|_{2} \leq R_{i}, i = 1, 2 \). Then for \( t \in [0, 1] \),

\[ (1-t)A_{1} + tA_{2} \in \text{dom}(\vartheta^{*}) \]

\[ \| \vartheta^{*}((1-t)A_{1} + tA_{2}) \|_{2} \leq (1-t)\| \vartheta^{*}(A_{1}) \|_{2} + t\| \vartheta^{*}(A_{2}) \|_{2} \leq (1-t)R_{1} + tR_{2} . \]

Hence

\[ \Sigma_{(1-t)R_{1} + tR_{2}}^{*}(X) \leq \| (1-t)A_{1} + tA_{2} \|_{HS} = (1-t)\| A_{1} \|_{HS} + t\| A_{2} \|_{HS} . \]

Taking the infimum over \( A_{1} \) and \( A_{2} \) completes the proof.

\[ \square \]

In light of Theorem 2.5, the following result is not at all surprising. Nevertheless, we proceed by appealing to the original definition of free Stein information, rather than its characterization as a dimension measurement.

Proposition 2.7. For \( X = (x_{1}, \ldots, x_{n}) \) and \( Y = (y_{1}, \ldots, y_{m}) \) algebraically free self-adjoint operators in \( M \), we have

\[ \Sigma^{*}(X)^{2} + \Sigma^{*}(Y)^{2} \leq \Sigma^{*}(X, Y)^{2} , \]

with equality if \( X \) and \( Y \) are freely independent in \( (M, \tau) \).

Proof. For \( Q \in \mathbb{C} \langle X, Y \rangle^{n+m} \), suppose \( A \) satisfies

\[ \langle Q, P \rangle_{2} = \langle A, \vartheta P \rangle_{HS} \]

\[ P \in \mathbb{C} \langle X, Y \rangle^{n+m} . \]

Let \( P_{1} \in \mathbb{C} \langle X \rangle^{n} \) and define \( P := (P_{1}, 0) \in \mathbb{C} \langle X, Y \rangle^{n+m} \). Then the above equation applied to this \( P \) reduces to

\[ \langle Q_{1}, P_{1} \rangle_{2} = \langle A_{1}, \vartheta P_{1} \rangle_{HS} , \]

where \( Q_{1} \) is the \( n \)-tuple formed via the first \( n \) entries of \( Q \), and \( A_{1} \) is the \( n \times n \) matrix formed from the top-left \( n \times n \) corner of \( A \). This same equation holds after replacing \( Q_{1} \) and \( A_{1} \) with their projections into \( L^{2}(W^{*}(X))^{n} \) and \( L^{2}(M_{n}(W^{*}(X) \otimes W^{*}(X)^{o})) \), respectively, which we will continue to denote in the same way. Thus \( A_{1} \) is a free Stein kernel for \( X \) relative to \( Q_{1} \) so that

\[ \Sigma^{*}(X) \leq \| A_{1} - 1 \|_{HS} . \]

Proceeding in the same way for \( Y \), we obtain \( A_{2} \in L^{2}(M_{n}(W^{*}(Y) \otimes W^{*}(Y)^{o})) \) such that

\[ \Sigma^{*}(Y) \leq \| A_{2} - 1 \|_{HS} . \]

6
Thus
\[ \Sigma^*(X)^2 + \Sigma^*(Y)^2 \leq \|A_1 - 1\|^2_{HS} + \|A_2 - 1\|^2_{HS} \leq \|A - 1\|^2_{HS}. \]
Since $A$ was an arbitrary free Stein kernel for $(X, Y)$, the desired inequality holds.

Now, suppose $X$ and $Y$ are free inside $(M, \tau)$. Let $A_1$ be a free Stein kernel for $X$ relative to some $n$-tuple $Q_1 \in \mathbb{C}\langle X\rangle^n$, and let $A_2$ be a free Stein kernel for $Y$ relative to some $n$-tuple $Q_2 \in \mathbb{C}\langle Y\rangle^m$. We claim that
\[ A := \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \]
is a free Stein kernel for $(X, Y)$ relative to $Q := (Q_1, Q_2)$. It suffices to show that for $P_1 \in \mathbb{C}\langle X, Y\rangle^n$ and $P_2 \in \mathbb{C}\langle X, Y\rangle^m$,
\[ \langle Q_1, P_1 \rangle_2 = \langle A_1, \mathcal{J}X P_1 \rangle_{HS} \quad \text{and} \quad \langle Q_2, P_2 \rangle_2 = \langle A_2, \mathcal{J}Y P_2 \rangle_{HS}. \]
We show only the first equality, as the second will follow mutatis mutandis. Fix $i \in [n]$, let $q := [Q_1]_i$, and $\eta = ([A]_{i1}, \ldots, [A]_{in})$. Note that $\langle q, 1 \rangle_2 = \langle \eta, \partial_X(1) \rangle_2 = 0$. Then it further suffices to show for any $p \in \mathbb{C}\langle X, Y\rangle$ that
\[ \langle q, p \rangle_2 = \langle \eta, \partial_X p \rangle_2. \]
(2)
Assume $p = c_0 P_1 c_1 \cdots P_n c_n$ where $c_0, c_1, \ldots, c_n \in \mathbb{C}\langle Y\rangle$ with $\tau(c_j) = 0$ for $j = 1, \ldots, n-1$, and $P_1, \ldots, P_n \in \mathbb{C}\langle X\rangle$ with $\tau(P_k) = 0$ for $k = 1, \ldots, n$. The case $n = 0$, where $p = c_0$, follows immediately by freeness, and for the case $n \geq 2$ one easily checks that both sides of (2) are zero. For $n = 1$, we have
\begin{align*}
\langle q, c_0 P_1 c_1 \rangle_2 &= \tau(c_0^*) \langle q, P_1 c_1 \rangle_2 \\
&= \tau(c_0^*) \langle q, P_1 \rangle_2 (1, c_1_2) \\
&= \tau(c_0^*) \langle \eta, \partial_X(P_1) \rangle_2 (1, c_1_2) \\
&= \langle \eta, c_0 \cdot \partial_X(P_1) \cdot c_1_2 angle \\
&= \langle \eta, \partial_X(c_0 P_1 c_1) \rangle_2
\end{align*}
Thus (2) holds for all elements of the form $p = c_0 P_1 c_1 \cdots P_n c_n$ as above, which span $\mathbb{C}\langle X, Y\rangle$.

