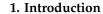*Article*

Hypercontractive Inequalities for the Second Norm of Highly Concentrated Functions, and Mrs. Gerber's-Type Inequalities for the Second Rényi Entropy

Niv Levhari [1,2,*] and Alex Samorodnitsky [1,*]

[1]   School of Engineering and Computer Science, The Hebrew University of Jerusalem, Jerusalem 9103401, Israel
[2]   School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
*   Correspondence: nlevhari@gmail.com (N.L.); salex@cs.huji.ac.il (A.S.)

Abstract: Let T_ϵ, $0 \le \epsilon \le 1/2$, be the noise operator acting on functions on the boolean cube $\{0,1\}^n$. Let f be a distribution on $\{0,1\}^n$ and let $q > 1$. We prove tight Mrs. Gerber-type results for the second Rényi entropy of $T_\epsilon f$ which take into account the value of the qth Rényi entropy of f. For a general function f on $\{0,1\}^n$ we prove tight hypercontractive inequalities for the ℓ_2 norm of $T_\epsilon f$ which take into account the ratio between ℓ_q and ℓ_1 norms of f.

Keywords: entropy; hypercontractivity; Rényi entropy; Mrs. Gerber's inequality

1. Introduction

This paper considers the problem of quantifying the decrease in the ℓ_2 norm of a function on the boolean cube when this function is acted on by the noise operator.

Given a noise parameter $0 \le \epsilon \le 1/2$, the noise operator T_ϵ acts on functions on the boolean cube as follows: for $f : \{0,1\}^n \to \mathbb{R}$, $T_\epsilon f$ at a point x is the expected value of f at y, where y is a random binary vector whose i^{th} coordinate is x_i with probability $1 - \epsilon$ and $1 - x_i$ with probability ϵ, independently for different coordinates. Namely, $(T_\epsilon f)(x) = \sum_{y \in \{0,1\}^n} \epsilon^{|y-x|}(1-\epsilon)^{n-|y-x|}f(y)$, where $|\cdot|$ denotes the Hamming distance. We will write f_ϵ for $T_\epsilon f$, for brevity.

Note that f_ϵ is a convex combination of shifted copies of f. Hence, the noise operator decreases norms. Recall that the ℓ_q norm of a function is given by $\|f\|_q = (\mathbb{E}|f|^q)^{\frac{1}{q}}$ (the expectations here and below are taken w.r.t. the uniform measure on $\{0,1\}^n$). The norms $\{\|f\|_q\}_q$ increase with q. An effective way to quantify the decrease of ℓ_q norm under noise is given by the hypercontractive inequality [1–3] (see also, e.g., [4] for background), which upperbounds the ℓ_q norm of the noisy version of a function by a smaller norm of the original function.

$$\|f_\epsilon\|_q \le \|f\|_{1+(1-2\epsilon)^2(q-1)}. \tag{1}$$

This inequality is essentially tight in the following sense. For any $p < 1 + (q-1)(1-2\epsilon)^2$ there exists a non-constant function $f : \{0,1\}^n \to \mathbb{R}$ with $\|f_\epsilon\|_q > \|f\|_p$.

Entropy provides another example of a convex homogeneous functional on (nonnegative) functions on the boolean cube. For a nonnegative function f let the entropy of f be given by $Ent(f) = \mathbb{E} f \log_2 f - \mathbb{E} f \log_2 \mathbb{E} f$. The entropy of f is closely related to Shannon's entropy of the corresponding distribution $f/\Sigma f$ on $\{0,1\}^n$, and similarly the entropy of f_ϵ is related to Shannon's entropy of the output of a binary symmetric channel with error probability ϵ on input distributed according to $f/\Sigma f$ (see below and, e.g., the discussion in the introduction of [5]). The decrease in entropy (or, correspondingly, the increase in Shannon's entropy) after noise is quantified in the "Mrs. Gerber's Lemma" [6]:

$$Ent(f_\epsilon) \le n \, \mathbb{E} f \cdot \psi\left(\frac{Ent(f)}{n \, \mathbb{E} f}, \epsilon\right), \tag{2}$$

where $\psi = \psi(x, \epsilon) = H\big((1 - 2\epsilon) \cdot H^{-1}(1 - x) + \epsilon\big)$ is an explicitly given function on $[0, 1] \times [0, 1/2]$, which is increasing and strictly concave in its first argument for any $0 < \epsilon < \frac{1}{2}$. Here and below we write $H(t) = t \log_2\left(\frac{1}{t}\right) + (1 - t) \log_2\left(\frac{1}{1-t}\right)$ for the binary entropy function.

