Improved log-Sobolev inequalities, hypercontractivity and uncertainty principle on the hypercube

Yury Polyanskiy[a], Alex Samorodnitsky[b]

[a]*Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02139 USA.*

[b]*School of Engineering and Computer Science, The Hebrew University of Jerusalem, Jerusalem 91904, Israel.*

Abstract

We develop a new class of log-Sobolev inequalities (LSIs) that provide a non-linear comparison between the entropy and the Dirichlet form. For the hypercube, these LSIs imply a new version of the hypercontractivity for functions of small support. As a consequence, we derive a sharp form of the uncertainty principle for the hypercube: a function whose energy is concentrated on a set of small size, and whose Fourier energy is concentrated on a small Hamming ball must be zero. The tradeoff we derive is asymptotically optimal. We observe that for the Euclidean space, an analogous (asymptotically optimal) tradeoff follows from the sharp form of Young's inequality due to Beckner. As an application, we show how uncertainty principle implies a new estimate of the metric properties of linear maps $\mathbb{F}_2^k \to \mathbb{F}_2^n$.

Keywords: Hamming space, log-Sobolev inequality, hypercontractivity, Fourier transform on the hypercube, uncertainty principle, coding theory

---

Contents

Email addresses: `yp@mit.edu` (Yury Polyanskiy), `salex@cs.huji.ac.il` (Alex Samorodnitsky)

June 23, 2016

1. Introduction

1.1. Definitions, background

We introduce some standard notions for continous-time semigroups on finite state spaces, e.g. [1, Section 1.7.1]. Consider a finite alphabet $\mathcal{X}$ and a matrix $(L_{x,y})_{x,y\in\mathcal{X}}$ such that 1) $L_{x,y} \ge 0$ for $x \neq y$; and 2) $\sum_{y\in\mathcal{X}} L_{x,y} = 0$ for all x. Then $T_t = e^{tL}$ is a stochastic semigroup, for which we assume that π is a stationary measure. We define $\|f\|_p \triangleq \mathbb{E}^{\frac{1}{p}}[|f|^p]$ and $(f,g) = \mathbb{E}[fg]$ with expectation over π. The Dirichlet form of semigroup T_t is

$$\mathcal{E}(f,g) \triangleq -\sum_{x,y} L_{x,y} f(y)g(x)\pi(x) = \mathbb{E}_\pi[(-Lf)g]\,.$$

We also define $T_t^{\otimes n}$ – a product semigroup on $\mathcal{X}^n$ – and notice that its Dirichlet form is given by

$$\mathcal{E}_n(f,g) \triangleq \sum_{k=1}^{n} \sum_{x_{\hat{k}}\in\mathcal{X}^{n-1}} h(x_{\hat{k}}) \prod_{j\neq k} \pi(x_j)\,, \tag{1}$$

where $x_{\hat{k}} = (x_1,\dots,x_{k-1},x_{k+1},\dots,x_n)$ and $h(x_{\hat{k}}) = \mathcal{E}(f(x_{\hat{k}},\cdot), g(x_{\hat{k}},\cdot))$ is the action of Dirichlet form $\mathcal{E}$ on k-th coordinate of f and g with other coordinates held frozen.

We say that a semigroup admits a p-logarithmic Sobolev inequality (p-LSI for short) if for some constant α_p

$$\mathrm{Ent}_\pi(f^p) \le \frac{1}{\alpha_p}\mathcal{E}(f, f^{p-1})\,, \tag{2}$$

where

$$\mathrm{Ent}_\pi(g) \triangleq \mathbb{E}_\pi[g(X)\ln\frac{g}{\mathbb{E}[g]}] = \mathbb{E}[g]D(\pi^{(g)}\|\pi)\,,$$

where $\pi^{(g)}(x) \triangleq \frac{g(x)\pi(x)}{\mathbb{E}[g]}$ and $D(\cdot\|\cdot)$ is the Kullback-Leibler divergence.

We note that $\mathcal{E}(f, f^{p-1}) \ge 0$ for $p > 1$ and $\mathcal{E}(f, f^{p-1}) < 0$ for $p < 1$ and this implies corresponding signs for constants α_p. As $p \to 1$ we have $\mathcal{E}(f, f^{p-1}) \to 0$ and so we need to renormalize by $\frac{1}{(p-1)}$ in this limit. Consequently, we define 1-LSI as

$$\mathrm{Ent}_\pi(f) \le \frac{1}{\alpha_1}\mathcal{E}(f, \ln f)\,,$$

which is required to hold for all $f > 0$ on $\mathcal{X}$.

We will mostly deal in this paper with a special case of a hypercube. Namely, we set $\mathcal{X} = \{0, 1\}$, $L_{x,y} = -1\{x = y\} + 1/2$, $\pi = \text{Bern}(1/2)$ and

$$T_t f(x) = f(x)\frac{1 + e^{-t}}{2} + f(1 - x)\frac{1 - e^{-t}}{2} . \tag{3}$$

For this case the best LSI constants are $\alpha_p = \frac{2(p-1)}{p^2}$, see [2, Theorem 2.2.8], and the Dirichlet form takes particularly simple form:

$$\mathcal{E}_n(f, g) = -\frac{1}{2}(\triangle f, g), \qquad \triangle f(x) \triangleq \sum_{y:y\sim x} (f(y) - f(x)) \tag{4}$$

$$\mathcal{E}_n(f, f) = \frac{1}{4}\sum_{x\sim y}(f(x) - f(y))^2 2^{-n} , \tag{5}$$

where $x \sim y$ denotes any $x, y \in \{0, 1\}^n$ differing in precisely one coordinate.

We also quote one inequality known as Mrs. Gerber's lemma [3], or MGL:

$$\frac{1}{n}\frac{\text{Ent}_{\pi^n}(T_t^{\otimes n} f)}{\mathbb{E}[f]} \le \ln 2 - m\left(t, \frac{1}{n}\frac{\text{Ent}_{\pi^n}(f)}{\mathbb{E}[f]}\right) , \tag{6}$$

where $m(t, x) = h(h^{-1}(\ln 2 - x) * \frac{1-e^{-t}}{2})$, where $h(x) = -x \ln x - (1 - x) \ln(1 - x)$ is the binary entropy function, $h^{-1} : [0, \ln 2] \to [0, 1/2]$ is its functional inverse and $a * b = (1 - a)b + (1 - b)a$ is the binary convolution.

A less cryptic restatement of MGL is the following:

$$\text{Ent}_{\pi^n}(T_t^{\otimes n} f) \le \text{Ent}_{\pi^n}(T_t^{\otimes n} f_{iid}) , \tag{7}$$

where $f_{iid}(x) = \prod_{k=1}^n f_1(x_k)$ with $f_1(\cdot)$ selected so that a) in (7) the equality holds for $t = 0$; b) $\mathbb{E}[f_{iid}] = \mathbb{E}[f]$. In other words, MGL states that among all functions f on the hypercube, Ent decreases slowest for product functions.

1.2. Motivation and Organization

We motivate our investigation by the following two questions:

- Log-Sobolev inequality implies an estimate of the form

  $$\text{Ent}(T_t^{\otimes n} f) \le e^{-Ct}\text{Ent}(f) .$$

  However, for the hypercube a stronger estimate is given by the MGL (6). *Can MGL be derived from some strengthening of LSI?*

  Note that by a method of comparison of Dirichlet forms, results derived from log-Sobolev inequalities can then be extended to semigroups other than $T_t^{\otimes n}$. As an illustration, note that [4, Example 3.3] estimates speed of convergence of the Metropolis chain on $\{0, \dots n\}$ by comparing to $T_t^{\otimes n}$. Our methods allow to show better estimates, similar to (6).

- Hypercontractivity inequality for the hypercube (known as Bonami or Bonami-Beckner theorem) says

$$\|T_t^{\otimes n} f\|_p \le \|f\|_{p_0}\,, \qquad p(t) = 1 + (p_0 - 1)e^{2t},\, p_0 \ge 1\,. \tag{8}$$

  This is well known to be tight in the sense that for any $q > p(t)$ we can find f s.t. $\|T_t f\|_q > \|f\|_{p_0}$. However, such f will be very close to identity (for this particular semigroup). *Is it possible to improve the range of (p,q) in (8) provided f is far from identity (say in the sense of $|\operatorname{supp} f| \ll 2^n$)?*

  For example, it is clear that $\|T_t^{\otimes n}\|_{1\to\infty} = (1+e^{-t})^n$. If f has small support, we have $\|f\|_{p_0} \ge e^{n\rho_0}\|f\|_1$, where $\rho_0 = (1-\frac{1}{p_0})\frac{1}{n}\ln\frac{2^n}{|\operatorname{supp} f|}$ and thus

$$\|T_t f\|_\infty \le \|f\|_{p_0}, \qquad \forall t \ge \ln\frac{1}{e^{\rho_0}-1} \tag{9}$$

  which is a significant improvement of (8) for large times t.

Both of these questions will be answered positively.

The structure of the paper is the following. In Section 2 we define the main concept of this paper – the non-linear LSIs and prove some of its consequences, such as refined hypercontractivity and general MGL. In Section 3 we switch from general theory to a particular case of the hypercube. We establish explicit forms of new LSIs and new hypercontractive estimates for functions of small support. In Section 4 we derive apply the latter to establish a sharp version of the uncertainty principle on the hypercube (i.e. tradeoff between the sizes of supports of the function and its Fourier image). Finally, Section 5 applies the uncertainty principle to showing an intriguing property of linear maps $\mathbb{F}_2^k \to \mathbb{F}_2^n$.

2. Non-linear log-Sobolev inequalities

Let us introduce a family of non-linear log-Sobolev inequalities.

Definition 1. *For $p \ge 1$ and a concave, non-negative function $\Phi_p : [0,\infty) \to [0,\infty)$ with $\Phi_p(0) = 0$, let us define a (p,Φ_p)-LSI as*

$$\frac{\operatorname{Ent}(f^p)}{\mathbb{E}[f^p]} \le \Phi_p\left(\frac{\mathcal{E}(f,f^{p-1})}{\mathbb{E}[f^p]}\right), \tag{10}$$

where for $p = 1$ we understand $\mathcal{E}(f,f^{p-1}) = \mathcal{E}(f,\ln f)$. For $p < 1$ and a concave, non-negative function $\Phi_p : (-\infty,0] \to [0,\infty)$ we define (p,Φ_p)-LSI as (10). When convenient, we will restate (p,Φ_p)-LSI in the form

$$\pm\frac{\mathcal{E}(f,f^{p-1})}{\mathbb{E}[f^p]} \ge b_p\left(\frac{\operatorname{Ent}(f^p)}{\mathbb{E}[f^p]}\right), \tag{11}$$

where $b_p : [0,\infty) \to [0,\infty]$ is convex and increasing with $b_p(0) = 0$ (i.e. $b_p = \pm\Phi_p^{-1}$ with $b_p = \infty$ if Φ_p is constant for $x \ge x_0$). The $\pm$ is taken to be $+$ for $p \ge 1$ and $-$ for $p < 1$.

