# Shannon's monotonicity problem for free and classical entropy

## After D. Shlyakhtenko and H. Schultz

Brent Nelson

UCLA

October 9, 2013

Recall that for a real valued random variable $X$ with density $p$, its entropy is defined as

$$H(X) = -\int_{\mathbb{R}} p(x) \log p(x) \, dx.$$

Amongst random variables with $E(X) = 0$ and $E(X^2) = 1$, entropy is maximized by the standard Gaussian random variables $G$ with variance 1.

Recall that for a real valued random variable $X$ with density $p$, its entropy is defined as

$$H(X) = -\int_{\mathbb{R}} p(x) \log p(x) \, dx.$$

Amongst random variables with $E(X) = 0$ and $E(X^2) = 1$, entropy is maximized by the standard Gaussian random variables $G$ with variance 1. Given a sequence $X_1, X_2, \ldots$ of independent, identically distributed random variables with $E(X_n) = 0$ and $\mathrm{Var}(X_n) = 1$ then the central limit states that their central limit sums

$$Z_N = \frac{X_1 + \cdots + X_N}{\sqrt{N}},$$

converge in law to $G$. Moreover, the entropy of this sequence is nondecreasing; a result due to Artstein, Ball, Barthe, and Naor [1].

**Setup:**

- Let $\mathcal{H}$ be a vector space and denote by $\mathcal{B}(\mathcal{H})$ the bounded operators on $\mathcal{H}$.

**Setup:**

- Let $\mathcal{H}$ be a vector space and denote by $\mathcal{B}(\mathcal{H})$ the bounded operators on $\mathcal{H}$.
- We fix a unital $*$-subalgebra $A \subset \mathcal{B}(\mathcal{H})$, and a linear functional $\tau \colon A \to \mathbb{C}$.

**Setup:**

- Let $\mathcal{H}$ be a vector space and denote by $\mathcal{B}(\mathcal{H})$ the bounded operators on $\mathcal{H}$.
- We fix a unital $*$-subalgebra $A \subset \mathcal{B}(\mathcal{H})$, and a linear functional $\tau \colon A \to \mathbb{C}$.
- We require that $\tau$ is

**Setup:**

- Let $\mathcal{H}$ be a vector space and denote by $\mathcal{B}(\mathcal{H})$ the bounded operators on $\mathcal{H}$.
- We fix a unital $*$-subalgebra $A \subset \mathcal{B}(\mathcal{H})$, and a linear functional $\tau \colon A \to \mathbb{C}$.
- We require that $\tau$ is
  1. *positive*: $\tau(X^*X) \geq 0$ for all $X \in A$;

**Setup:**

- Let $\mathcal{H}$ be a vector space and denote by $\mathcal{B}(\mathcal{H})$ the bounded operators on $\mathcal{H}$.
- We fix a unital $*$-subalgebra $A \subset \mathcal{B}(\mathcal{H})$, and a linear functional $\tau \colon A \to \mathbb{C}$.
- We require that $\tau$ is
    1. *positive*: $\tau(X^*X) \geq 0$ for all $X \in A$;
    2. *tracial*: $\tau(XY) = \tau(YX)$ for all $X, Y \in A$; and

**Setup:**

- Let $\mathcal{H}$ be a vector space and denote by $\mathcal{B}(\mathcal{H})$ the bounded operators on $\mathcal{H}$.
- We fix a unital $*$-subalgebra $A \subset \mathcal{B}(\mathcal{H})$, and a linear functional $\tau \colon A \to \mathbb{C}$.
- We require that $\tau$ is
  1. *positive*: $\tau(X^*X) \geq 0$ for all $X \in A$;
  2. *tracial*: $\tau(XY) = \tau(YX)$ for all $X, Y \in A$; and
  3. $\tau(1) = 1$.

**Setup:**

- Let $\mathcal{H}$ be a vector space and denote by $\mathcal{B}(\mathcal{H})$ the bounded operators on $\mathcal{H}$.
- We fix a unital $*$-subalgebra $A \subset \mathcal{B}(\mathcal{H})$, and a linear functional $\tau\colon A \to \mathbb{C}$.
- We require that $\tau$ is
  1. *positive*: $\tau(X^*X) \geq 0$ for all $X \in A$;
  2. *tracial*: $\tau(XY) = \tau(YX)$ for all $X, Y \in A$; and
  3. $\tau(1) = 1$.
- Then $(A, \tau)$ is a *non-commutative probability space*, and any element $X \in A$ is a *non-commutative random variable*.

**Setup:**

- Let $\mathcal{H}$ be a vector space and denote by $\mathcal{B}(\mathcal{H})$ the bounded operators on $\mathcal{H}$.
- We fix a unital $*$-subalgebra $A \subset \mathcal{B}(\mathcal{H})$, and a linear functional $\tau \colon A \to \mathbb{C}$.
- We require that $\tau$ is
  1. *positive*: $\tau(X^*X) \geq 0$ for all $X \in A$;
  2. *tracial*: $\tau(XY) = \tau(YX)$ for all $X, Y \in A$; and
  3. $\tau(1) = 1$.
- Then $(A, \tau)$ is a *non-commutative probability space*, and any element $X \in A$ is a *non-commutative random variable*.
- $\tau(X)$ is the expectation, or first moment, and in general the *law* of $X$ refers to its moments $\{\tau(X^n) \colon n \in \mathbb{N}\}$.

**Setup:**

- Let $\mathcal{H}$ be a vector space and denote by $\mathcal{B}(\mathcal{H})$ the bounded operators on $\mathcal{H}$.
- We fix a unital $*$-subalgebra $A \subset \mathcal{B}(\mathcal{H})$, and a linear functional $\tau\colon A \to \mathbb{C}$.
- We require that $\tau$ is
  1. *positive*: $\tau(X^*X) \geq 0$ for all $X \in A$;
  2. *tracial*: $\tau(XY) = \tau(YX)$ for all $X, Y \in A$; and
  3. $\tau(1) = 1$.
- Then $(A, \tau)$ is a *non-commutative probability space*, and any element $X \in A$ is a *non-commutative random variable*.
- $\tau(X)$ is the expectation, or first moment, and in general the *law* of $X$ refers to its moments $\{\tau(X^n)\colon n \in \mathbb{N}\}$.
- Can think of the law of $X$ as a linear functional on polynomials $\mu_X\colon \mathbb{C}[t] \to \mathbb{C}$ so that $\mu_X(p(t)) = \tau(p(X))$.

**A commutative example:**

- Let $\mathcal{H} = L^2([0,1], m)$ with $m$ the Lebesgue measure.

**A commutative example:**

- Let $\mathcal{H} = L^2([0,1], m)$ with $m$ the Lebesgue measure.
- Let $A = L^\infty([0,1], m) \subset \mathcal{B}(\mathcal{H})$, and $\tau(f) = \int_{[0,1]} f \, dm$ for $f \in A$.

**A commutative example:**

- Let $\mathcal{H} = L^2([0,1], m)$ with $m$ the Lebesgue measure.
- Let $A = L^\infty([0,1], m) \subset \mathcal{B}(\mathcal{H})$, and $\tau(f) = \int_{[0,1]} f \, dm$ for $f \in A$.
- This example is more important to the non-commutative case than it first seems.

If $X \in A$ is self-adjoint, then there exists a measure $\mu$ supported on the spectrum of $X$ so that

$$\tau(X^n) = \int_{\mathbb{R}} t^n \, d\mu_X(t),$$

in which case we may refer to $d\mu$ as the law of $X$.

If $X \in A$ is self-adjoint, then there exists a measure $\mu$ supported on the spectrum of $X$ so that

$$\tau(X^n) = \int_{\mathbb{R}} t^n \, d\mu_X(t),$$

in which case we may refer to $d\mu$ as the law of $X$. In fact, the spectral theorem states that given any Borel measurable function $f$ supported on the spectrum of $X$, there exists an operator $f(X) \in \mathcal{B}(\mathcal{H})$ so that

$$\tau(f(X)) = \int_{\mathbb{R}} f(t) \, d\mu_X(t).$$

Moreover, given two such functions $f$ and $g$, $f(X)g(X) = (f \cdot g)(X)$.

If $X \in A$ is self-adjoint, then there exists a measure $\mu$ supported on the spectrum of $X$ so that

$$\tau(X^n) = \int_{\mathbb{R}} t^n \, d\mu_X(t),$$

in which case we may refer to $d\mu$ as the law of $X$. In fact, the spectral theorem states that given any Borel measurable function $f$ supported on the spectrum of $X$, there exists an operator $f(X) \in \mathcal{B}(\mathcal{H})$ so that

$$\tau(f(X)) = \int_{\mathbb{R}} f(t) \, d\mu_X(t).$$

Moreover, given two such functions $f$ and $g$, $f(X)g(X) = (f \cdot g)(X)$. For the remainder of the talk we assume all non-commutative random variables are self-adjoint.

Given several non-commutative random variables $X_1, \ldots, X_k$, their *joint law* can be thought of as a linear functional on non-commutative polynomials $\mu_{X_1,\ldots,X_k} : \mathbb{C}\langle t_1, \ldots, t_k \rangle \to \mathbb{C}$ such that $\mu_{X_1,\ldots,X_k}(p(t_1, \ldots, t_k)) = \tau(p(X_1, \ldots, X_k))$.

Given several non-commutative random variables $X_1, \ldots, X_k$, their *joint law* can be thought of as a linear functional on non-commutative polynomials $\mu_{X_1,\ldots,X_k} : \mathbb{C}\langle t_1, \ldots, t_k \rangle \to \mathbb{C}$ such that $\mu_{X_1,\ldots,X_k}(p(t_1, \ldots, t_k)) = \tau(p(X_1, \ldots, X_k))$.
There is no longer a single moment of each degree, and because of the non-commutativity we cannot identify the moment associated to $t_1 t_2 t_3$ to that of $t_1 t_3 t_2$, for example.

