TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 369, Number 7, July 2017, Pages 4843–4863
http://dx.doi.org/10.1090/tran/6916
Article electronically published on February 13, 2017

STRONG CONTRACTION AND INFLUENCES IN TAIL SPACES

STEVEN HEILMAN, ELCHANAN MOSSEL, AND KRZYSZTOF OLESZKIEWICZ

ABSTRACT. We study contraction under a Markov semi-group and influence bounds for functions in L^2 tail spaces, i.e., functions all of whose low level Fourier coefficients vanish. It is natural to expect that certain analytic inequalities are stronger for such functions than for general functions in L^2. In the positive direction we prove an L^p Poincaré inequality and moment decay estimates for mean 0 functions and for all $1 < p < \infty$, proving the degree one case of a conjecture of Mendel and Naor as well as the general degree case of the conjecture when restricted to Boolean functions. In the negative direction, we answer negatively two questions of Hatami and Kalai concerning extensions of the Kahn-Kalai-Linial and Harper Theorems to tail spaces. That is, we construct a function $f\colon \{-1,1\}^n \to \{-1,1\}$ whose Fourier coefficients vanish up to level $c\log n$, with all influences bounded by $C\log n/n$ for some constants $0 < c, C < \infty$. We also construct a function $f\colon \{-1,1\}^n \to \{0,1\}$ with nonzero mean whose remaining Fourier coefficients vanish up to level $c'\log n$, with the sum of the influences bounded by $C'(\mathbb{E}f)\log(1/\mathbb{E}f)$ for some constants $0 < c', C' < \infty$.

1. INTRODUCTION

Consider the uniform measure on $\{-1,1\}^n$. Any $f\colon \{-1,1\}^n \to \mathbb{R}$ can be written as $f = \sum_{S\subseteq\{1,\ldots,n\}} \widehat{f}(S)W_S$, where for all $x = (x_1,\ldots,x_n) \in \{-1,1\}^n$, $W_S(x) := \prod_{i\in S} x_i$ and $\widehat{f}(S) := 2^{-n}\sum_{x\in\{-1,1\}^n} f(x)W_S(x)$. For any $t \geq 0$, define $P_t f := \sum_{S\subseteq\{1,\ldots,n\}} e^{-t|S|}\widehat{f}(S)W_S$, and define $Lf := \sum_{S\subseteq\{1,\ldots,n\}} |S|\,\widehat{f}(S)W_S$.

Our interest in this paper is in tail spaces. For the case of the uniform measure on $\{-1,1\}^n$, we are interested in the linear subspace of all functions satisfying $\widehat{f}(S) = 0$ for all $S \subseteq \{1,\ldots,n\}$ with $|S| \leq k$. Our interest in understanding such functions follows recent conjectures by Mendel and Naor and by Hatami and Kalai.

1.1. **Heat smoothing.**

In their study of a general notion of expander (with respect to all uniformly convex spaces), Mendel and Naor made the following conjecture.

Received by the editors November 13, 2014 and, in revised form, July 27, 2015.

2010 *Mathematics Subject Classification.* Primary 60E15, 47D07, 06E30.

The first author was supported by NSF Graduate Research Fellowship DGE-0813964 and a Simons-Berkeley Research Fellowship. Part of this work was completed while the first author was visiting the Network Science and Graph Algorithms program at ICERM.

The second author was supported by NSF grant DMS-1106999, NSF grant CCF 1320105 and DOD ONR grant N000141110140 and grant 328025 from the Simons Foundation.

The third author was supported by NCN grant DEC-2012/05/B/ST1/00412. Part of this work was carried out while the authors were visiting the Real Analysis in Computer Science program at the Simons Institute for the Theory of Computing.

Conjecture 1.1 (Heat smoothing [14, Remark 5.5]). *Let $1 < p < \infty$. Let $f\colon \{-1,1\}^n \to \mathbb{R}$ with $\mathbb{E} f W_S = 0$ for all $S \subseteq \{1,\ldots,n\}$ with $|S| < k$. Then*

$$\forall t > 0, \quad \|P_t f\|_p \le e^{-tkc(p)} \|f\|_p . \tag{1.1}$$

In our main result we prove a special case of their conjecture for $k = 1$:

Theorem 1.2 (Heat smoothing). *For every $p \in (1,\infty)$ and every $f : \{-1,1\}^n \to \mathbb{R}$ with $\mathbb{E} f = 0$, for every $t > 0$,*

$$\|P_t f\|_p \le \exp\left(-\frac{(2p-2)t}{(p^2-2p+2)}\right) \cdot \|f\|_p.$$

The proof of this theorem covers all Markov operators satisfying an L^2 Poincaré inequality. We also show that if we restrict to $\{-1,0,1\}$-valued functions, then (1.1) always holds.

Theorem 1.3 (Conjecture 1.1 for $\{-1,0,1\}$-valued functions). *Let $1 < p < \infty$ and let $f\colon \{-1,1\}^n \to \{-1,0,1\}$ with $\mathbb{E} f W_S = 0$ for all $S \subseteq \{1,\ldots,n\}$ with $|S| < k$. Then for all $t > 0$,*

$$\|P_t f\|_p \le e^{-2tk\min\left(\frac{p-1}{p},\frac{1}{p}\right)} \|f\|_p . \tag{1.2}$$

The constant in Theorem 1.3 for $k = 1$, which comes from an application of Hölder's inequality, is strictly worse than that of Theorem 1.2. Again the proof of Theorem 1.3 extends to cover P_t being any symmetric Markov semigroup as long as $f\colon \Omega \to \{-1,0,1\}$ satisfies $\|P_t f\|_2 \le e^{-tk} \|f\|_2$.

Our results in Theorem 1.2 and Theorem 1.3 should be compared to the following result of Mendel-Naor, which they attributed to P. A. Meyer [15].

Theorem 1.4 ([14, Lemma 5.4]). *Let $2 \le p < \infty$. Then there exists $c(p) > 0$ such that the following holds. Let $f\colon \{-1,1\}^n \to \mathbb{R}$ with $\mathbb{E} f W_S = 0$ for all $|S| < k$. Then*

$$\forall\, t > 0, \quad \|P_t f\|_p \le e^{-k\min(t,t^2)c(p)} \|f\|_p ,$$

$$\|Lf\|_p \ge c(p)\sqrt{k} \|f\|_p .$$

The second inequality can be considered a "higher-order" Poincaré inequality, and it follows from the first by writing $f = \int_0^\infty e^{-tL} Lf dt$ and then applying the $L^p(\{-1,1\}^n)$ triangle inequality.

One should also compare our results to the following result of Hino (in a much more general setting), which is also briefly mentioned at the end of the proof of Theorem 1 in [15].

Theorem 1.5 ([7, Theorem 3.6(ii)b]). *Let $1 < p < \infty$. Then there exists $\infty > M(n), \delta(n) > 0$ such that, for any $f\colon \{-1,1\}^n \to \mathbb{R}$ with $\mathbb{E} f = 0$, for any $t > 0$,*

$$\|P_t f\|_p \le M(n) e^{-\delta(n)t} \|f\|_p .$$

The dependence of the constants $M(n)$ and $\delta(n)$ on the dimension makes this inequality weaker than the previous two in settings where dimension independent inequalities are desired.

1.2. **Poincaré inequalities.** This heat smoothing estimate in Theorem 1.2 is equivalent to the following Poincaré inequality.

Theorem 1.6 (Poincaré inequality). *Under the above assumptions for every* $p \in (1, \infty)$ *and every* $f : \{-1,1\}^n \to \mathbb{R}$ *with* $\mathbb{E}f = 0$ *there is*

$$\mathbb{E}\,|f|^{p-1}\,\mathrm{sign}(f)Lf \geq \frac{2p-2}{(p^2-2p+2)} \cdot \mathbb{E}\,|f|^p\,.$$

The usual Poincaré inequality corresponds to the case $p = 2$ of Theorem 1.6. Theorem 1.6 should be contrasted with Beckner's Poincaré inequality.

Theorem 1.7 ([1]). *Let* $f\colon \{-1,1\}^n \to \mathbb{R}$. *For all* $1 \leq p \leq 2$,

$$(2-p)\mathbb{E}fLf \geq \mathbb{E}\,|f|^2 - (\mathbb{E}\,|f|^p)^{2/p}.$$

Specifically, Beckner notes that, for $t > 0$ with $e^{-2t} = p - 1$, $(2-p)\mathbb{E}fLf \geq \mathbb{E}\,|f|^2 - \mathbb{E}\,|P_t f|^2$ by Fourier analysis. He then adds the hypercontractive inequality [3, 5, 16] to this inequality to prove Theorem 1.7. However, Theorem 1.6 does not seem to follow from hypercontractivity, so we need to apply different methods.

1.3. **The KKL, Talagrand and Harper theorems in tail spaces.** The KKL theorem and its strengthening by Talagrand are two of the most fundamental theorems in the theory of Boolean functions. Harper's theorem is an edge-isoperimetric inequality on the hypercube. Recent questions by Hatami and Kalai asked if the KKL and Harper theorems could be improved for functions in tail spaces. It is natural to ask the same question for Talagrand's theorem.

We recall some standard definitions.

Definition 1.8 (Influences). Let $f\colon \{-1,1\}^n \to \{-1,1\}$ and let $i \in \{1,\ldots,n\}$. Define the i'th *influence* $I_i(f) \in \mathbb{R}$ of a function $f : \{-1,1\}^n \to \{-1,1\}$ by

$$I_i(f) := P[f(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_n) \neq f(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n)],$$

where x_i, y are i.i.d. uniform random variables on $\{-1,1\}$ for all $i = 1,\ldots,n$.

Since the range of a Boolean function is restricted to $\{-1,1\}$, its Fourier coefficients should satisfy some constraints that general real-valued functions with $\|f\|_2 = 1$ do not satisfy. For instance, the influences of a Boolean function could be slightly larger than expected. For example, the non-Boolean function $f = (n(n-1)/2)^{-1/2}\sum_{S\subseteq\{1,\ldots,n\}\colon |S|=2} W_S$ satisfies $\|f\|_2 = 1$, where $I_i f = 2/n$ for all $i = 1,\ldots,n$. At the opposite extreme, the Boolean function $f = W_{\{1,\ldots,n\}}$ satisfies $\|f\|_2 = 1$, where $I_i f = 1$ for all $i = 1,\ldots,n$. With these examples in mind, we may be led to believe that Boolean functions have larger influences than arbitrary functions with $\|f\|_2 = 1$. Indeed, Ben-Or and Linial proved the following proposition, and they conjectured that their bound on influences was the best possible.

Proposition 1.9 ([2, Theorem 3]). *There exists a universal constant* $c' > 0$, *and there exists a Boolean function* $f\colon \{-1,1\}^n \to \{-1,1\}$ *with* $\mathbb{E}f = 0$ *such that* $\max_{i=1,\ldots,n} I_i(f) \leq c'(\log n)/n$.