With the claim established, we obtain
\[ \Sigma^*(X, Y)^2 \leq \|A - 1\|^2_{HS} = \|A_1 - 1\|^2_{HS} + \|A_2 - 1\|^2_{HS}. \]
Since $A_1$ and $A_2$ were arbitrary, we obtain the desired equality. \qed

We will next show that $\Sigma^*(X)$ is an algebra invariant. We first require a change of variables estimate.

**Lemma 2.8.** Let $Y = (y_1, \ldots, y_n) \in \mathbb{C}\langle X\rangle^n$ be such that $\mathbb{C}\langle y_1, \ldots, y_n \rangle$ is dense in $L^2(M)$. Then
\[ \Sigma^*(Y)^2 \leq \Sigma^*(X)^2 + \frac{1}{n} \dim_{M \otimes M^*}(\ker(J \mathcal{J}Y J)), \]
where $J = J_{(r \otimes r)_{0T}}$ is the Tomita conjugation operator on $M_n(L^2(M \otimes M^*))$.

**Proof.** Suppose $A$ is a free Stein kernel for $X$. Then $\mathcal{J}^*(A) \in L^2(M) = L^2(W^*(y_1, \ldots, y_n))$, and for any $P \in \mathbb{C}\langle Y\rangle^n$ we have
\[ \langle \mathcal{J}^*(A), P(Y) \rangle_2 = \langle A, [\mathcal{J}P](Y) \# \mathcal{J}Y \rangle_{HS} = \langle A\#(\mathcal{J}Y)^*, [\mathcal{J}P](Y) \rangle_{HS}. \]
Thus $A\#(\mathcal{J}Y)^*$ is a free Stein kernel for $Y$. It follows that
\[ \overline{\text{dom}(\mathcal{J}^*) \#(\mathcal{J}Y)^*} \subset \overline{\text{dom}(\mathcal{J}Y^*)}, \]
where $\mathcal{J}Y$ is the non-commutative Jacobian with respect to $Y$. Denote
\[ T = J \mathcal{J}Y J \]
so that $A\#(\mathcal{J}Y)^* = TA$. The above inclusion and the rank–nullity theorem implies
\[ \dim_{M \otimes M^*}(\overline{\text{dom}(\mathcal{J}^*)}) - \dim_{M \otimes M^*}(\ker T) = \dim_{M \otimes M^*}(\overline{\text{dom}(\mathcal{J}^*) \#(\mathcal{J}Y)^*}) \leq \dim_{M \otimes M^*}(\overline{\text{dom}(\mathcal{J}Y^*)}). \]
So by Theorem 2.5 we have
\[ \Sigma^*(Y)^2 \leq \Sigma^*(X)^2 + \frac{1}{n} \dim_{M \otimes M^*}(\ker T). \] \qed
**Theorem 2.9.** If \( Y = (y_1, \ldots, y_n) \in \mathbb{C} \langle x_1, \ldots, x_n \rangle^n \) is such that \( \mathbb{C} \langle y_1, \ldots, y_n \rangle = \mathbb{C} \langle x_1, \ldots, x_n \rangle \), then
\[
\Sigma^*(Y) = \Sigma^*(X).
\]
That is, \( \Sigma^*(X) \) is an algebra invariant.

**Proof.** We have \( X \in \mathbb{C} \langle y_1, \ldots, y_n \rangle^n \) and hence
\[
1 = \mathcal{J} X = \mathcal{J} Y(X) \# \mathcal{J} Y.
\]
Thus, \( \mathcal{J} \mathcal{J} Y \) has left-inverse \( \mathcal{J} \mathcal{J} Y \mathcal{J} \). This implies \( \ker(\mathcal{J} \mathcal{J} Y) = \{0\} \) so that \( \Sigma^*(Y) \leq \Sigma^*(X) \) by the previous lemma. The reverse inequality follows by swapping the roles of \( X \) and \( Y \).

Using the concepts from Subsection 1.2, we can upgrade this slightly and see that is in fact an analytic invariant. The proof is exactly the same.

**Corollary 2.10.** Let \( R \geq \max_i \|x_i\| \) and let \( Y = (y_1, \ldots, y_n) \in \mathbb{C} \langle X \rangle^n_R \) be such that \( \mathbb{C} \langle y_1, \ldots, y_n \rangle \) is dense in \( L^2(M) \). Then
\[
\Sigma^*(Y)^2 \leq \Sigma^*(X)^2 + \frac{1}{n} \dim_{M \otimes M^\circ}(\ker(\mathcal{J} \mathcal{J} Y)),
\]
where \( J = J_{(r \otimes r^*) Y} \) is the Tomita conjugation operator on \( M_n(L^2(M \otimes M^\circ)) \). In particular, if \( R \geq \max_i \|y_i\| \) and if \( \mathbb{C} \langle X \rangle_R^\# \) is the image of \( \mathcal{A} \|1, \ldots, t_n; R \) under the homomorphism that sends \( t_i \) to \( y_i \), then \( \Sigma^*(Y) = \Sigma^*(X) \).