Remark 1. *For convenience we define Φ_p and b_p on $[0,\infty)$ even though the arguments in (10) and (11) may be universally bounded by constants $< \infty$. Note also that concave non-negative function on $[0,\infty)$ must be increasing on $[0,a)$ and then constant on $[a,\infty)$ (either interval could be empty).*

It is clear from concavity of Φ_p that the linear-LSIs (2) are obtained by taking $\frac{1}{\alpha_p} = \frac{d}{dx}\Big|_{x=0} \Phi_p$. We briefly review history of such inequalities:

- For a Lebesgue measure on $\mathbb{R}^n$ and $\mathcal{E}(f,g) = \int (\nabla f, \nabla g)$ the $p = 2$ inequality takes the form:

  $$\int_{\mathbb{R}^n} f^2(x) \ln f^2 dx \le \frac{n}{2} \ln \left(\frac{2}{n\pi e} \int_{\mathbb{R}^n} \|\nabla f\|^2 dx \right), \qquad \int f^2 = 1\,.$$

  It appeared in information theory [5, (2.3)] and [6] as a solution to the problem of minimizing Fisher information subject to differential entropy constraint (the minimizer is Gaussian density).

- More generally, the $p = 2$ inequalities were introduced into operator theory by Davies and Simon [7] under the name of entropy-energy inequalities; see [8] for a survey.

- A $p = 2$ inequality for the hypercube was proved in [9] for the purpose of showing that the Faber-Krahn problem on the hypercube is asymptotically solved by a Hamming ball. Same reference mentioned [9, paragraph after (11)] a tightening of hypercontractivity (8) for $p_0 = 2$ and functions of large entropy, although no proof was published at the time.

Here we prove general results about non-linear LSIs.

Theorem 1. *(Tensorization) Suppose that (p,Φ_p)-LSI holds for semigroup $(\mathcal{X},\pi,T_t,\mathcal{E})$. Then for all $n \ge 1$ the $(p, n\Phi_p(\frac{1}{n}\cdot))$-LSI holds for $(\mathcal{X}^n, \pi^n, T_t^{\otimes n}, \mathcal{E}_n)$. In other words, for all $f : \mathcal{X}^n \to \mathbb{R}_+$ we have*

$$\frac{1}{n}\frac{\mathrm{Ent}_{\pi^n}(f^p)}{\mathbb{E}_{\pi^n}[f^p]} \le \Phi_p\left(\frac{1}{n}\frac{\mathcal{E}_n(f, f^{p-1})}{\mathbb{E}_{\pi^n}[f^p]}\right), \tag{12}$$

where $\pi^n = \prod_{k=1}^n \pi$ – a product measure on $\mathcal{X}^n$ and $\mathcal{E}_n$ is the Dirichlet form associated to the product semigroup (1).

Theorem 2. *(General MGL) Suppose a semigroup T_t admits a $(1,\Phi_1)$-LSI. Let $b_1 = \Phi_1^{-1}$ be a convex, strictly increasing inverse of Φ_1 and assume that the differential equation*

$$\frac{d}{dt}\tilde{\rho}(t) = -b_1(\tilde{\rho}(t))$$

has a $\mathcal{C}^1$-solution $\tilde{\rho}(t)$ on $[0,t_0)$ with $\tilde{\rho}(0) > 0$. Then for any $f : \mathcal{X}^n \to \mathbb{R}_+$ with $\frac{1}{n}\frac{\mathrm{Ent}(f)}{\mathbb{E}[f]} \le \tilde{\rho}(0)$ we have

$$\mathrm{Ent}(T_t^{\otimes n} f) \le n\tilde{\rho}(t)\mathbb{E}[f] \qquad \forall 0 \le t < t_0\,.$$

Theorem 3. *(Hypercontractivity) Fix a non-constant function $f : \mathcal{X}^n \to \mathbb{R}^+$ and $p_0 \in (1, \infty)$. Then there is a finite t_0 and a unique function $p(t)$ on $[0, t_0)$ satisfying $\|T_t^{\otimes n} f\|_{p(t)} = \|f\|_{p_0}$. This function is $\mathcal{C}^\infty$-smooth, strictly increasing and surjective onto $[p_0, \infty)$ with $p(0) = p_0$. Furthermore, if a semigroup T_t admits a (p, Φ_p)-LSI for each $p \ge p_0$, then*

$$\frac{d}{dt}p(t) \ge \frac{p(t)(p(t)-1)}{\rho_0} b_{p(t)}\left(\frac{p(t)\rho_0}{p(t)-1}\right), \qquad \rho_0 = \frac{1}{n}\ln\frac{\|f\|_{p_0}}{\|f\|_1}. \tag{13}$$

Proof of Theorem 1. Let us consider the case $n = 2$ only. For a function $f(x_1, x_2)$ denote by $\mathrm{Ent}_{\pi_i}(f^p)$ the entropy evaluated only along x_i, $i = 1, 2$. Then, from standard chain-rule and convexity of Ent we have

$$\mathrm{Ent}_{\pi_1\times\pi_2}(f^p) = \mathbb{E}_{X_1}[\mathrm{Ent}_{\pi_2}(f^p)] + \mathrm{Ent}_{\pi_1}(\mathbb{E}_{X_2}[f^p]) \tag{14}$$
$$\le \mathbb{E}_{X_1}[\mathrm{Ent}_{\pi_2}(f^p)] + \mathbb{E}_{X_2}[\mathrm{Ent}_{\pi_1}(f^p)] \tag{15}$$

Now, we apply Φ_p-LSI to each term (not forgetting appropriate normalization). For example, for the first term we get

$$\mathbb{E}_{X_1}[\mathrm{Ent}_{\pi_2}(f^p)] \le \mathbb{E}_{X_1}\left[\Phi_p\left(\frac{\mathcal{E}_{\pi_2}(f, f^{p-1})}{\mathbb{E}_{X_2}[f^p]}\right)\mathbb{E}_{X_2}[f^p]\right] \tag{16}$$
$$\le \Phi_p\left(\frac{\mathbb{E}_{X_1}[\mathcal{E}_{\pi_2}(f, f^{p-1})]}{\mathbb{E}_{X_1,X_2}[f^p]}\right)\mathbb{E}_{X_1,X_2}[f^p], \tag{17}$$

where in the second step we used Jensen's and the fact that

$$(x, y) \mapsto \Phi(\frac{x}{y})y$$

is jointly concave for any concave function Φ. Now plugging (17) (and its analog for the second term) into (15) and after applying Jensen's again we get

$$\frac{1}{2}\frac{\mathrm{Ent}_{\pi_1\times\pi_2}(f^p)}{\mathbb{E}_{X_1,X_2}[f^p]} \le \Phi_p\left(\frac{1}{2}\frac{\mathbb{E}_{X_1}[\mathcal{E}_{\pi_2}(f, f^{p-1})] + \mathbb{E}_{X_2}[\mathcal{E}_{\pi_1}(f, f^{p-1})]}{\mathbb{E}_{X_1,X_2}[f^p]}\right),$$

which is precisely (12). The $n > 2$ is treated similarly. □

Proof of Theorem 2. Since the statement is scale-invariant, we assume $\mathbb{E}[f] = 1$. Define $\rho(t) \triangleq \frac{1}{n}\mathrm{Ent}(T_t^{\otimes n} f)$. Consider the identity

$$\frac{d}{dt}\mathrm{Ent}(T_t^{\otimes n} f) = -\mathcal{E}(T_t^{\otimes n} f, \ln T_t^{\otimes n} f).$$

From tensorizing the $(1, \Phi_1)$-LSI we get

$$\frac{1}{n}\mathcal{E}(T_t^{\otimes n} f, \ln T_t^{\otimes n} f) \ge b_1(\rho(t)),$$

and hence

$$\rho'(t) \le -b_1(\rho(t)).$$

Let us introduce $\alpha(t) = \ln \rho(t) - \ln \tilde{\rho}(t)$, then we have for $\alpha(t)$ the following

$$\alpha'(t) \le -e^{\alpha(t)} \left(\Psi(\tilde{\rho}(t) e^{\alpha(t)}) - \Psi(\tilde{\rho}(t)) \right) , \tag{18}$$

where $\Psi(x) = b_1(x)/x$ is a non-decreasing function of $x \ge 0$. We know $\alpha(0) \le 0$. Suppose that for some $t_0 > 0$ we have $\alpha(t_0) > 0$. Let $t_1 = \sup\{0 \le t < t_0 : \alpha(t) = 0\}$. From continuity of α we have $t_1 < t_0$, $\alpha(t_1) = 0$ and $\alpha(t) > 0$ for all $t \in (t_1, t_0]$. From mean value theorem, we have for some $t_2 \in (t_1, t_0)$ that $\alpha'(t_2) > 0$. But then from monotonicity of Ψ, we have

$$\Psi(\tilde{\rho}(t_2) e^{\alpha(t_2)}) - \Psi(\tilde{\rho}(t_2)) \ge 0 ,$$

contradicting (18). Hence $\alpha(t_0) \le 0$ for all $t_0 > 0$. □

Proof of Theorem 3. Since all the statements are scale-invariant, we assume $\mathbb{E}[f] = 1$. To avoid clutter, we will write T_t instead of $T_t^{\otimes n}$. We define on $\mathbb{R}_+^2$ the following function

$$\phi(t, \xi) \stackrel{\triangle}{=} \ln \|T_t f\|_{\frac{1}{\xi}} .$$

It is clear that ϕ is monotonically decreasing in ξ. Steepness of ϕ in ξ encodes information about non-uniformity of $T_t f$. *As time progresses, $\xi \mapsto \phi(t, \xi)$ converges to an all-zero function.* MGL, LSI and hypercontractivity are estimates on the speed of this relaxation.

We summarize the information we have about $\phi(t, \xi)$ assuming f is non-constant:

- A consequence of Hölder's inequality, cf. [10, Theorems 196-197], implies $\xi \mapsto \ln \|g\|_{\frac{1}{\xi}}$ is strictly convex, unless $g = c 1_S$ (a scaled indicator), in which case the function of is linear in ξ. Thus $\phi(t, \xi)$ is convex in ξ.