Kahn, Kalai and Linial then showed that the influence bound in Proposition 1.9 is in fact the best possible, thereby proving the conjecture of Ben-Or and Linial.

Theorem 1.10 (KKL [9, Theorem 3.1])**.** *There exists a universal constant $c > 0$ such that, for any $f\colon \{-1,1\}^n \to \{-1,1\}$,*

$$\max_{i=1,\ldots,n} I_i(f) \geq c(\mathbb{E}(f - \mathbb{E}f)^2)(\log n)/n.$$

If a function $f\colon \{-1,1\}^n \to \{-1,1\}$ not only has mean zero but also has many Fourier coefficients which are zero, it similarly seems that even more special structure should exist within the Fourier coefficients of f. That is, perhaps this function should have a larger influence than a mean zero function. Hatami and Kalai therefore asked the following question, which would improve upon Theorem 1.10.

Question 1.11. Suppose $k = k(n) \to \infty$ as $n \to \infty$. Does there exist $\omega(k) > 0$ such that $\omega(k) \to \infty$ as $k \to \infty$, such that the following statement holds? Let $f\colon \{-1,1\}^n \to \{-1,1\}$ with $\mathbb{E}fW_S = 0$ for all $S \subseteq \{1,\ldots,n\}$ with $|S| \leq k$. Then $\max_{i=1,\ldots,n} I_i f \geq ((\log n)/n) \cdot \omega(k)$.

Hatami speculated that a positive answer to the question above may help in proving the Entropy Influence Conjecture. Here we prove that the answer to the question is negative by showing that

Theorem 1.12 (Question 1.11 for $k = \log n$)**.** *There exists $0 < C, c < \infty$ such that, for infinitely many $n \in \mathbb{N}$, there exists $f\colon \{-1,1\}^n \to \{-1,1\}$ with $\mathbb{E}fW_S = 0$ for all $S \subseteq \{1,\ldots,n\}$ with $|S| \leq c\log n$ such that $\max_{i=1,\ldots,n} I_i f \leq C(\log n)/n$.*

In other words, there is a phase transition for the maximum influence of Boolean functions with vanishing Fourier coefficients. This phase transition occurs when we require the first $k(n)$ Fourier coefficients to vanish where $k(n)/\log n$ is either bounded or unbounded, as $n \to \infty$. We note that the functions constructed in Theorem 1.12 do not provide a counterexample to the Entropy Influence Conjecture, as their entropy is of the same order as for the standard Tribes function.

We also note that if $k = g(n)\log n$, where $g(n) \to \infty$, then it is trivial to improve the KKL estimate since

$$\sum_{i=1}^{n} I_i f = \sum_{S \subseteq \{1,\ldots,n\}} |S|\,|\widehat{f}(S)|^2,$$

which implies $\max_{i=1,\ldots,n} I_i f \geq g(n)(\log n)/n$.

With a similar motivation to Question 1.11, Kalai also asked whether or not the following isoperimetric inequality could be improved.

Theorem 1.13 (Harper's inequality [17, Theorem 2.39], [6])**.** *For any $f\colon \{-1,1\}^n \to \{0,1\}$, $\sum_{i=1}^{n} I_i f \geq (2/\log 2)(\mathbb{E}f)\log(1/\mathbb{E}f)$.*

To see the isoperimetric content of Theorem 1.13, we consider the hypercube $\{-1,1\}^n$ as the vertices of a graph, where an edge connects $(x_1,\ldots,x_n),(y_1,\ldots,y_n) \in \{-1,1\}^n$ if and only if $|\{i \in \{1,\ldots,n\}\colon x_i \neq y_i\}| = 1$. Then $(1/n)\sum_{i=1}^{n} I_i f$ is equal to the fraction of edges of $\{-1,1\}^n$ between the sets $\{x \in \{-1,1\}^n\colon f(x) = 0\}$ and $\{x \in \{-1,1\}^n\colon f(x) = 1\}$, and the quantity $(\mathbb{E}f)\log(1/\mathbb{E}f)$ measures the volume of the set $\{x \in \{-1,1\}^n\colon f(x) = 1\}$.

If $f(x_1,\ldots,x_n) := (x_1+1)/2$, then equality nearly holds in Theorem 1.13. Note that, in this case, f has Fourier coefficients only of degrees zero and one. It therefore seems sensible that if f has only Fourier coefficients of higher order, then f will oscillate, so the perimeter of its level sets should be much larger than the volume of its level sets. Kalai therefore asked if the constant 2 in Theorem 1.13 would become large when a large number of Fourier coefficients of the function are zero.

Question 1.14. Suppose $k = k(n) \to \infty$ as $n \to \infty$. Does there exist $\omega(k) > 0$ such that $\omega(k) \to \infty$ as $k \to \infty$, such that the following statement holds? Let $f\colon \{-1,1\}^n \to \{0,1\}$ with $\mathbb{E} f W_S = 0$ for all $S \subseteq \{1,\ldots,n\}$ with $1 \leq |S| \leq k$. Then $\sum_{i=1}^n I_i f \geq \omega(k)(\mathbb{E}f)\log(1/\mathbb{E}f)$.

A simplification of the function from Theorem 1.12 shows that Question 1.14 has a negative answer.

Theorem 1.15 (Negative answer to Question 1.14)**.** *There exists* $0 < C, c < \infty$ *such that, for infinitely many* $n \in \mathbb{N}$, *there exists* $f\colon \{-1,1\}^n \to \{0,1\}$ *with* $\mathbb{E} f W_S = 0$ *for all* $S \subseteq \{1,\ldots,n\}$ *with* $1 \leq |S| \leq c\log n$ *such that* $\sum_{i=1}^n I_i f \leq C(\mathbb{E}f)\log(1/\mathbb{E}f)$.

A less trivial argument allows us to extend Talagrand's theorem to tail spaces.

Theorem 1.16 (Talagrand inequality for tail space)**.** *Let* $k \geq 1$. *Let* $f\colon \{-1,1\}^n \to \mathbb{R}$ *with* $\mathbb{E} f W_S = 0$ *for all* $S \subseteq \{1,\ldots,n\}$ *with* $|S| < k$. *For* $i = 1,\ldots,n$, *define* $D_i f(x) := [f(x_1,\ldots,x_n) - f(x_1,\ldots,x_{i-1},-x_i,x_{i+1},\ldots,x_n)]/2$. *Then*

$$\begin{aligned}\mathbb{E}f^2 &\leq \frac{3}{k}\sum_{i=1}^n \mathbb{E}(D_i f)^2/\max(1,\log(\|D_i f\|_2/\|D_i f\|_{1+e^{-2/k}}))\\ &\leq 8\sum_{i=1}^n \frac{\mathbb{E}(D_i f)^2}{k+\log\left(\|D_i f\|_2/\|D_i f\|_1\right)}.\end{aligned} \tag{1.3}$$

Note that the usual form of Talagrand's inequality is obtained by setting $k = 1$ and substituting $f - \mathbb{E}f$ in place of f on the left side of (1.3) (which is redundant for $k \geq 1$).

We note that in principle Theorem 1.16 may indicate that the answer to Question 1.11 is positive, since the usual Talagrand inequality implies the Kahn-Kalai-Linial Theorem 1.10. However, the improvement of Theorem 1.16 over Theorem 1.10 only occurs for k of the form $k = g(n)\log n$ where $g(n) \to \infty$ as $n \to \infty$.

1.4. **General setting.** Theorem 1.2 and Theorem 1.6 are proven in the following general setting:

$$\left\{\begin{array}{l}\bullet\ (\Omega, 2^\Omega, \mu) \text{ is a finite probability space.}\\ \bullet\ (P_t)_{t\geq 0} \text{ is a symmetric Markov semigroup on } L^2(\Omega,\mu).\\ \bullet\ (P_t)_{t\geq 0} \text{ has generator } L = -\frac{\mathrm{d}}{\mathrm{d}t}P_t\Big|_{t=0^+}.\\ \bullet\ \mathbb{E} \text{ denotes the expectation with respect to the invariant measure } \mu.\\ \bullet\ \|f\|_p \text{ will stand for } \left(\mathbb{E}|f|^p\right)^{1/p}.\end{array}\right. \tag{1.4}$$

We assume additionally that L satisfies the Poincaré inequality with a positive constant C, i.e.,

$$\mathbb{E}f^2 - (\mathbb{E}f)^2 \leq C \cdot \mathbb{E} f L f, \tag{1.5}$$

for every $f : \Omega \to \mathbb{R}$ or, equivalently, $\mathbb{E}(P_t f)^2 \leq e^{-2t/C} \cdot \mathbb{E}f^2$ for every $t \geq 0$ and every mean-zero f. Theorem 1.2 is a special case of the following theorem.

Theorem 1.17 (Heat smoothing)**.** *Assume that* (1.4) *and* (1.5) *hold. Then for every* $p \in (1,\infty)$ *and every* $f : \Omega \to \mathbb{R}$ *with* $\mathbb{E}f = 0$*, for every* $t > 0$*,*

$$\|P_t f\|_p \leq \exp\left(-\frac{(2p-2)t}{(p^2-2p+2)C}\right) \cdot \|f\|_p.$$

Theorem 1.6 is a special case of the following result.

Theorem 1.18 (Poincaré inequality)**.** *Assume that* (1.4) *and* (1.5) *hold. Then for every* $p \in (1,\infty)$ *and every* $f\colon \Omega \to \mathbb{R}$ *with* $\mathbb{E}f = 0$ *there is*

$$\mathbb{E}\,|f|^{p-1}\operatorname{sign}(f) Lf \geq \frac{2p-2}{(p^2-2p+2)C} \cdot \mathbb{E}\,|f|^p\,.$$

Theorem 1.18 is equivalent to Theorem 1.17. Below, we will first prove Theorem 1.18 and then deduce Theorem 1.17 as a consequence.

For diffusion semigroups, a variant of Theorem 1.17 was proven in [4].

After the proof of Theorem 1.17 we briefly discuss how Theorem 1.17 and Theorem 1.18 can be extended to infinite spaces.

1.5. **Organization.** We prove Theorem 1.17 in Section 2. Theorem 1.18 is then derived as a corollary in Section 3, where Theorem 1.3 is also shown. A complex interpolation proof of Theorem 1.17 is given in Section 4, albeit with worse constants. A semigroup proof of Theorem 1.17 is given in Section 5. Theorem 1.16 is proven in Section 6, and Theorem 1.12 is proven in Section 7.

2. Poincaré inequalities

In this section we prove Theorem 1.17 and Theorem 1.18.

Let $x \in \mathbb{R}$. In what follows, we use the standard notation $x_+ := \max(x,0)$ and $x_- := \max(-x,0)$, so that $x = x_+ - x_-$ and $|x|^p = x_+^p + x_-^p$ for any $x \in \mathbb{R}$ and $p > 0$. Also, for $s > 0$ we will denote by ϕ_s the function $\phi_s(x) := \operatorname{sign}(x)\cdot|x|^s$, so that $\phi_s(x) = x_+^s - x_-^s$ for every $x \in \mathbb{R}$.