Equality holds iff f is a product function with equal marginals. That is, there exists a function $g : \{0, 1\} \to \mathbb{R}$, such that for any $x = (x_1, ..., x_n) \in \{0, 1\}^n$ holds $f(x) = \prod_{i=1}^n g(x_i)$.

One has $\psi(0, \epsilon) = 0$ and $\frac{\partial \psi}{\partial x}\big|_{x=0} = (1 - 2\epsilon)^2$. Hence $\psi(x, \epsilon) \le (1 - 2\epsilon)^2 \cdot x$, with equality only at $x = 0$. Hence the inequality (2) has the following weaker linear approximation version

$$Ent(f_\epsilon) \le (1 - 2\epsilon)^2 \cdot Ent(f), \tag{3}$$

in which equality holds if and only if f is a constant function.

Rényi entropies. There is a well-known connection between ℓ_q norms of a nonnegative function f and its entropy (see, e.g., [7]): Assume, as we may by homogeneity, that $\mathbb{E} f = 1$. Then $Ent(f) = \lim_{q \to 1} \frac{1}{q-1} \log_2 \|f\|_q^q$. (The quantity $Ent_q(f) = \frac{1}{q-1} \log_2 \|f\|_q^q$ is known as the qth Rényi entropy of f ([8])). (Note that this notion is defined for all, not necessarily nonnegative, functions on $\{0, 1\}^n$.) The entropies $\{Ent_q(f)\}_q$ increase with q. Restating the inequality (1) in terms of Rényi entropies gives

$$Ent_q(f_\epsilon) \le \frac{(1 - 2\epsilon)^2 q}{(1 - 2\epsilon)^2(q - 1) + 1} \cdot Ent_{1 + (1 - 2\epsilon)^2(q-1)}(f).$$

Note that taking $q \to 1$ in this inequality recovers only the (weaker) linear approximation version (3) of Mrs. Gerber's inequality (2). This highlights an important difference between inequalities (1) and (2). Mrs. Gerber's lemma takes into account the distribution of a function, specifically the ratio between its entropy and its ℓ_1 norm. When this ratio is exponentially large in n, which typically holds in the information theory contexts in which this inequality is applied, (2) is significantly stronger than (3). On the other hand, hypercontractive inequalities seem to be typically applied in contexts in which the ratio between different norms of the function is subexponential in n, and there are examples of such functions for which (1) is essentially tight. With that, there are several recent results [9–11] which show that (1) can be strengthened, if the ratio $\frac{\|f\|_q}{\|f\|_1}$, for some $q > 1$, is exponentially large in n. In the framework of Rényi entropies, the possibility of a result analogous to (2) for higher Rényi entropies was discussed in [12].

Our results. This paper proves a Mrs. Gerber type result for the second Rényi entropy, and a hypercontractive inequality for the ℓ_2 norm of f_ϵ which take into account the ratio between ℓ_q and ℓ_1 norms of f. We try to pattern the results below after (2). We start with a Mrs. Gerber type inequality.

Proposition 1. *Let $q > 1$, and let f be a nonnegative function on $\{0, 1\}^n$ such that $\mathbb{E} f = 1$. Then*

$$\frac{Ent_2(f_\epsilon)}{n} \le \psi_{2,q}\left(\frac{Ent_q(f)}{n}, \epsilon\right), \tag{4}$$

where $\psi_{2,q}$ is an explicitly given function on $[0, 1] \times [0, 1/2]$, which is increasing and concave in its first argument. The function $\psi_{2,q}$ is defined in Definition 1 below.

This inequality is essentially tight in the following sense. For any $0 < x < 1$ and $0 < \epsilon < \frac{1}{2}$, and for any $y < \psi_{2,q}(x, \epsilon)$ there exists a sufficiently large n and a nonnegative function f on $\{0, 1\}^n$ with $\mathbb{E} f = 1$, $\frac{Ent_q(f)}{n} \le x$ and $\frac{Ent_2(f_\epsilon)}{n} > y$.

