The Annals of Probability
2016, Vol. 44, No. 2, 867–882
DOI: 10.1214/14-AOP990

A NONCOMMUTATIVE MARTINGALE CONVEXITY INEQUALITY[1]

BY ÉRIC RICARD* AND QUANHUA XU[†,‡]

Université de Caen Basse-Normandie,*
Wuhan University[†] and Université de Franche-Comté[‡]

Let $\mathcal{M}$ be a von Neumann algebra equipped with a faithful semifinite normal weight ϕ and $\mathcal{N}$ be a von Neumann subalgebra of $\mathcal{M}$ such that the restriction of ϕ to $\mathcal{N}$ is semifinite and such that $\mathcal{N}$ is invariant by the modular group of ϕ. Let $\mathcal{E}$ be the weight preserving conditional expectation from $\mathcal{M}$ onto $\mathcal{N}$. We prove the following inequality:

$$\|x\|_p^2 \geq \|\mathcal{E}(x)\|_p^2 + (p-1)\|x - \mathcal{E}(x)\|_p^2, \qquad x \in L_p(\mathcal{M}), 1 < p \leq 2,$$

which extends the celebrated Ball–Carlen–Lieb convexity inequality. As an application we show that there exists $\varepsilon_0 > 0$ such that for any free group $\mathbb{F}_n$ and any $q \geq 4 - \varepsilon_0$,

$$\|P_t\|_{2\to q} \leq 1 \quad \Leftrightarrow \quad t \geq \log\sqrt{q-1},$$

where (P_t) is the Poisson semigroup defined by the natural length function of $\mathbb{F}_n$.

1. Introduction. Let $\mathcal{M}$ be a von Neumann algebra equipped with a faithful semifinite normal weight ϕ. The associated noncommutative L_p-spaces will be simply denoted by $L_p(\mathcal{M})$. We refer to [11] for information on noncommutative integration. Recall that if $\mathcal{N}$ is a von Neumann subalgebra of $\mathcal{M}$ such that the restriction of ϕ to $\mathcal{N}$ is semifinite and such that $\mathcal{N}$ is σ^ϕ-invariant [i.e., $\sigma_t^\phi(\mathcal{N}) = \mathcal{N}$ for all $t \in \mathbb{R}$], then there exists a unique ϕ-preserving conditional expectation $\mathcal{E}$ from $\mathcal{M}$ onto $\mathcal{N}$ such that

$$\mathcal{E}(axb) = a\mathcal{E}(x)b, \qquad a, b \in \mathcal{N}, x \in \mathcal{M}.$$

Here σ^ϕ denotes the modular group of ϕ. Moreover, $\mathcal{E}$ extends to a contractive projection from $L_p(\mathcal{M})$ onto $L_p(\mathcal{N})$ for any $1 \leq p < \infty$. Below is our main result.

THEOREM 1. *Let $\mathcal{M}$, $\mathcal{N}$ and $\mathcal{E}$ be as above. If $1 < p \leq 2$, then*

$$\|x\|_p^2 \geq \|\mathcal{E}(x)\|_p^2 + (p-1)\|x - \mathcal{E}(x)\|_p^2, \qquad x \in L_p(\mathcal{M}). \tag{1}$$

If $2 < p < \infty$, the inequality is reversed.

Received May 2014; revised November 2014.
[1] Supported by ANR-2011-BS01-008-01 and NSFC Grant No. 11271292.
MSC2010 subject classifications. Primary 46L51, 47A30; secondary 60G42, 81S25.
Key words and phrases. Noncommutative L_p-spaces, martingale convexity inequality, hypercontractivity, free groups.

Inequality (1) is a martingale convexity inequality. It is closely related to the celebrated convexity inequality of Ball, Carlen and Lieb [2] for the Schatten classes S_p. Namely, for $1 < p \le 2$, we have

$$\|x+y\|_p^2 + \|x-y\|_p^2 \ge 2\|x\|_p^2 + 2(p-1)\|y\|_p^2, \qquad x, y \in S_p. \tag{2}$$

In fact, it is easy to see that (2) is a special case of (1) by considering $\mathcal{M} = B(\ell_2) \oplus B(\ell_2)$. Conversely, the validity of (2) for any noncommutative L_p-spaces implies (1). Indeed, we will deduce (1) from the following:

THEOREM 2. *Let $\mathcal{M}$ be any von Neumann algebra. If $1 < p \le 2$, then*

$$\|x+y\|_p^2 + \|x-y\|_p^2 \ge 2\|x\|_p^2 + 2(p-1)\|y\|_p^2, \qquad x, y \in L_p(\mathcal{M}). \tag{3}$$

If $2 < p < \infty$, the inequality is reversed.

What is new and remarkable in (2) or (3) is the fact that $(p-1)$ is the best constant. In fact, if one allows a constant depending on p in place of $(p-1)$, then (3) is equivalent to the well-known results on the 2-uniform convexity of $L_p(\mathcal{M})$. We refer to [2] for more discussion on this point. The optimality of the constant $(p-1)$ has important applications to hypercontractivity in the noncommutative case. It is the key to the solution of Gross's longstanding open problem about the optimal hypercontractivity for Fermi fields by Carlen and Lieb [4]. It plays the same role in [3] and [9]. Note that the optimality of $(p-1)$ in (3) implies $(p-1)$ is also the best constant in (1). It seems that (1) with this best constant is new even in the commutative case.

Clearly, (2) implies (3) for injective $\mathcal{M}$ (or more generally, QWEP $\mathcal{M}$) since then $L_p(\mathcal{M})$ is finitely representable in S_p. The proof of (2) in [2] goes through a differentiation argument for the function $t \mapsto \|x+ty\|_p^p$ with self-adjoint x and y. It seems difficult to directly extend their argument to finite von Neumann algebras. The subtle point is the fact that to be able to differentiate the above function, one needs the invertibility of $x+ty$ for all $t \in [0,1]$ except possibly countably many of them. This invertibility is easily achieved in the matrix algebra case, that is, for $\mathcal{M} = \mathbb{M}_n$, the algebra of $n \times n$-matrices. Instead, we will use a pseudo-differentiation argument which is much less rigid than that of [2]. The main novelty in our argument can be simply explained as follows. We first cut the operator $x+ty$ by its spectral projections in order to reduce the general case to the invertible one; to do so we need $x+ty$ to be of full support for all $t \in [0,1]$. We then get this full support property for all t by adding to $x+ty$ an independent operator with diffuse spectral measure. Note that by standard perturbation argument, it is easy to insure the full support (or even the invertibility) of $x+ty$ for one t.