Lemma 2.1. *Let* $p > 1$*, and let* X *be a real random variable with* $\mathbb{E}\,|X|^p < \infty$ *and* $\mathbb{E}X = 0$*. For every* $p \in (1,\infty)\setminus\{2\}$*,*

$$\begin{aligned}(\mathbb{E}X_+^{p/2})^2 + (\mathbb{E}X_-^{p/2})^2 + (\mathbb{E}X_+^{p/2})^{-\frac{2}{p-2}}(\mathbb{E}X_-^{p/2})^{\frac{2p-2}{p-2}}&\\ + (\mathbb{E}X_-^{p/2})^{-\frac{2}{p-2}}(\mathbb{E}X_+^{p/2})^{\frac{2p-2}{p-2}} &\leq \mathbb{E}\,|X|^p\,.\end{aligned} \tag{2.1}$$

Proof. Since $\mathbb{E}X = 0$,

$$\mathbb{E}X_+ = \mathbb{E}X_- = (1/2)\mathbb{E}\,|X|\,. \tag{2.2}$$

Assume $p > 2$. Note that Jensen's inequality implies that

$$\mathbb{E}X_+^{p/2} = \mathbb{E}X_+^{\frac{p^2-2p}{2p-2}}X_+^{\frac{p}{2p-2}} \leq (\mathbb{E}X_+^p)^{\frac{p-2}{2p-2}}\cdot(\mathbb{E}X_+)^{\frac{p}{2p-2}} \overset{(2.2)}{=} (\mathbb{E}X_+^p)^{\frac{p-2}{2p-2}}\cdot(\mathbb{E}\,|X|/2)^{\frac{p}{2p-2}}. \tag{2.3}$$

Also, by Hölder's inequality,

$$\mathbb{E}\,|X|/2 \overset{(2.2)}{=} \mathbb{E}X_+ = \mathbb{E}X_+1_{\{X>0\}} \leq (\mathbb{E}X_+^{p/2})^{2/p}\cdot\mathbb{P}(X>0)^{\frac{p-2}{p}}. \tag{2.4}$$

Applying (2.3) to X and $(-X)$ separately, exponentiating both sides to the power $(2p-2)/(p-2)$, and then adding the results, we obtain

$$(\mathbb{E}X_+^{p/2})^{\frac{2p-2}{p-2}} + (\mathbb{E}X_-^{p/2})^{\frac{2p-2}{p-2}} \overset{(2.3)}{\leq} 2^{-\frac{p}{p-2}}(\mathbb{E}\,|X|)^{\frac{p}{p-2}}(\mathbb{E}X_+^p + \mathbb{E}X_-^p) = 2^{-\frac{p}{p-2}}(\mathbb{E}\,|X|)^{\frac{p}{p-2}}\mathbb{E}\,|X|^p. \tag{2.5}$$

Applying (2.4) to X and $(-X)$ separately, exponentiating both sides to the power $-p/(p-2)$, and then adding the results, we obtain

$$(\mathbb{E}X_+^{p/2})^{-\frac{2}{p-2}} + (\mathbb{E}X_-^{p/2})^{-\frac{2}{p-2}} \overset{(2.4)}{\leq} 2^{\frac{p}{p-2}}[\mathbb{P}(X>0)+\mathbb{P}(X<0)](\mathbb{E}\,|X|)^{-\frac{p}{p-2}} \leq 2^{\frac{p}{p-2}}(\mathbb{E}\,|X|)^{-\frac{p}{p-2}}. \tag{2.6}$$

Finally, multiplying (2.5) and (2.6) gives (2.1), if $p>2$.

Assume $1<p<2$. Then (2.5), (2.6) and (2.1) also hold. To see this, we use the following two consequences of Hölder's inequality:

$$\mathbb{E}X_+^{p/2} = \mathbb{E}X_+^{p/2}1_{\{X>0\}} \leq (\mathbb{E}X_+)^{p/2}\cdot\mathbb{P}(X>0)^{\frac{2-p}{2}} \overset{(2.2)}{=} (\mathbb{E}\,|X|/2)^{p/2}\cdot\mathbb{P}(X>0)^{\frac{2-p}{2}}, \tag{2.7}$$

$$\mathbb{E}\,|X|/2 \overset{(2.2)}{=} \mathbb{E}X_+ = \mathbb{E}X_+^{p-1}X_+^{2-p} \leq (\mathbb{E}X_+^{p/2})^{\frac{2p-2}{p}}\cdot(\mathbb{E}X_+^p)^{\frac{2-p}{p}}. \tag{2.8}$$

Applying (2.7) to X and $(-X)$ separately, exponentiating both sides by the power $2/(2-p)$, and then adding the results, we obtain (2.6)

$$(\mathbb{E}X_+^{p/2})^{\frac{2}{2-p}} + (\mathbb{E}X_-^{p/2})^{\frac{2}{2-p}} \overset{(2.7)}{\leq} 2^{-\frac{p}{2-p}}(\mathbb{E}\,|X|)^{\frac{p}{2-p}}\cdot[\mathbb{P}(X>0)+\mathbb{P}(X<0)] \leq 2^{-\frac{p}{2-p}}(\mathbb{E}\,|X|)^{\frac{p}{2-p}}.$$

Applying (2.8) to X and $(-X)$ separately, exponentiating both sides to the power $-p/(2-p)$, and then adding the results, we obtain (2.5)

$$(\mathbb{E}X_+^{p/2})^{-\frac{2p-2}{2-p}} + (\mathbb{E}X_-^{p/2})^{-\frac{2p-2}{2-p}} \overset{(2.8)}{\leq} 2^{\frac{p}{2-p}}(\mathbb{E}\,|X|)^{-\frac{p}{2-p}}(\mathbb{E}X_+^p + \mathbb{E}X_-^p) = 2^{\frac{p}{2-p}}(\mathbb{E}\,|X|)^{-\frac{p}{2-p}}\mathbb{E}\,|X|^p.$$

□

Lemma 2.2. *Let $p\in(1,\infty)\setminus\{2\}$. For any $a,b>0$,*

$$(a-b)^2 \leq \frac{p^2-4p+4}{2p^2-4p+4}\cdot(a^2+b^2+a^{\frac{2}{2-p}}\cdot b^{\frac{2p-2}{p-2}}+b^{\frac{2}{2-p}}\cdot a^{\frac{2p-2}{p-2}}).$$

Proof. Without loss of generality, $a\geq b$. Define $s:=p/(p-2)$. Since $|s|>1$, for every $x\geq 0$ we have

$$e^{sx}+e^{-sx} \geq e^{|s|x}-e^{-|s|x} = 2\sum_{j=0}^{\infty}\frac{|s|^{2j+1}x^{2j+1}}{(2j+1)!} \geq 2\,|s|\sum_{j=0}^{\infty}\frac{x^{2j+1}}{(2j+1)!} = |s|\,(e^x-e^{-x}). \tag{2.9}$$

Set $x:=(1/2)\log(a/b)$ and square the inequality (2.9) to get

$$a^sb^{-s}+b^sa^{-s}+2 \geq s^2(ab^{-1}+ba^{-1}-2). \tag{2.10}$$

Multiplying both sides of (2.10) by ab, then adding $(a-b)^2$ to both sides, we obtain

$$a^2+b^2+a^{1-s}b^{1+s}+b^{1-s}a^{1+s}\geq(1+s^2)(a-b)^2. \tag{2.11}$$

This completes the lemma. $\square$

Lemma 2.3. *Let $p>1$ and let X be a real random variable such that $\mathbb{E}\,|X|^p<\infty$ and $\mathbb{E}X=0$. Then*

$$\left(\mathbb{E}X_+^{p/2}-\mathbb{E}X_-^{p/2}\right)^2\leq\left(1-\frac{p^2}{2(p^2-2p+2)}\right)\cdot\mathbb{E}\,|X|^p\,. \tag{2.12}$$

Proof. If $p=2$ or if $X=0$ almost surely, then both sides are zero. So, we may assume $p\in(1,\infty)\setminus\{2\}$ and X is nonzero on a set of positive measure. In this case, set $a:=\mathbb{E}X_+^{p/2}$, $b:=\mathbb{E}X_-^{p/2}$, and apply Lemma 2.2 and then (2.1). $\square$

Lemma 2.4 (Stroock-Varopoulos [18,20])**.** *Let $a,b\in\mathbb{R}$, $p>1$. Then*

$$(\phi_{p-1}(a)-\phi_{p-1}(b))(a-b)\geq\frac{4(p-1)}{p^2}(\phi_{p/2}(a)-\phi_{p/2}(b))^2. \tag{2.13}$$

Proof. Applying the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\frac{4}{p^2}(|a|^{p/2}\operatorname{sign}(a)-|b|^{p/2}\operatorname{sign}(b))^2&=\left(\int_a^b|t|^{(p/2)-1}\,dt\right)^2\leq\int_a^b|t|^{p-2}\,dt\cdot\int_a^b dt\\&=\frac{1}{p-1}(|a|^{p-1}\operatorname{sign}(a)-|b|^{p-1}\operatorname{sign}(b))(a-b).\end{aligned}$$

$\square$

Proof of Theorem 1.18. Recall that for any $g,h:\Omega\to\mathbb{R}$ we have

$$\mathbb{E}gLh=\frac{1}{2}\sum_{x,y\in\Omega}\left(-\mathbb{E}\mathbf{1}_{\{x\}}L\mathbf{1}_{\{y\}}\right)\cdot(g(x)-g(y))(h(x)-h(y)), \tag{2.14}$$

and that for $x\neq y$ there is

$$-\mathbb{E}\mathbf{1}_{\{x\}}L\mathbf{1}_{\{y\}}=\mu(\{y\})\cdot\frac{\mathrm{d}}{\mathrm{d}t}(P_t\mathbf{1}_{\{x\}})(y)\Big|_{t=0^+}\geq0. \tag{2.15}$$

Therefore,

$$\begin{aligned}\mathbb{E}\phi_{p-1}(f)Lf&\overset{(2.14)\wedge(2.15)\wedge(2.13)}{\geq}\frac{4(p-1)}{p^2}\mathbb{E}\phi_{p/2}(f)L\phi_{p/2}(f)\\&\overset{(1.5)}{\geq}\frac{4(p-1)}{Cp^2}\cdot\left(\mathbb{E}\phi_{p/2}(f)^2-(\mathbb{E}\phi_{p/2}(f))^2\right)\\&=\frac{4(p-1)}{Cp^2}\left(\mathbb{E}\,|f|^p-\left(\mathbb{E}f_+^{p/2}-\mathbb{E}f_-^{p/2}\right)^2\right)\\&\overset{(2.12)}{\geq}\frac{4(p-1)}{Cp^2}\left(\frac{p^2}{2(p^2-2p+2)}\right)\mathbb{E}\,|f|^p\,.\end{aligned}$$

$\square$

3. Heat smoothing

We now show that Theorem 1.18 implies Theorem 1.17.