The following result is simply [Shl04, Theorem 2.7]. We state (and prove) it here using our notation and terminology.

**Proposition 2.11.** Let \( S \) be a free semicircular family, free from \( X \). Then
\[
\lim_{t \to 0} \sup_t \Phi^*(X + \sqrt{t} S) \leq \Sigma^*(X)^2.
\]

**Proof.** Let \( f: (0, \infty) \to (0, \infty) \) be any decreasing function such that
\[
\lim_{t \to 0} f(t) = +\infty, \text{ but } \lim_{t \to 0} \sqrt{t} f(t) = 0.
\]
(E.g. \( f(t) = t^{-\frac{1}{2}} \)). Then
\[
\lim_{t \to 0} \Sigma_{f(t)}(X) = \Sigma^*(X).
\]
For each \( t > 0 \), let \( Q_t \in \mathbb{C} \langle X \rangle^n \) be such that \( \|Q_t\|_2 \leq f(t) \) and such that there exists a free Stein kernel \( A_t \) for \( X \) relative to \( Q_t \) such that
\[
\|A_t - 1\|_{HS} \leq \Sigma^*_{f(t)}(X) + t.
\]
Recall that the conjugate variables to \( X + \sqrt{t} S \) are \( \mathcal{E}_t(\frac{1}{\sqrt{t}} S) \) where \( \mathcal{E}_t: W^*(X, S) \to W^*(X + \sqrt{t} S) \) is the conditional expectation. It follows from [Shl04, Lemma 2.3] that
\[
\mathcal{E}_t(Q_t) = \mathcal{E}_t \left( A_t \# \frac{1}{\sqrt{t}} S \right).
\]
Thus
\[
\sqrt{t} \Phi^*(X + \sqrt{t} S)^\frac{1}{2} = \sqrt{t} \left\| \mathcal{E}_t \left( \frac{1}{\sqrt{t}} S \right) \right\|_2 \leq \sqrt{t} \left\| \mathcal{E}_t \left( (1 - A_t) \# \frac{1}{\sqrt{t}} S \right) \right\|_2 + \sqrt{t} \|\mathcal{E}_t(Q_t)\|_2 \leq \sqrt{t} \left\| (1 - A_t) \# \frac{1}{\sqrt{t}} S \right\|_2 + \sqrt{t} f(t) = \|1 - A_t\|_{HS} + \sqrt{t} f(t) \leq \Sigma^*_{f(t)}(X) + t + \sqrt{t} f(t).
\]
This tends to \( \Sigma^*(X) \) as \( t \to 0 \). \( \square \)
We remind the reader that the non-microstates free entropy of $X$ is defined as the quantity

$$\chi^\ast(X) := \frac{1}{2} \int_0^\infty \frac{n}{1 + t} - \Phi^\ast(X + \sqrt{t}S) \, dt + \frac{n}{2} \log(2\pi e),$$

where $S$ is a free semicircular family free from $X$ (cf. [Voi98, Definition 7.1]). The following is [Shl04, Corollary 2.8]. As with the previous result, we state (and prove) it using our notation and terminology.

**Proposition 2.12.** Let $S$ be a free semicircular family free from $X$. Then

$$\limsup_{\epsilon \to 0} \frac{\chi^\ast(X + \sqrt{\epsilon}S)}{\frac{1}{2} \log \epsilon} \leq \Sigma^\ast(X)^2.$$

**Proof.** Using [Voi98, Corollary 6.14] and implementing the change of variable $t \to t - \epsilon$ in the integral appearing in the above definition of $\chi^\ast$, we obtain

$$\limsup_{\epsilon \to 0} \frac{\chi^\ast(X + \sqrt{\epsilon}S)}{\frac{1}{2} \log \epsilon} = \limsup_{\epsilon \to 0} \frac{1}{\log \epsilon} \int_{\epsilon}^1 \frac{n}{1 + t - \epsilon} - \Phi^\ast(X + \sqrt{t}S) \, dt.$$

Now, for any free Stein kernel $A$ relative to some $Q$ we have

$$\Phi^\ast(X + \sqrt{\epsilon}S) \leq \left( \| \mathcal{E}_t((1 - A)\# \frac{1}{\sqrt{t}} S) \|_2 + \| \mathcal{E}_t(A\# \frac{1}{\sqrt{t}} S) \|_2 \right)^2 \leq \frac{1}{t} \| 1 - A \|_{HS}^2 + \frac{2}{\sqrt{t}} \| 1 - A \|_{HS} \| Q \|_2 + \| Q \|_2^2.$$

Thus

$$\int_{\epsilon}^1 \frac{n}{1 + t - \epsilon} - \Phi^\ast(X + \sqrt{t}S) \, dt \geq n \log(2 - \epsilon) + \log(\epsilon) \| 1 - A \|_{HS}^2 - 4(1 - \sqrt{\epsilon}) \| 1 - A \|_{HS} \| Q \|_2 - (1 - \epsilon) \| Q \|_2^2.$$

Since $\log(\epsilon) < 0$ for $\epsilon < 1$, this in turn implies

$$\limsup_{\epsilon \to 0} \frac{\chi^\ast(X + \sqrt{\epsilon}S)}{\frac{1}{2} \log \epsilon} \leq \| 1 - A \|_{HS}^2.$$