An iteration of Theorem 1 immediately implies the following inequality on noncommutative martingales.

COROLLARY 3. *Let $(\mathcal{M}_n)_{n\geq 0}$ be an increasing sequence of von Neumann subalgebras of $\mathcal{M}$ with w*-dense union in $\mathcal{M}$. Assume that each $\mathcal{M}_n$ is σ^ϕ-invariant, and $\phi|_{\mathcal{M}_n}$ is semifinite. Let $\mathcal{E}_n$ be the conditional expectation with respect to $\mathcal{M}_n$. Then for $1 < p \leq 2$,*

$$\|x\|_p^2 \geq \|\mathcal{E}_0(x)\|_p^2 + (p-1)\sum_{n\geq 1}\|\mathcal{E}_n(x) - \mathcal{E}_{n-1}(x)\|_p^2, \qquad x \in L_p(\mathcal{M}).$$

For $2 < p < \infty$, the inequality is reversed.

Another possible iteration is the following:

COROLLARY 4. *Let $(\mathcal{M}_n)_{1\leq n\leq N}$ be a family of von Neumann subalgebras of $\mathcal{M}$. Assume that each $\mathcal{M}_n$ is σ^ϕ-invariant and $\phi|_{\mathcal{M}_n}$ is semifinite. Let $\mathcal{E}_n^+$ be the conditional expectation with respect to $\mathcal{M}_n$ and $\mathcal{E}_n^- = \mathrm{Id} - \mathcal{E}_n^+$. Then for $1 < p \leq 2$,*

$$\|x\|_p^2 \geq \sum_{(\varepsilon_i)\in\{+,-\}^N} (p-1)^{|\{i|\varepsilon_i = -\}|}\left\|\left(\prod_{i=1}^N \mathcal{E}_i^{\varepsilon_i}\right)(x)\right\|_p^2, \qquad x \in L_p(\mathcal{M}).$$

For $2 < p < \infty$, the inequality is reversed.

Applying it to the case where $\mathcal{M} = L_\infty(\{\pm 1\}^N)$ and $\mathcal{M}_n$ is the subalgebra of functions independent of the nth variable, we deduce the classical hypercontractivity for the Walsh system (with operator valued coefficients). Similarly, taking $\mathcal{M}$ to be the Clifford algebra with N generators, we obtain the optimal hypercontractivity for Fermi fields as pointed out in [2, 4].

We end the paper with some applications to hypercontractivity for group von Neumann algebras. In particular for the Poisson semigroup of a free group, we obtain the optimal time for the hypercontractivity from L_2 to L_q for $q \geq 4$.

2. The proofs. We will prove Theorems 1 and 2. Using the Haagerup reduction theorem as in [7], one can reduce both theorems to the finite case. Thus throughout this section $\mathcal{M}$ will denote a von Neumann algebra equipped with a faithful tracial normal state τ. $L_p(\mathcal{M})$ is then constructed with respect to τ. We will first prove (3), then deduce (1) from it. $1 < p < 2$ will be fixed in the sequel.

As explained before, the proof of (3) will be done by a pseudo-differentiation argument. Recall that for a continuous function f from an interval I to $\mathbb{R}$ its pseudo-derivative of second order at t is

$$D^2 f(t) = \liminf_{h\to 0^+} \frac{f(t+h) + f(t-h) - 2f(t)}{h^2}.$$

This pseudo-derivative shares many properties of the second derivative. For instance, if $D^2 f$ is nonnegative on I, then f is convex. Indeed, by adding εt^2 to f

(with $\varepsilon > 0$), we can assume that $D^2 f(t)$ is positive for all t. If f was not convex, there would exist $t_0 < t_1$ in I such that the function $f - g$ takes a positive value at some point of (t_0, t_1), where g is the straight line joining the two points $(t_0, f(t_0))$ and $(t_1, f(t_1))$. So $f - g$ achieves a local maximum at a point $s \in (t_0, t_1)$. Consequently, $D^2 f(s) = D^2(f - g)(s) \leq 0$, which is a contradiction.

Our pseudo-differentiation argument consists in proving the following inequality for $x, y \in L_p(\mathcal{M})$:

$$D^2\|x + ty\|_p^2(0) \geq 2(p-1)\|y\|_p^2. \tag{$D^2_{x,y}$}$$

Here the differentiation is, of course, taken with respect to the variable t. The arguments from [2] can be adapted to give:

LEMMA 5. *Let $a, b \in \mathcal{M}$ be self-adjoint elements with a invertible. Then $(D^2_{a,b})$ holds.*

PROOF. As a is invertible in $\mathcal{M}$, $a + tb$ is also invertible for small t. Introduce an auxiliary function ψ on $\mathbb{R}$,

$$\psi(t) = \|a + tb\|_p^p = \tau\big(\big(a^2 + t(ab + ba) + t^2 b^2\big)^{p/2}\big).$$

ψ is differentiable in a neighborhood of the origin and

$$\psi'(t) = \frac{p}{2}\tau\big[\big(a^2 + t(ab + ba) + t^2 b^2\big)^{p/2-1}\big((ab + ba) + 2tb^2\big)\big].$$

As in [2] by functional calculus, the operator $(a^2 + t(ab + ba) + t^2 b^2)^{p/2-1}$ admits the following integral representation:

$$\begin{aligned} &\big(a^2 + t(ab + ba) + t^2 b^2\big)^{p/2-1} \\ &\qquad = c_p \int_0^\infty s^{p/2-1} \frac{1}{s + a^2 + t(ab + ba) + t^2 b^2}\, ds, \end{aligned} \tag{4}$$

where

$$c_p^{-1} = \int_0^\infty s^{p/2-1} \frac{1}{s + 1}\, ds.$$

Thus ψ is twice differentiable at $t = 0$ and

$$\begin{aligned} \psi''(0) &= p\tau\big(|a|^{p-2} b^2\big) \\ &\quad - c_p \int_0^\infty s^{p/2-1} \tau\left[\frac{1}{s + a^2}(ab + ba)\frac{1}{s + a^2}(ab + ba)\right] ds. \end{aligned} \tag{5}$$