Proof of Theorem 1.17. Note that $\mathbb{E}f = 0$ implies $\mathbb{E}P_t f = 0$ for all $t \geq 0$. So, by Theorem 1.18,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}&\left(\exp\left(\frac{(2p-2)pt}{(p^2-2p+2)C}\right)\cdot \mathbb{E}|P_t f|^p\right)\\ &= \exp\left(\frac{(2p-2)pt}{(p^2-2p+2)C}\right)\left(\frac{(2p-2)p}{(p^2-2p+2)C}\,\mathbb{E}|P_t f|^p - p\cdot \mathbb{E}\phi_{p-1}(P_t f)LP_t f\right) \leq 0.\end{aligned}$$

□

Remark 3.1. One easily extends Theorem 1.18 from real-valued functions to f taking values in a Euclidean space, with the same constant. In particular, we get the same statement for complex-valued functions. Indeed, it suffices to apply Theorem 1.17 to $f_v(x) := \langle f(x), v\rangle$ and average over v's from the unit sphere. Note that $|\langle w, v\rangle|^p$ averaged over the unit sphere (with respect to the uniform measure) is proportional to $\|w\|^p$, and the proportionality constant will cancel out.

Remark 3.2. Let $\kappa(p) = \inf_{u>1} \frac{u^{\frac{p}{p-2}} + u^{-\frac{p}{p-2}}}{u-u^{-1}}$. Note that $\kappa(p) = \kappa(p')$ since $\frac{p}{p-2} = \frac{pp'}{p-p'}$. A simple analysis of the proof shows that we may strengthen the assertion of Theorem 1.18 to

$$\|P_t f\|_p \leq \exp\left(-\frac{(4p-4)t}{Cp^2(1+\kappa(p)^{-2})}\right)\cdot \|f\|_p.$$

We have established (in the proof of Lemma 2.2) the estimate $\kappa(p) \geq |\frac{p}{p-2}|$. One can do better, however. For example, there is $\kappa(4) = \kappa(4/3) = 2\sqrt{2}$ and $\kappa(6) = \kappa(6/5) = 2$, so that for every mean-zero f we have

$$\|P_t f\|_4 \leq e^{-\frac{2t}{3C}}\|f\|_4, \quad \|P_t f\|_{4/3} \leq e^{-\frac{2t}{3C}}\|f\|_{4/3},$$

$$\|P_t f\|_6 \leq e^{-\frac{4t}{9C}}\|f\|_6, \quad \|P_t f\|_{6/5} \leq e^{-\frac{4t}{9C}}\|f\|_{6/5}.$$

Also, one can easily strengthen the lower bound to $\kappa(p) \geq \sqrt{\frac{p^2+4p-4}{p^2-4p+4}}$. Indeed, for $s = \frac{p}{p-2}$ we have $|s| > 1$, so that $u^{|s|/2} - u^{-|s|/2} \geq |s|(u^{1/2} - u^{-1/2})$ for every $u > 1$. Squaring this inequality, we get $u^s + u^{-s} \geq 2 + s^2(u + u^{-1} - 2)$, so that

$$\kappa(p) = \inf_{u>1}\frac{u^s + u^{-s}}{u-u^{-1}} \geq \inf_{u>1}\frac{2+s^2(u+u^{-1}-2)}{u-u^{-1}} = \sqrt{2s^2-1} = \sqrt{\frac{p^2+4p-4}{p^2-4p+4}},$$

which, for $p > 1$ and mean-zero functions f, yields

$$\|P_t f\|_p \leq e^{-\left(\frac{2}{pp'}+\frac{8}{p^2p'^2}\right)t/C}\|f\|_p.$$

Corollary 3.3. *Let $p \in (1,\infty)\setminus\{2\}$ and let X be a mean-zero real random variable with $\mathbb{E}|X|^p < \infty$. Then*

$$\left(\kappa(p)^2 + 1\right)\cdot\left(\mathbb{E}X_+^{p/2} - \mathbb{E}X_-^{p/2}\right)^2 \leq \mathbb{E}|X|^p,$$

where $\kappa(p) = \inf_{u>1} \frac{u^{\frac{p}{p-2}}+u^{-\frac{p}{p-2}}}{u-u^{-1}}$. *Moreover, the constant* $\kappa(p)^2+1$ *is optimal. Furthermore, under the assumptions of Theorem* 1.18*,*

$$\mathbb{E}\phi_{p/2}(f)L\phi_{p/2}(f) \geq \frac{C^{-1}}{1+\kappa(p)^{-2}} \cdot \mathbb{E}|f|^p,$$

and the constant $C^{-1}/\left(1+\kappa(p)^{-2}\right)$ *is optimal.*

Proof. The above bounds follow from an obvious strengthening of the proof of Theorem 1.18 as explained in Remark 3.2. To check the optimality of the constants, choose $v>1$ such that $v^{\frac{p}{p-2}}+v^{-\frac{p}{p-2}} = \kappa(p)(v-v^{-1})$ and set $\alpha = 1/(1+v^{\frac{4}{p-2}})$ and $\beta = 1-\alpha$, so that $\beta/\alpha = v^{\frac{4}{p-2}}$. Then, for a mean-zero random variable X such that $\mathbb{P}(X=\beta)=\alpha$ and $\mathbb{P}(X=-\alpha)=\beta$, we have

$$\mathbb{E}|X|^p = (\alpha+\beta)\mathbb{E}|X|^p = (\alpha+\beta)(\alpha\beta^p+\beta\alpha^p) = \alpha^2\beta^p+\beta^2\alpha^p+\alpha\beta^{p+1}+\beta\alpha^{p+1},$$

and thus

$$\begin{aligned}
(\alpha\beta)^{-\frac{p+2}{2}}\mathbb{E}|X|^p &= (\alpha/\beta)^{\frac{p-2}{2}} + (\beta/\alpha)^{\frac{p-2}{2}} + (\alpha/\beta)^{p/2} + (\beta/\alpha)^{p/2} \\
&= v^2+v^{-2}+v^{\frac{2p}{p-2}}+v^{-\frac{2p}{p-2}} = \left(v^{\frac{p}{p-2}}+v^{-\frac{p}{p-2}}\right)^2 + \left(v-v^{-1}\right)^2 \\
&= \left(\kappa(p)^2+1\right)\left(v-v^{-1}\right)^2 = \left(\kappa(p)^2+1\right)\left((\alpha/\beta)^{\frac{p-2}{4}} - (\beta/\alpha)^{\frac{p-2}{4}}\right)^2 \\
&= (\alpha\beta)^{-\frac{p+2}{2}}\left(\kappa(p)^2+1\right)\left(\alpha\beta^{p/2}-\beta\alpha^{p/2}\right)^2 \\
&= (\alpha\beta)^{-\frac{p+2}{2}}\left(\kappa(p)^2+1\right)\left(\mathbb{E}X_+^{p/2}-\mathbb{E}X_-^{p/2}\right)^2.
\end{aligned}$$

To see that the constant $C^{-1}/\left(1+\kappa(p)^{-2}\right)$ in the second claim is also optimal, consider $\Omega = \{-\alpha,\beta\}$ with $\mu = \alpha\delta_\beta+\beta\delta_{-\alpha}$, $L = C^{-1}\cdot(Id-\mathbb{E})$, and $f(x)=x$. With this definition of L, equality holds in (1.5) for any function. So, arguing as in the proof of Theorem 1.18, we have $\mathbb{E}\phi_{p/2}(f)L\phi_{p/2}(f) = \mathbb{E}|f|^p - (\mathbb{E}f_+^{p/2}-\mathbb{E}f_-^{p/2})^2$, and we then use the equality that holds from the first part of the present corollary. $\square$

Remark 3.4. We now show that the dependence on p of Theorem 1.18 is of optimal order as $p\to 1$ or as $p\to\infty$. In what follows, L is the generator of the standard one-dimensional Ornstein-Uhlenbeck semigroup and γ is the standard $\mathcal{N}(0,1)$ Gaussian measure.

For $\varepsilon>0$, let $g_\varepsilon : \mathbb{R}\to\mathbb{R}$ be an increasing 2-Lipschitz function, smooth on $\mathbb{R}\setminus\{0\}$ and such that $g_\varepsilon(x)=x$ for $|x|>\varepsilon$ and $g_\varepsilon = \phi_{3p/2}$ on some neighbourhood of zero. Furthermore, let $f_\varepsilon = \phi_{2/p}\circ g_\varepsilon$. Then f_ε is a smooth function and it belongs to the domain of L. We have

$$\mathbb{E}_\gamma|f_\varepsilon|^p = \mathbb{E}_\gamma g_\varepsilon^2 \geq \int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]} x^2\,\mathrm{d}\gamma(x) \overset{\varepsilon\to 0^+}{\longrightarrow} 1$$

and

$$\frac{p^2}{4(p-1)}\cdot\mathbb{E}_\gamma\phi_{p-1}(f_\varepsilon)Lf_\varepsilon = \frac{p^2}{4(p-1)}\int_{\mathbb{R}}(\phi_{p-1}\circ f_\varepsilon)'f_\varepsilon'\,\mathrm{d}\gamma$$

$$= \int_{\mathbb{R}}\left(\left(\phi_{p/2}\circ f_\varepsilon\right)'\right)^2\mathrm{d}\gamma = \int_{\mathbb{R}}(g_\varepsilon')^2\mathrm{d}\gamma \leq \gamma(\mathbb{R}\setminus[-\varepsilon,\varepsilon]) + 4\gamma([-\varepsilon,\varepsilon]) \overset{\varepsilon\to 0^+}{\longrightarrow} 1.$$

Remark 3.5. The proof of Theorem 1.17 used the finiteness of the space Ω in a nonessential way. Infinite spaces require a bit more care, so we have chosen the above presentation. However, the reader can find a proof of Theorem 1.17 in Section 5 which applies to infinite spaces and which only uses the semigroup itself.

To conclude the section, we prove Theorem 1.3.

Proof of Theorem 1.3. Recall that

$$\forall\, 1 \le q \le \infty, \quad \|P_t f\|_q \le \|f\|_q \,. \tag{3.1}$$

Also, if $\mathbb{E} f W_S = 0$ for all $S \subseteq \{1, \dots, n\}$ with $|S| < k$, then for all $t > 0$,

$$\|P_t f\|_2 \le e^{-tk} \|f\|_2 \,. \tag{3.2}$$

Let $f \colon \{-1,1\}^n \to \{-1,0,1\}$. For such f we have that $\mathbb{E}[|f|^p] = \mathbb{E}[|f|]$ for all p and $\|P_t f\|_\infty \le 1$ for all $t \ge 0$. Now, if $p > 2$, then

$$\mathbb{E}|P_t f|^p \le \mathbb{E}|P_t f|^2 \le e^{-2tk}\mathbb{E} f^2 = e^{-2tk}\mathbb{E}|f|^p.$$

If $1 < p < 2$, then from Hölder's inequality,

$$\begin{aligned}\mathbb{E}\,|P_t f|^p &\le (\mathbb{E}\,|P_t f|)^{2-p}(\mathbb{E}\,|P_t f|^2)^{p-1}\\ &\overset{(3.1)\wedge(3.2)}{\le} (\mathbb{E}\,|f|)^{2-p} e^{-tk2(p-1)} (\mathbb{E}\,|f|^2)^{p-1} = e^{-2tk(p-1)}\mathbb{E}\,|f|^p \,.\end{aligned}$$

□

4. An interpolation proof of heat smoothing

After the results of this paper were presented at some seminars, Fedor Nazarov informed the authors about an alternative complex interpolation proof of the strong contractivity bounds. With his kind permission, we present a simple version of such an approach, strongly inspired by Nazarov's proof but different from it. Nazarov proved that a perturbation of a Markov operator P by an appropriately chosen small multiplicity of the expectation, i.e., $T = P - \delta \cdot \mathbb{E}$, is a contraction (a trick he ascribed to Bernstein), and his interpolation bounds were more elaborate than ours.