Let us make some comments about this result.

– The functions $\{\psi_{2,q}\}_q$ are somewhat cumbersome to describe, and hence we relegate their precise definition to Definition 1 below.

- Inequality (4) upper bounds $Ent_2(f_\epsilon)$ in terms of $Ent_q(f)$ for $q > 1$, and ϵ. Taking $q = 2$ gives an upper bound on $Ent_2(f_\epsilon)$ in terms of $Ent_2(f)$ and ϵ, in analogy to (2).
- Recall that for a point $x \in \{0,1\}^n$ and $0 \leq r \leq n$, the Hamming sphere of radius r around x is the set $\{y \in \{0,1\}^n : |y - x| = r\}$. As will be seen from the proof of Proposition 1, (4) is essentially tight for a certain convex combination of the uniform distribution on $\{0,1\}^n$ and the characteristic function of a Hamming sphere of an appropriate radius (depending on q, ϵ, and the required value of $Ent_q(f)$).
- In information theory one typically considers a slightly different notion of Rényi entropies: For a probability distribution P on Ω, the qth Renyi entropy of P is given by $H_q(P) = -\frac{1}{q-1} \log_2(\sum_{\omega \in \Omega} P^q(\omega))$. To connect notions, if f is a nonnegative (non-zero) function on $\{0,1\}^n$ with expectation 1, then $P = \frac{f}{2^n}$ is a probability distribution, and $Ent_q(f) = n - H_q(P)$. Furthermore, $Ent_q(f_\epsilon) = n - H_q(X \oplus Z)$, where X is a random variable on $\{0,1\}^n$ distributed accordinng to P and Z is an independent noise vector corresponding to a binary symmetric channel with crossover probability ϵ. Hence, (2) can be restated as

$$H(X \oplus Z) \geq n \cdot \varphi\left(\frac{H(X)}{n}, \epsilon\right),$$

  and Proposition 1 can be restated as

$$H_2(X \oplus Z) \geq n \cdot \varphi_{2,q}\left(\frac{H_q(X)}{n}, \epsilon\right)$$

  Here φ is an explicitly given function on $[0,1] \times [0,1/2]$, which is increasing and convex in its first argument ($\varphi(x, \epsilon) = 1 - \psi(1 - x, \epsilon)$), and similarly for $\varphi_{2,q}$.
  Next, we describe our main result, a hypercontractive inequality for the ℓ_2 norm of f_ϵ which takes into account the ratio between ℓ_q and ℓ_1 norms of f, and more specifically $Ent_q\left(\frac{f}{\|f\|_1}\right) = \frac{q}{q-1} \log_2\left(\frac{\|f\|_q}{\|f\|_1}\right)$.

Theorem 1. *Let $q > 1$, and let f be a non-zero function on $\{0,1\}^n$. Then*

$$\|f_\epsilon\|_2 \leq \|f\|_\kappa, \tag{5}$$

where $\kappa = \kappa_{2,q}\left(\frac{Ent_q\left(\frac{f}{\|f\|_1}\right)}{n}, \epsilon\right)$, and $\kappa_{2,q}$ is an explicitly given function on $[0,1] \times [0,1/2]$, which is decreasing in its first argument and which satisfies $\kappa_{2,q}(0, \epsilon) = 1 + (1 - 2\epsilon)^2$, for all $0 \leq \epsilon \leq \frac{1}{2}$. The function $\kappa_{2,q}$ is defined in Definition 1 below.
This inequality is essentially tight in the following sense. For any $0 < x < 1$ and $0 < \epsilon < \frac{1}{2}$, and for any $y < \kappa_{2,q}(x, \epsilon)$ there exists a sufficiently large n and a function f on $\{0,1\}^n$ with $\frac{Ent_q(f/\|f\|_1)}{n} \geq x$ and $\|f_\epsilon\|_2 > \|f\|_y$.

Some comments (see also Lemma 10 below).