Since $A$ was an arbitrary free Stein kernel, we obtain the desired inequality. \qed

We remind the reader that the there are two versions of the non-microstates free entropy dimension of $X$:

$$\delta^\ast(X) := n - \liminf_{\epsilon \to 0} \frac{\chi^\ast(X + \sqrt{\epsilon}S)}{\frac{1}{2} \log \epsilon} \quad \delta^\ast(X) := n - \liminf_{\epsilon \to 0} e\Phi^\ast(X + \sqrt{\epsilon}S),$$

where $S$ is a free semicircular family free from $X$ (cf. [CS05, Section 4.1.1]). Thus, from Propositions 2.11 and 2.12 we obtain:

**Corollary 2.13.** For algebraically free $X$,

$$n - \Sigma^\ast(X)^2 \leq \delta^\ast(X) \leq \delta^\ast(X).$$

**2.3 Computations.** We provide some examples in which the free Stein information can be explicitly computed.

**Proposition 2.14.** Let $x \in (M, \tau)$ be self-adjoint and algebraically free, with distribution $\mu$ on $\mathbb{R}$. Then

$$\Sigma^\ast(x)^2 = \sum_{t \in \mathbb{R}} \mu(\{t\})^2.$$

Consequently, $\Sigma^\ast(x) = 0$ if and only if $x$ has no atoms.

**Proof.** Recall that in the one-variable case

$$\liminf_{t \to 0} t\Phi^\ast(x + \sqrt{ts}) = \sum_{t \in \mathbb{R}} \mu(\{t\})^2.$$
Thus Proposition 2.11 implies
\[ \sum_{t \in \mathbb{R}} \mu(\{t\})^2 \leq \Sigma^*(x)^2. \]

To see the reverse inequality, consider for \( \epsilon > 0 \) the function
\[ g_\epsilon(t) := 2 \int_{\mathbb{R}} \frac{(t - s)}{(t - s)^2 + \epsilon^2} \, d\mu(s). \]

Observe that \( |g_\epsilon(t)| \leq \frac{2}{\epsilon}(|t| + \tau(|x|)) \in L^2(\mu) \). In particular, for any polynomial \( p \) we have
\[
\int g_\epsilon(t)p(t) \, d\mu(t) = 2 \int_{\mathbb{R}} \frac{(t - s)p(t)}{(t - s)^2 + \epsilon^2} \, d\mu(s) \, d\mu(t)
= \int_{\mathbb{R}} \frac{(t - s)[p(t) - p(s)]}{(t - s)^2 + \epsilon^2} \, d\mu(s) \, d\mu(t)
= \int_{\mathbb{R}} \frac{(t - s)^2}{(t - s)^2 + \epsilon^2} \frac{p(t) - p(s)}{t - s} \, d\mu(s) \, d\mu(t).
\]

That is, \( A_\epsilon(t, s) := \frac{(t-s)^2}{(t-s)^2 + \epsilon^2} \) is a free Stein kernel for \( x \) relative to \( g_\epsilon \). So we compute for \( \delta > 0 \)
\[
\Sigma^*(x)^2 \leq \|A_\epsilon - 1\|^2_{L^2(\mu)} = \iint |A_\epsilon(t, s) - 1|^2 \, d\mu(t) \, d\mu(s)
\leq \iint_{|t-s| \geq \delta} \epsilon^4 \, d\mu(t) \, d\mu(s) + \iint_{|t-s| < \delta} 1 \, d\mu(t) \, d\mu(s)
\leq \frac{\epsilon^4}{\delta^4} + (\mu \otimes \mu)(\{(t, s) \in \mathbb{R}^2 : |t - s| < \delta\}).
\]

Letting first \( \epsilon \) tend to zero and then \( \delta \), we obtain the other inequality. \( \square \)

The next result concerns Atiyah’s \( L^2 \)-Betti numbers for discrete groups (cf. [Ati76, CG86]). Also see [Lüc02, Chapter 1] for the definition considered here, and [MS05] for the connection to free entropy dimension.

**Proposition 2.15.** Let \( \Gamma \) be a discrete group and let \( x_1, \ldots, x_n \in \mathbb{C}[\Gamma] \)\,a.e. generate the group algebra. Then
\[ n - \Sigma^*(x_1, \ldots, x_n)^2 = \beta_1^{(2)}(\Gamma) - \beta_0^{(2)}(\Gamma) + 1, \]
where \( \beta_0^{(2)}(\Gamma) \) and \( \beta_1^{(2)}(\Gamma) \) are the \( L^2 \)-Betti numbers of \( \Gamma \).

**Proof.** Recall that by [MS05, Theorem 4.1]
\[ \delta^*(x_1, \ldots, x_n) = \delta^*(x_1, \ldots, x_n) = \beta_1^{(2)}(\Gamma) - \beta_0^{(2)}(\Gamma) + 1. \]

So by Corollary 2.13, it suffices to show
\[ n - \Sigma^*(x_1, \ldots, x_n)^2 \geq \beta_1^{(2)}(\Gamma) - \beta_0^{(2)}(\Gamma) + 1. \]

Recall the following space from [Shl06, Section 2]: \( H_1 = \overline{H_1^{HS}} \) where
\[ H_1^{HS} := \text{span}\{(T_1, \ldots, T_n) \in \text{HS}(L^2(M))^n : \exists Y = Y^* \text{ unbounded, densely defined with}
\quad 1 \in \text{dom}(Y), \quad [Y, X_j] = T_j \text{ for each } j = 1, \ldots, n\}. \]