The L^p spaces considered in this section are complex.

Proposition 4.1. *Let (Ω, μ) be a probability space (for brevity, we leave simple measurability considerations to the reader) and let $P : L^1(\Omega,\mu) \to L^1(\Omega,\mu)$ be a linear operator with $P\mathbf{1} = \mathbf{1}$, $\|P\|_{L^1(\Omega,\mu)\to L^1(\Omega,\mu)} \le 1$, and $\|P\|_{L^\infty(\Omega,\mu)\to L^\infty(\Omega,\mu)} \le 1$. Also, we assume that P is mean preserving, i.e., $\mathbb{E}Pf = \mathbb{E}f$ for every $f \in L^1(\Omega,\mu)$. In particular, these conditions are satisfied by every, not necessarily symmetric, Markov operator P for which μ is an invariant measure. Furthermore, let as assume that there exists $\varepsilon \in [0,1)$ such that $\|Pg\|^2_{L^2(\Omega,\mu)} \le (1-\varepsilon)\|g\|^2_{L^2(\Omega,\mu)}$ for every mean-zero g. Then, for $p \in (1,\infty)$,*

$$\|Pg\|_{L^p(\Omega,\mu)} \le \left(1 - 2^{2-p^*}\varepsilon\right)^{1/p^*} \cdot \|g\|_{L^p(\Omega,\mu)}$$

for every mean-zero g, where $p^* = \max\left(p, \frac{p}{p-1}\right)$.

Proof. Let c be a positive constant, to be specified later. We define a measure $\tilde{\mu}$ on a new space $\tilde{\Omega} = \Omega \times \{0,1\}$ by setting $\tilde{\mu}(A \times \{0\}) = \mu(A)$ and $\tilde{\mu}(A \times \{1\}) = c^2\varepsilon \cdot \mu(A)$, and a linear operator $T : L^1(\Omega,\mu) \to L^1(\tilde{\Omega},\tilde{\mu})$ by $(Tf)(\omega,0) = c \cdot (Pf)(\omega)$ and $(Tf)(\omega,1) = f(\omega) - \mathbb{E}f$ for $\omega \in \Omega$. Clearly, $\|T\|_{L^1(\Omega,\mu)\to L^1(\tilde{\Omega},\tilde{\mu})} \leq c + 2c^2\varepsilon$ and $\|T\|_{L^\infty(\Omega,\mu)\to L^\infty(\tilde{\Omega},\tilde{\mu})} \leq \max(c,2)$.

For $f \in L^2(\Omega,\mu)$, let $g = f - \mathbb{E}f$, so that $\mathbb{E}g = 0$. Therefore also $\mathbb{E}Pg = 0$ and thus

$$\begin{aligned}\|Pf\|^2_{L^2(\Omega,\mu)} &= \|Pg + \mathbb{E}f\|^2_{L^2(\Omega,\mu)} = \|Pg\|^2_{L^2(\Omega,\mu)} + |\mathbb{E}f|^2 \\ &\leq (1-\varepsilon)\|g\|^2_{L^2(\Omega,\mu)} + |\mathbb{E}f|^2 = \|g + \mathbb{E}f\|^2_{L^2(\Omega,\mu)} - \varepsilon\|g\|^2_{L^2(\Omega,\mu)} \\ &= \|f\|^2_{L^2(\Omega,\mu)} - \varepsilon\|g\|^2_{L^2(\Omega,\mu)},\end{aligned}$$

so that

$$\|Tf\|^2_{L^2(\tilde{\Omega},\tilde{\mu})} = c^2\|Pf\|^2_{L^2(\Omega,\mu)} + c^2\varepsilon\|g\|^2_{L^2(\Omega,\mu)} \leq c^2\|f\|^2_{L^2(\Omega,\mu)}.$$

We have proved that $\|T\|_{L^2(\Omega,\mu)\to L^2(\tilde{\Omega},\tilde{\mu})} \leq c$.

- For $p \in (1,2]$, by the Riesz-Thorin theorem,

$$\begin{aligned}\|T\|_{L^p(\Omega,\mu)\to L^p(\tilde{\Omega},\tilde{\mu})} &\leq \|T\|^{(2-p)/p}_{L^1(\Omega,\mu)\to L^1(\tilde{\Omega},\tilde{\mu})}\|T\|^{(2p-2)/p}_{L^2(\Omega,\mu)\to L^2(\tilde{\Omega},\tilde{\mu})} \\ &\leq (c+2c^2\varepsilon)^{\frac{2-p}{p}} \cdot c^{\frac{2p-2}{p}} = c(1+2c\varepsilon)^{\frac{2-p}{p}}.\end{aligned}$$

Thus for any $g \in L^p(\Omega,\mu)$ with $\mathbb{E}g = 0$,

$$c^p\|Pg\|^p_{L^p(\Omega,\mu)} + c^2\varepsilon\|g\|^p_{L^p(\Omega,\mu)} = \|Tg\|^p_{L^p(\tilde{\Omega},\tilde{\mu})} \leq c^p(1+2c\varepsilon)^{2-p}\|g\|^p_{L^p(\Omega,\mu)}.$$

Note that for $c = \left(2^{\frac{1}{p-1}} - 2\varepsilon\right)^{-1}$, we have

$$(1+2c\varepsilon)^{2-p} - c^{2-p}\varepsilon = c^{2-p}((c^{-1}+2\varepsilon)^{2-p} - \varepsilon) = \frac{c^{1-p}}{2},$$

and therefore

$$\|Pg\|_{L^p(\Omega,\mu)} \leq \left((1+2c\varepsilon)^{2-p} - c^{2-p}\varepsilon\right)^{1/p}\|g\|_{L^p(\Omega,\mu)} = \left(1 - 2^{2-p^*}\varepsilon\right)^{1/p^*}\|g\|_{L^p(\Omega,\mu)}.$$

- For $p \in [2,\infty)$, we simply set $c = 2$. Then we have $\|T\|_{L^2(\Omega,\mu)\to L^2(\tilde{\Omega},\tilde{\mu})} \leq 2$ and $\|T\|_{L^\infty(\Omega,\mu)\to L^\infty(\tilde{\Omega},\tilde{\mu})} \leq 2$, and thus, by the Riesz-Thorin theorem, also $\|T\|_{L^p(\Omega,\mu)\to L^p(\tilde{\Omega},\tilde{\mu})} \leq 2$. Therefore, for any mean-zero $g \in L^p(\Omega,\mu)$,

$$2^p\|Pg\|^p_{L^p(\Omega,\mu)} + 4\varepsilon\|g\|^p_{L^p(\Omega,\mu)} = \|Tg\|^p_{L^p(\tilde{\Omega},\tilde{\mu})} \leq 2^p\|g\|^p_{L^p(\Omega,\mu)},$$

so that $\|Pg\|_{L^p(\Omega,\mu)} \leq \left(1 - 2^{2-p}\varepsilon\right)^{1/p}\|g\|_{L^p(\Omega,\mu)} = \left(1 - 2^{2-p^*}\varepsilon\right)^{1/p^*}\|g\|_{L^p(\Omega,\mu)}$. □

Remark 4.2. Obviously, if $P : L^2(\Omega,\mu) \to L^2(\Omega,\mu)$ is a Markov operator such that $\|Pg\|^2_{L^2(\Omega,\mu)} \leq (1-\varepsilon)\|g\|^2_{L^2(\Omega,\mu)}$ for every real-valued mean-zero function g, then we also have $\|Pg\|^2_{L^2(\Omega,\mu)} \leq (1-\varepsilon)\|g\|^2_{L^2(\Omega,\mu)}$ for every complex-valued mean-zero function g.

5. A semigroup proof of heat smoothing

We were asked by experts in operator theory if the results of Section 2 and Section 3 can be extended to the setting of infinite probability spaces. The proof in this setting, which does not use the notion of the semigroup generator, is outlined in the following section.

We will need some standard and simple bounds.

Lemma 5.1. *Let $(P_s)_{s\geq 0}$ be a semigroup of symmetric linear contractions on some inner product space $(\mathcal{H}, \|\cdot\|)$. Then for any $\varepsilon, t > 0$ we have $\|P_{\varepsilon+t} - P_\varepsilon\|_{\mathcal{H}\to\mathcal{H}} \leq 2t/\varepsilon$.*

Proof. Indeed, for $f \in \mathcal{H}$, let $h = P_t f - f$, so that $P_{\varepsilon+t} f - P_\varepsilon f = P_\varepsilon h$. Then

$$\begin{aligned}
(\varepsilon/t)^2 \|P_{\varepsilon+t} f - P_\varepsilon f\|^2 &\leq \sum_{k=0}^{[\varepsilon/t]} \sum_{l=0}^{[\varepsilon/t]} \|P_{\varepsilon - \frac{k+l}{2}t} P_{\frac{k+l}{2}t} h\|^2 \leq \sum_{k=0}^{[\varepsilon/t]} \sum_{l=0}^{[\varepsilon/t]} \|P_{\frac{k+l}{2}t} h\|^2 \\
&= \sum_{k=0}^{[\varepsilon/t]} \sum_{l=0}^{[\varepsilon/t]} \langle P_{kt} h, P_{lt} h \rangle = \left\langle \sum_{k=0}^{[\varepsilon/t]} \left(P_{(k+1)t} f - P_{kt} f\right), \sum_{l=0}^{[\varepsilon/t]} \left(P_{(l+1)t} f - P_{lt} f\right) \right\rangle \\
&= \left\| P_{t[\varepsilon/t]+t} f - f \right\|^2 \leq (2\|f\|)^2.
\end{aligned}$$

□

The following is a weak version of the Stroock-Varopoulos inequality.