Given several non-commutative random variables $X_1, \ldots, X_k$, their *joint law* can be thought of as a linear functional on non-commutative polynomials $\mu_{X_1,\ldots,X_k} : \mathbb{C}\langle t_1, \ldots, t_k \rangle \to \mathbb{C}$ such that $\mu_{X_1,\ldots,X_k}(p(t_1, \ldots, t_k)) = \tau(p(X_1, \ldots, X_k))$.

There is no longer a single moment of each degree, and because of the non-commutativity we cannot identify the moment associated to $t_1 t_2 t_3$ to that of $t_1 t_3 t_2$, for example.

Also, the law is no longer encoded by a measure.

**The Fock space example:**

- Let $\mathcal{H}$ be a Hilbert space.

**The Fock space example:**

- Let $\mathcal{H}$ be a Hilbert space.
- $\mathcal{H}^{\otimes k}$ represents the $k$-fold Hilbert space tensor product of $\mathcal{H}$. It is spanned by elements of the form $e_1 \otimes \cdots \otimes e_k$, with $e_1, \ldots, e_k \in \mathcal{H}$ and has an inner product defined by

$$\langle e_1 \otimes \cdots \otimes e_k, f_1 \otimes \cdots \otimes f_k \rangle = \langle e_1, f_1 \rangle \cdots \langle e_k, f_k \rangle .$$

**The Fock space example:**

- Let $\mathcal{H}$ be a Hilbert space.
- $\mathcal{H}^{\otimes k}$ represents the $k$-fold Hilbert space tensor product of $\mathcal{H}$. It is spanned by elements of the form $e_1 \otimes \cdots \otimes e_k$, with $e_1, \ldots, e_k \in \mathcal{H}$ and has an inner product defined by

$$\langle e_1 \otimes \cdots \otimes e_k, f_1 \otimes \cdots \otimes f_k \rangle = \langle e_1, f_1 \rangle \cdots \langle e_k, f_k \rangle .$$

- The *Fock space* is defined as $\mathcal{F}(\mathcal{H}) = \mathbb{C}\Omega \oplus \bigoplus_{k=1}^{\infty} \mathcal{H}^{\otimes k}$, where $\Omega$ is the vacuum vector (think "zero length tensor product"). Its inner product is the extension of the above where tensor products of different lengths are orthogonal.

- Fix a vector $e \in \mathcal{H}$ with $\|e\| = 1$. The *left creation operator* is defined by

$$l(e)\Omega = e$$
$$l(e)e_1 \otimes \cdots \otimes e_k = e \otimes e_1 \otimes \cdots \otimes e_k.$$

- Fix a vector $e \in \mathcal{H}$ with $\|e\| = 1$. The *left creation operator* is defined by

$$l(e)\Omega = e$$
$$l(e)e_1 \otimes \cdots \otimes e_k = e \otimes e_1 \otimes \cdots \otimes e_k.$$

- Its adjoint, the *left annihilation operator*, is defined by

$$l(e)^*\Omega = 0$$
$$l(e)^* e_1 \otimes \cdots \otimes e_k = \langle e, e_1 \rangle\, e_2 \otimes \cdots \otimes e_k.$$

- Fix a vector $e \in \mathcal{H}$ with $\|e\| = 1$. The *left creation operator* is defined by

$$l(e)\Omega = e$$
$$l(e)e_1 \otimes \cdots \otimes e_k = e \otimes e_1 \otimes \cdots \otimes e_k.$$

- Its adjoint, the *left annihilation operator*, is defined by

$$l(e)^*\Omega = 0$$
$$l(e)^* e_1 \otimes \cdots \otimes e_k = \langle e, e_1 \rangle \, e_2 \otimes \cdots \otimes e_k.$$

- Define $c(e) = l(e) + l(e)^* \in \mathcal{B}(\mathcal{F}(\mathcal{H}))$.

- Define a linear functional $\tau\colon \mathcal{B}(\mathcal{F}(\mathcal{H})) \to \mathbb{C}$ by $\tau(X) = \langle \Omega, X\Omega \rangle$, then $\tau$ is a positive trace.

- Define a linear functional $\tau\colon \mathcal{B}(\mathcal{F}(\mathcal{H})) \to \mathbb{C}$ by $\tau(X) = \langle \Omega, X\Omega \rangle$, then $\tau$ is a positive trace.
- Consider the law of $c(e)$ with respect to $\tau$:

$$\tau(c(e)^n) = \left\langle \Omega, \underbrace{(l(e) + l^*(e)) \cdots (l(e) + l(e)^*)}_{n} \Omega \right\rangle$$

- Define a linear functional $\tau\colon \mathcal{B}(\mathcal{F}(\mathcal{H})) \to \mathbb{C}$ by $\tau(X) = \langle \Omega, X\Omega \rangle$, then $\tau$ is a positive trace.
- Consider the law of $c(e)$ with respect to $\tau$:

$$\tau(c(e)^n) = \left\langle \Omega, \underbrace{(l(e) + l^*(e)) \cdots (l(e) + l(e)^*)}_{n} \Omega \right\rangle$$

- Then

$$\tau(c(e)^n) = \begin{cases} \frac{1}{k+1}\binom{2k}{k} & \text{if } n = 2k \\ 0 & \text{if } n = 2k+1 \end{cases},$$

where $C_k = \frac{1}{k+1}\binom{2k}{k}$ are the Catalan numbers.

- Recall the *semicircle distribution* is defined as $d\mu = \chi_{[-2,2]}(t)\frac{1}{2\pi}\sqrt{4-t^2}\ dt$.

- Recall the *semicircle distribution* is defined as $d\mu = \chi_{[-2,2]}(t)\frac{1}{2\pi}\sqrt{4-t^2}\ dt$.
- The moments of this distribution are precisely those of $c(e)$:

$$\int_{[-2,2]} t^n \frac{1}{2\pi}\sqrt{4-t^2}\ dt = \begin{cases} \frac{1}{k+1}\binom{2k}{k} & \text{if } n = 2k \\ 0 & \text{if } n = 2k+1 \end{cases}$$

- Recall the *semicircle distribution* is defined as $d\mu = \chi_{[-2,2]}(t)\frac{1}{2\pi}\sqrt{4-t^2}\ dt$.
- The moments of this distribution are precisely those of $c(e)$:

$$\int_{[-2,2]} t^n \frac{1}{2\pi}\sqrt{4-t^2}\ dt = \begin{cases} \frac{1}{k+1}\binom{2k}{k} & \text{if } n = 2k \\ 0 & \text{if } n = 2k+1 \end{cases}$$

- Thus we say that $c(e)$ has the *semicircle law* or is *semicircular*.

Recall that the classical notion of independence says that two random variables $X$ and $Y$ are independent if their joint density is the product of their individual densities: $p_{X,Y}(s,t) = p_X(s)p_Y(t)$.

Recall that the classical notion of independence says that two random variables $X$ and $Y$ are independent if their joint density is the product of their individual densities: $p_{X,Y}(s,t) = p_X(s)p_Y(t)$.
Consequently, any joint moment can be expressed as a product of their individual moments:

$$\iint_{\mathbb{R}^2} s^m t^n p_{X,Y}(s,t)\ ds\ dt = \int_{\mathbb{R}} s^m p_X(s)\ ds \int_{\mathbb{R}} t^n p_Y(t)\ dt.$$

Recall that the classical notion of independence says that two random variables $X$ and $Y$ are independent if their joint density is the product of their individual densities: $p_{X,Y}(s,t) = p_X(s)p_Y(t)$.
Consequently, any joint moment can be expressed as a product of their individual moments:

$$\iint_{\mathbb{R}^2} s^m t^n p_{X,Y}(s,t)\ ds\ dt = \int_{\mathbb{R}} s^m p_X(s)\ ds \int_{\mathbb{R}} t^n p_Y(t)\ dt.$$

In particular, if $W \in \mathbf{Alg}(1, X)$ and $Z \in \mathbf{Alg}(1, Y)$ are two random variables so that $E(W) = E(Z) = 0$, then $E(WZ) = 0$.

Recall that the classical notion of independence says that two random variables $X$ and $Y$ are independent if their joint density is the product of their individual densities: $p_{X,Y}(s,t) = p_X(s)p_Y(t)$.
Consequently, any joint moment can be expressed as a product of their individual moments:

$$\iint_{\mathbb{R}^2} s^m t^n p_{X,Y}(s,t)\ ds\ dt = \int_{\mathbb{R}} s^m p_X(s)\ ds \int_{\mathbb{R}} t^n p_Y(t)\ dt.$$

In particular, if $W \in \mathbf{Alg}(1,X)$ and $Z \in \mathbf{Alg}(1,Y)$ are two random variables so that $E(W) = E(Z) = 0$, then $E(WZ) = 0$.
Free independence captures this idea in the non-commutative case.

Let $F_1, F_2 \subset (A, \tau)$ be two families of non-commutative random variables. Then we say that these families are *freely independent* if

$$\tau(W_1 W_2 \cdots W_n) = 0$$

when $W_j \in \textbf{Alg}(1, F_{i(j)})$ are such that $\tau(W_j) = 0$ and $i(j) \neq i(j+1)$, with $j = 1, \ldots, n$ and $i(j) \in \{1, 2\}$.

Let $F_1, F_2 \subset (A, \tau)$ be two families of non-commutative random variables. Then we say that these families are *freely independent* if

$$\tau(W_1 W_2 \cdots W_n) = 0$$

when $W_j \in \mathbf{Alg}(1, F_{i(j)})$ are such that $\tau(W_j) = 0$ and $i(j) \neq i(j+1)$, with $j = 1, \ldots, n$ and $i(j) \in \{1, 2\}$.
In order to manage the non-commutativity, we must consider these lengthier products.