- We have

  $$\phi(0, \xi_2) \ge \phi(0, \xi_1) + (\xi_1 - \xi_2) \ln \frac{1}{\pi^n[\operatorname{supp} f]} \qquad \forall \xi_2 < \xi_1$$

  with equality iff $f = c 1_S$ (scaled indicator).

- Note that $T_t f = 0$ has only $f = 0$ as solution (indeed, $\det e^{tL} = e^{\operatorname{tr} L} \neq 0$ since $\mathcal{X}$ is finite). So $\phi(t, \xi)$ is finite and infinitely differentiable in (t, ξ).

- The function $t \mapsto \phi(t, \xi)$ is strictly decreasing from $\phi(0, \xi)$ to 0 for any $\xi < 1$ and strictly increasing from $\phi(0, \xi)$ to 0 for $\xi > 1$. Indeed, $\|T_t f\|_p = \|f\|_p$ implies f is constant. Furthermore, $\|T_t f\|_{\frac{1}{\xi}} \to \mathbb{E}[f] = 1$ since $T_t f \to \mathbb{E}[f]$ as $t \to \infty$.

- Consequently, for each ξ_0 the fiber

  $$\{t : \phi(t, \xi_0) = c\} \tag{19}$$

consists of at most one point. Define t_0 as the unique solution of

$$\phi(t_0, 0) = n\rho_0 \,.$$

Solution exists from continuity of ϕ and the fact that $\phi(0,0) = \ln \|f\|_\infty > \rho_0 > \phi(+\infty, 0) = 0$.

- We have the standard identities:

$$\frac{\partial \phi}{\partial \xi} = -\frac{\mathrm{Ent}(f^{\frac{1}{\xi}})}{\mathbb{E}\,[f^{\frac{1}{\xi}}]} \tag{20}$$

$$\frac{\partial \phi}{\partial t} = \begin{cases} 0, & \xi = 1 \\ -\frac{\mathcal{E}(T_t f, (T_t f)^{\frac{1}{\xi}-1})}{\mathbb{E}\,[(T_t f)^{\frac{1}{\xi}}]}\,, & \xi \neq 1 \end{cases} \tag{21}$$

- Since f is non-constant, so is $T_t f$ for all $t \ge 0$ (for otherwise $f - \mathbb{E}\,[f]$ is in the kernel of T_t). Therefore, $\frac{\partial \phi}{\partial \xi} < 0$ for all (ξ, t). Thus, for any $t \in [0, t_0]$ there is at most one solution ξ of

$$\phi(t, \xi(t)) = \phi(0, \frac{1}{p_0}) = n\rho_0 \tag{22}$$

  It is clear that $\xi(t)$ is non-increasing. Since fibers (19) are singletons, we also conclude that $\xi(t)$ is strictly decreasing. From implicit function theorem and $\frac{\partial \phi}{\partial \xi} \neq 0$, we infer that solution $\xi(t)$ of (22) is $\mathcal{C}^\infty$-smooth.

- As we mentioned $\xi \mapsto \phi(t, \xi)$ is convex and strictly decreasing. Furthermore, it is strictly convex for $t > 0$. From this convexity and (20) we infer the following important consequences:

$$r \mapsto \frac{\mathrm{Ent}(f^r)}{\mathbb{E}\,[f^r]} \quad \text{–increasing in } r \in (0, \infty)\text{; strictly unless } f = c1_S \tag{23}$$

$$\frac{\mathrm{Ent}(f^r)}{\mathbb{E}\,[f^r]} \ge \frac{\ln \|f\|_r - \ln \|f\|_1}{1 - \frac{1}{r}} \qquad r > 1 \tag{24}$$

We now set $p(t) = \frac{1}{\xi(t)}$, where $\xi(t)$ was found from solving (22). From observation (22) we already inferred that $t \mapsto p(t)$ is well-defined, strictly increasing and $\mathcal{C}^\infty$-smooth. The fact that $p(t)$ is surjective follows from $\xi(t) \to 0$ as $t \to t_0$.

It remains to show (13). This follows from differentiating (22):

$$\xi'(t) = -\frac{\mathcal{E}(t)}{E(t)}\,,$$

where we defined

$$\mathcal{E}(t) \triangleq \frac{1}{n} \frac{\mathcal{E}(T_t f, (T_t f)^{p(t)-1})}{\mathbb{E}\,[(T_t f)^{p(t)}]} \tag{25}$$

$$E(t) \triangleq \frac{1}{n} \frac{\mathrm{Ent}((T_t f)^{p(t)})}{\mathbb{E}\,[(T_t f)^{p(t)}]}\,. \tag{26}$$

From (p, Φ_p)-LSI we get then

$$\xi'(t) \le -\frac{b_{p(t)}(E(t))}{E(t)}\,. \tag{27}$$

Finally, from (24) we get

$$E(t) \ge \frac{1}{n}\frac{\phi(t,\xi(t))}{1-\xi(t)} = \frac{\rho_0}{1-\xi(t)} \tag{28}$$

From convexity of $b_p(\cdot)$, the function $\frac{b_p(E)}{E}$ is increasing in E and so we can further upper-bound $\xi'(t)$ as

$$\xi'(t) \le -\frac{b_{p(t)}(\frac{p(t)\rho_0}{p(t)-1})}{\frac{p(t)\rho_0}{p(t)-1}}\,.$$

Noticing that $\xi'(t) = -\frac{p'(t)}{p^2(t)}$ we get (13). □

3. New LSIs and hypercontractivity for the hypercube

An elegant idea of Stroock and Varopoulos allows to compare Dirichlet forms $\mathcal{E}(f, f^{p-1})$ with $\mathcal{E}(f^{\frac{p}{2}}, f^{\frac{p}{2}})$ immediately leading to the conclusion that for any reversible semigroup (i.e. $T_t^* = T_t$ in $L_2(\pi)$) we have

$$b_p(x) \ge \frac{4|p-1|}{p^2} b_2(x) \qquad \forall x \ge 0\, \forall p \in (-\infty, \infty) \setminus \{1\} \tag{29}$$

$$b_p(x) \ge \frac{1-p}{p^2} b_1(x) \qquad \forall x \ge 0\, \forall p < 1\,. \tag{30}$$

(see, e.g., [8] for $p = 2$ and [11] for $p = 1$). Thus, we can get non-trivial non-linear p-LSIs by only establishing $p = 1, 2$ cases (of which $p = 2$ was already done in [9]). However, we can also find the sharpest non-linear LSIs for all p explicitly.

Theorem 4 (1-LSI for the hypercube)**.** *For all $f : \{0,1\}^n \to (0,\infty)$ with $\mathbb{E}[f] = 1$ we have*

$$\frac{1}{n}\mathcal{E}(f, \ln f) \ge b_1\left(\frac{1}{n}\mathrm{Ent}(f)\right)\,, \tag{31}$$

where Dirichlet form is given by (4), all expectations and Ent *are with respect to uniform probability measure on $\{0,1\}^n$ and $b_1 : [0, \ln 2) \to [0, \infty)$ is a convex increasing function given by*

$$b_1(\ln 2 - h(y)) = (\frac{1}{2} - y)\ln\frac{1-y}{y}\,, \qquad y \in (0, 1/2]\,. \tag{32}$$

Proof. By Theorem 1, we only need to work out the case $n = 1$. Then, the space of all f can be parameterized by $f(0) = 2 - f(1) = 2y$ with $y \in (0, 1/2]$. With such a choice we have $\mathrm{Ent}(f) = \ln 2 - h(y)$ and $\mathcal{E}(f, \ln f) = (\frac{1}{2} - y) \ln \frac{1-y}{y}$. Thus, we only need to verify that b_1 is increasing and convex. Monotonicity follows from the fact that both factors in the RHS of (32) are decreasing in y. For convexity, we have

$$b'(\ln 2 - h(y)) = 1 - \frac{1 - 2y}{2y(1-y) \ln \frac{y}{1-y}} \tag{33}$$

$$b''(\ln 2 - h(y)) = -\frac{1}{2 \ln \frac{y}{1-y}} \cdot \frac{d}{dy} \frac{1 - 2y}{y(1-y) \ln \frac{y}{1-y}} \tag{34}$$

Since $\ln \frac{y}{1-y} < 0$, it suffices to prove $\frac{d}{dy} \frac{1-2y}{y(1-y) \ln \frac{y}{1-y}} \geq 0$. Computing the derivative and rearranging, we need to show $\ln \frac{1-y}{y} \geq \frac{1-2y}{1-2y+2y^2}$. Substituting $z = \ln \frac{1-y}{y}$ this is the same as

$$\ln z \geq \frac{z^2 - 1}{z^2 + 1} \tag{35}$$

for $z \geq 1$. Both sides of (35) vanish at $z = 1$. Comparing the derivatives, we need to verify $1/z \geq \frac{4z}{(z^2+1)^2}$, which is the same as $\left(z^2 + 1\right)^2 \geq 4z^2$ and hence is true. □

Corollary 5. *Classical MGL (6) holds.*

Proof. We only need to invoke Theorem 2 with b_1 taken from (31). Indeed, notice that setting $\tilde{\rho}(t) = \ln 2 - m\left(t, \frac{\mathrm{Ent}(f)}{n\mathbb{E}[f]}\right)$ to equal the RHS of (6) solves $\tilde{\rho}(t)' = -b_1(\tilde{\rho}(t))$. □

Next, we proceed to LSI's with $p \neq 1$.

Theorem 6 (p-LSI for the hypercube)**.** *Fix $p \in (-\infty, \infty) \setminus \{1, 0\}$. For all $f : \{0,1\}^n \to \mathbb{R}_+$ with $\mathbb{E}[f^p] = 1$ (and $f > 0$ if $p < 1$) we have*

$$\frac{1}{n} \mathrm{sgn}(p-1) \mathcal{E}(f, f^{p-1}) \geq b_p\left(\frac{1}{n} \mathrm{Ent}(f^p)\right),$$

where Dirichlet form is given by (4), all expectations and Ent *are with respect to uniform probability measure on $\{0,1\}^n$ and $b_p : [0, \ln 2] \to [0, \infty)$ is a convex increasing function given by*

$$b_p(\ln 2 - h(y)) = \frac{\mathrm{sgn}(p-1)}{2} \left(1 - y^{\frac{1}{p}}(1-y)^{1-\frac{1}{p}} - y^{1-\frac{1}{p}}(1-y)^{\frac{1}{p}}\right), \tag{36}$$

with $0 < y \leq \frac{1}{2}$.