Lemma 5.2. *Let P be a symmetric Markov operator on a probability space (Ω, μ). Then, for any $f \in L^\infty(\Omega, \mu)$ and $p \in (1, \infty)$,*

$$(p-2)^2\, \mathbb{E}|f|^p + 4(p-1)\mathbb{E}\phi_{p/2}(f) P\left(\phi_{p/2}(f)\right) \geq p^2\, \mathbb{E}\phi_{p-1}(f) P f.$$

Proof. Upon simple algebraic transformations, Lemma 2.4 implies that, μ-a.e.,

$$\forall_{a\in\mathbb{Q}}\ \ (p-2)^2(|a|^p + |f|^p) + 8(p-1)\phi_{p/2}(a)\phi_{p/2}(f) \geq p^2\left(a\phi_{p-1}(f) + \phi_{p-1}(a) f\right).$$

Since P is linear and positivity preserving, μ-a.e. we also have

$$\begin{aligned}
\forall_{a\in\mathbb{Q}}\ \ (p-2)^2\left(|a|^p + P(|f|^p)\right) + 8(p-1)\phi_{p/2}(a) P\left(\phi_{p/2}(f)\right)& \\
\geq p^2\left(a P\left(\phi_{p-1}(f)\right) + \phi_{p-1}(a) P f\right)&.
\end{aligned}$$

By the continuity in a, the same holds true μ-a.e. when $\forall_{a\in\mathbb{Q}}$ is replaced by $\forall_{a\in\mathbb{R}}$. In particular, μ-a.e.,

$$\begin{aligned}
(p-2)^2\left(|f|^p + P(|f|^p)\right) + 8(p-1)\phi_{p/2}(f) P\left(\phi_{p/2}(f)\right)& \\
\geq p^2\left(f P\left(\phi_{p-1}(f)\right) + \phi_{p-1}(f) P f\right)&.
\end{aligned}$$

We finish by taking the expectation of both sides and using the symmetry of P (together with the fact that symmetric Markov operators are mean-preserving). □

Definition 5.3. Let $(P_t)_{t\geq 0}$ be a symmetric Markov semigroup on a probability space (Ω, μ). For $M, \varepsilon > 0$ we will denote by $\mathcal{C}(M, \varepsilon)$ the class of all functions of the form $P_\varepsilon \varphi$, where $\varphi \in L^\infty(\Omega, \mu)$ and $\|\varphi\|_\infty \leq M$. Furthermore, for $p \in (1, \infty)$ we set

$$\alpha_{p,M,\varepsilon}(t) = \sup_{f\in\mathcal{C}(M,\varepsilon)} \left((p-1)\mathbb{E}|f|^p + \mathbb{E}|P_t f|^p - p\mathbb{E}\phi_{p-1}(f) P_t f\right).$$

Lemma 5.4. *For any $p \in (1, \infty)$ and $M, \varepsilon > 0$, we have $\alpha_{p,M,\varepsilon}(t)/t \to 0$ as $t \to 0^+$.*

Proof. Let $f\in\mathcal{C}(M,\varepsilon)$, i.e., $f=P_\varepsilon\varphi$ and $\|\varphi\|_\infty\le M$. Let $\rho_p=\sup_{s\neq 1}\frac{|s|^p-ps+p-1}{(|s|+1)^{p-2}(s-1)^2}$ for $p>2$, while for $p\in(1,2]$ let $\rho_p=\sup_{s\neq1}\frac{|s|^p-ps+p-1}{|s-1|^p}$. Note that $\rho_p<\infty$. By the homogeneity,

$$\begin{aligned}(p-1)\mathbb{E}|f|^p+\mathbb{E}|P_tf|^p-p\mathbb{E}\phi_{p-1}(f)P_tf &\le \rho_p\,\mathbb{E}(|P_tf|+|f|)^{p-2}(P_tf-f)^2\\ &\le (2M)^{p-2}\rho_p\cdot\mathbb{E}(P_tf-f)^2,\end{aligned}$$

for $p>2$, and for $p\in(1,2]$ we have

$$(p-1)\mathbb{E}|f|^p+\mathbb{E}|P_tf|^p-p\mathbb{E}\phi_{p-1}(f)P_tf\le\rho_p\,\mathbb{E}|P_tf-f|^p\le\rho_p\cdot\left(\mathbb{E}(P_tf-f)^2\right)^{p/2}.$$

We finish by Lemma 5.1: $\mathbb{E}(P_tf-f)^2=\mathbb{E}(P_{t+\varepsilon}\varphi-P_\varepsilon\varphi)^2\le(2t/\varepsilon)^2\mathbb{E}\varphi^2\le(2M/\varepsilon)^2t^2$. ∎

Now we are in a position to recover Remark 3.2 by a purely semigroup approach:

Proposition 5.5. *Let $(P_t)_{t\ge0}$ be a symmetric Markov semigroup on a probability space (Ω,μ). We also assume that it is a C_0-semigroup on $L^q(\Omega,\mu)$ for some $q\in[1,\infty)$, i.e., $\|P_tg-g\|_q\to0$ as $t\to0^+$ for every $g\in L^q(\Omega,\mu)$. Furthermore, let us assume that there exists a positive constant C such that $\|P_tf\|_2\le e^{-t/C}\|f\|_2$ for every mean-zero function $f\in L^2(\Omega,\mu)$ and every $t>0$. Then, for every $p\in(1,\infty)$, $t>0$, and any mean-zero f,*

$$\|P_tf\|_p\le\exp\left(-\frac{(4p-4)t}{Cp^2(1+\kappa(p)^{-2})}\right)\cdot\|f\|_p. \tag{5.1}$$

Proof. Let $\mathcal{C}_0(M,\varepsilon)=\{f\in\mathcal{C}(M,\varepsilon):\mathbb{E}f=0\}$ and let $L^p_0(\Omega,\mu)$ be the subspace of mean-zero functions in $L^p(\Omega,\mu)$. The C_0-semigroup condition implies that $\bigcup_{M,\varepsilon>0}\mathcal{C}_0(M,\varepsilon)$ is dense in $L^p_0(\Omega,\mu)$. Indeed, for $\varphi\in L^\infty_0(\Omega,\mu)$ by assumption we have $P_\varepsilon\varphi\overset{\varepsilon\to0^+}{\to}\varphi$ in $L^q(\Omega,\mu)$, and thus also in $L^p(\Omega,\mu)$, since $\mathbb{E}|P_\varepsilon\varphi-\varphi|^p\le(2\|\varphi\|_\infty)^{p-q}\mathbb{E}|P_\varepsilon\varphi-\varphi|^q$ if $p>q$ and $\|P_\varepsilon\varphi-\varphi\|_p\le\|P_\varepsilon\varphi-\varphi\|_q$ if $p\le q$. Now it suffices to note that bounded mean-zero functions are dense in $L^p_0(\Omega,\mu)$ and $P_\varepsilon\varphi\in\mathcal{C}_0(\|\varphi\|_\infty,\varepsilon)$.

Since, as a contraction, P_t is uniformly continuous on $L^p(\Omega,\mu)$, it is enough to prove the assertion for $f\in\mathcal{C}_0(M,\varepsilon)$ for every $M,\varepsilon>0$. By assumption,

$$\begin{aligned}\mathbb{E}\phi_{p/2}(f)P_t\left(\phi_{p/2}(f)\right)-\left(\mathbb{E}\phi_{p/2}(f)\right)^2&=\left\|P_{t/2}\left(\phi_{p/2}(f)-\mathbb{E}\phi_{p/2}(f)\right)\right\|_2^2\\ \le e^{-t/C}\left\|\phi_{p/2}(f)-\mathbb{E}\phi_{p/2}(f)\right\|_2^2&=e^{-t/C}\cdot\left(\mathbb{E}|f|^p-\left(\mathbb{E}\phi_{p/2}(f)\right)^2\right).\end{aligned}$$

By Corollary 3.3, $\left(\mathbb{E}\phi_{p/2}(f)\right)^2\le\mathbb{E}|f|^p/\left(\kappa(p)^2+1\right)$. These inequalities, together with Lemma 5.2 and Definition 5.3, yield

$$\mathbb{E}|P_tf|^p\le\left(1-\frac{(4p-4)(1-e^{-t/C})}{(1+\kappa(p)^{-2})p}\right)\cdot\mathbb{E}|f|^p+\alpha_{p,M,\varepsilon}(t).$$

Thus, for positive integers k and n, by a simple induction on k, we have

$$\mathbb{E}|P_{kt/n}f|^p\le\left(1-\frac{(4p-4)(1-e^{-(t/n)/C})}{(1+\kappa(p)^{-2})p}\right)^k\cdot\mathbb{E}|f|^p+k\alpha_{p,M,\varepsilon}(t/n);$$

it suffices to consider t/n instead of t and note that $f\in\mathcal{C}_0(M,\varepsilon)$ implies $P_sf\in\mathcal{C}_0(M,\varepsilon)$ for all $s\ge0$. Taking $k=n$ and $n\to\infty$ ends the proof since, by Lemma 5.4, $n\alpha_{p,M,\varepsilon}(t/n)\to0$. ∎

Remark 5.6. Some authors include the C_0-semigroup assumption for $q = 1$ into the very definition of Markov semigroups. It is easy to prove that a Markov semigroup is a C_0-semigroup on $L^q(\Omega, \mu)$ for some $q \in [1, \infty)$ if and only if it is a C_0-semigroup on $L^q(\Omega, \mu)$ for every $q \in [1, \infty)$. In many cases it is convenient to test the property for $q = 2$. In particular, if $(P_t)_{t\geq 0}$ can be represented, by means of functional calculus, as e^{-tL} for some positive semidefinite self-adjoint operator L on $L^2(\Omega, \mu)$, this property (for $q = 2$) easily follows from the spectral theorem.

For f belonging to the domain of the generator L, we obtain the L^p Poincaré inequality simply by differentiating (5.1) at zero.

6. Talagrand's inequality for tail space

Proof of Theorem 1.16. The argument follows the one in [11]. Let $f\colon \{-1, 1\}^n \to \mathbb{R}$ with $\mathbb{E} f W_S = 0$ for all $S \subseteq \{1, \ldots, n\}$ with $|S| < k$. Hence $\|P_{1/k} f\|_2 \leq e^{-1}\|f\|_2$ and thus

(6.1)
$$(1 - e^{-2})\mathbb{E}f^2 \leq \mathbb{E}f^2 - \mathbb{E}(P_{1/k}f)^2 = -\int_0^{1/k} \frac{d}{dt}\mathbb{E}(P_t)^2\,dt = 2\int_0^{1/k} \mathbb{E}P_t f L P_t f\,dt$$

$$= 2\int_0^{1/k}\sum_{i=1}^n \mathbb{E}(D_i P_t f)^2\,dt = 2\sum_{i=1}^n\int_0^{1/k}\mathbb{E}(P_t D_i f)^2\,dt \leq 2\sum_{i=1}^n\int_0^{1/k}\|D_i f\|^2_{1+e^{-2t}}\,dt,$$

where the last inequality is the usual hypercontractive bound [3, 5, 16].