Let $F_1, F_2 \subset (A, \tau)$ be two families of non-commutative random variables. Then we say that these families are *freely independent* if

$$\tau(W_1 W_2 \cdots W_n) = 0$$

when $W_j \in \mathbf{Alg}(1, F_{i(j)})$ are such that $\tau(W_j) = 0$ and $i(j) \neq i(j+1)$, with $j = 1, \ldots, n$ and $i(j) \in \{1, 2\}$.
In order to manage the non-commutativity, we must consider these lengthier products.
We say $W_j$ is *centered* when $\tau(W_j) = 0$.

Let $F_1, F_2 \subset (A, \tau)$ be two families of non-commutative random variables. Then we say that these families are *freely independent* if

$$\tau(W_1 W_2 \cdots W_n) = 0$$

when $W_j \in \mathbf{Alg}(1, F_{i(j)})$ are such that $\tau(W_j) = 0$ and $i(j) \neq i(j+1)$, with $j = 1, \ldots, n$ and $i(j) \in \{1, 2\}$.
In order to manage the non-commutativity, we must consider these lengthier products.
We say $W_j$ is *centered* when $\tau(W_j) = 0$.
More generally, $F_1, \ldots, F_k \subset (A, \tau)$ are freely independent if the above holds but now we simply take $i(j) \in \{1, \ldots, k\}$.

We say non-commutative random variables $X_1, \ldots, X_k$ are freely independent if the families $\{X_1\}, \ldots, \{X_k\}$ are freely independent.

We say non-commutative random variables $X_1, \ldots, X_k$ are freely independent if the families $\{X_1\}, \ldots, \{X_k\}$ are freely independent.
That is, if $p_1, \ldots, p_n \in \mathbb{C}[t]$ are polynomials, $i(1), \ldots, i(n) \in \{1, \ldots, k\}$ is a sequence with non-equal adjacent terms, and $\tau(p_j(X_{i(j)})) = 0$ for each $j$, then

$$\tau(p_1(X_{i(1)})p_2(X_{i(2)}) \cdots p_n(X_{i(n)})) = 0.$$

We say non-commutative random variables $X_1, \ldots, X_k$ are freely independent if the families $\{X_1\}, \ldots, \{X_k\}$ are freely independent.
That is, if $p_1, \ldots, p_n \in \mathbb{C}[t]$ are polynomials, $i(1), \ldots, i(n) \in \{1, \ldots, k\}$ is a sequence with non-equal adjacent terms, and $\tau(p_j(X_{i(j)})) = 0$ for each $j$, then

$$\tau(p_1(X_{i(1)})p_2(X_{i(2)}) \cdots p_n(X_{i(n)})) = 0.$$

For $i = 1, \ldots, k$, let $X_i = (X_i^{(1)}, \ldots, X_i^{(p)}) \in A^p$ be a $p$-tuple of non-commutative random variables. Then we say these $p$-tuples are freely independent if the families $\{X_1^{(1)}, \ldots, X_1^{(p)}\}, \ldots, \{X_k^{(1)}, \ldots, X_k^{(p)}\}$ are freely independent.

Suppose $X_1$ and $X_2$ are freely independent, then the moments of their joint law can be computed in terms of their individual moments (just as in the classical case).

Suppose $X_1$ and $X_2$ are freely independent, then the moments of their joint law can be computed in terms of their individual moments (just as in the classical case).
The trick is *centering*.

Suppose $X_1$ and $X_2$ are freely independent, then the moments of their joint law can be computed in terms of their individual moments (just as in the classical case).
The trick is *centering*.
For example, to compute $\tau(X_1X_2)$. First write $X_j = \mathring{X}_j + \tau(X_j)1$, where $\mathring{X}_j = X_j - \tau(X_j)1$. Then $\mathring{X}_j \in \mathbf{Alg}(1, X_j)$ and is centered. Thus we have

Suppose $X_1$ and $X_2$ are freely independent, then the moments of their joint law can be computed in terms of their individual moments (just as in the classical case).
The trick is *centering*.
For example, to compute $\tau(X_1X_2)$. First write $X_j = \mathring{X}_j + \tau(X_j)1$, where $\mathring{X}_j = X_j - \tau(X_j)1$. Then $\mathring{X}_j \in \mathbf{Alg}(1, X_j)$ and is centered. Thus we have

$$\begin{aligned}\tau(X_1X_2) &= \tau((\mathring{X}_1 + \tau(X_1)1)(\mathring{X}_2 + \tau(X_2)1)) \\ &= \tau(\mathring{X}_1\mathring{X}_2) + \tau(\mathring{X}_1)\tau(X_2) + \tau(X_1)\tau(\mathring{X}_2) + \tau(X_1)\tau(X_2) \\ &= \tau(X_1)\tau(X_2),\end{aligned}$$

Suppose $X_1$ and $X_2$ are freely independent, then the moments of their joint law can be computed in terms of their individual moments (just as in the classical case).

The trick is *centering*.

For example, to compute $\tau(X_1X_2)$. First write $X_j = \mathring{X}_j + \tau(X_j)1$, where $\mathring{X}_j = X_j - \tau(X_j)1$. Then $\mathring{X}_j \in \mathbf{Alg}(1, X_j)$ and is centered. Thus we have

$$\begin{aligned}\tau(X_1X_2) &= \tau((\mathring{X}_1 + \tau(X_1)1)(\mathring{X}_2 + \tau(X_2)1)) \\ &= \tau(\mathring{X}_1\mathring{X}_2) + \tau(\mathring{X}_1)\tau(X_2) + \tau(X_1)\tau(\mathring{X}_2) + \tau(X_1)\tau(X_2) \\ &= \tau(X_1)\tau(X_2),\end{aligned}$$

where the first term vanishes because of free independence and the other two vanish because the $\mathring{X}_j$ are centered.

**Other examples:**

**Other examples:**

$$\tau(X_1X_2^2X_1)=\tau(X_1^2)\tau(X_2^2)$$

$$\tau(X_1X_2X_1X_2)=\tau(X_1^2)\tau(X_2)^2+\tau(X_1)^2\tau(X_2^2)-\tau(X_1)^2\tau(X_2)^2$$

**Other examples:**

$$\tau(X_1X_2^2X_1) = \tau(X_1^2)\tau(X_2^2)$$

$$\tau(X_1X_2X_1X_2) = \tau(X_1^2)\tau(X_2)^2 + \tau(X_1)^2\tau(X_2^2) - \tau(X_1)^2\tau(X_2)^2$$

Given the above two examples, if $X_1X_2 = X_2X_1$ then it would follow that

$$\tau((X_1 - \tau(X_1)1)^2)\tau((X_2 - \tau(X_2)1)^2) = 0,$$

i.e. the variance of $X_1$ or $X_2$ must vanish.

Returning to the Fock space example, let $e, f \in \mathcal{H}$ be orthogonal unit vectors.

Returning to the Fock space example, let $e, f \in \mathcal{H}$ be orthogonal unit vectors.
Then $c(e)$ and $c(f)$ are freely independent.

Returning to the Fock space example, let $e, f \in \mathcal{H}$ be orthogonal unit vectors.
Then $c(e)$ and $c(f)$ are freely independent.
In fact, it is true that the families $\{l(e), l(e)^*\}$ and $\{l(f), l(f)^*\}$ are freely independent:

Returning to the Fock space example, let $e, f \in \mathcal{H}$ be orthogonal unit vectors.
Then $c(e)$ and $c(f)$ are freely independent.
In fact, it is true that the families $\{l(e), l(e)^*\}$ and $\{l(f), l(f)^*\}$ are freely independent:

- Let $l_1 = l(e)$ and $l_2 = l(f)$.

Returning to the Fock space example, let $e, f \in \mathcal{H}$ be orthogonal unit vectors.
Then $c(e)$ and $c(f)$ are freely independent.
In fact, it is true that the families $\{l(e), l(e)^*\}$ and $\{l(f), l(f)^*\}$ are freely independent:

- Let $l_1 = l(e)$ and $l_2 = l(f)$.
- Since $l_i^* l_i = 1$, any monomial in $l_i, l_i^*$ reduces to $l_i^n l_i^{*m}$, with $n + m > 0$.

Returning to the Fock space example, let $e, f \in \mathcal{H}$ be orthogonal unit vectors.
Then $c(e)$ and $c(f)$ are freely independent.
In fact, it is true that the families $\{l(e), l(e)^*\}$ and $\{l(f), l(f)^*\}$ are freely independent:

- Let $l_1 = l(e)$ and $l_2 = l(f)$.
- Since $l_i^* l_i = 1$, any monomial in $l_i, l_i^*$ reduces to $l_i^n l_i^{*m}$, with $n + m > 0$.
- $\tau(l_i^n l_i^{*m}) = 0$ unless $m = 0 = n$, hence a polynomial $p_k \in \mathbf{Alg}(1, l_i)$ has $\tau(p_k) = 0$ iff it can be written as a sum of monomials (each having zero expectation) and no constant term.

Returning to the Fock space example, let $e, f \in \mathcal{H}$ be orthogonal unit vectors.
Then $c(e)$ and $c(f)$ are freely independent.
In fact, it is true that the families $\{l(e), l(e)^*\}$ and $\{l(f), l(f)^*\}$ are freely independent:

- Let $l_1 = l(e)$ and $l_2 = l(f)$.
- Since $l_i^* l_i = 1$, any monomial in $l_i, l_i^*$ reduces to $l_i^n l_i^{*m}$, with $n + m > 0$.
- $\tau(l_i^n l_i^{*m}) = 0$ unless $m = 0 = n$, hence a polynomial $p_k \in \mathbf{Alg}(1, l_i)$ has $\tau(p_k) = 0$ iff it can be written as a sum of monomials (each having zero expectation) and no constant term.
- Suffices to show $\tau(p_1 \cdots p_r) = 0$ for monomials $p_k = l_{i_k}^{n_k} l_{i_k}^{*m_k}$, with $n_k + m_k > 0$ and $i_k \neq i_{k+1}$.