We can identify \( H_1 \) with a closed subspace in \( L^2(M \otimes M^o)^n \) using the identification
\[ L^2(M \otimes M^o) \cong \text{HS}(L^2(M)) \]
\[ a \otimes b^o \mapsto aP_1 b, \]
where \( P_1 \) is the rank one projection onto \( 1 \in L^2(M) \). By [Shl06, Theorem 1], for every \( T := (T_1, \ldots, T_n) \in H_1^{HS} \)
we have \( 1 \otimes 1 \in \text{dom}(\partial_T) \) where \( \partial_T : \mathbb{C}(X) \to L^2(M \otimes M^o) \) is the derivation defined by
\[ \partial_T(p) = \sum_{j=1}^n \partial_j(p) \# T_j \quad p \in \mathbb{C}(X). \]
Observe that if \( J = J_{\otimes \tau_0} \) is the Tomita conjugation operator on \( L^2(M \otimes M^0) \), then for \( p \in \mathbb{C} \langle X \rangle \) we have
\[
\langle 1 \otimes 1, \partial_T(p) \rangle_2 = \sum_{j=1}^n \langle 1 \otimes 1, \partial_j(p)J_T \rangle_2 = \sum_{j=1}^n \langle J_T, \partial_j(p) \rangle_2 = \langle J_T, \partial(p) \rangle_2.
\]
Consequently, \( 1 \otimes 1 \in \text{dom}(\partial_T^*) \) if and only if \( JT \in \text{dom}(\partial^*) \). It follows that \( JH_1 \subset \text{dom}(\partial^*) \) and so
\[
\dim_{M \otimes M^0}(\text{dom}(\partial^*)) \geq \dim_{M \otimes M^0}(H_1),
\]
where the latter dimension is as a right \( M \otimes M^0 \)-module. In the proof of [Shl06, Corollary 4] it was shown that the latter dimension is \( \beta_1^{(2)}(\Gamma) - \beta_0^{(2)}(\Gamma) + 1 \), and so Theorem 2.5 completes the proof. \( \square \)

### 2.4 Regularity hierarchy.
Let us relate the condition \( \Sigma^*(X) = 0 \) to other well-studied regularity conditions. We have the following picture:

\[
\Phi^*(X) < -\infty \quad \Rightarrow \quad \chi^*(X) > -\infty \quad \Rightarrow \quad \Sigma^*(X) = 0 \quad \Rightarrow \quad \delta^*(X) = n \]

The top two arrows are of course well-known results: the first is [Voi98, Proposition 7.9] while the second follows from [Vol98, Proposition 7.5] and the definition of \( \delta^* \) in [CS05, Section 4.1.1]. The bottom two arrows follow from Theorem 2.3 and Corollary 2.13, respectively. Thus it is natural to ask what the relationship is between having finite non-microstates free entropy and having zero free Stein information. In the case \( n = 1 \), we see that the former implies the latter by Proposition 2.14. We therefore suspect this to be the case for \( n \geq 2 \), but are so far unable to prove this.

In order to begin analyzing the relationship between these two conditions, consider the following quantity:
\[
\alpha := \limsup_{R \to \infty} \frac{\ln \Sigma_R^*(X)}{\ln R} \in [-\infty, 0].
\]
That is, \( \alpha \) compares how quickly \( \Sigma_R^*(X) \) decays as \( R \) grows. Note that if \( \Sigma^*(X) \neq 0 \) we have \( \alpha = 0 \); however, it may be that \( \alpha = 0 \) even when \( \Sigma^*(X) = 0 \). Indeed, consider the Example B.1 below.

**Proposition 2.16.** With \( \alpha \) as above, if \( \alpha < 0 \) then \( \chi^*(X) > -\infty \).

**Proof.** Let \( \alpha < \beta < 0 \). Then there exists \( R_0 > 0 \) such that for all \( R \geq R_0 \) we have
\[
\Sigma_R^*(X) \leq R^\beta.
\]
Let \( \gamma \in (0, 1) \). Then substituting \( R = \frac{1}{t^{\gamma/2}} \) we have
\[
\Sigma_{1/t^{\gamma/2}}^*(X) \leq t^{-\gamma \beta/2} \quad \forall t < t_0 := \frac{1}{R_0^{2/\gamma}}.
\]
Using Equation (3) we therefore have
\[
\Phi^*(X + \sqrt{t}S) \leq \frac{1}{t} \left( t^{-\gamma \beta/2} + t^{(1-\gamma)/2} \right)^2 = \left( t^{(-\gamma \beta-1)/2} + t^{-\gamma/2} \right)^2 \quad \forall t < t_0.
\]
Since \( (-\gamma \beta - 1)/2 > -1/2 \) and \( -\gamma/2 > -1/2 \) we have that the above quantity is integrable on \([0, t_0]\). \( \square \)

### Appendix A.
In this appendix, we will demonstrate that the Mai kernel \( B_{\Xi} \), given in Proposition 1.1, satisfies
\[
\|B_{\Xi} - 1\|_{HS} = \Sigma^*(X \mid \Xi)
\]
if and only if \( \Xi = 0 \). We emphasize that any free Stein kernel attaining the free Stein discrepancy of \( X \) is necessarily contained in the closure of the range of \( \mathcal{F} \).