It then follows that $\varphi = \psi^{2/p}$ is also twice differentiable at $t = 0$ and

$$\varphi''(0) = \frac{2}{p}\left(\frac{2}{p} - 1\right)\|a\|_p^{2-2p}\psi'(0)^2 + \frac{2}{p}\|a\|_p^{2-p}\psi''(0) \geq \frac{2}{p}\|a\|_p^{2-p}\psi''(0).$$

Hence $(D^2_{a,b})$ will be a consequence of

$$\frac{1}{p}\|a\|_p^{2-p}\psi''(0) \geq (p-1)\|b\|_p^2. \tag{6}$$

To prove the last inequality we claim that $\psi''(0)$ increases when a is replaced by $|a|$. Indeed, the trace inside the integral in (5) is equal to twice the following sum:

$$\tau\left[\frac{a}{s+a^2}b\frac{a}{s+a^2}b\right]+\tau\left[\frac{a^2}{s+a^2}b\frac{1}{s+a^2}b\right].$$

The second term above depends only on $|a|$ (recalling that a is self-adjoint). It remains to show that the first one increases when a is replaced by $|a|$. By decomposing a into its positive and negative parts, we see that the first term is equal to

$$\tau\left[\frac{a_+}{s+a_+^2}b\frac{a_+}{s+a_+^2}b\right]+\tau\left[\frac{a_-}{s+a_-^2}b\frac{a_-}{s+a_-^2}b\right]-2\tau\left[\frac{a_+}{s+a_+^2}b\frac{a_-}{s+a_-^2}b\right].$$

All above traces are nonnegative. Therefore, the above quantity increases when the subtraction is replaced by addition. Then tracing back the argument and noting that $|a| = a_+ + a_-$, we get the desired inequality

$$\tau\left[\frac{a}{s+a^2}b\frac{a}{s+a^2}b\right]\leq\tau\left[\frac{|a|}{s+a^2}b\frac{|a|}{s+a^2}b\right].$$

Returning back to (5), we deduce the claim. Thus in the following we will assume that a is a positive invertible element of $\mathcal{M}$.

The positivity of a will facilitate the calculation of $\psi''(0)$ as explained in [2]. Since $a+tb$ is positive for small t, we have

$$\psi(t) = \tau\big((a+tb)^p\big).$$

Thus for t close to 0,

$$\psi'(t) = p\tau\big((a+tb)^{p-1}b\big).$$

To calculate the second derivative we use the following integral representation:

$$(a+tb)^{p-1} = d_p\int_0^\infty s^{p-1}\left[\frac{1}{s}-\frac{1}{s+a+tb}\right]ds.$$

Consequently,

$$\psi''(0) = p\,d_p\int_0^\infty s^{p-1}\tau\left[\frac{1}{s+a}b\frac{1}{s+a}b\right]ds.$$

As shown in [2], the function

$$F : z\mapsto \tau\left[\frac{1}{s+z}b\frac{1}{s+z}b\right]$$

is convex on the positive cone of $\mathcal{M}$.

Let u be the unitary operator in the polar decomposition of b (as $\mathcal{M}$ is finite, the usual partial isometry in this decomposition can be chosen to be a self-adjoint unitary). Then clearly

$$F(z) = \frac{1}{2}\big(F(z) + F(uzu)\big) \geq F\Big(\frac{z + uzu}{2}\Big).$$

Now $z' = \frac{z+uzu}{2}$ commutes with u, so

$$F(z) \geq F(z') = \tau\Big[\frac{1}{s+z'}|b|\frac{1}{s+z'}|b|\Big].$$

Let $\mathcal{B}$ be the Abelian von Neumann subalgebra of $\mathcal{M}$ generated by b, and let $\mathcal{E}_b$ be the associated trace preserving conditional expectation. Then

$$F(z') = \tau\Big[\mathcal{E}_b\Big(\frac{1}{s+z'}|b|\frac{1}{s+z'}\Big)|b|\Big].$$

However, the Kadison–Schwarz inequality implies

$$\mathcal{E}_b\Big(\frac{1}{s+z'}|b|\frac{1}{s+z'}\Big) \geq \mathcal{E}_b\Big(\frac{1}{s+z'}|b|^{1/2}\Big)\mathcal{E}_b\Big(|b|^{1/2}\frac{1}{s+z'}\Big) = \mathcal{E}_b\Big(\frac{1}{s+z'}\Big)^2|b|.$$

Hence, by the positivity of the trace on products of positive elements, we deduce

$$F(z) \geq \tau\Big[\mathcal{E}_b\Big(\frac{1}{s+z'}\Big)^2|b|^2\Big].$$

Then by the operator convexity of $\frac{1}{t}$, we have

$$\mathcal{E}_b\Big(\frac{1}{s+z'}\Big) \geq \frac{1}{s+\mathcal{E}_b(z')} = \frac{1}{s+\mathcal{E}_b(z)}.$$

Letting $\widetilde{a} = \mathcal{E}_b(a)$, we have just shown

$$\psi''(0) \geq p\, d_p \int_0^\infty s^{p-1}\tau\Big[\frac{1}{s+\widetilde{a}}b\frac{1}{s+\widetilde{a}}b\Big]ds = \widetilde{\psi}''(0), \tag{7}$$

where

$$\widetilde{\psi}(t) = \tau\big((\widetilde{a}+tb)^p\big).$$

Finally, (6) immediately follows from (7). Indeed, by (7) and the Hölder inequality,

$$\frac{1}{p}\|a\|_p^{2-p}\psi''(0) \geq \frac{1}{p}\|\widetilde{a}\|_p^{2-p}\widetilde{\psi}''(0) = (p-1)\|\widetilde{a}\|_p^{2-p}\tau\big(\widetilde{a}^{p-2}b^2\big) \geq (p-1)\|b\|_p^2.$$

This finishes the proof of the lemma. □

For $a \in \mathcal{M}$ self-adjoint, we denote by $s(a) = \mathbb{1}_{(0,\infty)}(|a|)$. $s(a)$ is the support of a, that is, the least projection e of $\mathcal{M}$ such that $ea = a$. We say that a has full support if $s(a) = 1$.