- The precise definition of the functions $\{\kappa_{2,q}\}_q$ will be given in Definition 1 below. At this point let us just observe that since the sequence $\{Ent_q(f)\}_q$ increases with q, we would expect the fact that $Ent_q(f)$ is large to become less significant as q increases. This is expressed in the properties of the functions $\{\kappa_{2,q}\}_q$ in the following manner: If $q \geq 2$ then for any $0 < \epsilon < \frac{1}{2}$ the function $\kappa_{2,q}(x, \epsilon)$ starts as a constant-$(1 + (1 - 2\epsilon)^2)$ function up to some $x = x(q, \epsilon) > 0$, and becomes strictly decreasing after that.

  In other words $x(q, \epsilon)$ is the largest possible value of $\frac{Ent_q\left(\frac{f}{\|f\|_1}\right)}{n}$ for which Theorem 1 provides no new information compared to (1). For $1 < q < 2$ there is a value $0 < \epsilon(q) < \frac{1}{2}$, such that for all $\epsilon \leq \epsilon(q)$ the function $\kappa_{2,q}(x, \epsilon)$ is strictly decreasing (in which case we say that $x(q, \epsilon) = 0$). However, $x(q, \epsilon) > 0$ for all $\epsilon > \epsilon(q)$. The function

$\epsilon(q)$ decreases with q (in particular, $\epsilon(q) = 0$ for $g \geq 2$). The function $x(q, \epsilon)$ increases both in q and in ϵ.

– Notably, taking $q \to 1$ in Theorem 1 gives (see Corollary 1)

$$\|f_\epsilon\|_2 \leq \|f\|_\kappa,$$

where $\kappa = \kappa_{2,1}\left(Ent\left(\frac{f}{\|f\|_1}\right)/n, \epsilon\right) = -\frac{Ent\left(\frac{f}{\|f\|_1}\right)/n}{\phi_\epsilon\left(1 - Ent\left(\frac{f}{\|f\|_1}\right)/n\right)}$. The function $\kappa_{2,1}(x, \epsilon) = -\frac{x}{\phi_\epsilon(1-x)}$ is strictly decreasing in x for any $0 < \epsilon < \frac{1}{2}$. It satisfies $\kappa_{2,1}(0, \epsilon) = \lim_{x \to 0} \kappa_{2,1}(x, \epsilon) = 1 + (1 - 2\epsilon)^2$, for all $0 \leq \epsilon \leq \frac{1}{2}$. Hence, this is stronger than (1) for any non-constant function f and for any $0 < \epsilon < \frac{1}{2}$, with the difference between the two inequalities becoming significant when $Ent\left(\frac{f}{\|f\|_1}\right)/n$ is bounded away from 0.

– As will be seen from the proof of Theorem 1, (5) is essentially tight for a certain convex combination of the uniform distribution on $\{0, 1\}^n$ and characteristic functions of one or two Hamming spheres of appropriate radii (the number of the spheres and their radii depend on q, ϵ, and the required value of $Ent_q\left(\frac{f}{\|f\|_1}\right)$).

– Let f be a non-constant function and let $0 < \epsilon < \frac{1}{2}$ be fixed. Consider the function $F(q) = F_{f,\epsilon}(q) = \kappa_{2,q}\left(\frac{Ent_q\left(\frac{f}{\|f\|_1}\right)}{n}, \epsilon\right)$. It will be seen that there is a unique value $1 < q(f, \epsilon) \leq 1 + (1 - 2\epsilon)^2$ of q for which $F(q) = q$. Furthermore, $q(f, \epsilon) = \min_{q \geq 1} F(q)$. Hence it provides the best possible value for κ in Theorem 1. With that, determining $q(f, \epsilon)$ might in principle require knowledge of all the Renyi entropies $Ent_q(f)$, for $1 \leq q \leq 1 + (1 - 2\epsilon)^2$, while typically we are in possession of one of the "easier" Rényi entropies, such as $Ent(f)$ or $Ent_2(f)$.