Returning to the Fock space example, let $e, f \in \mathcal{H}$ be orthogonal unit vectors.
Then $c(e)$ and $c(f)$ are freely independent.
In fact, it is true that the families $\{l(e), l(e)^*\}$ and $\{l(f), l(f)^*\}$ are freely independent:

- Let $l_1 = l(e)$ and $l_2 = l(f)$.
- Since $l_i^* l_i = 1$, any monomial in $l_i, l_i^*$ reduces to $l_i^n l_i^{*m}$, with $n + m > 0$.
- $\tau(l_i^n l_i^{*m}) = 0$ unless $m = 0 = n$, hence a polynomial $p_k \in \mathbf{Alg}(1, l_i)$ has $\tau(p_k) = 0$ iff it can be written as a sum of monomials (each having zero expectation) and no constant term.
- Suffices to show $\tau(p_1 \cdots p_r) = 0$ for monomials $p_k = l_{i_k}^{n_k} l_{i_k}^{*m_k}$, with $n_k + m_k > 0$ and $i_k \neq i_{k+1}$.
- Considering $\langle \Omega, p_1 \cdots p_r \Omega \rangle$, it is easy to see $m_k \neq 0$ for any $k$ implies this is zero, but $m_k = 0$ for all $k$ implies $n_k > 0$ for all $k$ and hence $p_1 \cdots p_r \Omega \in \mathcal{H}^{\otimes(n_1 + \cdots + n_r)} \perp \mathbb{C}\Omega$.

With the notion of free independence, we can state one of the first parallels to the classical case:

**Theorem 1 (Free central limit theorem, [6])**

*Let $X_1, X_2, \dots$ be a sequence of freely independent random variables in some non-commutative probability space $(A, \tau)$. Assume $\tau(X_n) = 0$ and $\tau(X_n^2) = 1$ for all $n$, and that $\sup_n |\tau(X_n^p)| < \infty$ for all $p$. Then the laws of the sequence*

$$Z_N = \frac{1}{\sqrt{N}}(X_1 + \cdots + X_N)$$

*converge in moments to the semicircle law $d\mu = \frac{1}{2\pi}\sqrt{4 - t^2}\, dt$.*

Using a standard construction in operator algebras (the Gelfand-Naimark-Segal construction), it is possible to associate a Hilbert space to $(A, \tau)$.

Using a standard construction in operator algebras (the Gelfand-Naimark-Segal construction), it is possible to associate a Hilbert space to $(A, \tau)$.
For $X, Y \in A$ let

$$\langle X, Y \rangle_{L^2(A,\tau)} = \langle X, Y \rangle_2 := \tau(X^* Y).$$

This defines a sesquilinear form on the complex vector space $A$, which is complex linear in the second coordinate.

Using a standard construction in operator algebras (the Gelfand-Naimark-Segal construction), it is possible to associate a Hilbert space to $(A, \tau)$.
For $X, Y \in A$ let

$$\langle X, Y \rangle_{L^2(A,\tau)} = \langle X, Y \rangle_2 := \tau(X^* Y).$$

This defines a sesquilinear form on the complex vector space $A$, which is complex linear in the second coordinate.
That is, it is an inner product except for the possibility that $\langle X, X \rangle_2 = 0$ for some non-zero $X \in A$.

Using a standard construction in operator algebras (the Gelfand-Naimark-Segal construction), it is possible to associate a Hilbert space to $(A, \tau)$.
For $X, Y \in A$ let

$$\langle X, Y\rangle_{L^2(A,\tau)} = \langle X, Y\rangle_2 := \tau(X^* Y).$$

This defines a sesquilinear form on the complex vector space $A$, which is complex linear in the second coordinate.
That is, it is an inner product except for the possibility that $\langle X, X\rangle_2 = 0$ for some non-zero $X \in A$.
To correct this we consider the set $N = \{X \in A \colon \langle X, X\rangle_2 = 0\}$. We want to mod out by $N$, but in order for $A/N$ to still be a vector space we need $N$ to be a vector subspace.

Using a standard construction in operator algebras (the Gelfand-Naimark-Segal construction), it is possible to associate a Hilbert space to $(A, \tau)$.
For $X, Y \in A$ let

$$\langle X, Y \rangle_{L^2(A,\tau)} = \langle X, Y \rangle_2 := \tau(X^* Y).$$

This defines a sesquilinear form on the complex vector space $A$, which is complex linear in the second coordinate.
That is, it is an inner product except for the possibility that $\langle X, X \rangle_2 = 0$ for some non-zero $X \in A$.
To correct this we consider the set $N = \{X \in A \colon \langle X, X \rangle_2 = 0\}$. We want to mod out by $N$, but in order for $A/N$ to still be a vector space we need $N$ to be a vector subspace.
This follows from the fact that $N = \{X \in A \colon \langle Y, X \rangle_2 = 0 \ \forall Y \in A\}$.

Using a standard construction in operator algebras (the Gelfand-Naimark-Segal construction), it is possible to associate a Hilbert space to $(A, \tau)$.
For $X, Y \in A$ let

$$\langle X, Y \rangle_{L^2(A,\tau)} = \langle X, Y \rangle_2 := \tau(X^* Y).$$

This defines a sesquilinear form on the complex vector space $A$, which is complex linear in the second coordinate.
That is, it is an inner product except for the possibility that $\langle X, X \rangle_2 = 0$ for some non-zero $X \in A$.
To correct this we consider the set $N = \{X \in A\colon \langle X, X \rangle_2 = 0\}$. We want to mod out by $N$, but in order for $A/N$ to still be a vector space we need $N$ to be a vector subspace.
This follows from the fact that $N = \{X \in A\colon \langle Y, X \rangle_2 = 0 \ \forall Y \in A\}$.
Now $A/N$ is a vector space on which $\langle \cdot, \cdot \rangle_2$ is an inner product. Let $L^2(A, \tau)$ be the Hilbert space obtained by taking the completion of $A/N$ with respect to the norm induced by $\langle \cdot, \cdot \rangle_2$.

To each $X \in A$ we have the associated vector $\hat{X} \in A/N \subset L^2(A, \tau)$.

To each $X \in A$ we have the associated vector $\hat{X} \in A/N \subset L^2(A, \tau)$. Moreover, each element of $X \in A$ defines a (bounded) operator $\pi(X)$ on $L^2(A, \tau)$ via the dense definition:

$$\pi(X)\hat{Y} = \widehat{XY}.$$

In particular, $\hat{X} = \pi(X)\hat{1}$ for all $X \in A$.

To each $X \in A$ we have the associated vector $\hat{X} \in A/N \subset L^2(A, \tau)$. Moreover, each element of $X \in A$ defines a (bounded) operator $\pi(X)$ on $L^2(A, \tau)$ via the dense definition:

$$\pi(X)\hat{Y} = \widehat{XY}.$$

In particular, $\hat{X} = \pi(X)\hat{1}$ for all $X \in A$.
The positivity of $\tau$ allows us to define the *non-commutative $L^p$ spaces*:

$$L^p(A, \tau) = \{X \in A \colon \tau((X^*X)^p) < \infty\}, \qquad \|X\|_p = \tau((X^*X)^p)^{\frac{1}{p}}$$

To each $X \in A$ we have the associated vector $\hat{X} \in A/N \subset L^2(A, \tau)$. Moreover, each element of $X \in A$ defines a (bounded) operator $\pi(X)$ on $L^2(A, \tau)$ via the dense definition:

$$\pi(X)\hat{Y} = \widehat{XY}.$$

In particular, $\hat{X} = \pi(X)\hat{1}$ for all $X \in A$.
The positivity of $\tau$ allows us to define the *non-commutative $L^p$ spaces*:

$$L^p(A, \tau) = \{X \in A \colon \tau((X^*X)^p) < \infty\}, \qquad \|X\|_p = \tau((X^*X)^p)^{\frac{1}{p}}$$

The *von Neumann algebra generated by $A$* is then what plays the role of the non-commutative $L^\infty$ space:

$$W^*(A) = \overline{\pi(A)}^{SOT} = \overline{\pi(A)}^{WOT} = \pi(A)'' \cap \mathcal{B}(L^2(A, \tau)).$$

More generally, given a family $F \subset A$ we let

$$W^*(F) = \overline{\pi(\mathrm{Alg}_*(1, F))}^{SOT} \subset W^*(A).$$

More generally, given a family $F \subset A$ we let

$$W^*(F) = \overline{\pi(\mathrm{Alg}_*(1, F))}^{SOT} \subset W^*(A).$$

Let that $L^2(W^*(F), \tau) = \overline{W^*(F) \cdot \hat{1}}^{\|\cdot\|_2} \subset L^2(A, \tau)$. Then we define the orthogonal projection $E_{W^*(F)}\colon L^2(A, \tau) \to L^2(W^*(F), \tau)$ onto this subspace.

Given non-commutative random variables $X_1, \ldots, X_n \in (A, \tau)$ we define Voiculescu's free difference quotients

$$\partial_{X_j : X_1, \ldots, \hat{X}_j, \ldots, X_n} = \partial_j \colon \mathbf{Alg}(1, X_1, \ldots, X_n) \to L^2(A, \tau) \bar{\otimes} L^2(A, \tau)$$

by $\partial_j(X_k) = \delta_{j=k} 1 \otimes 1$ and the Leibniz rule:

$$\partial_j(WZ) = \partial_j(W) \cdot Z + W \cdot \partial_j(Z).$$

Given non-commutative random variables $X_1, \ldots, X_n \in (A, \tau)$ we define Voiculescu's free difference quotients

$$\partial_{X_j : X_1, \ldots, \hat{X}_j, \ldots, X_n} = \partial_j \colon \mathbf{Alg}(1, X_1, \ldots, X_n) \to L^2(A, \tau) \bar{\otimes} L^2(A, \tau)$$

by $\partial_j(X_k) = \delta_{j=k} 1 \otimes 1$ and the Leibniz rule:

$$\partial_j(WZ) = \partial_j(W) \cdot Z + W \cdot \partial_j(Z).$$

For example,

$$\begin{aligned}
\partial_2(X_1 X_2 X_3) &= \partial_2(X_1) \cdot (X_2 X_3) + X_1 \cdot \partial_2(X_2 X_3) \\
&= 0 + X_1 \cdot [\partial_2(X_2) \cdot X_3 + X_2 \cdot \partial_2(X_3)] \\
&= X_1 \cdot (1 \otimes 1) \cdot X_3 + 0 = X_1 \otimes X_3.
\end{aligned}$$

If we think of $\partial_j$ as a map on $L^2(W^*(X_1,\ldots,X_n),\tau)\subset L^2(A,\tau)$, then we can consider its adjoint $\partial_j^*\colon L^2(A,\tau)\bar{\otimes}L^2(A,\tau)\to L^2(A,\tau)$.