LEMMA 6. *Let $a, b \in \mathcal{M}$ be self-adjoint with $s(a) = 1$. Then $(D^2_{a,b})$ holds.*

PROOF. We will reduce this lemma to the previous one by cutting $a + tb$ with the spectral projections of a. Let e be a nonzero spectral projection of a, and put $a_e = eae$ and $b_e = ebe$. Since a is of full support, a_e is invertible in the reduced von Neumann algebra $\mathcal{M}_e = e\mathcal{M}e$. Thus Lemma 5 can be applied to the couple (a_e, b_e) in $\mathcal{M}_e$. Let $\psi_e(t) = \|a_e + tb_e\|_p^2$ as before. ψ_e is twice differentiable at $t = 0$, and (6) holds with ψ_e in place of ψ.

Let $e^\perp = 1 - e$. Then for t in a neighborhood of the origin, we have [recalling that $\varphi(t) = \|a + tb\|_p^2$]

$$\varphi(t) \geq \big(\|e(a + tb)e\|_p^p + \|e^\perp(a + tb)e^\perp\|_p^p\big)^{2/p} \stackrel{\text{def}}{=} \big(\psi_e(t) + \gamma_e(t)\big)^{2/p}.$$

However,

$$\psi_e(t) = \|a_e\|_p^p + t\psi_e'(0) + \frac{t^2}{2}\psi_e''(0) + \mathrm{o}(t^2) \qquad \text{as } t \to 0.$$

Let

$$\alpha(t) = \|a_e\|_p^p + \gamma_e(t) = \|a_e + e^\perp(a + tb)e^\perp\|_p^p = \|a + te^\perp be^\perp\|_p^p.$$

Then

$$\begin{aligned}
\varphi(t) &\geq \left(\alpha(t) + t\psi_e'(0) + \frac{t^2}{2}\psi_e''(0) + \mathrm{o}(t^2)\right)^{2/p} \\
&= \alpha(t)^{2/p}\left(1 + \frac{2t}{p}\frac{\psi_e'(0)}{\alpha(t)} + \frac{t^2}{p}\frac{\psi_e''(0)}{\alpha(t)} + \frac{1}{p}\left(\frac{2}{p} - 1\right)t^2\frac{\psi_e'(0)^2}{\alpha(t)^2} + \mathrm{o}(t^2)\right) \\
&= \alpha(t)^{2/p} + \frac{2t}{p}\psi_e'(0)\alpha(t)^{2/p-1} + \frac{t^2}{p}\psi_e''(0)\alpha(t)^{2/p-1} \\
&\quad + \frac{1}{p}\left(\frac{2}{p} - 1\right)t^2\psi_e'(0)^2\alpha(t)^{2/p-2} + \mathrm{o}(t^2) \\
&\geq \alpha(t)^{2/p} + \frac{2t}{p}\psi_e'(0)\alpha(t)^{2/p-1} + \frac{t^2}{p}\psi_e''(0)\alpha(t)^{2/p-1} + \mathrm{o}(t^2).
\end{aligned}$$

By convexity of norms,

$$\alpha(t)^{2/p} + \alpha(-t)^{2/p} \geq 2\|a\|_p^2 = 2\varphi(0).$$

We then deduce that

$$\begin{aligned}
&\frac{\varphi(t) + \varphi(-t) - 2\varphi(0)}{t^2} \\
&\quad \geq \frac{2}{p}\psi_e'(0)\frac{\alpha(t)^{2/p-1} - \alpha(-t)^{2/p-1}}{t} \\
&\qquad + \frac{1}{p}\psi_e''(0)\big[\alpha(t)^{2/p-1} + \alpha(-t)^{2/p-1}\big] + \mathrm{o}(1).
\end{aligned}$$

The uniform smoothness of the norm $\| \|_p$ implies that the function $\alpha^{2/p-1}$ is differentiable at $t = 0$, and its derivative is equal to

$$(2-p)\|a\|_p^{1-p}\tau\big(v|a|^{p-1}e^{\perp}be^{\perp}\big)\overset{\text{def}}{=}\delta_e,$$

where v is the unitary in the polar decomposition of a. It then follows that

$$D^2\varphi(0) \geq \frac{4}{p}\psi_e'(0)\delta_e + \frac{2}{p}\psi_e''(0)\|a\|_p^{2-p}.$$

Hence by (6),

$$D^2\varphi(0) \geq \frac{4}{p}\psi_e'(0)\delta_e + 2(p-1)\|b_e\|_p^2.$$

Thanks to the full support assumption of a, we can let $e \to 1$ in the above inequality. This limit procedure removes the first extra term, so we finally get

$$D^2\varphi(0) \geq 2(p-1)\|b\|_p^2. \qquad \square$$

Now we are ready to show (3).

PROOF OF THEOREM 2. First by density, we need only to show (3) for $x, y \in \mathcal{M}$. Then notice that it suffices to do it for self-adjoint elements using a classical 2×2-matrix trick. Indeed, let $\widetilde{\mathcal{M}} = \mathbb{M}_2 \otimes \mathcal{M}$ equipped with the tensor trace. Given $x, y \in \mathcal{M}$ let

$$a = \begin{pmatrix} 0 & x \\ x^* & 0 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 0 & y \\ y^* & 0 \end{pmatrix}.$$

Then a and b are self-adjoint. Moreover, by easy computations, (3) for x and y is equivalent to the same inequality for a and b.

To use Lemma 6, we require that $a + tb$ have full support for any $t \in \mathbb{R}$. This is achieved by a tensor product argument. Choose a positive element $c \in L_\infty([0, 1])$ whose spectral measure with respect to Lebesgue measure is diffuse (atomless), say $c(t) = t$ for $t \in [0, 1]$. In other words, considered as a random variable in the probability space $[0, 1]$, the law of c is diffuse. On the other hand, for $t \in \mathbb{R}$, composing the spectral resolution of $a+tb$ with the trace τ, we can view $a+tb$ as a random variable in another probability space (Ω, P). Now, consider the tensor von Neumann algebra $L_\infty([0, 1]) \overline{\otimes} \mathcal{M}$; it is finite. $\mathcal{M}$ and $L_\infty([0, 1])$ are identified as subalgebras of $L_\infty([0, 1]) \overline{\otimes} \mathcal{M}$ in the usual way. Then for any $\varepsilon > 0$, the law of $a + tb + \varepsilon c$ is the convolution of the laws of $a + tb$ and εc. It is atomless since the law of c is atomless. Consequently, the support of $a + tb + \varepsilon c$ is full in $L_\infty([0, 1]) \overline{\otimes} \mathcal{M}$.