Proof. By Theorem 1, we only need to work out the case $n = 1$. Then, the space of all f can be parameterized by $f(0) = (2y)^{\frac{1}{p}}, f(1) = (2-2y)^{\frac{1}{p}}$ with $y \in [0, 1/2]$. Thus we only need verify monotonicity and convexity.

First, consider the case $p > 1$. Let $q = \frac{p}{p-1}$. Taking the first derivative, we get

$$b_p'(\ln 2 - h(y)) = -\frac{1}{2}\frac{1}{\ln \frac{y}{1-y}} \cdot \left(\frac{1}{p}\left(\left(\frac{1-y}{y}\right)^{1/q} - \left(\frac{y}{1-y}\right)^{1/q}\right) + \frac{1}{q}\left(\left(\frac{1-y}{y}\right)^{1/p} - \left(\frac{y}{1-y}\right)^{1/p}\right)\right) \tag{37}$$

From here, monotonicity of b_p follows from the fact that the RHS is negative $(\frac{1-y}{y} > \frac{y}{1-y})$. We proceed to showing convexity. Let $z = \frac{y}{1-y}$. Then $0 < z \leq 1$ and, taking another derivative, we have

$$b_p''(\ln 2 - h(y)) = -\frac{1}{2}\frac{1}{(1-y)^2} \cdot \frac{1}{\ln z} \cdot \frac{d}{dz}\left[\frac{1}{\ln z} \cdot \left(\frac{1}{p}\left(\left(\frac{1}{z}\right)^{1/q} - z^{1/q}\right) + \frac{1}{q}\left(\left(\frac{1}{z}\right)^{1/p} - z^{1/p}\right)\right)\right] \tag{38}$$

Since $\ln z < 0$ for $z < 1$, it would suffice to argue that the derivative w.r.t. z on RHS is nonnegative. Let $r(z) = \frac{1}{p}\left(\left(\frac{1}{z}\right)^{1/q} - z^{1/q}\right) + \frac{1}{q}\left(\left(\frac{1}{z}\right)^{1/p} - z^{1/p}\right)$. We need to show $z \ln \frac{1}{z} \cdot (-r') \geq r$.

Making another substitution of variables, let $w = \ln z$, that is $-\infty < w \leq 0$. Let $t(w) = r(z) = r(e^w)$. Substituting and simplifying, we need to show $wt'(w) \geq t(w)$.

We have $t(w) = \frac{1}{p}\left(e^{-w/q} - e^{w/q}\right) + \frac{1}{q}\left(e^{-w/p} - e^{w/p}\right)$. Hence

$$t'(w) = -\frac{1}{pq}\left(e^{-w/q} + e^{w/q} + e^{-w/p} + e^{w/p}\right) \quad \text{and}$$

$$t''(w) = -\frac{1}{pq}\left(-\frac{1}{q}e^{-w/q} + \frac{1}{q}e^{w/q} - \frac{1}{p}e^{-w/p} + \frac{1}{p}e^{w/p}\right)$$

In particular, t is a decreasing convex function on $(-\infty, 0]$ which vanishes at 0, and $wt'(w) \geq t(w)$ is satisfied.

Next, consider the case $0 < p < 1$. We repeat the computation above, multiplying throughout by $-1 = \operatorname{sgn}(p-1)$. Since in this case $q < 0$, the sign change cancels out, and the convexity argument, with minor changes as needed, goes through. For monotonicity, observe that again the signs of both terms in the RHS of (37) are negative (the front $-$ sign is canceled by $\operatorname{sgn}(p-1)$).

Finally, for the case $p < 0$, observe that we can set $g = f^{p-1}$ and apply the already proven inequality to the pair $(g, \frac{p}{p-1})$ since $\frac{p}{p-1} \in (0, 1)$.

□

Our chief goal is to derive hypercontractivity inequality tighter than Bonami's (8) for functions with small support. We will replace constraint on the support $|\operatorname{supp} f| \le 2^{nR}$ with an analytical proxy:

$$\|f\|_{p_0} \ge e^{n\rho_0}\|f\|_1 \,, \quad \rho_0 = (1-p_0^{-1})(1-R)\ln 2 \,,$$

as discussed in (9). We get the following result:

Theorem 7. *Fix $1 < p_0 < \infty$ and $0 \le \rho_0 \le (1-p_0^{-1})\ln 2$. Then differential equation*

$$u'(t) = C\left(\rho_0(1+e^{-u(t)})\right) \,, \quad C(\ln 2 - h(y)) = \frac{2-4\sqrt{y(1-y)}}{\ln 2 - h(y)} \tag{39}$$

has unique solution on $[0,\infty)$ with $u(0) = \ln(p_0-1)$. Furthermore, for any $f : \{0,1\}^n \to \mathbb{R}_+$ with $\|f\|_{p_0} \ge e^{n\rho_0}\|f\|_1$ we have

$$\|T_t^{\otimes n} f\|_{p(t)} \le \|f\|_{p_0} \,, \qquad p(t) = 1 + e^{u(t)} \,. \tag{40}$$

Remark 2. *Ref. [9] showed that $C(x) : [0,\ln 2] \to [2, 2/\ln 2]$ is a smooth, convex and strictly increasing bijection. Consequently, the $p(t)$ in (40) is smooth and satisfies*

$$p(t) > 1 + (p_0-1)e^{2t} \qquad \forall t > 0$$

thereby strictly improving Bonami's hypercontractivity inequality (8). Furthermore, it satisfies

$$p(t) = p_0 + p'(0)t + \frac{1}{2}p''(0)t^2 + o(t^2), \qquad t \to 0 \,, \tag{41}$$

$$p'(0) = (p_0-1)C(x_0) \tag{42}$$

$$p''(0) = (p_0-1)\left(C(x_0)^2 - C'(x_0)C(x_0)\frac{x_0}{p_0}\right) \tag{43}$$

$$x_0 = \frac{\rho_0 p_0}{p_0 - 1} \,. \tag{44}$$

Remark 3. *Our estimate is locally optimal at $t = 0$ in the following sense: for every $q(t)$ such that $q(0) = p_0$ and $q'(0) < p'(0)$ there exists a function f with $\|f\|_{p_0} \ge e^{n\rho_0}\|f\|_1$ and $\|T_t f\|_{q(t)} > \|f\|_{p_0}$ for a sequence of $t \to 0$. This follows from the fact that had a counter-example $q(t)$ existed, it would imply that the second half of the proof of Theorem 9 could be improved to contradict the first half.*

Proof. First, notice that $C(x) = \frac{4b_2(x)}{x}$, where b_2 was defined in Theorem 6. Let $p_1(t)$ be the function defined by

$$\|T_t^{\otimes n} f\|_{p_1(t)} = \|f\|_{p_0} \,.$$

Theorem 3 showed this function to be smooth and growing faster than (13). From (13) and using (29) to lower-bound $b_p(\cdot)$ via $b_2(\cdot)$ we get that

$$p_1'(t) \ge (p_1(t)-1)C\left(\frac{p_1(t)\rho_0}{p_1(t)-1}\right) \,,$$

or introducing $u_1(t) = \ln(p_1(t) - 1)$ that

$$u_1'(t) \geq C\left(\rho_0(1 + e^{-u_1(t)})\right) .$$

The case of $\rho_0 = 1 - p_0^{-1}$ corresponds to f supported on a single point and can be dealt with separately. So we assume $\rho_0 < 1 - p_0^{-1}$, in which case the map

$$u \mapsto C\left(\rho_0(1 + e^{-u})\right)$$

is smooth on some interval $(\ln(p_0 - 1) - \epsilon, \infty)$. Consequently, (39) possesses unique solution with $u(0) = \ln(p_0 - 1)$ and Chaplygin-type theorem, e.g. [12, Theorem 4.1], implies

$$u_1(t) \geq u(t) .$$

□

For $p_0 = 2$, we also prove an alternative estimate $p(t)$ via a method tailored to the hypercube.

Theorem 8. *In the setting of Theorem 7 assume $p_0 = 2$. Then (40) holds with $p(t)$ given as*

$$p(t) = 1 + e^{\int_0^t C(\tilde{\rho}(s)\vee 0)ds} \tag{45}$$

$$\tilde{\rho}(s) = \frac{\rho_0 p_0}{p_0 - 1} - \ln\left(\frac{2}{1 + e^{-2s}}\right) . \tag{46}$$

Remark 4. *Using convexity of C we get $C(x \vee 0) \geq C(x_0) + (x - x_0)C'(x_0)$. Similarly, $\ln \frac{1+e^{-2t}}{2} \geq -t$. Therefore, altogether we get an explicit estimate:*

$$p(t) \geq 1 + e^{C(x_0)t - \frac{C'(x_0)}{2}t^2} , \tag{47}$$

where $x_0 = \tilde{\rho}(0)$ is from (44). The t^2 term here is, however, worse than that of (41).

Proof. We return to (27). Recalling that $\xi(t) = \frac{1}{p(t)}$ and lower-bounding b_p by b_2 via (29) we get

$$\frac{d}{dt}\ln(p(t) - 1) \geq C(E(t)) ,$$

(with $E(t)$ from (26)), which implies via Gronwall inequality

$$p(t) \geq 1 + e^{\int_0^s C(E(s))ds} .$$

Thus, it only suffices to prove that $E(s) \geq \tilde{\rho}(s) \vee 0$, or equivalently (since $E \geq 0$ by definition) the estimate

$$E(t) \geq (1 - R)\ln 2 - \ln\left(\frac{2}{1 + e^{-2t}}\right) \tag{48}$$

Next, we obtain a lower bound on $\|T_t^{\otimes n} f\|_2$. To that end introduce a function $\Lambda_t^{\otimes n}$ on $\{0,1\}^n$ with the property $T_t^{\otimes n} f = \Lambda_t^{\otimes n} * f$. Note that $\Lambda_t^{\otimes n}(x) = \left(1-e^{-t}\right)^{|x|}\left(1+e^{-t}\right)^{n-|x|}$, where $|x|$ denotes the Hamming weight of x. Clearly $\Lambda_t^{\otimes n} \geq 0$. Furthermore, without loss of generality we may assume $f \geq 0$. Then, we have

$$\|T_t^{\otimes n} f\|_2^2 = \left\langle T_t^{\otimes n} f, T_t^{\otimes n} f\right\rangle = \left\langle \Lambda_t^{\otimes n} * f, \Lambda_t^{\otimes n} * f\right\rangle =$$

$$\left\langle \Lambda_t^{\otimes n} * \Lambda_t^{\otimes n}, f * f\right\rangle = \left\langle \Lambda_{2t}^{\otimes n}, f * f\right\rangle \geq \frac{1}{2^n}\Lambda_{2t}^{\otimes n}(0)\cdot(f*f)(0) = \left(\frac{1+e^{-2t}}{2}\right)^n \cdot \|f\|_2^2$$