If we think of $\partial_j$ as a map on $L^2(W^*(X_1,\ldots,X_n),\tau) \subset L^2(A,\tau)$, then we can consider its adjoint $\partial_j^*\colon L^2(A,\tau)\bar{\otimes}L^2(A,\tau) \to L^2(A,\tau)$.
When $1\otimes 1$ is in the domain of the $\partial_j^*$, we can define the *conjugate variables*:

$$\xi_j = J(X_j\colon X_1,\ldots,\hat{X}_j,\ldots,X_n) = \partial_j^*(1\otimes 1).$$

If we think of $\partial_j$ as a map on $L^2(W^*(X_1,\ldots,X_n),\tau) \subset L^2(A,\tau)$, then we can consider its adjoint $\partial_j^*\colon L^2(A,\tau)\bar{\otimes}L^2(A,\tau) \to L^2(A,\tau)$.
When $1 \otimes 1$ is in the domain of the $\partial_j^*$, we can define the *conjugate variables*:

$$\xi_j = J(X_j : X_1,\ldots,\hat{X}_j,\ldots,X_n) = \partial_j^*(1 \otimes 1).$$

That is, if $Y \in L^2(W^*(X_1,\ldots,X_n),\tau)$ then

$$\langle Y, \xi_j \rangle_2 = \langle \partial_j(Y), 1 \otimes 1 \rangle_{L^2(A,\tau)\bar{\otimes}L^2(A,\tau)}.$$

If we think of $\partial_j$ as a map on $L^2(W^*(X_1,\dots,X_n),\tau)\subset L^2(A,\tau)$, then we can consider its adjoint $\partial_j^*\colon L^2(A,\tau)\bar{\otimes}L^2(A,\tau)\to L^2(A,\tau)$.
When $1\otimes 1$ is in the domain of the $\partial_j^*$, we can define the *conjugate variables*:

$$\xi_j = J(X_j : X_1,\dots,\hat{X}_j,\dots,X_n) = \partial_j^*(1\otimes 1).$$

That is, if $Y\in L^2(W^*(X_1,\dots,X_n),\tau)$ then

$$\langle Y,\xi_j\rangle_2 = \langle \partial_j(Y), 1\otimes 1\rangle_{L^2(A,\tau)\bar{\otimes}L^2(A,\tau)}.$$

The *free Fisher information* of the $n$-tuple $(X_1,\dots,X_n)$ is then defined as

$$\Phi^*(X_1,\dots,X_n) := \sum_{j=1}^n \|\xi_j\|_{L^2(A,\tau)}^2.$$

The *free entropy* of the $n$-tuple $(X_1, \ldots, X_n)$ is defined as

$$\chi^*(X_1, \ldots, X_n) = \frac{1}{2} \int_0^\infty \left[ \frac{n}{1+t} - \Phi^*(X_1^t, \ldots, X_n^t) \right] \, dt + \frac{n}{2} \log 2\pi e,$$

where $X_j^t = X_j + \sqrt{t} S_j$ and $S_1, \ldots, S_n$ are freely independent, identically distributed, centered, semicircular variables of variance 1, which are also freely independent from $X_1, \ldots, X_n$.

The *free entropy* of the $n$-tuple $(X_1, \ldots, X_n)$ is defined as

$$\chi^*(X_1, \ldots, X_n) = \frac{1}{2} \int_0^\infty \left[ \frac{n}{1+t} - \Phi^*(X_1^t, \ldots, X_n^t) \right] \, dt + \frac{n}{2} \log 2\pi e,$$

where $X_j^t = X_j + \sqrt{t} S_j$ and $S_1, \ldots, S_n$ are freely independent, identically distributed, centered, semicircular variables of variance 1, which are also freely independent from $X_1, \ldots, X_n$.

When $n = 1$, the above definition is equivalent to

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

where $\mu_X$ is the law of $X$.

## Theorem 2

*Let* $(A, \tau)$ *be a non-commutative probability space. Let* $X_j = (X_j^{(1)}, \ldots, X_j^{(p)}) \in A^p$, $j = 1, 2, \ldots$ *be a sequence of p-tuples of random variables, such that* $X_1, X_2, \ldots$ *are freely independent, identically distributed, and have finite second moments. Define* $Z_N = N^{-1/2}(X_1 + \cdots + X_N)$. *Then the function* $N \mapsto \chi^*(Z_N^{(1)}, \ldots, Z_N^{(p)})$ *is monotone nondecreasing.*

## Theorem 2

*Let $(A, \tau)$ be a non-commutative probability space. Let $X_j = (X_j^{(1)}, \ldots, X_j^{(p)}) \in A^p$, $j = 1, 2, \ldots$ be a sequence of p-tuples of random variables, such that $X_1, X_2, \ldots$ are freely independent, identically distributed, and have finite second moments. Define $Z_N = N^{-1/2}(X_1 + \cdots + X_N)$. Then the function $N \mapsto \chi^*(Z_N^{(1)}, \ldots, Z_N^{(p)})$ is monotone nondecreasing.*

To show this we first show that the free Fisher information is monotone nonincreasing.

### Lemma 3

*Assume* $Z \in (A, \tau)$ *is freely independent from* $X, Y_1, \ldots, Y_n \in (A, \tau)$. *Then* $J(X\colon Y_1, \ldots, Y_n)$ *exists iff* $J(X\colon Y_1, \ldots, Y_n, Z)$ *exists, in which case we have*

$$J(X\colon Y_1, \ldots, Y_n) = J(X\colon Y_1, \ldots, Y_n, Z).$$

## Lemma 3

*Assume $Z \in (A, \tau)$ is freely independent from $X, Y_1, \ldots, Y_n \in (A, \tau)$. Then $J(X\colon Y_1, \ldots, Y_n)$ exists iff $J(X\colon Y_1, \ldots, Y_n, Z)$ exists, in which case we have*

$$J(X\colon Y_1, \ldots, Y_n) = J(X\colon Y_1, \ldots, Y_n, Z).$$

## Proof.

Let $\partial_0 = \partial_{X\colon Y_1,\ldots,Y_n}$ and $\partial_1 = \partial_{X\colon Y_1,\ldots,Y_n,Z}$. Then we can think of these as maps on $B_0 := W^*(X, Y_1, \ldots, Y_n)$ and $B_1 := W^*(X, Y_1, \ldots, Y_n, Z)$.

## Lemma 3

*Assume* $Z \in (A, \tau)$ *is freely independent from* $X, Y_1, \ldots, Y_n \in (A, \tau)$. *Then* $J(X\colon Y_1, \ldots, Y_n)$ *exists iff* $J(X\colon Y_1, \ldots, Y_n, Z)$ *exists, in which case we have*

$$J(X\colon Y_1, \ldots, Y_n) = J(X\colon Y_1, \ldots, Y_n, Z).$$

## Proof.

Let $\partial_0 = \partial_{X\colon Y_1,\ldots,Y_n}$ and $\partial_1 = \partial_{X\colon Y_1,\ldots,Y_n,Z}$. Then we can think of these as maps on $B_0 := W^*(X, Y_1, \ldots, Y_n)$ and $B_1 := W^*(X, Y_1, \ldots, Y_n, Z)$. Suppose $\xi_1 := J(X\colon Y_1, \ldots, Y_n, Z)$ exists (i.e. $1 \otimes 1$ is in the domain of $\partial_1^*$). If $E_0 = E_{B_0}$, then by considering inner products against $\eta \in L^2(B_0, \tau)$ it is clear that

$$\xi_0 := J(X\colon Y_1, \ldots, Y_n) = E_0(\xi_1).$$

## Proof of Lemma 3 (cont.)

Thus it suffices to show that if $\xi_0$ exists, then $\xi_1$ exists and $\xi_0 = \xi_1$.

## Proof of Lemma 3 (cont.)

Thus it suffices to show that if $\xi_0$ exists, then $\xi_1$ exists and $\xi_0 = \xi_1$. Consider an arbitrary element of $\mathbf{Alg}(1, X, Y_1, \ldots, Y_n, Z)$:

$$R = Q_0 P_1 Q_1 \cdots P_r Q_r$$

with $P_k \in \mathbf{Alg}(1, X, Y_1, \ldots, Y_n)$, $Q_k \in \mathbf{Alg}(1, Z)$.

## Proof of Lemma 3 (cont.)

Thus it suffices to show that if $\xi_0$ exists, then $\xi_1$ exists and $\xi_0 = \xi_1$.
Consider an arbitrary element of $\mathbf{Alg}(1, X, Y_1, \ldots, Y_n, Z)$:

$$R = Q_0 P_1 Q_1 \cdots P_r Q_r$$

with $P_k \in \mathbf{Alg}(1, X, Y_1, \ldots, Y_n)$, $Q_k \in \mathbf{Alg}(1, Z)$.
We must show $\langle \xi_0, R \rangle_2 = \langle 1 \otimes 1, \partial_1(R) \rangle_{L^2(A,\tau)\bar{\otimes}L^2(A,\tau)}$. We proceed by induction on $r$.