Thus by Lemma 6 applied to the pair $(a + \varepsilon c, b)$, the function $f(t) = \|a + tb + \varepsilon c\|_p^2 - (p-1)t^2\|b\|_p^2$ satisfies $D^2(f)(t) \geq 0$, so it is convex. Hence $f(1) + f(-1) \geq 2f(0)$; this is (3) for $a + \varepsilon c$ and b. Letting $\varepsilon \to 0$ gives the desired result. □

Finally, we deduce (1) from (3).

PROOF OF THEOREM 1. Given $x \in L_p(\mathcal{M})$ let $a = \mathcal{E}(x)$ and $b = x - \mathcal{E}(x)$. Consider again the function f defined by

$$f(t) = \|a + tb\|_p^2 - (p-1)t^2\|b\|_p^2.$$

Then (3) implies $D^2 f \geq 0$, so f is convex. On the other hand, the function $g(t) = \|a + tb\|_p^2$ is also convex and by the contractivity of $\mathcal{E}$ on $L_p(\mathcal{M})$,

$$\|a + tb\|_p \geq \big\|\mathcal{E}(a + tb)\big\|_p = \|a\|_p.$$

Hence we conclude that the right derivative $g_r'(0) \geq 0$, so that $f_r'(0) \geq 0$, too. Consequently, f is increasing on $\mathbb{R}^+$. In particular, $f(1) \geq f(0)$, which is nothing but (1). □

3. Applications to hypercontractivity. We give in this section some applications to hypercontractivity inequalities on group von Neumann algebras. Let G be a discrete group and $vN(G)$ the associated group von Neumann algebra. Recall that $vN(G)$ is the von Neumann algebra generated by the left regular representation λ: $vN(G) = \lambda(G)'' \subset B(\ell_2(G))$. It is equipped with a canonical trace τ, that is, $\tau(x) = \langle xe, e\rangle$, where e is the identity of G. Given a function $\psi : G \to \mathbb{R}_+$ with $\psi(e) = 0$, we consider the associated Fourier–Schur multiplier initially defined on the family $\mathbb{C}[G]$ of polynomials on G:

$$P_t : \sum_{g\in G} x(g)\lambda(g) \mapsto \sum_{g\in G} e^{-t\psi(g)} x(g)\lambda(g), \qquad t > 0.$$

We will assume that P_t extends to a contraction on $L_p(vN(G))$ for every $1 \leq p \leq \infty$. Schoenberg's classical theorem asserts that if ψ is symmetric and conditionally negative, P_t is a completely positive map on $vN(G)$. Since it is trace preserving, P_t defines a contraction on $L_p(vN(G))$ for every $1 \leq p \leq \infty$. Thus in this case our assumption is satisfied.

The hypercontractivity problem for the semigroup $(P_t)_{t>0}$ and for $1 < p < q < \infty$, consists in determining the optimal time $t_{p,q} > 0$ such that

$$\|P_t\|_{p\to q} \leq 1 \qquad \forall t \geq t_{p,q}.$$

We refer to [8, 9] for more information and historical references. It is easy to check that if such a time $t_{p,q}$ exists, then ψ has a spectral gap, namely $\inf_{g\in G\setminus\{e\}} \psi(g) > 0$. After rescaling, we will assume that $\inf_{g\in G\setminus\{e\}} \psi(g) = 1$.

In most-known cases the expected optimal time $t_{p,q}$ is attained, namely,

$$t_{p,q} = \log\sqrt{\frac{q-1}{p-1}}.$$

It is a particularly interesting problem of determining the optimal time $t_{p,q}$ when $G = \mathbb{F}_n$ is the free group on n generators with $n \in \mathbb{N} \cup \{\infty\}$, and ψ is its natural length function. Some partial results are obtained in [8, 9]. For instance, by embedding $vN(\mathbb{F}_n)$ into a free product of Clifford algebras, it is proved in [9] that for any $q > 2$,

$$t_{2,q} \leq \log\sqrt{q-1} + \left(\frac{1}{2} - \frac{1}{q}\right)\log\sqrt{2}.$$

On the other hand, Junge et al. [8] show that for any finite n there exists $q(n)$ such that if $q \geq q(n)$ is an even integer, then

$$t_{2,q} = \log\sqrt{q-1}.$$

The proof is combinatoric and based on lengthy calculations.

Here we provide an improvement. We will use Haagerup-type inequalities [6]. Letting S_k be the set of words of length k in $\mathbb{F}_n$ and for any $x \in vN(\mathbb{F}_n)$ supported on S_k, the original Haagerup inequality is

$$\|x\|_\infty \leq (k+1)\|x\|_2. \tag{8}$$

For $q > 2$ and $k \in \mathbb{N}$, let $K_{k,q}$ be the best constant in the following Khintchine inequality for homogeneous polynomials x of degree k:

$$\left\|\sum_{g\in S_k} x(g)\lambda(g)\right\|_q \leq K_{k,q}\left\|\sum_{g\in S_k} x(g)\lambda(g)\right\|_2.$$

We will need the following:

LEMMA 7. *We have* $K_{k,4} \leq (k+1)^{1/4}$.