Let \( d : L^2(M) \to L^2(M \otimes M^0) \) be the derivation given by commutation against \( 1 \otimes 1 \): \( \zeta \mapsto \zeta \otimes 1 - 1 \otimes \zeta \). Given \( Z = (\zeta_1, \ldots, \zeta_n) \in L^2(M)^n \), let \( D : L^2(M)^n \to L^2(M \otimes M^0)^n \) be given by applying \( d \) to each coordinate: \( D(Z) = (d(\zeta_1), \ldots, d(\zeta_n)) \).
Lemma A.1. Suppose that \( A \in \partial \mathbb{C} \langle X \rangle \subseteq L^2(M \otimes M^\circ)^n \). If \((p_k)_{k \in \mathbb{N}}\) is a sequence in \( \mathbb{C} \langle X \rangle \) so that
\[
A = \lim_{k \to \infty} \partial p_k,
\]
then
\[
A \cdot D(X) = \lim_{k \to \infty} (p_k \otimes 1 - 1 \otimes p_k).
\]

Proof. Observe that \( D(X) \in (M \otimes M^\circ)^n \), so that it has a bounded right action on \( L^2(M \otimes M)^n \). Thus the equation follows from a straightforward computation:
\[
A \cdot D(X) = \lim_{k \to \infty} (\partial p_k) \cdot D(X) = \lim_{k \to \infty} \sum_{j=1}^{n} \partial_j p_k \#(x_j \otimes 1 - 1 \otimes x_j) = \lim_{k \to \infty} p_k \otimes 1 - 1 \otimes p_k. \quad \square
\]

The lemma applies, in particular, to the rows of any free Stein kernel that attains the free Stein discrepancy of \( X \).

Proposition A.2. Suppose \( \Xi = (\xi_1, \ldots, \xi_n) \in L^2(M)^n \otimes \mathbb{C}^n \), and let \( B_{\Xi} \) be as in Proposition 1.1:
\[
B_{\Xi} := \left[ \frac{1}{2} (\xi_i \otimes 1 - 1 \otimes \xi_i) \#(x_j \otimes 1 - 1 \otimes x_j) \right]_{i,j=1}^{n} \in M_n(L^2(M \otimes M^\circ)).
\]
If \( \|B_{\Xi} - 1\|_{HS} = \Sigma^*(X \mid \Xi) \), then \( \Xi = 0 \).

Proof. First note that it suffices to assume that \( \tau(x_1) = \cdots = \tau(x_n) = 0 \). Indeed, let
\[
\hat{X} = (x_1, \ldots, x_n) = (x_1 - \tau(x_1), \ldots, x_n - \tau(x_n)).
\]
Then clearly \( \mathbb{C} \langle \hat{X} \rangle = \mathbb{C} \langle X \rangle \) and consequently \( \Xi \in L^2(W^*(\hat{X}))^n \). Moreover, \( B_{\Xi} \) is unchanged when replacing \( X \) with \( \hat{X} \). Now for any
\[
A \in M_n(L^2(W^*(X) \otimes W^*(X)^\circ)) = M_n(L^2(W^*(\hat{X}) \otimes W^*(\hat{X})^\circ)),
\]
if \( A \) is a free Stein kernel for \( X \) relative to \( \Xi \), then by the chain rule it is also a free Stein kernel for \( \hat{X} \) relative to \( \Xi \), and vice versa. Hence \( \Sigma^*(X \mid \Xi) = \Sigma^*(\hat{X} \mid \Xi) \) and so, replacing \( X \) with \( \hat{X} \) if necessary, we may assume \( \tau(x_1) = \cdots = \tau(x_n) = 0 \).

Note that the \( i \)-th row of \( B_{\Xi} \) is given by \( \frac{1}{2} d(\xi_i) \# D(X) =: r_i \), and from the assumption that \( \Sigma^*(X \mid \Xi) = \|B_{\Xi} - 1\|_{HS} \) we have that \( r_i \in \partial \mathbb{C} \langle X \rangle \subseteq L^2(M \otimes M^\circ)^n \). Now, pick \((p_k)_{k \in \mathbb{N}}\) in \( \mathbb{C} \langle X \rangle \) so that \( \partial p_k \to r_i \); since \( \mathbb{C}1 \in \ker \partial \), we may assume \( \tau(p_k) = 0 \), replacing \( p_k \) by \( p_k - \tau(p_k) \) if needed. Then from Lemma A.1, we have
\[
(1 \otimes \tau)(r_i \cdot D(X)) = \lim_{k \to \infty} p_k \quad \text{and} \quad (\tau \otimes 1)(r_i \cdot D(X)) = - \lim_{k \to \infty} p_k.
\]
We compute
\[
\sum_{j=1}^{n} \xi_i x_j^2 \otimes 1 - 2\xi_i x_j \otimes x_j + \xi_i \otimes x_j^2 - x_j^2 \otimes \xi_i + 2x_j \otimes x_j \xi_i - 1 \otimes x_j^2 \xi_i
\]
\[
= 2r_i \cdot D(X)
\]
\[
= 2 \lim_{k \to \infty} p_k \otimes 1 - 1 \otimes p_k
\]
\[
= [(1 \otimes \tau)(2r_i \cdot D(X))] \otimes 1 + 1 \otimes [(\tau \otimes 1)(2r_i \cdot D(X))]
\]
\[
= \left( \sum_{j=1}^{n} \xi_i (x_j^2 + \tau(x_j^2)) + 2x_j \tau(x_j \xi_i) - \tau(x_j^2 \xi_i) \right) \otimes 1 + 1 \otimes \left( \sum_{j=1}^{n} \tau(\xi_i x_j^2) - 2\tau(\xi_i x_j)x_j - (\tau(x_j^2) + x_j^2)\xi_i \right)
\]
\[
= \sum_{j=1}^{n} \xi_i x_j^2 \otimes 1 + \tau(x_j^2) d(\xi_i) + 2\tau(x_j \xi_i) d(x_j) - 1 \otimes x_j^2 \xi_i.
\]
Subtracting common terms on each side, we find
\[
\sum_{j=1}^{n} \xi_i [x_j^2 - \tau(x_j^2)] + 2x_j \otimes [x_j \xi_i - \tau(x_j \xi_i)] - 2[\xi_i x_j - \tau(\xi_i x_j)] \otimes x_j - [x_j^2 - \tau(x_j^2)] \otimes \xi_i = 0. \quad (4)
\]
As $X$ is algebraically free, we may find polynomials $p$ and $q$ such that $\langle x_1^2, p \rangle = 1$ while $p$ is orthogonal to all other monomials of degree at most two, and $\langle x_1, q \rangle = 1$ while $q$ is orthogonal to all other monomials of degree at most three. Applying the map $1 \otimes \langle \cdot \rangle_2$ to the above equality yields