To prove (48), observe that by (23) and (24) and by the preceding calculation,

$$E(t) \geq \frac{1}{n}\ln\frac{\|T_t^{\otimes n} f\|_2^2}{\|T_t^{\otimes n} f\|_1^2} = \frac{1}{n}\ln\frac{\|T_t^{\otimes n} f\|_2^2}{\|f\|_1^2} \geq \frac{1}{n}\ln\frac{\|f\|_2^2}{\|f\|_1^2} - \ln\left(\frac{2}{1+e^{-2t}}\right) =$$

$$(1-R)\ln 2 - \ln\left(\frac{2}{1+e^{-2t}}\right)$$

□

4. Uncertainty principle on the hypercube

4.1. Background

Uncertainty principle asserts that a function and its Fourier transform cannot be simultaneously narrowly concentrated. There are several approaches to quantifying this statement, and here we adopt the Hilbert space point of view, cf. [13, Chapter 3]. Namely, for a pair of subspaces V_1, V_2 of a Hilbert space with inner product $(\cdot,\cdot)$ and $\|f\|_2^2 \triangleq (f,f)$ we define

$$\cos\angle(V_1, V_2) \triangleq \sup_{f_1\in V_1, f_2\in V_2} \frac{|(f_1,f_2)|}{\|f_1\|_2\|f_2\|_2}\,.$$

For the uncertainty principle, we will select sets S and Σ and define subspaces

$$V_S \triangleq \{f : \operatorname{supp} f \subset S\} \tag{49}$$

$$\hat{V}_\Sigma \triangleq \{f : \operatorname{supp} \hat{f} \subset \Sigma\}\,, \tag{50}$$

where $\hat{f}$ denotes the corresponding Fourier transform (we will define it precisely). Uncertainty principle corresponds to bounding $\cos\angle(V_S, \hat{V}_\Sigma)$ away from 1, thus establishing to what extent functions can simultaneously concentrate on (S,Σ).

There is a number of equivalent ways to think of $\cos\angle(V_1, V_2)$. Letting P_i be an orthogonal projection on V_i and $P_i^\perp$ projection on $V_i^\perp$, it can be shown [13, Chapter 3]:

$$\begin{aligned}
\cos\angle(V_1,V_2)\le\theta &\iff \forall f\in V_1:\quad \|P_2 f\|_2\le\theta\|f\|_2 && (51)\\
&\iff \forall f\in V_1:\quad \|P_2^\perp f\|_2\ge\sqrt{1-\theta^2}\|f\|_2 && (52)\\
&\iff \lambda_{\max}(P_1P_2P_1)\le\theta && (53)\\
&\iff \|P_1P_2\|_{2\to 2}\le\sqrt{\theta} && (54)\\
&\iff \forall f:\|f\|_2^2\le\frac{1}{1-\theta}\left(\|P_1^\perp f\|_2^2+\|P_2^\perp f\|_2^2\right) && (55)
\end{aligned}$$

Furthermore, there is also a simple criterion:

$$\cos\angle(V_1,V_2)<1\iff V_1\cap V_2=\{0\}\text{ and }(V_1+V_2)\text{ — closed}\,,$$

where for finite-dimensional V_i's the closedness condition is vacuous (but not in general).

Finally, as shown in Fuchs [14] and Landau-Pollak [15], knowledge of $\cos\angle(V_1,V_2)$ is sufficient for completely characterizing the two-dimensional region

$$\{(\|P_1 f\|_2^2,\|P_2 f\|_2^2)\}$$

Before proceeding to our own results, we briefly review the history of results for $\mathbb{R}^n$. First, Slepian and Landau [16] computed $\cos\angle(V_S,\hat V_\Sigma)$ for S,Σ being two balls (in fact they computed $\lambda_{max}(P_1P_2P_1)$ and named eigenfunctions of the latter prolate spheroidal functions). Next, Benedicks in 1974 (published as [17]) showed that

$$\mathrm{vol}(S),\mathrm{vol}(\Sigma)<\infty\qquad\Longrightarrow\qquad V_S\cap\hat V_\Sigma=\{0\}\,.$$

Later, Amrein and Berthier [18] strengthened this to

$$\mathrm{vol}(S),\mathrm{vol}(\Sigma)<\infty\qquad\Longrightarrow\qquad \cos\angle(V_S,V_\Sigma)<1\,.$$

Finally, for $n=1$ Nazarov [19] showed

$$\mathrm{vol}(S),\mathrm{vol}(\Sigma)<\infty\qquad\Longrightarrow\qquad \cos\angle(V_S,V_\Sigma)<1-ce^{-c\,\mathrm{vol}(S)\,\mathrm{vol}(\Sigma)}\,.$$

Lately, there were a number of extensions and improvements of Nazarov's theorem for $n>1$, e.g. [20].

4.2. Sharp uncertainty principle on $\mathbb{F}_2^n$

Define the characters, indexed by $v\in\mathbb{F}_2^n$,

$$\chi_v(x)\triangleq\prod_{j:v_j=1}\chi_j(x)=(-1)^{<v,x>}\,,$$

where $<v,x>=\sum_{j=1}^n v_jx_j$ is a non-degenerate bilinear form on $\mathbb{F}_2^n$. The Fourier transform of $f:\mathbb{F}_2^n\to\mathbb{C}$ is

$$\hat f(\omega)\triangleq\sum_{x\in\mathbb{F}_2^n}\chi_\omega(x)f(x)=2^n(f,\chi_\omega)\,,\qquad \omega\in\mathbb{F}_2^n\,.$$

We denote by $|x|$ the Hamming weight of $x \in \mathbb{F}_2^n$ and by $B_r = \{x : |x| \le r\}$ – Hamming ball.

Theorem 9. *For any $\rho_1, \rho_2 \in [0, 1/2]$ satisfying*

$$(1 - 2\rho_1)^2 + (1 - 2\rho_2)^2 > 1 \,. \tag{56}$$

there exist an $\epsilon > 0$ and n_0 such that for any $n \ge n_0$, any $S \subset \mathbb{F}_2^n$ with $|S| \le e^{nh(\rho_1)}$ and $\Sigma = B_{\rho_2 n}$ we have

$$\cos\angle(V_S, \hat{V}_\Sigma) \le e^{-n\epsilon} \,, \tag{57}$$

where $V_S, \hat{V}_\Sigma$ are defined in (49)-(50).

Conversely, for $\rho_1, \rho_2 \in [0, 1/2]$ satisfying[1]

$$(1 - 2\rho_1)^2 + (1 - 2\rho_2)^2 < 1 \,, \tag{58}$$

there exist $\epsilon > 0$ and n_0 such that for all $n \ge n_0$ we have

$$\cos\angle(V_S, \hat{V}_\Sigma) \ge 1 - e^{-n\epsilon}, \qquad S = B_{\rho_1 n}, \Sigma = B_{\rho_2 n} \,. \tag{59}$$

Proof. For the case (58), fix $\alpha \in \mathbb{R}$, $\alpha \ne \pm 1$ and consider the following Fourier pair:

$$f(x) = \alpha^{|x|} \tag{60}$$

$$\hat{f}(\omega) = c \left(\frac{1-\alpha}{1+\alpha}\right)^{|\omega|}, \qquad c = (1 + \alpha)^n \,. \tag{61}$$

Then, it is easy to see that the L_2-energy of f is concentrated around $|x| \approx \frac{1}{1+\alpha^{-2}} n$. Thus, whenever radius $\rho_1 > \frac{1}{1+\alpha^{-2}}$, we have for some $\epsilon > 0$

$$\sum_{x:|x|>\rho_1 n} f(x)^2 \le e^{-\epsilon n} \sum_{x \in \mathbb{F}_2^n} f(x)^2 \,.$$

Similarly, whenever $\rho_2 > \frac{1}{1+\beta^{-2}}$, where $\beta = \frac{1-\alpha}{1+\alpha}$, we have

$$\sum_{\omega:|\omega|>\rho_2 n} \hat{f}(\omega)^2 \le e^{-\epsilon n} \sum_{\omega \in \mathbb{F}_2^n} \hat{f}(\omega)^2 \,.$$

Whenever, (58) holds, it is not hard to see that there exists a choice of $\alpha \in (0, 1)$ satisfying both $\rho_1 > \frac{1}{1+\alpha^{-2}}$ and $\rho_2 > \frac{1}{1+\beta^{-2}}$. Thus, taking corresponding f and using (55) we get (59).

Next, we assume (56). We define the Fourier projection operators Π_a as

$$\widehat{\Pi_a f}(\omega) \triangleq \hat{f}(\omega) 1\{|\omega| = a\} \,, \qquad a = 0, 1, \ldots, n \,. \tag{62}$$

[1] When $\rho_1 > \frac{1}{2}$ (or $\rho_2 > \frac{1}{2}$), the result (59) also holds by reducing to $\rho_1 = \frac{1}{2}$. This is possible since $\cos\angle(V_S, \hat{V}_\Sigma)$ is monotone in S, Σ.

and set $\Pi_{\le r} = \sum_{a=0}^{r} \Pi_a$. By (51) we need to show that for any function f with support $|\operatorname{supp} f| \le e^{nh(\rho_1)}$ we have

$$\|\Pi_{\le \rho_2 n} f\|_2 \le e^{-n\epsilon} \|f\|_2 \,,$$

for some $\epsilon > 0$ independent of n and f.

Note that $\widehat{T_t f}(\omega) = e^{-t|\omega|} \hat{f}(\omega)$. Thus, comparing eigenvalues we have $e^{ta} T_t \succeq \Pi_a$ (in the sense of positive-semidefiniteness). Consequently,

$$\|\Pi_a f\|_2^2 = (\Pi_a f, f) \le e^{at} (T_t f, f) \le e^{at} \|f\|_q \|T_t f\|_p \,, \tag{63}$$

where p and q are Hölder conjugates. Since $|\operatorname{supp} f| \le e^{nh(\rho_1)}$ we have from Theorem 7 with $p_0 = 2$ and $\rho_0 = \frac{\ln 2 - h(\rho_1)}{2}$:

$$\|T_t f\|_{p(t)} \le \|f\|_2 \,, \qquad p(t) = 2 + p'(0) t + o(t) \,,$$

where the value of $p'(0)$ is given in (42)[2] Taking $p = p(t) > 2$ in (63) we need upper-bound $\|f\|_q$, which we again do by invoking the bound on support

$$\|f\|_q \le \|f\|_2 2^{-n(1-R)(\frac{1}{q} - \frac{1}{2})} \,, \qquad \forall 1 \le q \le 2 \,. \tag{64}$$

Overall, we have shown for all a and t that

$$\|\Pi_a f\|_2^2 \le e^{at} e^{-n(\ln 2 - h(\rho_1))(\frac{1}{2} - \frac{1}{p(t)})} \|f\|_2^2 \,.$$

Analyzing this inequality for t close to 0 we conclude that whenever

$$\rho_2 < \frac{p'(0)}{4} (\ln 2 - h(\rho_1)) \tag{65}$$

we necessarily have for some $\epsilon > 0$ (depending on the gap in the inequality above and on the local bound for $p''(t)$ at 0) that for all $a \le \rho_2 n$

$$\|\Pi_a f\|_2 \le e^{-n\epsilon} \|f\|_2 \,.$$

Using expression for $p'(0)$ in (42), we see that (65) is equivalent to

$$2\rho_2 < 1 - 2\sqrt{\rho_1(1-\rho_1)} \,, \tag{66}$$

which is in turn equivalent to (56).