1.1. Full Statements of Proposition 1 and Theorem 1

We now define the functions $\{\psi_{2,q}\}_q$ in Proposition 1 and $\{\kappa_{2,q}\}_q$ in Theorem 1, completing the statements of these claims. We start with introducing yet another function on $[0, 1] \times [0, 1/2]$ which will play a key role in what follows (we remark that this function was studied in [9]). For $0 \leq x \leq 1$ and $0 \leq \epsilon \leq \frac{1}{2}$, let $\sigma = H^{-1}(x)$ and let $y = y(x, \epsilon) = \frac{-\epsilon^2 + \epsilon\sqrt{\epsilon^2 + 4(1 - 2\epsilon)\sigma(1 - \sigma)}}{2(1 - 2\epsilon)}$. Let

$$\Phi(x, \epsilon) = \frac{1}{2} \cdot \left(x - 1 + \sigma H\left(\frac{y}{\sigma}\right) + (1 - \sigma)H\left(\frac{y}{1 - \sigma}\right) + 2y\log_2(\epsilon) + (1 - 2y)\log_2(1 - \epsilon)\right).$$

The function Φ is nonpositive. It is increasing and concave in its first argument. Additional relevant properties of Φ are listed in Lemma 3 below. For a fixed ϵ, it will be convenient to write $\phi_\epsilon(x) = \Phi(x, 2\epsilon(1 - \epsilon))$, viewing ϕ_ϵ as a univariate function on $[0, 1]$.

Definition 1. *Let $0 \leq x \leq 1$ and $0 \leq \epsilon \leq \frac{1}{2}$.*

- *If $\phi'_\epsilon(1 - x) < \frac{1}{q}$, let $\alpha_0 = (\phi'_\epsilon)^{-1}\left(\frac{1}{q}\right)$. Define*

$$\psi_{2,q}(x, \epsilon) = 2 \cdot \begin{cases} \frac{q-1}{q} \cdot x + \left(\phi_\epsilon(\alpha_0) + \frac{1-\alpha_0}{q}\right) & \text{if} & \phi'_\epsilon(1 - x) < \frac{1}{q} \\ \phi_\epsilon(1 - x) + x & \text{otherwise} \end{cases}$$

- *Let $y = \frac{q-1}{q} \cdot x + \frac{1}{q}$. Let $q_0 = 1 + (1 - 2\epsilon)^2$. If $y \geq \frac{1}{q_0}$, let α_0 be determined by $1 - \alpha_0 - \frac{\alpha_0 \phi_\epsilon(\alpha_0)}{1 - \alpha_0} = y$. If $x = 0$, define $\kappa_{2,q}(x, \epsilon) = q_0$. Otherwise, define*

$$\kappa_{2,q}(x,\epsilon) = \begin{cases} q_0 & \text{if} & y \le \frac{1}{q_0} \\ -\frac{x}{\phi_\epsilon(1-x)} & \text{if} & y > \frac{1}{q_0} \text{ and } -\frac{x}{\phi_\epsilon(1-x)} \ge q \\ \frac{\alpha_0-1}{\phi_\epsilon(\alpha_0)} & \text{if} & y > \frac{1}{q_0} \text{ and } -\frac{x}{\phi_\epsilon(1-x)} < q \end{cases}$$

We remark that it is not immediately obvious that the functions $\psi_{2,q}$ and $\kappa_{2,q}$ are well-defined. This will be clarified in the proofs of Proposition 1 and Theorem 1.

We state explicitly some special cases of Theorem 1, which seem to be the most relevant for applications. They describe the improvement over (1), given non-trivial information about $Ent(f)$ and $\|f\|_2$.

Corollary 1.

1. *Taking $q \to 1$ in Theorem 1 gives:*

$$\|f_\epsilon\|_2 \le \|f\|_\kappa, \quad \text{with} \quad \kappa = -\frac{Ent\left(\frac{f}{\|f\|_1}\right)/n}{\phi_\epsilon\left(1 - Ent\left(\frac{f}{\|f\|_1}\right)/n\right)}.$$

2. *Taking $q = 2$ in Theorem 1 gives, for $x = \frac{Ent_2\left(\frac{f}{\|f\|_1}\right)}{n}$ and $q_0 = 1 + (1-2\epsilon)^2$*

$$\|f_\epsilon\|_2 \le \|f\|_\kappa, \quad \text{with} \quad \kappa = \begin{cases} q_0 & \text{if} & \frac{x+1}{2} \le \frac{1}{q_0} \\ \frac{\alpha-1}{\phi_\epsilon(\alpha)} & \text{otherwise} \end{cases}$$

In the second case α is determined by $1 - \alpha - \frac{\alpha\phi_\epsilon(\alpha)}{1-\alpha} = \frac{x+1}{2}$.

We observe that both Proposition 1 and Theorem 1 are based on the following claim ([9], Corollary 3.2). This claim also explains the relevance of function Φ.