$$\xi_i + \sum_{j=1}^n 2x_j \langle x_j \xi_i, p \rangle - [x_j^2 - \tau(x_j^2)] \langle \xi_i, p \rangle = 0,$$

whence $\xi_i$ is a polynomial in $X$ of degree at most two. Now, applying $1 \otimes \langle \cdot \rangle_2$ to Equation 4 and using the fact that $x_j \xi_i$ is a polynomial of degree at most three, we find

$$2x_1 \langle x_1 \xi_i, q \rangle - 2(\xi_1 x_1 - \tau(\xi_1 x_1)) - \langle \xi_i, q \rangle \sum_{j=1}^n (x_j^2 - \tau(x_j^2)) = 0.$$

From this it follows that $\xi_1 x_1$ is a linear combination of $1, x_1, x_1^2, x_1^3, \ldots, x_1^2$. But then $\xi_i$ must be a linear combination of $1$ and $x_1$; say $\xi_i = s + tx_1$. Looking at the coefficient of $x_1^2$ in the above equation, we find that $-2t - \langle \xi_i, q \rangle = 0$; since $\langle \xi_i, q \rangle = t$, we have $t = 0$, whence $\xi_i \in \mathbb{C}$. As $\xi_i \in L^2(M) \otimes \mathbb{C}$, $\xi_i = 0$. \hfill \Box

Appendix B.

In this appendix we consider a few informative examples. The first example demonstrates the fact that full free entropy dimension is strictly weaker than finite free entropy, by explicitly constructing a probability measure with no atoms and infinite logarithmic energy. While this result is already known, we are not aware of an explicit example in the literature. The example was concocted to demonstrate explicitly that $\alpha = 0$ does not imply $\Sigma^*(x) > 0$.

**Example B.1.** Let $I_n \subset [0, 1]$ be a disjoint sequence of intervals such that the Lebesgue measure $\lambda(I_n) < e^{-12^n}$. Define a function $f$ as follows:

$$f : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0},$$

$$t \mapsto \sum_{n=1}^{\infty} \frac{1}{2^n \lambda(I_n)} \chi_{I_n}(t).$$

By construction $f$ is non-negative, integrable, and has mass 1, so it is a probability density; let $\mu$ be the measure with density given by $f$. We claim that the (negative) logarithmic energy of $\mu$ is infinite. Indeed,

$$\begin{align*}
\int \int_{\mathbb{R}^2} \log |x - y| \, d\mu(x) \, d\mu(y) &\leq \sum_{n=1}^{\infty} \int \int_{I_n^2} \log |x - y| \, d\mu(x) \, d\mu(y) \\
&\leq \sum_{n=1}^{\infty} \log(e^{-12^n}) 4^{-n} \\
&= -\infty.
\end{align*}$$

Now, since supp($\mu$) is bounded and has a diffuse component, there exists a bounded, self-adjoint, algebraically free operator $x$ with spectral measure $\mu$. It follows from Proposition 2.16 that $\alpha = 0$, and by Proposition 2.14 we have $\Sigma^*(x) = 0$. \hfill \blacksquare

As a decreasing convex function, if $R \mapsto \Sigma^*_R(X)$ ever plateaus it remains constant forever. This happens, for example, when conjugate variables actually exist: $\Sigma^*_R(X) = 0$ for $R \geq \sqrt{\Phi^*_R(X)}$. One may wonder, then, if this behaviour can occur when $\Sigma^*(X) > 0$; we provide a family of examples to show that it can.

**Example B.2.** Let $\mu = \frac{1}{2} \sigma + \frac{1}{2} \delta_a$ where $d\sigma = \chi_{[-2,2]}(t) \frac{1}{\sqrt{4 - t^2}} \, dt$ is the semicircle law. Then we will show that if $|a| > 2$ and $x$ has spectral measure $\mu$, then there is $R > 0$ so that $\Sigma^*_R(x) = \Sigma^*_R(x) = \frac{\pi}{2}$.

As in the proof of Proposition 2.14, define the functions

$$g_s(t) := 2 \int_{\mathbb{R}} \frac{t - s}{(t - s)^2 + \epsilon^2} \, d\mu(s).$$
As before, we have a free Stein kernel for $x$ relative to $g$, given by

$$A_\epsilon(t, s) := \frac{(t-s)^2}{(t-s)^2 + \epsilon^2}$$

Notice that as $\epsilon \to 0$, $A_\epsilon(t, s) \to \chi_{t\neq s} = A(s, t)$ which has $\|A - 1\|_{L^2(\mu \times \mu)}^2 = \mu(\{a\})^2$; so it suffices to show that $A$ is a free Stein kernel.