## Proof of Lemma 3 (cont.)

Thus it suffices to show that if $\xi_0$ exists, then $\xi_1$ exists and $\xi_0 = \xi_1$.
Consider an arbitrary element of $\mathbf{Alg}(1, X, Y_1, \ldots, Y_n, Z)$:

$$R = Q_0 P_1 Q_1 \cdots P_r Q_r$$

with $P_k \in \mathbf{Alg}(1, X, Y_1, \ldots, Y_n)$, $Q_k \in \mathbf{Alg}(1, Z)$.
We must show $\langle \xi_0, R \rangle_2 = \langle 1 \otimes 1, \partial_1(R) \rangle_{L^2(A,\tau)\bar{\otimes}L^2(A,\tau)}$. We proceed by induction on $r$.
For $r = 0$, note that

$$\tau(\xi_0) = \langle \xi_0, 1 \rangle_2 = \langle 1 \otimes 1, \partial_0(1) \rangle_{L^2(A,\tau)\bar{\otimes}L^2(A,\tau)} = 0,$$

and so by free independence...

## Proof of Lemma 3 (cont.)

$$\begin{aligned}\langle \xi_0, Q_0\rangle_2 &= \tau(\xi_0 Q_0)\\ &= \tau(\xi_0(Q_0 - \tau(Q_0))) + \tau(\xi_0)\tau(Q_0) = 0.\end{aligned}$$

On the other hand, $\partial_1(Q_0) = 0$, so the base case holds.

## Proof of Lemma 3 (cont.)

$$\begin{aligned}\langle \xi_0, Q_0\rangle_2 &= \tau(\xi_0 Q_0)\\ &= \tau(\xi_0(Q_0 - \tau(Q_0))) + \tau(\xi_0)\tau(Q_0) = 0.\end{aligned}$$

On the other hand, $\partial_1(Q_0) = 0$, so the base case holds.
For $r > 0$, we note that we can assume $P_1, Q_1, \ldots, P_{r-1}, Q_{r-1}, P_r$ are centered.

## Proof of Lemma 3 (cont.)

$$\begin{aligned}\langle \xi_0, Q_0\rangle_2 &= \tau(\xi_0 Q_0)\\ &= \tau(\xi_0(Q_0 - \tau(Q_0))) + \tau(\xi_0)\tau(Q_0) = 0.\end{aligned}$$

On the other hand, $\partial_1(Q_0) = 0$, so the base case holds.
For $r > 0$, we note that we can assume $P_1, Q_1, \ldots, P_{r-1}, Q_{r-1}, P_r$ are centered.
Indeed, by expanding each term in $P_1 Q_1 \cdots Q_{r-1} P_r$ into its centered part and scalar part, the resulting products have either all centered terms or are covered by the induction hypothesis.

## Proof of Lemma 3 (cont.)

Thus, for $r = 1$ we have

$$\begin{aligned}\tau(\xi_0 Q_0 P_1 Q_1) =& \tau(\xi_0 \mathring{Q}_0 P_1 \mathring{Q}_1) + \tau(Q_0)\tau(\xi_0 P_1 \mathring{Q}_1) \\ &+ \tau(Q_1)\tau(\xi_0 \mathring{Q}_0 P_1) + \tau(Q_0)\tau(Q_1)\tau(\xi_0 P_1) \\ =& \tau(Q_0)\tau(Q_1)\tau(\xi_0 P_1) \\ =& \tau(Q_0)\tau(Q_1) \langle 1 \otimes 1, \partial_0 P_1 \rangle_{L^2(A,\tau)\bar{\otimes} L^2(A,\tau)},\end{aligned}$$

## Proof of Lemma 3 (cont.)

Thus, for $r = 1$ we have

$$\begin{aligned}\tau(\xi_0 Q_0 P_1 Q_1) =&\tau(\xi_0 \mathring{Q}_0 P_1 \mathring{Q}_1) + \tau(Q_0)\tau(\xi_0 P_1 \mathring{Q}_1)\\ &+ \tau(Q_1)\tau(\xi_0 \mathring{Q}_0 P_1) + \tau(Q_0)\tau(Q_1)\tau(\xi_0 P_1)\\ =&\tau(Q_0)\tau(Q_1)\tau(\xi_0 P_1)\\ =&\tau(Q_0)\tau(Q_1)\left\langle 1\otimes 1, \partial_0 P_1\right\rangle_{L^2(A,\tau)\bar{\otimes} L^2(A,\tau)},\end{aligned}$$

while on the other hand

$$\begin{aligned}\langle \xi_1, Q_0 P_1 Q_1\rangle_2 &= \tau\otimes\tau(\partial_1(Q_0 P_1 Q_1))\\ &= \tau\otimes\tau(Q_0\cdot\partial_1(P_1)\cdot Q_1)\\ &= \tau\otimes\tau(Q_0\cdot\partial_0(P_1)\cdot Q_1)\\ &= \tau(Q_0)\tau(Q_1)\tau\otimes\tau(\partial_0 P_1),\end{aligned}$$

where the last equality follows from free independence.

## Proof of Lemma 3 (cont.)

For $r \geq 2$ we have

$$\begin{aligned}\tau(\xi_0 Q_0 P_1 \cdots P_r Q_r) \\ =&\tau(\xi_0 \mathring{Q}_0 P_1 \cdots P_r \mathring{Q}_r) + \tau(Q_0)\tau(\xi_0 P_1 \cdots P_r \mathring{Q}_r) \\ &+ \tau(Q_r)\tau(\xi_0 \mathring{Q}_0 P_1 \cdots P_r) + \tau(Q_0)\tau(Q_r)\tau(\xi_0 P_1 \cdots P_r) = 0.\end{aligned}$$

And

$$\begin{aligned}\tau \otimes \tau(\partial_1(Q_0 P_1 \cdots P_r Q_r)) \\ = \sum_{l=1}^{r} \tau \otimes \tau([Q_0 P_1 \cdots Q_{l-1}] \cdot \partial_1(P_l) \cdot [Q_l \cdots P_r Q_r]) = 0.\end{aligned}$$

□

## Lemma 4

*Let* $\{X_j^{(k)}\} \subset (A, \tau)$, $k = 1, \ldots, p$, $j = 1, 2, \ldots$ *be non-commutative random variables. Fix* $N \in \mathbb{N}$, $j = 1, \ldots, N+1$, *and* $k = 1, \ldots, p$. *Then one has*

$$J\left(\sum_{i=1}^{N+1} X_i^{(k)} : \left\{\sum_{i=1}^{N+1} X_i^{(r)}\right\}_{r \neq k}\right)$$

$$= E_{W^*\left(\left\{\sum_{i=1}^{N+1} X_i^{(r)}\right\}_{r=1}^{p}\right)} J\left(\sum_{i \neq j} X_i^{(k)} : \left\{\sum_{i \neq j} X_i^{(r)}\right\}_{r \neq k}, \left\{X_j^{(r)}\right\}_{r=1}^{p}\right)$$

*assuming the conjugate variables on the right-hand side exists.*

## Lemma 4

*Let* $\{X_j^{(k)}\} \subset (A, \tau)$, $k = 1, \ldots, p$, $j = 1, 2, \ldots$ *be non-commutative random variables. Fix* $N \in \mathbb{N}$, $j = 1, \ldots, N+1$, *and* $k = 1, \ldots, p$. *Then one has*

$$J\left(\sum_{i=1}^{N+1} X_i^{(k)} : \left\{\sum_{i=1}^{N+1} X_i^{(r)}\right\}_{r \neq k}\right)$$

$$= E_{W^*\left(\left\{\sum_{i=1}^{N+1} X_i^{(r)}\right\}_{r=1}^{p}\right)} J\left(\sum_{i \neq j} X_i^{(k)} : \left\{\sum_{i \neq j} X_i^{(r)}\right\}_{r \neq k}, \left\{X_j^{(r)}\right\}_{r=1}^{p}\right)$$

*assuming the conjugate variables on the right-hand side exists.*

The basic idea, is that if $y = \tilde{y} + x$ then $\partial_y(p(y)) = \partial_{\tilde{y}}(p(\tilde{y} + x))$.

## Proof.

Let $Y_k = \sum_{i=1}^{N+1} X_i^{(k)}$ and $Y_k' = \sum_{i \neq j} X_i^{(k)}$, so that $Y_k = Y_k' + X_j^{(k)}$. Then a polynomial $P$ in $Y_1, \ldots, Y_p$ can be viewed as a polynomial in $Y_1', \ldots, Y_p', X_j^{(1)}, \ldots, X_j^{(p)}$.

## Proof.

Let $Y_k = \sum_{i=1}^{N+1} X_i^{(k)}$ and $Y_k' = \sum_{i\neq j} X_i^{(k)}$, so that $Y_k = Y_k' + X_j^{(k)}$. Then a polynomial $P$ in $Y_1, \ldots, Y_p$ can be viewed as a polynomial in $Y_1', \ldots, Y_p', X_j^{(1)}, \ldots, X_j^{(p)}$.
In particular,

$$\partial_{Y_k' \colon \{Y_r' \colon r\neq k\},\{X_j^{(r)} \colon r=1,\ldots,p\}} P = \partial_{Y_k \colon \{Y_r \colon r\neq k\}} P,$$

since the derivation is determined by the Leibniz rule and the values

$$\partial_{Y_k' \colon \{Y_r' \colon r\neq k\},\{X_j^{(r)} \colon r=1,\ldots,p\}} (Y_q' + X_j^{(q)}) = \partial_{Y_k \colon \{Y_r \colon r\neq k\}} (Y_q) = \delta_{k=q} 1 \otimes 1.$$

## Proof of Lemma 4 (cont.)

Hence

$$\left\langle P, J\left(Y_k' : \{Y_r' : r \neq k\}, \{X_j^{(r)} : r = 1, \ldots, p\}\right)\right\rangle_2 \\ = \langle P, J(Y_k : \{Y_r : r \neq k\})\rangle_2 ,$$

which concludes the proof as $P \in W^*(Y_1, \ldots, Y_p)$. □

## Theorem 5

*Let* $(A, \tau)$ *be a non-commutative probability space. Let* $X_j = (X_j^{(1)}, \ldots, X_j^{(p)}) \in A^p$, $j = 1, 2, \ldots$ *be a sequence of p-tuples of random variables, such that* $X_1, X_2, \ldots$ *are freely independent, identically distributed, and have finite second moments. Define* $Z_N = N^{-1/2}(X_1 + \cdots + X_N)$. *Then the function* $N \mapsto \Phi^*(Z_N^{(1)}, \ldots, Z_N^{(p)})$ *is monotone nonincreasing.*

## Theorem 5

*Let $(A, \tau)$ be a non-commutative probability space. Let $X_j = (X_j^{(1)}, \ldots, X_j^{(p)}) \in A^p$, $j = 1, 2, \ldots$ be a sequence of p-tuples of random variables, such that $X_1, X_2, \ldots$ are freely independent, identically distributed, and have finite second moments. Define $Z_N = N^{-1/2}(X_1 + \cdots + X_N)$. Then the function $N \mapsto \Phi^*(Z_N^{(1)}, \ldots, Z_N^{(p)})$ is monotone nonincreasing.*

## Proof.

First note that

$$J(cX : bY_1, \ldots, bY_n) = c^{-1} J(X : Y_1, \ldots, Y_n);$$

this follows from the observation $\partial_{ct}(t) = c^{-1}\partial_{ct}(ct) = c^{-1}\partial_t(t)$.