□

For completeness, we also provide a criterion for when two subspaces have a common element (for the special case of two balls as S, Σ. It demonstrates that there is a "discontinuity" between the regime $\cos \ge 1 - e^{O(n)}$ and $\cos = 1$.

[2]For extracting explicit constants, one may invoke (47) instead.

Proposition 10. *Let $S = B_{r_1}$ and $\Sigma = B_{r_2}$ in $\mathbb{F}_2^n$. Then*

$$V_S \cap \hat{V}_\Sigma \neq \{0\} \iff \cos\angle(V_S, \hat{V}_\Sigma) = 1 \iff r_1 + r_2 \geq n\,.$$

Proof. If $r_1 + r_2 \geq n$, then take $f(x) = 1\{x_{r_1+1} = \cdots = x_n = 0\}$. Its Fourier transform is supported on $\{\omega : \omega_1 = \cdots = \omega_{r_1} = 0\}$. Thus $f \in V_S \cap \hat{V}_\Sigma$. On the other hand, suppose there is $f \in V_S \cap \hat{V}_\Sigma$. By averaging over permutations of coordinates (both subspaces are invariant to such), we conclude that $f(x) = f_1(|x|)$. As such, it can be expanded in terms of Krawtchouk polynomials:

$$f_1(|x|) = \sum_{k=0}^{n} a_k K_k(|x|)\,,$$

where each $K_k(\cdot)$ is a degree k univariate polynomial. Constraint $\operatorname{supp} \hat{f} \subset B_{r_2}$ is equivalent to requiring $a_k = 0$ for $k > r_2$. Thus, we conclude that f_1 on integers inside $[0, n]$ coincides with a degree r_2 polynomial, and hence has $\leq r_2$ zeros. Thus, $r_1 \geq n - r_2$ as claimed. □

4.3. Discussion

In this section we collect some remarks regarding Theorem 9. First, it is instructive to provide an equivalent statement as follows:

Theorem 11 (Restatement of the uncertainty principle)**.** *For any $\delta < 1/2$ and $0 < E < R_{LP1}(\delta) \triangleq h\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right)$ there is $\epsilon > 0$ with the following property. Let $f(x_1, \ldots, x_n)$ be polynomial of total degree at most δn. Then, for any $S \subset \{\pm 1\}^n$ of size $|S| \leq e^{nE}$ we have*

$$\sum_{x \in S} f(x)^2 \leq e^{-n\epsilon} \sum_{x \in \{\pm 1\}^n} f(x)^2\,.$$

This shows that any low-degree polynomial (restricted to the hypercube) smears its energy so evenly that one needs to sum $e^{nR_{LP1}(\delta)}$ top values in order to obtain a sizable fraction of its overall energy. It is interesting to compare this with [21] showing that any f that is a) a degree $\leq \delta n$ polynomial and b) $f \geq 0$ satisfies

$$\max_{x \in \{\pm 1\}^n} |f(x)| \leq e^{-n(\ln 2 - h(\frac{\delta}{2})) + o(n)} \sum_{x \in \{\pm 1\}^n} |f(x)|\,. \tag{67}$$

We *conjecture* that (67) holds for all $f : \{\pm 1\}^n \to \mathbb{R}$ of degree $\leq \delta n$. This could be called an L_1-version of the uncertainty principle. If true, it would imply that the sum of any $e^{nR_{Ham}(\delta)}$, $R_{Ham}(\delta) = \ln 2 - h(\frac{\delta}{2})$ values of $|f(x)|$ is negligible compared to the sum over all of $\{\pm 1\}^n$.

Finally, we discuss to what extent one can relax condition that $\Sigma = B_{\rho_2 n}$ in Theorem 9. First, notice that clearly the same conclusion holds for Σ which is an image of a ball $B_{\rho_2 n}$ under a linear isomorphism $\mathbb{F}_2^n \to \mathbb{F}_2^n$. This provides a wealth of examples of Σ that are less "contiguous" than $B_{\rho_2 n}$.

At the same time, we cannot extend Theorem 9 to Σ being an arbitrary subset of the same cardinality as $B_{\rho_2 n}$. Indeed, a simple computation shows that when S and Σ are linear subspaces of $\mathbb{F}_2^n$ we have

$$\cos\angle(V_S, \hat{V}_\Sigma) = \sqrt{\frac{|\Sigma \cap S^\perp|}{|S^\perp|}}\,, \tag{68}$$

where $S^\perp \triangleq \{x :< x, v >= 0\,, \forall v \in S\}$ is the dual of S. Thus, if we take S to be a linear subspace of dimension αn, $0 < \alpha < 1$, and $\Sigma = S^\perp$ (of dimension $(1-\alpha)n$) and solve for ρ_1 and ρ_2 in

$$h(\rho_1) = \alpha \ln 2, h(\rho_2) = (1-\alpha)\ln 2$$

we conclude that these ρ_1 and ρ_2 always satisfy

$$(1-2\rho_1)^2 + (1-2\rho_2)^2 > 1\,,$$

while from (68) we have $\cos\angle(V_S, \hat{V}_\Sigma) = 1$.

Consequently, we leave open the question of determining the more general uncertainty principle, i.e. characterizing the best pairs (E_1, E_2) for which one can prove implication

$$|S| \le e^{nE_1}, |\Sigma| \le e^{nE_2} \qquad \Longrightarrow \qquad \cos\angle(V_S, \hat{V}_\Sigma) \le \epsilon\,.$$

A partial result easily follows from the Hausdorff-Young inequality:

Proposition 12. *For any $E_1, E_2 \in (0, \ln 2)$ satisfying $E_1 + E_2 < \ln 2$ there exist $\epsilon > 0$ and n_0 such that for all $n \ge n_0$, all $S, \Sigma \subset \mathbb{F}_2^n$ with $|S| = e^{nE_1}, |\Sigma| = e^{nE_2}$ we have*

$$\cos\angle(V_S, \hat{V}_\Sigma) \le 1 - \epsilon\,.$$

Conversely, for any positive integers $k_1, k_2 \le n$ such that $k_1 + k_2 \ge n$ there exist $|S| = 2^{k_1}$ and $|\Sigma| = 2^{k_2}$ such that

$$\cos\angle(V_S, \hat{V}_\Sigma) = 1\,.$$

Proof. Second part follows from (68). For the first part, let $\theta = \cos\angle(V_S, \hat{V}_\Sigma)$ and $E_1 + E_2 = \ln 2 - \delta$ for $\delta > 0$. We will show that

$$\theta^2 \le 1 - \frac{\delta - \frac{1}{n}\ln 2}{\ln 2 - \max(E_1, E_2)}\,. \tag{69}$$

Without loss of generality, suppose $E_2 \ge E_1$. Recall that a simple consequence of the Hausdorff-Young inequality is the Hirschmann (or entropic) uncertainty principle [22, Exercise 4.2.10]: For any $f : \mathbb{F}_2^n \to \mathbb{R}$ we have

$$\frac{\mathrm{Ent}(f^2)}{\mathbb{E}\,[f^2]} + \frac{\mathrm{Ent}(\hat{f}^2)}{\mathbb{E}\,[\hat{f}^2]} \le \ln|\mathbb{F}_2^n| = n\ln 2\,. \tag{70}$$

Thus, taking f supported on S we estimate (from Jensen's inequality)

$$\frac{\mathrm{Ent}(f^2)}{\mathbb{E}\,[f^2]} \geq n\ln 2 - \ln|S| = n(\ln 2 - E_1)\,. \tag{71}$$

Suppose that $\frac{\mathbb{E}\,[\hat{f}^2 1_\Sigma]}{\mathbb{E}\,[\hat{f}^2]} = \theta^2$, and introduce random variable U taking values in $\mathbb{F}_2^n$ with

$$\mathbb{P}[U = u] \triangleq \frac{\hat{f}^2(u)}{\sum_\omega \hat{f}^2(\omega)}\,.$$

Then, we have $\frac{\mathrm{Ent}(\hat{f}^2)}{\mathbb{E}\,[\hat{f}^2]} = n\ln 2 - H(U)$, with $H(\cdot)$ denoting the Shannon entropy. Introducing also $T = 1\{U \in \Sigma\}$ we get by the chain rule

$$n\ln 2 - \frac{\mathrm{Ent}(\hat{f}^2)}{\mathbb{E}\,[\hat{f}^2]} = H(U) = H(U,T) = H(T) + H(U|T) \tag{72}$$

$$\leq \ln 2 + \theta^2\ln|\Sigma| + (1-\theta^2)\ln|\Sigma^c| \tag{73}$$

$$\leq \ln 2 + n(\theta^2 E_2 + (1-\theta^2)\ln 2)\,. \tag{74}$$

Altogether, from (70), (71) and (74) we get (69). □

4.4. A similar result for Euclidean space

It is interesting to observe that a result in $(\mathbb{R}^n, \mathrm{Leb})$ analogous to Theorem 9 follows from the sharp form of Young's inequality due to Beckner [23]. This provokes us to hypothesize that the refined hypercontractivity result on the hypercube (Theorem 7) could play the role of the sharp Young inequality (and/or Babenko-Beckner inequality) in $\mathbb{R}^n$.

Notation: we define $B_r = \{x \in \mathbb{R}^n : \|x\| \leq r\}$, $\|x\|^2 = (x,x)$, $(x,y) = \sum_{k=1}^n x_k y_k$, $|S|$ – the Lebesgue measure of S, and for $f \in L_1 \cap L_2$

$$\hat{f}(\omega) = \int_{\mathbb{R}^n} e^{-2\pi i(\omega,x)} f(x)dx\,, \qquad \omega \in \mathbb{R}^n\,,$$

with the standard extension by continuity to all of $f \in L_2$.