By Hölder's inequality, for $0 < q < p < 2$ we have

$$\mathbb{E}|g|^p = \mathbb{E}|g|^{\frac{(2-p)q}{2-q}}|g|^{\frac{2(p-q)}{2-q}} \leq (\mathbb{E}|g|^q)^{\frac{2-p}{2-q}}\left(\mathbb{E}g^2\right)^{\frac{p-q}{2-q}}. \tag{6.2}$$

Applying this estimate to $g = D_i f$, $q = 1+e^{-2/k}$, and $p = 1+e^{-2t}$ with $t \in (0, 1/k)$,

$$\begin{aligned}\|D_i f\|^2_{1+e^{-2t}} &\leq \|D_i f\|_2^2\left(\|D_i f\|_{1+e^{-2/k}}/\|D_i f\|_2\right)^{\frac{2\tanh t}{\tanh(1/k)}}\\ &\leq \|D_i f\|_2^2\left(\|D_i f\|_{1+e^{-2/k}}/\|D_i f\|_2\right)^{2tk}\end{aligned}$$

since $t \mapsto \frac{\tanh t}{t}$ is decreasing on $(0, \infty)$. Therefore

$$\begin{aligned}\int_0^{1/k}\|D_i f\|^2_{1+e^{-2t}}\,dt &\leq \|D_i f\|_2^2\int_0^{1/k}\left(\|D_i f\|_{1+e^{-2/k}}/\|D_i f\|_2\right)^{2tk}\,dt\\ &= \|D_i f\|_2^2\frac{1-\left(\|D_i f\|_{1+e^{-2/k}}/\|D_i f\|_2\right)^2}{2k\log\left(\|D_i f\|_2/\|D_i f\|_{1+e^{-2/k}}\right)}\\ &\leq \frac{1}{k}\|D_i f\|_2^2\min\left(1, \frac{1}{2\log\left(\|D_i f\|_2/\|D_i f\|_{1+e^{-2/k}}\right)}\right),\end{aligned}$$

where we have used the fact that $1 - a^{-2} \leq 2\log a$ for $a \geq 1$. Together with (6.1) this ends the proof of the first inequality of Theorem 1.16.

Applying (6.2) to $g = D_i f$, $q = 1$, and $p = 1+e^{-2/k}$, we get $\|D_i f\|_2/\|D_i f\|_{1+e^{-2/k}} \geq \left(\|D_i f\|_2/\|D_i f\|_1\right)^{\tanh(1/k)}$. Since $k\tanh(1/k) \geq \tanh(1)$, the second inequality of Theorem 1.16 easily follows. □

7. The coding tribes function

Recall that in Proposition 1.9 Ben-Or and Linial constructed a Boolean function with mean zero and all of whose influences are $O(\log n/n)$. The results of KKL in Theorem 1.10 imply that it is impossible for the maximum influence of a mean-zero Boolean function to be of lower order. In Question 1.11 Hatami and Kalai asked if the KKL result can be strengthened if the function f satisfies additionally that $\mathbb{E}[fW_S]=0$ for all S with $|S|<k$ where $k(n)\to\infty$ as $n\to\infty$.

The KKL result in fact implies that mean-zero Boolean functions which are invariant under permutation of the inputs have an influence sum which is $\Omega(\log n)$. We first note that by taking the Ben-Or and Linial tribes function f and letting $g(x_1,\ldots,x_n,y_1,\ldots,y_k)=f(x)y_1\ldots y_k$, we obtain a function all of whose Fourier coefficients up to level k vanish and such that its sum of influences is $O(\log n+k)$. Thus one cannot improve on the KKL sum of influence result unless $k/\log n\to\infty$ as $n\to\infty$. In this section we will construct an example of a function all of whose Fourier coefficients up to level $\Omega(\log n)$ vanish and all of whose individual influences are at most $O(\log n/n)$, thus proving Theorem 1.12 and answering in the negative Question 1.11.

We denote by $L^{>k}(\{-1,1\}^n)$ the space of all functions $f\colon\{-1,1\}^n\to\mathbb{R}$ such that $\mathbb{E}fW_S=0$ for all $S\subseteq\{1,\ldots,n\}$ with $|S|\le k$. We denote by $L_+^{>k}(\{-1,1\}^n)$ the space of all functions $f\colon\{-1,1\}^n\to\{-1,1\}$ such that $\mathbb{E}fW_S=0$ for all $S\subseteq\{1,\ldots,n\}$ with $1\le|S|\le k$. The difference between the two families is that the latter functions are allowed to have nonzero expectation.

We will use the convention that 1 and -1 map to the logical values TRUE and FALSE, respectively. Thus for $x_1,\ldots,x_n\in\{-1,1\}$, we have $x_1\vee\cdots\vee x_n=-1$ if and only if $x_1=\cdots=x_n=-1$, and $x_1\wedge\cdots\wedge x_n=1$ if and only if $x_1=\cdots=x_n=1$.

Our strategy is to construct a function in $L_+^{>k}(\{-1,1\}^n)$ with low influences and small mean and then "correct" it so that it has mean zero.

The basic idea behind the construction is the following: we want to mimic the construction of the tribes function. Recall that the tribe function is given by

$$(x_1\wedge\cdots\wedge x_r)\vee\cdots\vee(x_{(b-1)r+1}\wedge\cdots\wedge x_{br}).$$

In our construction, which we call the *Coding Tribes* function, instead of substituting AND functions into the arguments of an OR function, we will substitute functions in $L_+^{>k}$ into the arguments of an OR function.

For example, for $k=1$, instead of the AND function on r bits we will take the function ALLEQ on $r+1$ bits, where $\mathrm{ALLEQ}(x_1,\ldots,x_{r+1})$ takes the value 1 exactly if the x_i are all 1 or all -1. Clearly the function ALLEQ is in $L_+^{>1}$ since it is not correlated with a single bit. To analyze this tribe-like construction we need the following.

Proposition 7.1. *Let $g\colon\{-1,1\}^r\to\{-1,1\}$. Consider a function $f\colon\{-1,1\}^n\to\{-1,1\}$ of the form*

$$f(x)=f_{b,r}(x):=g(x_1,\ldots,x_r)\vee g(x_{r+1},\ldots,x_{2r})\vee\cdots\vee g(x_{(b-1)r+1},\ldots,x_{br}),\ br=n,$$

and where $\mathbb{P}(g=1)\le 2^{-m}$ where $m\le r$. Then

$$\mathbb{E}f=2(1-(1-\mathbb{P}(g=1))^b)-1, \tag{7.1}$$

$$\max_{i=1,\ldots,n} I_i(f)\le 2\times 2^{-m}. \tag{7.2}$$

One can choose b so that

$$|\mathbb{E}f| \leq 2^{-m+1}. \tag{7.3}$$

Proof. Equation (7.1) is obvious, and (7.3) follows from the fact that

$$0 \leq \mathbb{E}[f_{b,r} - f_{b+1,r}] \leq 2^{-m+1}.$$

Equation (7.2) is also easy: for x_i to be pivotal where $i \in \{dr+1, dr+2, \ldots, (d+1)r\}$, we need that the g value of the other x_j in the block with $j \in \{dr+1, dr+2, \ldots, (d+1)r\}$, together with either $x_i = -1$ or $x_i = 1$, evaluate to 1. □

We will also need the following fact.

Proposition 7.2. *Consider a function of the form*

$$f(x) = F(g_1(x_1, \ldots, x_r), g_2(x_{r+1}, \ldots, x_{2r}) \ldots g_b(x_{(b-1)r+1}, \ldots, x_{br})), \quad br = n,$$

where $\{g_j\}_{j=1}^b$ are Boolean functions all taking the values $\{0,1\}$ or all taking the values $\{-1,1\}$. Assume further that $g_j \in L_+^{>k}(\{-1,1\}^r)$ for all $j = 1, \ldots, b$. Then $f \in L_+^{>k}(\{-1,1\}^n)$.

Proof. Since we can write F as a multilinear polynomial of its binary inputs, it suffices to show that each product of a subset of the g_i is in $L_+^{>k}$. By induction it suffices to show this for two functions, which is immediate. □

We are particularly interested in the case where g is an indicator of a linear code. Recall that a *linear code* is a linear subspace of $\{0,1\}^n$, where we treat $\{0,1\}$ as the field of two elements. The *minimal weight* $w(C)$ of a code C is defined by

$$w(C) = \min\{\|x\|_1 : 0 \neq x \in C\},$$

where $\|(x_1, \ldots, x_n)\|_1 := \sum_{i=1}^n |x_i|$ is the Hamming weight of x. The *dual* code of C, denoted $C^\perp \subseteq \{0,1\}^n$, is given by

$$C^\perp := \big\{y \in \{0,1\}^n : \sum_{i=1}^n x_i y_i = 0 \mod 2,\ \forall\, x \in C\big\}.$$

Given a code $C \subseteq \{0,1\}^n$, we will write $g_C \colon \{-1,1\}^n \to \{-1,1\}$ for the following Boolean function:

$$g_C(x_1, \ldots, x_n) := \begin{cases} 1, & \text{if } ((1-x_1)/2, \ldots, (1-x_n)/2) \in C, \\ -1, & \text{if } ((1-x_1)/2, \ldots, (1-x_n)/2) \notin C. \end{cases}$$

By the MacWilliams identities [13] (see e.g. [10, Lemma 3.3]) we have:

Proposition 7.3. *Let C be a linear code. Then $g_C \in L_+^{>k}$ if and only if $w(C^\perp) > k$.*

For example, for $C = \{(0, \ldots, 0), (1 \ldots 1)\}$, we have $g_C(x) = 1$ if and only if $x = \pm(1, \ldots, 1)$, and the code $C^\perp$ consists of all codewords x with $\|x\|_1$ even, so $C^\perp$ has minimal weight $w(C^\perp) = 2$.

Proposition 7.4. *There exists a constant $\gamma > 1$ such that for every $m > 0$, there exists a function $g : \{-1,1\}^{\lceil \gamma m \rceil} \to \{-1,1\}$ with $g \in L_+^{\geq m}$ and $2^{-3m} \leq \mathbb{P}[g = 1] \leq 2^{-m}$.*

Proof. The function g will be constructed via the dual of a "good code". It is well known that good codes exist [12]. Such (linear) codes $C \subseteq \{0,1\}^{m'}$ have the following properties (where δ is independent of m'):

- $(3/4)m' \geq \dim(C) \geq m'/4$,
- $w(C) \geq \delta m'$, where $\delta > 0$.

We let $g := g_{C^\perp}$. Then $\mathbb{P}[g = 1] = 2^{-\dim(C^\perp)}$, so

$$2^{-3m'/4} \leq \mathbb{P}[g = 1] \leq 2^{-m'/4},$$

and by Proposition 7.3, $g \in L_+^{\geq k}$ where

$$w(C^{\perp\perp}) = w(C) \geq \delta m' = k.$$

Setting $\gamma = \max(4, \delta^{-1})$, the proof follows. □

Propositions 7.1 and 7.4 are already enough to prove that Harper's inequality cannot be improved for tail spaces.