## Proof of Theorem 5 (cont.)

Fix $j$ and let $Y_k$ and $Y_k'$ be as in the previous proof. Then $Z_{N+1}^{(r)} = (N+1)^{-1/2} Y_r$. Let $B = W^*(Z_{N+1}^{(1)}, \ldots, Z_{N+1}^{(p)})$.

## Proof of Theorem 5 (cont.)

Fix $j$ and let $Y_k$ and $Y_k'$ be as in the previous proof. Then $Z_{N+1}^{(r)} = (N+1)^{-1/2} Y_r$. Let $B = W^*(Z_{N+1}^{(1)}, \ldots, Z_{N+1}^{(p)})$. Then

$$\begin{aligned} J(Z_{N+1}^{(k)} \colon \{Z_{N+1}^{(r)} \colon r \neq k\}) &= (N+1)^{\frac{1}{2}} J\left(Y_k \colon \{Y_r \colon r \neq k\}\right) \\ &= (N+1)^{\frac{1}{2}} E_B\ J\left(Y_k' \colon \{Y_r' \colon r \neq k\}, \left\{X_j^{(r)}\right\}_{r=1}^{n}\right) \\ &= (N+1)^{\frac{1}{2}} E_B\ J\left(Y_k' \colon \{Y_r' \colon r \neq k\}\right), \end{aligned}$$

where we have used (in order) our initial observation, Lemma 4, and Lemma 3. (Recall that the $X_j^{(r)}$ are freely independent from the $Y_r' = \sum_{i \neq j} X_i^{(r)}$.)

## Proof of Theorem 5 (cont.)

Thus we have

$$\begin{aligned}(N+1)^{\frac{1}{2}} J(Z_{N+1}^{(k)} : \{Z_{N+1}^{(r)} : r \neq k\}) \\ &= E_B (N+1) J\left( \sum_{i\neq j} X_i^{(k)} : \left\{ \sum_{i \neq j} X_i^{(r)} \right\}_{r\neq k} \right) \\ &= E_B \sum_{j=1}^{N+1} J\left( \sum_{i\neq j} X_i^{(k)} : \left\{ \sum_{i \neq j} X_i^{(r)} \right\}_{r\neq k} \right),\end{aligned}$$

where we have used the fact that our initial choice of $j$ was arbitrary.

## Proof of Theorem 5 (cont.)

Since $E_B$ is a contraction on $L^2(A,\tau)$ we then have

$$\left\| J(Z_{N+1}^{(k)} : \{Z_{N+1}^{(r)} : r \neq k\}) \right\|_2^2 \leq (N+1)^{-1} \left\| \sum_{j=1}^{N+1} \zeta_j \right\|_2^2, \qquad (1)$$

where

$$\zeta_j = J\left( \sum_{i \neq j} X_i^{(k)} : \left\{ \sum_{i \neq j} X_i^{(r)} \right\}_{r \neq k} \right).$$

To proceed, we need to appeal to a lemma from the classical proof:

To proceed, we need to appeal to a lemma from the classical proof:

## Lemma 6 ([1])

*Let* $E_1, \ldots, E_{N+1}$ *be commuting orthogonal projections in a Hilbert space. Assume that we have* $N+1$ *vectors* $\zeta_1, \ldots, \zeta_{N+1}$ *such that for every* $j$*,* $E_1 \cdots E_{N+1}\zeta_j = 0$*. Then*

$$\left\| \sum_{j=1}^{N+1} E_j \zeta_j \right\|^2 \leq N \sum_{j=1}^{N+1} \|\zeta_j\|^2.$$

To proceed, we need to appeal to a lemma from the classical proof:

## Lemma 6 ([1])

*Let $E_1, \ldots, E_{N+1}$ be commuting orthogonal projections in a Hilbert space. Assume that we have $N+1$ vectors $\zeta_1, \ldots, \zeta_{N+1}$ such that for every $j$, $E_1 \cdots E_{N+1}\zeta_j = 0$. Then*

$$\left\| \sum_{j=1}^{N+1} E_j \zeta_j \right\|^2 \leq N \sum_{j=1}^{N+1} \|\zeta_j\|^2.$$

## Proof of Theorem 5 (cont.)

We let $M_j = W^*(X_j)$, $M = W^*(X_1, \ldots, X_{N+1})$, $Q_j = W^*(X_i \colon i \neq j)$, and $E_j = E_{Q_j}$.

To proceed, we need to appeal to a lemma from the classical proof:

## Lemma 6 ([1])

*Let $E_1, \ldots, E_{N+1}$ be commuting orthogonal projections in a Hilbert space. Assume that we have $N+1$ vectors $\zeta_1, \ldots, \zeta_{N+1}$ such that for every $j$, $E_1 \cdots E_{N+1}\zeta_j = 0$. Then*

$$\left\| \sum_{j=1}^{N+1} E_j \zeta_j \right\|^2 \leq N \sum_{j=1}^{N+1} \|\zeta_j\|^2.$$

## Proof of Theorem 5 (cont.)

We let $M_j = W^*(X_j)$, $M = W^*(X_1, \ldots, X_{N+1})$, $Q_j = W^*(X_i \colon i \neq j)$, and $E_j = E_{Q_j}$.

We claim the $E_j$ commute, $E_j\zeta_j = \zeta_j$, and $E_1 \cdots E_{N+1}\zeta_j = \tau(\zeta_j) = 0$.

## Proof of Theorem 5 (cont.)

That $E_j\zeta_j = \zeta_j$ follows from the definition of $\zeta_j$, and $\tau(\zeta_j) = \langle 1, \zeta_j \rangle_2 = \langle \partial_j(1), 1 \otimes 1 \rangle = 0$.

# Proof of Theorem 5 (cont.)

That $E_j\zeta_j = \zeta_j$ follows from the definition of $\zeta_j$, and $\tau(\zeta_j) = \langle 1, \zeta_j \rangle_2 = \langle \partial_j(1), 1 \otimes 1 \rangle = 0$.
Let $\mathring{M}_j = M_j \ominus \mathbb{C}1$ (i.e. the centered elements in $M_j$). Then $M$ has the following orthogonal decomposition:

$$L^2(M, \tau) = \mathbb{C}1 \oplus \bigoplus_{n=1}^{\infty} \left[ \bigoplus_{j_1 \neq \cdots \neq j_n} \mathring{M}_{j_1} \mathring{M}_{j_2} \cdots \mathring{M}_{j_n} \right].$$

It is orthogonal precisely because of the free independence.

## Proof of Theorem 5 (cont.)

That $E_j\zeta_j = \zeta_j$ follows from the definition of $\zeta_j$, and $\tau(\zeta_j) = \langle 1, \zeta_j\rangle_2 = \langle \partial_j(1), 1 \otimes 1\rangle = 0$.
Let $\mathring{M}_j = M_j \ominus \mathbb{C}1$ (i.e. the centered elements in $M_j$). Then $M$ has the following orthogonal decomposition:

$$L^2(M,\tau) = \mathbb{C}1 \oplus \bigoplus_{n=1}^{\infty} \left[ \bigoplus_{j_1 \neq \cdots \neq j_n} \mathring{M}_{j_1} \mathring{M}_{j_2} \cdots \mathring{M}_{j_n} \right].$$

It is orthogonal precisely because of the free independence.
$Q_j$ has the same decomposition except that $j_k$ is never allowed to be $j$, and $E_j$ is determined by $E_j 1 = 1$ and

$$E_j \mid_{\mathring{M}_{j_1} \mathring{M}_{j_2} \cdots \mathring{M}_{j_n}} = \begin{cases} id & \text{if } j \notin \{j_1, \ldots, j_n\} \\ 0 & \text{otherwise} \end{cases} .$$

## Proof of Theorem 5 (cont.)

From this characterization is clear that the $E_j$ commute with one another and

$$E_jE_i\mid_{\mathring{M}_{j_1}\mathring{M}_{j_2}\cdots\mathring{M}_{j_n}}=\begin{cases} id & \text{if } \{i,j\}\cap\{j_1,\ldots,j_n\}=\emptyset \\ 0 & \text{otherwise}\end{cases}.$$

## Proof of Theorem 5 (cont.)

From this characterization is clear that the $E_j$ commute with one another and

$$E_j E_i \mid_{\mathring{M}_{j_1}\mathring{M}_{j_2}\cdots\mathring{M}_{j_n}} = \begin{cases} id & \text{if } \{i,j\} \cap \{j_1,\ldots,j_n\} = \emptyset \\ 0 & \text{otherwise} \end{cases} .$$

Moreover, $E_1 \cdots E_{N+1}$ is the orthogonal projection onto the scalars $\mathbb{C}1$. Hence we can determine $E_1 \cdots E_{N+1}\zeta_j$ by considering the inner products of the $\zeta_j$ against scalars:

$$\langle 1, \zeta_j \rangle_2 = \tau(\zeta_j),$$

so $E_1 \cdots E_{N+1}\zeta_j = \tau(\zeta_j)1$, as claimed.