Theorem 13. *For any $\rho_1, \rho_2 > 0$ satisfying*

$$\rho_1\rho_2 < \frac{1}{4\pi} \tag{75}$$

there exist an $\epsilon > 0$ and n_0 such that for any $n \geq n_0$, any $S \subset \mathbb{R}^n$ with $|S| = |B_{\rho_1\sqrt{n}}|$ and $\Sigma = B_{\rho_2\sqrt{n}}$ we have

$$\cos\angle(V_S, \hat{V}_\Sigma) \leq e^{-n\epsilon}\,, \tag{76}$$

where $V_S, \hat{V}_\Sigma$ are defined in (49)-(50).

Conversely, for $\rho_1, \rho_2 \ge 0$ *satisfying*

$$\rho_1 \rho_2 > \frac{1}{4\pi} \tag{77}$$

there exist $\epsilon > 0$ *and* n_0 *such that for all* $n \ge n_0$ *we have*

$$\cos\angle(V_S, \hat{V}_\Sigma) \ge 1 - e^{-n\epsilon}, \qquad S = B_{\rho_1\sqrt{n}}, \Sigma = B_{\rho_2\sqrt{n}}\,. \tag{78}$$

Proof. Since the statement is asymptotic, we will use the standard fact

$$\ln|B_1| = \frac{n}{2}\ln\frac{2\pi e}{n} - \frac{1}{2}\ln(\pi n) + O(\frac{1}{n})$$

and thus

$$\ln|B_{\rho\sqrt{n}}| = \frac{n}{2}\ln(2\pi e\rho^2) + O(\ln n)\,. \tag{79}$$

To prove the second part, consider the Fourier pair (for any $\sigma > 0$):

$$f(x) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{\|x\|^2}{2\sigma^2}} \tag{80}$$

$$\hat{f}(\omega) = e^{-2\pi^2\sigma^2\|\omega\|^2}\,. \tag{81}$$

Choose $\sigma > 0$ so that $\rho_1 > \frac{\sigma}{\sqrt{2}}$ and $\rho_2 > \frac{1}{\sqrt{8}\sigma\pi}$ (which is possible due to (77)). From concentration of Gaussian measure, it is easy to check that for some $\epsilon > 0$ we have

$$\|f 1_{B_{\rho_1\sqrt{n}}}\|_2 \ge (1 - e^{-n\epsilon})\|f\|_2 \tag{82}$$

$$\|\hat{f} 1_{B_{\rho_2\sqrt{n}}}\|_2 \ge (1 - e^{-n\epsilon})\|\hat{f}\|_2 \tag{83}$$

and therefore (78) follows from (55).

For the first part, recall a sharp form of the Young inequality on $\mathbb{R}^n$ (from [23])

$$\|f * g\|_r \le \left(\frac{C_p C_q}{C_r}\right)^n \|f\|_p \|g\|_q\,, \qquad C_s = e^{\frac{1}{2}(\frac{\ln s}{s} + \frac{s-1}{s}\ln(1-s^{-1}))} \tag{84}$$

valid for $1 \le p, q, r \le \infty$ and $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$. Consider the heat semigroup

$$e^{t\Delta} f \triangleq f * \phi_t, \qquad \phi_t(x) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{\|x\|^2}{4t}}\,.$$

Let $\gamma > 0$ be a constant to be specified later, and for a real $t < \frac{1}{2\gamma}$ we set $\frac{1}{p(t)} = \frac{1}{2} - \gamma t$. Then apply (84) with $r = p(t), p = 2$ and $q = q(t)$ given by $\frac{1}{q(t)} = 1 - \gamma t$ to get, after some calculations, a hypercontractive inequality

$$\|e^{t\Delta} f\|_{p(t)} \le e^{nE(t)}\|f\|_2\,, \tag{85}$$

where

$$E(t) = \frac{\gamma t}{2}\ln\frac{\gamma}{\pi e^2} + o(t)\,, \qquad t \to 0\,.$$

Now, we proceed as in the proof of Theorem 9 with (85) replacing the use of the more precise hypercontractivity for the cube.

Namely, we define the ball-multiplier operator

$$\widehat{\Pi_r f}(\omega) = \hat{f}(\omega) 1_{B_r}(\omega) \,.$$

Now consider a function f supported on S and note the chain

$$\|\Pi_r f\|_2^2 = (\Pi_r f, f) \le e^{4\pi^2 r^2 t}(e^{t\Delta} f, f) \tag{86}$$
$$\le e^{4\pi^2 r^2 t}\|f\|_{q(t)}\|e^{t\Delta} f\|_{p(t)} \tag{87}$$
$$\le e^{4\pi^2 r^2 t + nE(t)}\|f\|_2\|f\|_{q(t)} \tag{88}$$
$$\le e^{4\pi^2 r^2 t + nE(t) + \gamma t \ln|S|}\|f\|_2^2 \,, \tag{89}$$

where in (86) we used the fact that $\widehat{e^{t\Delta} f}(\omega) = e^{-4\pi^2\|\omega\|^2}\hat{f}(\omega)$, in (87) we used Hölder's inequality with $q(t)$ denoting the conjugate of $p(t)$, (88) is by (85), and (89) is by invoking the bound on the support of S via Hölder's inequality

$$\|f\|_q \le \|f\|_2 |S|^{\frac{1}{q} - \frac{1}{2}} \,.$$

Taking $r = \rho_2\sqrt{n}$ and using (79) to estimate $|S|$, we conclude that

$$\|\Pi_r f\|_2 \le e^{-n\epsilon}\|f\|_2 \,, \tag{90}$$

whenever there is a $\gamma > 0$ such that

$$4\pi^2\rho_2^2 + \frac{\gamma}{2}\ln\frac{2\rho_1^2\gamma}{e} < 0 \,.$$

Since $\min_{\gamma>0}\gamma\ln(a\gamma) = -\frac{1}{ea}$, we get that (90) holds whenever

$$4\pi^2\rho_2^2 - \frac{1}{4\rho_1^2} < 0 \,,$$

which is equivalent to (85). □

The structure of the proof for $\mathbb{R}^n$ suggests that perhaps it is worthwhile to look for a general inequality on the hypercube that could replace the use of hypercontractivity in the proof of Theorem 9, i.e. play a role similar to that of the sharp Young inequality on $\mathbb{R}^n$ (of course, the Young inequality itself cannot be sharpened on the hypercube, or on any finite group).

5. An application to coding theory

Theorem 14. *For any $0 < R' < R < 1$ there exists $\delta_n \to 0$ such that for any linear map $f : \mathbb{F}_2^k \to \mathbb{F}_2^n$ with $\frac{k}{n} = R$ there exists an $x \in \mathbb{F}_2^k$ s.t.*

$$\frac{1}{n}|f(x)| \le \delta_{LP1}(R') + \delta_n \tag{91}$$
$$\frac{1}{k}|x| \ge \delta_{LP1}\left(\frac{R'}{R}\right) - \delta_n \,, \tag{92}$$

where $\delta_{LP1}(h(\rho)) = \frac{1}{2} - \sqrt{\rho(1-\rho)}$.

Remark 5. *This estimate significantly outperforms previously best known bounds of this kind [24, Theorem 1], but only applies to linear maps.*

We give two different proofs, in two subsections below. Note that the two proofs take slightly different points of view. The first proof deals with linear maps $f : \mathbb{F}_2^k \to \mathbb{F}_2^n$, while the second proof looks rather at images of these maps, linear codes in $\mathbb{F}_2^n$. In particular, in the second proof we assume that the image of f is of dimension k (i.e. f is of full rank).

5.1. Method 1 - graph covers

Proof. To every linear map $f : \mathbb{F}_2^k \to \mathbb{F}_2^n$ we associate the following increasing sequence of numbers:

$$d_r(f) \stackrel{\triangle}{=} \min\{|f(x)| : |x| \geq r\} ,$$

where d_1 is just the minimum distance of f. Following, Friedman and Tillich [25] we also associate to f a Cayley graph Γ with vertices $\mathbb{F}_2^k$ and generators taken as columns c_i of $k \times n$ matrix describing f. (We will use freely facts from [25], perhaps in a somewhat different formulation, from now on.) Then

$$n - 2d_r = \max\left\{ \frac{(Ah, h)}{\|h\|_2^2} : \hat{h} = 0 \text{ on ball } B(0, r-1) \right\} , \tag{93}$$

where A is the adjacency matrix of Γ. Note that A is also a convolution operator on $\mathbb{F}_2^k$:

$$Ah = h * (\sum_{i=1}^{n} \delta_{c_i}) .$$

As in [25], select a covering map $\mathbb{F}_2^n \to \mathbb{F}_2^k$ and take $B \subset \mathbb{F}_2^n$ to be the Hamming ball of radius $n\rho_1$, with $0 < \rho_1 < \frac{1}{2}$ found as $h(\rho_1) = R'$. There exists a function $g_B : \mathbb{F}_2^n \to \mathbb{R}$, supported on B with the property:

$$(A_C \; g_B, g_B) \geq \lambda_B \|g_B\|_2^2 ,$$

where A_C is the adjacency matrix of the n-dimensional hypercube, and $\lambda_B = 2n\sqrt{\rho_1(1-\rho_1)} + o(n)$.

Hence, there exists a function $h_B : \mathbb{F}_2^k \to \mathbb{R}$ supported on the image of B under the covering map with the property:

$$(Ah_B, h_B) \geq \lambda_B \|h_B\|_2^2 ,$$

Then to get a lower bound on (93) we set

$$h = h_B - \Pi_{<r} h_B ,$$

where $\Pi_{<r} = \sum_{a<r} \Pi_a$ and Π_a is from (62).

Note that A and Π commute and so $(A\Pi h, h) \leq n(\Pi h, h)$, then we have :

$$(Ah, h) = (Ah_B, h_B) - (A\Pi_{<r} h_B, h_B) \geq \lambda_B \|h_B\|_2^2 - n(\Pi_{<r} h_B, h_B)$$

Thus,

$$n - 2d_r \geq 2n\sqrt{\rho_1(1-\rho_1)} + o(n) \tag{94}$$

whenever

$$\|\Pi_{<r} h_B\|_2^2 \leq \|h_B\|_2^2 \cdot o(1)\,.$$

Using the uncertainty principle for the k-dimensional cube (Theorem 9) we estimate

$$\|\Pi_{<r} h_B\| \ll \|h_B\|^2\,,$$

as long as

$$\frac{r}{k} < \frac{1}{2} - \sqrt{\rho_1'(1-\rho_1')}\,, \tag{95}$$

where ρ_1' is found from $h(\rho_1') = \frac{h(\rho_1)}{R} = \frac{R'}{R}$. After simple algebra, we see that (94)-(95) are equivalent to (91)-(92).