Here we will use the fact that $a \notin \text{supp}(\sigma)$ to conclude that $g_\epsilon$ converges in $L^2(\mu)$ as $\epsilon \to 0$. This can be checked by, for example, recognizing that $g_\epsilon$ converges in both $L^2(\delta_a)$ and $L^2(\sigma)$: in the former space,

$$g_\epsilon \to \int \frac{1}{a - s} d\sigma(s),$$

which converges since $a$ is outside the support of $\sigma$; in the latter,

$$g_\epsilon(t) \to Kt + \frac{1}{t - a},$$

where we have used the fact that the Hilbert transform of the semicircle distribution is $t$ while $a$ is, once again, outside of the support of $\sigma$. Let $g = \lim_{\epsilon \to 0} g_\epsilon$ with the limit in $L^2(\mu)$.

We claim that $A$, above, is a free Stein kernel for $x$ relative to $g$, whereupon $\Sigma^*_a(x) = \mu(\{a\}) = \frac{1}{2}$ for $R \geq \|g\|_{L^2(\mu)}$. (However, note that $\|g\|_{L^2(\mu)}$ diverges as $|a| \to 2$.) To see that, notice that $\partial^* \mu$ is closed since $\partial$ is densely defined. Since $A_\epsilon \in \text{dom}(\partial^*)$ with $\partial^*(A_\epsilon) = g_\epsilon$, (by virtue of being a free Stein kernel) we therefore have $A \in \text{dom}(\partial^*)$ with $\partial^*(A) = g$. That is, $A$ is a free Stein kernel for $x$ relative to $g$. \[\square\]

References

[Ati76] Michael F. Atiyah, Elliptic operators, discrete groups and von Neumann algebras, 43–72. Astérisque, No. 32–33. MR 0420729

[BM18] Marwa Banna and Tobias Mai, Hölder Continuity of Cumulative Distribution Functions for Noncommutative Polynomials under Finite Free Fisher Information, ArXiv e-prints (2018).

[CFM18] Guillaume Cébron, Max Fathi, and Tobias Mai, A note on existence of free Stein kernels, arXiv e-prints (2018), arXiv:1811.02926.

[CG86] Jeff Cheeger and Mikhail Gromov, $L_2$-cohomology and group cohomology, Topology 25 (1986), no. 2, 189–215. MR 837621

[CS05] Alain Connes and Dimitri Shlyakhtenko, $L^2$-homology for von Neumann algebras, J. Reine Angew. Math. 586 (2005), 125–168. MR 2180603 (2007b:46104)

[CS16] Ian Charlesworth and Dimitri Shlyakhtenko, Free entropy dimension and regularity of non-commutative polynomials, J. Funct. Anal. 271 (2016), no. 8, 2274–2292. MR 3539353

[Dab10] Yoann Dabrowski, A note about proving non-$F$ under a finite non-microstates free Fisher information assumption, J. Funct. Anal. 258 (2010), no. 11, 3662–3674. MR 2606808 (2011d:46135)

[FN17] Max Fathi and Brent Nelson, Free Stein kernels and an improvement of the free logarithmic Sobolev inequality, Adv. Math. 317 (2017), 193–223. MR 3682667

[Ge98] Liming Ge, Applications of free entropy to finite von Neumann algebras. II, Ann. of Math. (2) 147 (1998), no. 1, 143–157. MR 1609522

[Hay18] Ben Hayes, 1-Bounded entropy and regularity problems in von Neumann algebras, Int. Math. Res. Not. IMRN (2018), no. 1, 57–137. MR 3801429

[HS07] Don Hadwin and Junhao Shen, Free orbit dimension of finite von Neumann algebras, J. Funct. Anal. 249 (2007), no. 1, 75–91. MR 2338855

[LNP15] Michel Ledoux, Ivan Nourdin, and Giovanni Peccati, Stein’s method, logarithmic Sobolev and transport inequalities, Geom. Funct. Anal. 25 (2015), no. 1, 256–306. MR 3320893

[L üc02] Wolfgang 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 [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 44, Springer-Verlag, Berlin, 2002. MR 1926649

[MS05] Igor Mineyev and Dimitri Shlyakhtenko, Non-microstates free entropy dimension for groups, Geom. Funct. Anal. 15 (2005), no. 2, 476–490. MR 2159076

[MSW17] Tobias Mai, Roland Speicher, and Moritz Weber, Absence of algebraic relations and of zero divisors under the assumption of full non-microstates free entropy dimension, Adv. Math. 304 (2017), 1080–1107. MR 3558227

[Shl04] Dimitri Shlyakhtenko, Some estimates for non-microstates free entropy dimension with applications to $q$-semicircular families, Int. Math. Res. Not. (2004), no. 51, 2757–2772. MR 2136068

[Shl06] Remarks on free entropy dimension, Operator Algebras: The Abel Symposium 2004, Abel Symp., vol. 1, Springer, Berlin, 2006, pp. 249–257. MR 2265052
[Voi93] Dan-Virgil 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. MR 1228526

[Voi94] ———, *The analogues of entropy and of Fisher’s information measure in free probability theory. II*, Invent. Math. **118** (1994), no. 3, 411–440. MR 1296352 (96a:46117)

[Voi96] ———, *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. MR 1371236

[Voi98] ———, *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. MR 1618636 (99d:46087)