## Proof of Theorem 5 (cont.)

Applying Lemma 6 to (1) yields

$$\left\| J(Z_{N+1}^{(k)} : \{Z_{N+1}^{(r)} : r \neq k\}) \right\|_2^2 \leq (N+1)^{-1} N \sum_{j=1}^{N+1} \|\zeta_j\|_2^2.$$

## Proof of Theorem 5 (cont.)

Applying Lemma 6 to (1) yields

$$\left\| J(Z_{N+1}^{(k)} : \{Z_{N+1}^{(r)} : r \neq k\}) \right\|_2^2 \leq (N+1)^{-1} N \sum_{j=1}^{N+1} \|\zeta_j\|_2^2.$$

However, since the $p$-tuples are identically distributed we have

$$\begin{aligned} \sum_{j=1}^{N+1} \|\zeta_j\|_2^2 &= (N+1)\|\zeta_{N+1}\|_2^2 \\ &= (N+1) \left\| J\left( \sum_{i=1}^{N} X_i^{(k)} : \left\{ \sum_{i=1}^{N} X_i^{(r)} : r \neq k \right\} \right) \right\|_2^2 \\ &= \frac{N+1}{N} \|J(Z_N^{(k)} : \{Z_N^{(r)} : r \neq k\}\|_2^2 \end{aligned}$$

## Proof of Theorem 5 (cont.)

Combining the two previous equations then yields

$$\left\| J(Z_{N+1}^{(k)} : \{Z_{N+1}^{(r)} : r \neq k\}) \right\|_2^2 \leq \|J(Z_N^{(k)} : \{Z_N^{(r)} : r \neq k\}\|_2^2.$$

## Proof of Theorem 5 (cont.)

Combining the two previous equations then yields

$$\left\| J(Z_{N+1}^{(k)} : \{Z_{N+1}^{(r)} : r \neq k\}) \right\|_2^2 \leq \| J(Z_N^{(k)} : \{Z_N^{(r)} : r \neq k\} \|_2^2.$$

Finally, summing over $k$ yields

$$\Phi^*(Z_{N+1}^{(1)}, \ldots, Z_{N+1}^{(p)}) \leq \Phi^*(Z_N^{(1)}, \ldots, Z_N^{(p)}).$$

□

## Proof of Theorem 2.

We wish to show

$$\chi^*\left(Z_N^{(1)},\ldots,Z_N^{(p)}\right) \leq \chi^*\left(Z_{N+1}^{(1)},\ldots,Z_{N+1}^{(p)}\right),$$

## Proof of Theorem 2.

We wish to show

$$\chi^*\left(Z_N^{(1)},\ldots,Z_N^{(p)}\right) \leq \chi^*\left(Z_{N+1}^{(1)},\ldots,Z_{N+1}^{(p)}\right),$$

where for an arbitrary $p$-tuple $(W^{(1)},\ldots,W^{(p)})$

$$\chi^*\left(W^{(1)},\ldots,W^{(p)}\right) = \frac{1}{2}\int_0^\infty \left[\frac{p}{1+t} - \Phi^*\left(W^{(1,t)},\ldots,W^{(p,t)}\right)\right] dt + \frac{p}{2}\log 2\pi e,$$

## Proof of Theorem 2.

We wish to show

$$\chi^*\left(Z_N^{(1)},\dots,Z_N^{(p)}\right) \le \chi^*\left(Z_{N+1}^{(1)},\dots,Z_{N+1}^{(p)}\right),$$

where for an arbitrary $p$-tuple $(W^{(1)},\dots,W^{(p)})$

$$\begin{aligned}\chi^*\left(W^{(1)},\dots,W^{(p)}\right)\\ = \frac{1}{2}\int_0^\infty \left[\frac{p}{1+t} - \Phi^*\left(W^{(1,t)},\dots,W^{(p,t)}\right)\right] dt + \frac{p}{2}\log 2\pi e,\end{aligned}$$

where $W^{(k,t)} = W^{(k)} + \sqrt{t}S^{(k)}$ with $S^{(1)},\dots,S^{(p)}$ a freely iid centered semicircular random variables of variance 1, freely independent from $W^{(1)},\dots,W^{(p)}$.

## Proof of Theorem 2 (cont.)

Let $\{S_j^{(k)} : j = 1, \dots, N+1, k = 1, \dots, p\}$ be freely iid centered semicircular variables of variance 1, which are freely independent from $\{X_j^{(k)}\}_{j,k}$.

## Proof of Theorem 2 (cont.)

Let $\{S_j^{(k)} \colon j = 1, \ldots, N+1, k = 1, \ldots, p\}$ be freely iid centered semicircular variables of variance 1, which are freely independent from $\{X_j^{(k)}\}_{j,k}$.

Define $X_j^{(k,t)} = X_j^{(k)} + \sqrt{t} S_j^{(k)}$ and $Z_N^{(k,t)} = N^{-1/2}(X_1^{(k,t)} + \cdots + X_N^{(k,t)})$.

## Proof of Theorem 2 (cont.)

Let $\{S_j^{(k)} \colon j = 1, \ldots, N+1, k = 1, \ldots, p\}$ be freely iid centered semicircular variables of variance 1, which are freely independent from $\{X_j^{(k)}\}_{j,k}$.

Define $X_j^{(k,t)} = X_j^{(k)} + \sqrt{t} S_j^{(k)}$ and $Z_N^{(k,t)} = N^{-1/2}(X_1^{(k,t)} + \cdots + X_N^{(k,t)})$.

The $p$-tuples $X_j^t = (X_j^{(1,t)}, \ldots, X_j^{(p,t)})$ are freely iid with finite second moments (via the Cauchy-Schwarz inequality).

## Proof of Theorem 2 (cont.)

Let $\{S_j^{(k)} : j = 1, \ldots, N+1, k = 1, \ldots, p\}$ be freely iid centered semicircular variables of variance 1, which are freely independent from $\{X_j^{(k)}\}_{j,k}$.

Define $X_j^{(k,t)} = X_j^{(k)} + \sqrt{t} S_j^{(k)}$ and $Z_N^{(k,t)} = N^{-1/2}(X_1^{(k,t)} + \cdots + X_N^{(k,t)})$.

The $p$-tuples $X_j^t = (X_j^{(1,t)}, \ldots, X_j^{(p,t)})$ are freely iid with finite second moments (via the Cauchy-Schwarz inequality).

Hence we may apply Theorem 5 to obtain

$$\Phi^* \left( Z_N^{(1,t)}, \ldots, Z_N^{(p,t)} \right) \geq \Phi^* \left( Z_{N+1}^{(1,t)}, \ldots, Z_{N+1}^{(p,t)} \right).$$

## Proof of Theorem 2 (cont.)

Note that $Z_N^{(k,t)} = Z_N^{(k)} + \sqrt{t} S^{(N,k)}$ where for each fixed $N$, $S^{(N,k)} = N^{-1/2}(S_1^{(k)} + \cdots + S_N^{(k)})$, $k = 1, \ldots, p$ is a family of centered freely iid semicircular variables freely independent from $\{Z_N^{(k)}\}_{k=1}^p$ and having variance 1.

# Proof of Theorem 2 (cont.)

Note that $Z_N^{(k,t)} = Z_N^{(k)} + \sqrt{t}S^{(N,k)}$ where for each fixed $N$, $S^{(N,k)} = N^{-1/2}(S_1^{(k)} + \cdots + S_N^{(k)})$, $k = 1, \ldots, p$ is a family of centered freely iid semicircular variables freely independent from $\{Z_N^{(k)}\}_{k=1}^p$ and having variance 1.

The definition of $\chi^*$ and the free Fisher information inequality gives

$$\begin{aligned}
&\chi^*\left(Z_N^{(1)}, \ldots, Z_N^{(p)}\right) \\
&\quad = \frac{1}{2}\int_0^\infty \left[\frac{p}{1+t} - \Phi^*\left(Z_N^{(1,t)}, \ldots, Z_N^{(p,t)}\right)\right] dt + \frac{p}{2}\log 2\pi e \\
&\quad \leq \left[\frac{p}{1+t} - \Phi^*\left(Z_{N+1}^{(1,t)}, \ldots, Z_{N+1}^{(p,t)}\right)\right] dt + \frac{p}{2}\log 2\pi e \\
&\quad = \chi^*\left(Z_{N+1}^{(1)}, \ldots, Z_{N+1}^{(p)}\right)
\end{aligned}$$

□

- Artstein, S., Ball, K., Barthe, F., and Naor, A., *Solution of Shannon's Problem on the monotonicity of entropy.* Journal Amer. Math. Soc. **17** (2004), 975–982;
- Shlyakhtenko, D., *A free analogue of Shannon's problem on monotonicity of entropy.* Adv. Math. **208** (2007), no. 2, 824–833;
- Shlyakhtenko, D., *Notes on free probability theory.* Preprint, arXiv.org:0504063v1, 2005;
- Shlyakhtenko, D. and Schultz, H., *Shannon's monotonicity problem for free and classical entropy.* Proc. Nat. Acad. Sci. **104** No. 39 15254–15258 (2007);
- Voiculescu, D., *The analogues of entropy and of Fisher's information measure in free probability, V.* Invent. Math. **132** (1998), 189–227.
- Voiculescue, D., *Symmetries of some reduced free product $C^*$-algebras*, Operator Algebras and Their Connections with Topology and Ergodic Theory, Lecture Notes in Mathematics, vol. 1132, Springer Verlag, 1985, pp. 556-588.