□

5.2. Method 2 (analytic)

We start an uncertainty-type claim for subspaces of $\mathbb{F}_2^n$.

Let C be a k-dimensional linear subspace C of $\mathbb{F}_2^n$. Given a basis $\mathbf{v} = \left\{v_1, ..., v_k\right\}$ of C, denote the length of representation of a vector $x \in C$ in terms of V by $|x|_{\mathbf{v}}$.

Lemma 15. *Let f be a function supported on a subset $A \subseteq \mathbb{F}_2^n$. Let $0 \leq r \leq k \leq n$ be integer parameters such that $k \geq \log_2|A|$, and, moreover, writing $|A| = 2^{h(\rho_1)\cdot k}$, $\binom{k}{r} = 2^{h(\rho_2)\cdot k}$, we have $(1-2\rho_1)^2 + (1-2\rho_2)^2 \;>\; 1$.*

Then, for any k-dimensional subspace C of $\mathbb{F}_2^n$ and for any basis $\mathbf{v}$ of C holds

$$\sum_{\omega \in C, |\omega|_{\mathbf{v}} \leq r} \widehat{f}^2(\omega) \ll \sum_{\omega \in C} \widehat{f}^2(\omega)$$

Here the $\ll$ sign means that the LHS is exponentially smaller than the RHS.

Proof. Let $F = |C| \cdot f * 1_{C^\perp}$.

Note that F is constant on cosets of $C^\perp$ and that $\widehat{F}(\omega) = \begin{cases} \widehat{f}(\omega) & \text{if} \quad \omega \in C \\ 0 & \text{otherwise} \end{cases}$.

Let M be a $k \times n$ matrix with rows $v_1, ..., v_k$. We define a function g on $\mathbb{F}_2^k$ as follows. For $x \in \mathbb{F}_2^k$, the pre-image $\{y \in \mathbb{F}_2^n, My = x\}$ is a coset of $C^\perp$, and we set $g(x)$ to be the (fixed) value of F on this coset.

Next, we calculate the Fourier transform of g. Let $\alpha \in \mathbb{F}_2^k$. Let $\omega = \alpha^t M \in \mathbb{F}_2^n$. We claim that $\widehat{g}(\alpha) = \widehat{F}(\omega)$. To see this, note that for any y such that $My = x$ holds $(x, \alpha) = (My, \alpha) = (y, M^t\alpha) = (y, \omega)$. Using this we compute

$$\widehat{g}(\alpha) = \sum_{x \in \mathbb{F}_2^k} g(x)(-1)^{(x,\alpha)} = \sum_{x \in \mathbb{F}_2^k} \sum_{y : My = x} F(y)(-1)^{(y,\omega)} = \widehat{F}(\omega)$$

Next, we apply the uncertainty principle for g on $\mathbb{F}_2^k$. Observe that the cardinality of the support of g is given by the number of cosets of $C^\perp$ intersecting A, which is at most $|A|$. The constraints on $|A|$, k, and r imply $\sum_{|\alpha|\le r} \widehat{g}^2(\alpha) \ll \sum_\alpha \widehat{g}^2(\alpha)$, which is equivalent to the claim of the lemma.

□

We now prove Theorem 14, first restating it for linear codes rather than for linear maps.

Theorem 16. *Let $0 < R < 1$. Let $C \subseteq \mathbb{F}_2^n$ be a linear code of rate $k = Rn$. Let $\mathbf{v} = \left\{v_1, ..., v_k\right\}$ be a basis of C. Then for any $0 \le R' < R$ there is a vector $x \in C$ with*

$$\frac{1}{n}|x| \le \delta_{LP1}(R') + \delta_n \quad \textit{and} \quad \frac{1}{k}|x|_{\mathbf{v}} \ge \delta_{LP1}\left(\frac{R'}{R}\right) - \delta_n$$

Proof. Let $r = h^{-1}(R') \cdot n$. Let B be the Hamming ball of radius r around zero in $\mathbb{F}_2^n$. As in [25], let g_B be a function supported on B, with the property:

$$(A_C\ g_B, g_B) \ge \lambda_B \|g_B\|_2^2\,,$$

where A_C is the adjacency matrix of the n-dimensional hypercube, and $\lambda_B = 2n\sqrt{\frac{r}{n}(1-\frac{r}{n})} + o(n)$.

Let $d = \frac{n-\lambda_B+1}{2}$. Note that $d = \left(\frac{1}{2} - \sqrt{h^{-1}(R')\,(1-h^{-1}(R'))}\right) \cdot n + o(n) = \delta_{LP1}(R') \cdot n + o(n)$.

Note that $|B| = 2^{R' \cdot n} \le |C| = 2^k$. We introduce two additional parameters with a view towards using Lemma 15. Let ρ_1 be such that $|B| = 2^{h(\rho_1)\cdot k}$, and let ρ_2 satisfy $(1-2\rho_1)^2 + (1-2\rho_2)^2 = 1$. Computing explicitly,

$$\rho_1 = h^{-1}\left(\frac{R'}{R}\right) \quad \texttt{and} \quad \rho_2 = \delta_{LP1}\left(\frac{R'}{R}\right)$$

We proceed with the following computation, as in [26]. Let $F = |C| \cdot g_B * 1_{C^\perp}$. Compute $(A_C F, F)$ in two ways. On one hand, since A_C commutes with convolutions, we have $(A_C\ F, F) \ge \lambda_B \cdot (F, F)$. On the other hand, $(A_C\ F, F) = ((n-2|x|) \cdot \widehat{F}(x), \widehat{F}(x)) = \sum_x (n-2|x|) \cdot \widehat{F}^2(x)$.

Note that $n - 2d = \lambda_B - 1$. Hence, taking into account both expressions gives, after a simple calculation:

$$n \cdot \sum_{x\in C, |x|\le d} \widehat{g_B}^2(x) = n \cdot \sum_{|x|\le d} \widehat{F}^2(x) \ge \sum_x \widehat{F}^2(x) = \sum_{x \in C} \widehat{g_B}^2(x)$$

We now apply Lemma 15 for g_B. Let $\frac{r}{k} < \rho_2 = \delta_{LP1}\left(\frac{R'}{R}\right)$. By the lemma, ignoring negligible factors,

$$n \cdot \sum_{x \in C, |x|\le d, |x|_{\mathbf{v}} > r} \widehat{g_B}^2(x) \ge \sum_{x\in C} \widehat{g_B}^2(x) > 0$$

This means that there exists a vector $x \in C$ such that $|x| \leq d$ and $|x|_{\mathbf{v}} > r$, proving the claim of the theorem.

□

[1] D. Bakry, Functional inequalities for Markov semigroups, in: Probability measures on groups, Tata Institute of Fundamental Research, Mubai, 2006, pp. 91–147.

[2] L. Saloff-Coste, Lectures on finite Markov chains, in: Lectures on probability theory and statistics, Springer, 1997, pp. 301–413.

[3] A. D. Wyner, J. Ziv, A theorem on the entropy of certain binary sequences and applications–I, IEEE Trans. Inf. Theory 19 (6) (1973) 769–772.

[4] P. Diaconis, L. Saloff-Coste, Logarithmic Sobolev inequalities for finite Markov chains, Ann. Appl. Probab. 6 (3) (1996) 695–750.

[5] A. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon, Inf. Contr. 2 (2) (1959) 101–112.

[6] M. Costa, T. Cover, On the similarity of the entropy power inequality and the Brunn-Minkowski inequality (corresp.), IEEE Trans. Inf. Theory 30 (6) (1984) 837–839.

[7] E. B. Davies, B. Simon, Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians, J. Func. Anal. 59 (2) (1984) 335–395.

[8] D. Bakry, L'hypercontractivité et son utilisation en théorie des semigroupes, in: Lectures on probability theory, Springer, 1994, pp. 1–114.

[9] A. Samorodnitsky, A modified logarithmic sobolev inequality for the hamming cube and some applications, arXiv preprint arXiv:0807.1679.

[10] G. H. Hardy, J. E. Littlewood, G. Polya, Inequalities, Cambridge University Press, 1988.

[11] E. Mossel, K. Oleszkiewicz, A. Sen, On reverse hypercontractivity, Geometric and Functional Analysis 23 (3) (2013) 1062–1097.

[12] P. Hartman, Ordinary differential equation, John Wiley & Sons, New York, USA, 1964.

[13] V. Havin, B. Jöricke, The uncertainty principle in harmonic analysis, Springer, 1994.

[14] W. H. J. Fuchs, On the magnitude of fourier transforms, in: Proc. Int. Math. Cong., Amsterdam, 1954, pp. 106–107.

[15] H. J. Landau, H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty– II, Bell Syst. Tech. J. 40 (1) (1961) 65–84.

[16] D. Slepian, H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty – I, Bell Syst. Tech. J. 40 (1) (1961) 43–63.

[17] M. Benedicks, On Fourier transforms of functions supported on sets of finite Lebesgue measure, J. Math. Anal. Appl. 106 (1) (1985) 180–183.

[18] W. Amrein, A. Berthier, On support properties of l_p-functions and their Fourier transforms, J. Func. Anal. 24 (3) (1977) 258–267.

[19] F. L. Nazarov, Local estimates for exponential polynomials and their applications to inequalities of the uncertainty principle type, Algebra i Analiz (in Russian) 5 (4) (1993) 3–66.

[20] P. Jaming, Nazarov's uncertainty principles in higher dimension, J. Approximation Th. 149 (1) (2007) 30–41.

[21] A. Samorodnitsky, Extremal properties of solutions for Delsarte's linear program, preprint.
URL http://www.cs.huji.ac.il/~salex/papers/old_sq_measure.ps

[22] T. Tao, V. H. Vu, Additive combinatorics, Vol. 105, Cambridge University Press, 2006.

[23] W. Beckner, Inequalities in Fourier analysis, Ann. Math. 102 (1) (1975) 159–182.

[24] Y. Polyanskiy, On metric properties of maps between hamming spaces and related graph homomorphisms, Arxiv preprint arXiv:1503.02779.

[25] J. Friedman, J.-P. Tillich, Generalized alon–boppana theorems and error-correcting codes, SIAM Journal on Discrete Mathematics 19 (3) (2005) 700–718.

[26] M. Navon, A. Samorodnitsky, Linear programming bounds for codes via a covering argument, Discrete & Computational Geometry 41 (2) (2009) 199–207.