PROOF. Denote by g_i the generators of $\mathbb{F}_n$ with the convention that $g_{-i} = g_i^{-1}$. For a multi-index $\underline{i} = (i_1, \ldots, i_d)$ with $i_j + i_{j+1} \neq 0$, we let $g_{\underline{i}} = g_{i_1} \cdots g_{i_d}$ and $|\underline{i}| = d$. So we may write

$$x = \sum_{|\underline{i}|=k} \alpha_{\underline{i}}\lambda(g_{\underline{i}}).$$

We compute x^*x according to simplifications that may occur

$$x^*x = \sum_{0\leq d\leq k} \sum_{\substack{|\underline{\beta}|=d \\ |\underline{i}|=|\underline{j}|=k-d \\ i_{k-d}\neq j_{k-d}\neq -\beta_1}} \overline{\alpha_{\underline{j},\underline{\beta}}}\alpha_{\underline{i},\underline{\beta}}\lambda\big(g_{\underline{j}}^{-1}g_{\underline{i}}\big),$$

where $\underline{i}, \underline{\beta}$ denotes the multi-index obtained by superposing the two multi-indices $\underline{i}$ and $\underline{\beta}$. We then deduce that

$$\begin{aligned}
\|x\|_4^4 = \|x^*x\|_2^2 &= \sum_{0\le d\le k} \sum_{\substack{|\underline{i}|=|\underline{j}|=k-d \\ i_{k-d}\neq j_{k-d}}} \Biggl(\sum_{\substack{|\underline{\beta}|=d \\ i_{k-d}\neq j_{k-d}\neq -\beta_1}} \overline{\alpha_{\underline{j},\underline{\beta}}} \alpha_{\underline{i},\underline{\beta}} \Biggr)^2 \\
&\le \sum_{0\le d\le k} \sum_{\substack{|\underline{i}|=|\underline{j}|=k-d \\ i_{k-d}\neq j_{k-d}}} \Biggl(\sum_{\substack{|\underline{\beta}|=d \\ i_{k-d}\neq -\beta_1}} |\alpha_{\underline{i},\underline{\beta}}|^2 \Biggr) \cdot \Biggl(\sum_{\substack{|\underline{\beta}|=d \\ j_{k-d}\neq -\beta_1}} |\alpha_{\underline{j},\underline{\beta}}|^2 \Biggr) \\
&\le \sum_{0\le d\le k} \Biggl(\sum_{\substack{|\underline{i}|=k-d, |\underline{\beta}|=d \\ i_{k-d}\neq -\beta_1}} |\alpha_{\underline{i},\underline{\beta}}|^2 \Biggr) \cdot \Biggl(\sum_{\substack{|\underline{j}|=k-d, |\underline{\beta}|=d \\ j_{k-d}\neq -\beta_1}} |\alpha_{\underline{j},\underline{\beta}}|^2 \Biggr) \\
&= (k+1)\|x\|_2^4. \qquad \square
\end{aligned}$$

REMARK 8. Taking $\alpha_{\underline{i}} = 1$ and by the free central limit theorem as $n \to \infty$, one can see that the previous inequality is sharp. Thus $K_{k,4} = (k+1)^{1/4}$. This constant is the L_4-norm of the kth Chebyshev polynomial for the semi-circle law.

Using the Hölder inequality we deduce from (8) and the previous lemma that for any $k \ge 1$,

$$K_{k,q} \le (k+1)^{1-3/q}, \qquad q \ge 4, \tag{9}$$

$$K_{k,q} \le (k+1)^{1/2-1/q}, \qquad 2 \le q \le 4. \tag{10}$$

We will also use the following elementary folklore:

REMARK 9. Let $T : L_p(\mathcal{M}) \to L_q(\mathcal{M})$ be a bounded linear map. Assume that T is 2-positive in the sense that $\mathrm{Id}_{\mathbb{M}_2} \otimes T$ maps the positive cone of $L_p(\mathbb{M}_2 \otimes \mathcal{M})$ to that of $L_q(\mathbb{M}_2 \otimes \mathcal{M})$. Then

$$\|T(x)\|_q \le \|T(|x|)\|_q^{1/2} \|T(|x^*|)\|_q^{1/2}, \qquad x \in L_p(\mathcal{M}).$$

Consequently,

$$\|T\| = \sup\{\|T(x)\|_q : x \in L_p(\mathcal{M})^+, \|x\|_p \le 1\}.$$

Indeed, for any $x \in L_p(\mathcal{M})$,

$$\begin{pmatrix} |x| & x \\ x^* & |x^*| \end{pmatrix} \ge 0.$$

So the 2-positivity of T implies

$$\begin{pmatrix} T(|x|) & T(x) \\ T(x^*) & T(|x^*|) \end{pmatrix} \ge 0.$$

This yields a contraction $c \in \mathcal{M}$ such that $T(x) = T(|x|)^{1/2} c T(|x^*|)^{1/2}$. Then the Hölder inequality gives the assertion.

THEOREM 10. *There exists $\varepsilon_0 > 0$ such that for any free $\mathbb{F}_n$ and any $q \geq 4 - \varepsilon_0$,*

$$\|P_t\|_{2\to q} \leq 1 \quad \Leftrightarrow \quad t \geq \log\sqrt{q-1}.$$

PROOF. The necessity is clear. The proof of the sufficiency will rely on Remark 3.7 of [9]. Let σ be the automorphism of $vN(\mathbb{F}_n)$ given by $\sigma(\lambda(g_i)) = \lambda(g_i^{-1})$. Then P_t is hypercontractive from L_2 to L_q with optimal time on $vN(\mathbb{F}_n)^\sigma$, the fixed point algebra of σ. Let $\mathcal{E}$ be the conditional expectation onto $vN(\mathbb{F}_n)^\sigma$. Note that $\mathcal{E} = \frac{\mathrm{Id}+\sigma}{2}$ and it commutes with P_t.

Fix $q > 2$. To prove $\|P_t\|_{2\to q} \leq 1$ for $t \geq \log\sqrt{q-1}$, it suffices to show $\|P_t(x)\|_q \leq \|x\|_2$ for any positive $x \in \mathbb{C}[\mathbb{F}_n]$ by virtue of Remark 9. We need one more reduction. Given complex numbers ζ_i of modulus 1, there exists an automorphism π_ζ of $vN(\mathbb{F}_n)$ given by $\pi(g_i) = \zeta_i g_i$. It is an isometry on all L_p's. Note that π_ζ and P_t commute. Thus to prove $\|P_t(x)\|_q \leq \|x\|_2$, we may assume that $x(g_i)$ is real for every generator g_i. We will fix a positive $x \in \mathbb{C}[\mathbb{F}_n]$ with the last property.