Proof of Theorem 1.15. Let $b = 1$ in Proposition 7.1 and use g from Proposition 7.4. Setting $n := \lceil \gamma m \rceil$ we get $g\colon \{-1,1\}^n \to \{-1,1\}$ with $g \in L_+^{\geq m}$, $\mathbb{E}g = 2\mathbb{P}(g = 1) - 1$, $\max_{i=1,\ldots,n} I_i g \leq 2\mathbb{P}(g = 1)$. Then, the function $f := (1 + h)/2 = 1_{(h=1)}$ satisfies $f\colon \{-1,1\}^n \to \{0,1\}$, $\sum_{i=1}^n I_i f \leq 2n\mathbb{P}(g = 1) \leq \gamma m \mathbb{P}(g = 1)$, and $\mathbb{E}f = 1/2 + \mathbb{E}h/2 = \mathbb{P}(g = 1)$. From Proposition 7.4, $\mathbb{P}[g = 1] \leq 2^{-m}$. That is,

$$\frac{\sum_{i=1}^n I_i f}{(\mathbb{E}f)\log(1/\mathbb{E}f)} \leq \frac{\gamma m}{\log(1/\mathbb{E}f)} \leq \gamma \frac{m}{m} = \gamma.$$

□

Substituting g from Proposition 7.4 into Propositions 7.1 and 7.2, and letting $n = mb$, where b is chosen so that $\mathbb{E}[f]$ is as close to 0 as possible (so that $m = O(\log n)$), we obtain:

Theorem 7.5. *There exists a family of Boolean functions $f = f_n : \{-1,1\}^n \to \{-1,1\}$ such that*

- *$f \in L_+^{>\Omega(\log n)}(\{-1,1\}^n)$,*
- *for all $i \in \{1,\ldots,n\}$, $I_i(f) \leq O((\log n)/n)$,*
- *$|\mathbb{E}f| \leq O((\log n)/n)$.*

We now wish to find similar functions that have zero mean.

Corollary 7.6. *There exists a family of functions $g = g_n : \{-1,1\}^{2n} \to \{-1,0,1\}$ such that*

- *$g \in L^{>\Omega(\log n)}(\{-1,1\}^{2n})$,*
- *for all $i \in \{1,\ldots,n\}$, $I_i(g) \leq O(\log n/n)$,*
- *$\mathbb{P}[g = 1] = 1/4 - O((\log n)/n)$, $\mathbb{P}[g = -1] = 1/4 - O((\log n)/n)$.*

Proof. Take f from Theorem 7.5 and define

$$g(x_1, \dots, x_n, y_1, \dots, y_n) := \frac{1}{2}(f(x_1, \dots, x_n) - f(y_1, \dots, y_n)).$$

□

With a little more work we can construct functions with the desired properties taking only values 0 and 1. For this we note that Proposition 7.4 implies the following corollary.

Corollary 7.7. *There exists a constant $\gamma > 1$ such that for every n, there exists a function $g : \{-1,1\}^{\gamma n} \to \{0,1\}$ with $g \in L_+^{\geq n}$ and $\mathbb{P}[g = 1] = 2^{-n-d}$ for some nonnegative integer d. Moreover, g has the following property: For $y \in \{-1,1\}^{\gamma n}$, write $g_y(x) = g(y_1 x_1, \dots, y_n x_n)$. Then for all $y, y' \in \{-1,1\}^{\gamma n}$ we either have $g_y = g_{y'}$ or the function $g_y g_{y'}$ is identically* 0.

Proof. Let h be the function from Proposition 7.4 and let $g = 1_{(h=1)} = (h+1)/2$. Then all the stated properties but the last one clearly hold if γ is large enough. The last property follows from the fact that cosets of linear codes are either identical or disjoint. □

Lemma 7.8. *There exists a constant $\gamma > 1$ such that the following holds. Let $0 \leq t < 2^n$, $t \in \mathbb{Z}$. Then there exists a function $f : \{-1,1\}^{\gamma n} \to \{0,1\}$ such that $\mathbb{E} f = t/2^n$ and $f \in L_+^{\geq n}(\{-1,1\}^{\gamma n})$.*

Proof. From Corollary 7.7 in the case $t = 1$ we can find a function in $L_+^{\geq n}$ and $\mathbb{E}[f] = 2^{-n-d}$, where d is a nonnegative integer. The general case follows by taking $h = \sum_i g_{y^i}$ where y^i are chosen so that $g_{y^i} g_{y^j} = 0$ for $i \neq j$. □

Theorem 7.9. *There exists a family of Boolean functions $G = G_n : \{-1,1\}^n \to \{-1,1\}$ such that*

- *$G \in L^{>\Omega(\log n)}$,*
- *for all $i \in \{1, \dots, n\}$, $I_i(G) \leq O((\log n)/n)$.*

Proof. We revise the construction of Theorem 7.5 as follows. Using Lemma 7.8, choose $g_0, \dots, g_b \colon \{-1,1\}^{\lceil \gamma m \rceil} \to \{-1,1\}$ all in $L_+^{\geq m}$. Moreover, for $1 \leq i \leq b$, let $\mathbb{P}[g_i = 1] = 2^{-m}$ and for $i = 0$, let $\mathbb{P}[g_0 = 1] = 4 \times 2^{-m}$.

We choose b to be the largest integer so that

$$\mathbb{E} f = (1 - (1 - 2^{-m})^b (1 - 2^{-m+2})) - 1 > 0$$

and let $n = (b+1)\lceil \gamma m \rceil$. Note that $m = O(\log n)$ and that

$$0 \leq \mathbb{E} f \leq 2^{-m}.$$

By Lemma 7.8, let $h \colon \{-1,1\}^{\lceil \gamma n \rceil} \to \{0,1\}$ with

$$2\mathbb{E} h = \mathbb{E} f / \mathbb{P}[g_0 = 1] \tag{7.4}$$

and such that h is in $L_+^{\geq n}(\{-1,1\}^{\gamma n})$.

Let $G \colon \{-1,1\}^{\gamma n + n} \to \{-1,1\}$ be a function of the x and y given by

$$G(x, y) := f(x) - 2 \cdot g_0(x) \cdot h(y).$$

Then clearly $G(x,y) \in L^m_+(\{-1,1\}^{\lceil \gamma n\rceil + n})$, and moreover $\mathbb{E}g = 0$ by (7.4). So we have $G \in L^m(\{-1,1\}^{\lceil \gamma n\rceil + n})$. Finally, since $f(x)$ and $g_0(x)$ have all of their influences $O((\log n)/n)$, the same is true for all of the x variables in g. Moreover, a y variable can be influential if and only if $g_0(x) = 1$. Therefore the influence of all of the y variables is also $O((\log n)/n)$. The proof follows. □

Acknowledgements

The authors thank Fedor Nazarov for sharing his complex interpolation arguments of the main theorem, which strongly inspired the proof of Section 4. The authors also thank Ryan O'Donnell for helpful discussions and references, particularly [19] and [8]. Thanks also to Michel Ledoux, Camil Muscalu, Assaf Naor, and Bob Strichartz for helpful discussions. Thanks to Cyril Roberto for informing the authors about [4].

References

[1] William Beckner, *A generalized Poincaré inequality for Gaussian measures*, Proc. Amer. Math. Soc. **105** (1989), no. 2, 397–400, DOI 10.2307/2046956. MR954373

[2] Michael Ben-Or and Nathan Linial, *Collective coin flipping*, Randomness and Computation (S. Micali, ed.), Academic Press, New York, 1989, pp. 91–115.

[3] Aline Bonami, *Étude des coefficients de Fourier des fonctions de $L^p(G)$* (French, with English summary), Ann. Inst. Fourier (Grenoble) **20** (1970), no. fasc. 2, 335–402 (1971). MR0283496

[4] Patrick Cattiaux, Arnaud Guillin, and Cyril Roberto, *Poincaré inequality and the L^p convergence of semi-groups*, Electron. Commun. Probab. **15** (2010), 270–280, DOI 10.1214/ECP.v15-1559. MR2661206

[5] Leonard Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083. MR0420249

[6] L. H. Harper, *Optimal assignments of numbers to vertices*, J. Soc. Indust. Appl. Math. **12** (1964), 131–135. MR0162737

[7] Masanori Hino, *Exponential decay of positivity preserving semigroups on L^p*, Osaka J. Math. **37** (2000), no. 3, 603–624. MR1789439

[8] Svante Janson, *Gaussian Hilbert spaces*, Cambridge Tracts in Mathematics, vol. 129, Cambridge University Press, Cambridge, 1997. MR1474726

[9] Jeff Kahn, Gil Kalai, and Nathan Linial, *The influence of variables on Boolean functions*, Proc. of 29th Annual IEEE Symposium on Foundations of Computer Science, 1988, pp. 68–80.

[10] Subhash Khot and Assaf Naor, *Nonembeddability theorems via Fourier analysis*, Math. Ann. **334** (2006), no. 4, 821–852, DOI 10.1007/s00208-005-0745-0. MR2209259

[11] M. Ledoux, *Logarithmic Sobolev inequalities for unbounded spin systems revisited*, Séminaire de Probabilités, XXXV, Lecture Notes in Math., vol. 1755, Springer, Berlin, 2001, pp. 167–194, DOI 10.1007/978-3-540-44671-2_13. MR1837286

[12] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes. I*, North-Holland Mathematical Library, vol. 16, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977. MR0465509

[13] Jessie MacWilliams, *A theorem on the distribution of weights in a systematic code*, Bell System Tech. J. **42** (1963), 79–94. MR0149978

[14] Manor Mendel and Assaf Naor, *Nonlinear spectral calculus and super-expanders*, Publ. Math. Inst. Hautes Études Sci. **119** (2014), 1–95, DOI 10.1007/s10240-013-0053-2. MR3210176

[15] P.-A. Meyer, *Transformations de Riesz pour les lois gaussiennes* (French), Seminar on probability, XVIII, Lecture Notes in Math., vol. 1059, Springer, Berlin, 1984, pp. 179–193, DOI 10.1007/BFb0100043. MR770960

[16] Edward Nelson, *The free Markoff field*, J. Functional Analysis **12** (1973), 211–227. MR0343816

[17] Ryan O'Donnell, *Analysis of Boolean functions*, Cambridge University Press, New York, 2014. MR3443800

[18] D. W. Stroock, *An introduction to the theory of large deviations*, Universitext, Springer-Verlag, New York, 1984. MR755154

[19] Michel Talagrand, *A conjecture on convolution operators, and a non-Dunford-Pettis operator on L^1*, Israel J. Math. **68** (1989), no. 1, 82–88, DOI 10.1007/BF02764970. MR1035882

[20] N. Th. Varopoulos, *Hardy-Littlewood theory for semigroups*, J. Funct. Anal. **63** (1985), no. 2, 240–260, DOI 10.1016/0022-1236(85)90087-4. MR803094

Department of Mathematics, University of California Los Angeles, Los Angeles, California 90095-1555

E-mail address: heilman@cims.nyu.edu

Department of Statistics, University of Pennsylvania, Philadelphia, Pennsylvania 19104 — and — Departments of Statistics and Computer Science, University of California Berkeley, Berkeley, California 94720

Current address: Department of Mathematics & Statistics Group, IDSS, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

E-mail address: elmos@mit.edu

Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland

E-mail address: koles@mimuw.edu.pl