Then write $x = y + z$ where $y = \mathcal{E}(x)$. Since $x(g_i) \in \mathbb{R}$, we have that z does not have constant terms nor of degree 1. By Theorem 1 (or Theorem 3) and Remark 3.7 of [9],

$$\|P_t(x)\|_q^2 \leq \|P_t(y)\|_q^2 + (q-1)\|P_t(z)\|_q^2 \leq \|y\|_2^2 + (q-1)\|P_t(z)\|_q^2.$$

Then for $t = \log\sqrt{q-1}$, decomposing z according to its homogeneous components (z_k) and using the Khintchine and the Cauchy–Schwarz inequalities, we get

$$\|P_t(z)\|_q^2 \leq \left(\sum_{k\geq 2} e^{-tk}\|z_k\|_q\right)^2 \leq \sum_{k\geq 2} K_{k,q}^2 \frac{1}{(q-1)^k}\|z\|_2^2.$$

We aim to find those $q > 2$ for which

$$R_q = (q-1)\sum_{k\geq 2} K_{k,q}^2 \frac{1}{(q-1)^k} \leq 1.$$

For $q \geq 4$, by (9) we have

$$R_q \leq \sum_{k\geq 2}(k+1)^{2(1-3/q)}\frac{1}{(q-1)^{k-1}}.$$

The terms of the sum on the right-hand side are decreasing functions of q if their derivatives are negative, that is, if

$$\frac{6(q-1)}{q^2} \leq \frac{k-1}{\log(k+1)}.$$

Noting that the left-hand side of the above inequality is decreasing on q, one easily checks that this inequality is true for $q \ge 4$ and $k \ge 3$. However, it is true for $k = 2$ if and only if $q \ge q_0$, where

$$q_0 = \sqrt{3\log 3}\big(\sqrt{3\log 3} + \sqrt{3\log 3 - 2}\big) \approx 5.36244.$$

We have the following numerical estimates:

$$R_4 \le 0.92952 \quad \text{and} \quad \frac{3^{2(1-3/q_0)}}{q_0 - 1} - \frac{3^{1/2}}{3} \le 0.02613.$$

Hence if $q \in [4, q_0]$,

$$R_q \le R_4 + \frac{3^{2(1-3/q_0)}}{q_0 - 1} - \frac{3^{1/2}}{3} < 1.$$

We thus conclude that $R_q < 1$ for all $q \ge 4$.

Since R_q is dominated by a continuous function of q, using (10) we get a similar estimate for $q \ge 4 - \varepsilon_0$ for some ε_0. A numerical estimate gives $\varepsilon_0 \approx 0.18$. $\square$

REMARK 11. Instead of Remark 3.7 of [9], we can equally use Theorem A(iii) of [9] in the preceding proof. But the commutation of P_t and the conditional expectation onto the symmetric subalgebra $\mathcal{A}^n_{\mathrm{sym}}$ is less obvious.

It is likely that $\varepsilon_0 = 2$, but other methods would have to be developed.

Gross's pioneering work [5] shows that hypercontractivity is equivalent to the validity of log-Sobolev inequalities. In the present situation of free groups, the validity of the hypercontractivity with optimal time in full generality (or equivalently, $\varepsilon_0 = 2$) is equivalent to the following log-Sobolev inequality in L_q for any $q \ge 2$:

$$(\mathrm{SL}_q)\quad \tau\big(x^q \log x\big) \le \frac{q}{2(q-1)}\tau\big(x^{q-1}L(x)\big) + \|x\|_q^q \log \|x\|_q, \qquad x \in \mathcal{D}^+.$$

Here L denotes the negative generator of (P_t), and $\mathcal{D}$ is a core for L where the inequality makes sense. It is known that (SL_2) implies (SL_q) for all q; see [10]. In the same spirit we can show that (SL_p) implies (SL_q) if $q > p \ge 2$. Let us record this explicitly here since it might be of interest. The semigroup (P_t) can be any completely positive symmetric Markovian semigroup such that $\mathcal{D}$ is rich enough.

REMARK 12. Let $q > p \ge 2$. Then (SL_p) implies (SL_q).

To check the remark we rewrite (SL_q) in a symmetric form with respect to q and its conjugate index q' (provided that $\mathcal{D}$ is big enough):

$$(\mathrm{SL}_q^s)\qquad \tau(x\log x) \le \tfrac{1}{2}q'q\tau\big(x^{1/q'}L\big(x^{1/q}\big)\big) + \tau(x)\log\tau(x), \qquad x \in \mathcal{D}^+.$$

Recall that for $y \in \mathrm{Dom}(L)$, $\tau(zL(y)) = \lim_{r\to 0}\frac{1}{r}\tau(z(1-P_r)(y))$. Let $r > 0$ and $x \in \mathcal{M}^+$, and we will check that the function $q \mapsto q'q\tau(x^{1/q'}(1-P_r)(x^{1/q}))$ is increasing for $q \geq 2$; we put $\theta = \frac{1}{q}$. It is known from [1] that there exists a positive symmetric Borel measure μ_r on $\sigma(x) \times \sigma(x)$ such that

$$\tau\big(x^{1-\theta} P_r(x^\theta)\big) = \int_{\sigma(x)\times\sigma(x)} s^{1-\theta} t^\theta \, d\mu_r(s,t).$$

Hence, by symmetry, it suffices to show that

$$f : \theta \mapsto \frac{1 + u - u^\theta - u^{1-\theta}}{\theta(1-\theta)}$$

is convex on [0, 1] for $u > 0$ as $f(\theta) = f(1-\theta)$. One easily checks that

$$f(\theta) = \int_0^1 \log(u)\big(u^{\theta+(1-\theta)(1-t)} - u^{\theta t} + u^{1-\theta+\theta t} - u^{(1-\theta)(1-t)}\big)\, dt,$$

$$f''(\theta) = \int_0^1 \log(u)^3 \big(t^2\big(u^{\theta+(1-\theta)(1-t)} - u^{\theta t}\big) + (1-t)^2\big(u^{1-\theta+\theta t} - u^{(1-\theta)(1-t)}\big)\big)\, dt \geq 0.$$

Passing to the limit in r gives the result if $\mathcal{D}$ is big enough.

We end this section with application to more general groups (G, ψ). If ψ is symmetric and satisfies the exponential order growth

$$\big|\{g \in G : \psi(g) \leq R\}\big| \leq C\rho^R \qquad \forall R > 0 \tag{11}$$

for some $C > 0$ and $\rho > 1$, then one of the main results of [8] shows that for $2 < q < \infty$,

$$t_{2,q} \leq \eta \log\sqrt{q-1}$$

for any $\eta > 2$ when ρ is large compared to C. Their argument consists in first considering the case $q = 4$ by combinatoric methods and then using Gross's extrapolation. We will show that the martingale inequality in Theorem 1 easily implies a slight improvement. Note that our estimate on $t_{2,q}$ is as close as to the expected optimal time as when q is sufficiently large, compared to ρ and C.

PROPOSITION 13. *Assume* (11) *and* $2 < q < \infty$. *Then*

$$t_{2,q} \leq \left(\frac{q-2}{q}\log\sqrt{2C\rho} + \log\sqrt{q-1}\right) \vee \log\rho.$$

PROOF. By (11), the range of ψ is countable. Let $\psi(G) = \{n_0, n_1, n_2, \ldots\}$ with $n_0 < n_1 < n_2 < \cdots$. Then $n_0 = 0$ and $n_1 = 1$. Let $x \in vN(G)$ be a polynomial, $x = \sum x(g)\lambda(g)$, and let $y = x - x(e)$. By Theorem 1

$$\|P_t(x)\|_q^2 \leq |x(e)|^2 + (q-1)\|P_t(y)\|_q^2.$$

Let $B_k = \{g \in G : \psi(g) \leq n_k\}$, $S_k = B_k \setminus B_{k-1}$ and $y_k = \sum_{g \in S_k} x(g)\lambda(g)$. Then

$$\|P_t(y)\|_\infty^2 \leq \left(\sum_{k\geq 1} e^{-tn_k}\|y_k\|_\infty\right)^2 \leq \left(\sum_{k\geq 1} e^{-2tn_k}|S_k|\right)\cdot\left(\sum_{k\geq 1}\frac{\|y_k\|_\infty^2}{|S_k|}\right).$$

Obviously,

$$\|y_k\|_\infty^2 \leq \left(\sum_{g\in S_k}|x(g)|\right)^2 \leq |S_k|\sum_{g\in S_k}|x(g)|^2.$$

We get, using the Hölder inequality,

$$\|P_t(y)\|_q^2 \leq e^{-4t/q}\left(\sum_{k\geq 1} e^{-2tn_k}|S_k|\right)^{(q-2)/q}\|y\|_2^2.$$

Actually exchanging the arguments, one has the following, slightly better estimate that we will not use:

$$\|P_t(y)\|_q^2 \leq \sum_{k\geq 1} e^{-2tn_k}|S_k|^{2(q-2)/q}\|y\|_2^2.$$

By (11), for $t > \log\rho$,

$$\begin{aligned}\sum_{k\geq 1} e^{-2tn_k}|S_k| &\leq \sum_{k\geq 1}(e^{-2tn_k} - e^{-2tn_{k+1}})|B_k| \leq 2Ct\int_1^\infty e^{-(2t-\log\rho)s}\,ds\\ &= 2C\frac{t}{2t-\log\rho}e^{-(2t-\log\rho)} \leq 2Ce^{-(2t-\log\rho)}.\end{aligned}$$

Hence, if $2t \geq \frac{q-2}{q}\log(2C\rho) + \log(q-1)$, we deduce $\|P_t(x)\|_q \leq \|x\|_2$, whence the assertion. $\square$

REFERENCES

[1] ALBEVERIO, S. and HØEGH-KROHN, R. (1977). Dirichlet forms and Markov semigroups on C^*-algebras. *Comm. Math. Phys.* **56** 173–187. MR0461153

[2] BALL, K., CARLEN, E. A. and LIEB, E. H. (1994). Sharp uniform convexity and smoothness inequalities for trace norms. *Invent. Math.* **115** 463–482. MR1262940

[3] BIANE, P. (1997). Free hypercontractivity. *Comm. Math. Phys.* **184** 457–474. MR1462754

[4] CARLEN, E. A. and LIEB, E. H. (1993). Optimal hypercontractivity for Fermi fields and related noncommutative integration inequalities. *Comm. Math. Phys.* **155** 27–46. MR1228524

[5] GROSS, L. (1975). Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** 1061–1083. MR0420249

[6] HAAGERUP, U. (1978/79). An example of a nonnuclear C^*-algebra, which has the metric approximation property. *Invent. Math.* **50** 279–293. MR0520930

[7] HAAGERUP, U., JUNGE, M. and XU, Q. (2010). A reduction method for noncommutative L_p-spaces and applications. *Trans. Amer. Math. Soc.* **362** 2125–2165. MR2574890

[8] JUNGE, M., PALAZUELOS, C., PARCET, J. and PERRIN, M. (2013). Hypercontractivity in group von Neumann algebras. Preprint.

[9] JUNGE, M., PALAZUELOS, C., PARCET, J., PERRIN, M. and RICARD, É. (2015). Hypercontractivity for free products. *Ann. Sci. École Norm. Sup.* (4) **48** 861–889. MR3377067

[10] OLKIEWICZ, R. and ZEGARLINSKI, B. (1999). Hypercontractivity in noncommutative L_p spaces. *J. Funct. Anal.* **161** 246–285. MR1670230

[11] PISIER, G. and XU, Q. (2003). Non-commutative L^p-spaces. In *Handbook of the Geometry of Banach Spaces*, *Vol.* 2 (W. B. Johnson and J. Lindenstrauss, eds.) 1459–1517. North-Holland, Amsterdam. MR1999201

LABORATOIRE DE MATHÉMATIQUES NICOLAS ORESME
UNIVERSITÉ DE CAEN BASSE-NORMANDIE
14032 CAEN CEDEX
FRANCE
E-MAIL: eric.ricard@unicaen.fr

SCHOOL OF MATHEMATICS AND STATISTICS
WUHAN UNIVERSITY
WUHAN 430072
CHINA
AND
LABORATOIRE DE MATHÉMATIQUES
UNIVERSITÉ DE FRANCHE-COMTÉ
25030 BESANÇON CEDEX
FRANCE
E-MAIL: qxu@univ-fcomte.fr