Theorem 2. *Let $0 \le x \le 1$. Let f be a function on $\{0,1\}^n$ supported on a set of cardinality at most 2^{xn}. Then, for any $0 \le \epsilon \le \frac{1}{2}$ holds*

$$\langle f_\epsilon, f \rangle \le 2^{(2\Phi(x,\epsilon)+1-x)\cdot n} \cdot \|f\|_2^2,$$

Moreover, this is tight, up to a polynomial in n factor, if f is the characteristic function of a Hamming sphere of radius $H^{-1}(x) \cdot n$.

1.2. Applications

We describe some applications of the results above, related mainly to coding theory. We start with providing some relevant context.

Coding theory. A binary error-correcting code C of length n and minimal distance d is a subset of $\{0,1\}^n$ in which the distance between any two distinct points is at least d. Let $A(n,d)$ be the maximal size of such a code. A well-known open problem in coding theory is to determine, given $0 < \delta < \frac{1}{2}$, the *asymptotic maximal rate* $R(\delta) = \limsup_{n\to\infty} \frac{1}{n} \log_2 A(n, \lfloor \delta n \rfloor)$ of a code with relative distance δ. The best known lower bound on $R(\delta)$ is the Gilbert-Varshamov bound $R(\delta) \ge 1 - H(\delta)$ [13]. The best known upper bounds on $R(\delta)$ were obtained in [14] using the linear programming relaxation, constructed in [15], of the combinatorial problem of bounding $A(n,d)$. Let $A_{LP}(n,d)$ be the value of the appropriate linear program of [15] and let $R_{LP}(\delta) = \limsup_{n\to\infty} \frac{1}{n} \log_2 A_{LP}(n, \lfloor \delta n \rfloor)$. By construction, $A_{LP}(n,d) \ge A(n,d)$ for all n and d and hence $R_{LP}(\delta) \ge R(\delta)$. The *first JPL bound* of [14] is $R(\delta) \le R_{LP}(\delta) \le H\left(1/2 - \sqrt{\delta(1-\delta)}\right)$. This bound is the best known for a subrange of values of δ. The best known bound is the *second JPL bound* of [14]. It is better than the first bound for relatively small values of δ. However, it is more complicated to state explicitly and we omit it here. The second JPL bound is strictly larger than the Gilbert-Varshamov bound for all $0 < \delta < \frac{1}{2}$, and hence $R(\delta)$ is unknown for all these values of δ.

The value of $R_{LP}(\delta)$ is also unknown, for all $0 < \delta < \frac{1}{2}$. Clearly $R_{LP}(\delta) \geq R(\delta) \geq 1 - H(\delta)$. It was conjectured in [14] that $R_{LP}(\delta)$ lies strictly between the second JPL bound and the Gilbert-Varshamov bound. On the other hand, there is a convincing numeric evidence [16] that $R_{LP}(\delta)$ in fact coincides with the second JPL bound. A lower bound $R_{LP}(\delta) \geq \frac{1 - H(\delta) + H\left(1/2 - \sqrt{\delta(1-\delta)}\right)}{2}$ was shown in [17] (note that the RHS here is the arithmetic average of the Gilbert-Varshamov bound and the first JPL bound). It was improved, for a subrange of δ, in [18].

A different approach to obtain upper bounds on the cardinality of binary codes was presented in [19]. For a subset $D \subseteq \{0,1\}^n$, let M_D be the adjacency matrix of the subgraph of the discrete cube induced by the vertices of D. Let $\lambda(D)$ be the maximal eigenvalue of M_D. The following claim was proved in [19] for binary linear codes (and extended in [18] to general binary codes): Let D be subset of $\{0,1\}^n$ with $\lambda(D) \geq n - 2d + 1$. Let C be a code of length n and minimal distance d. Then $|C| \lesssim |D|$ (here we use the approximate inequality sign to indicate that the inequality holds up to lower order terms). Choosing for D the Hamming balls of different radii with their corresponding parameters leads to a simple proof of the first JPL bound on $R(\delta)$. Ref. [19] posed the natural problem of finding subsets of $\{0,1\}^n$ with the largest possible eigenvalue for their cardinality. This question was answered in [20], where is was shown that Hamming balls of radius $r = \rho n$, $0 < \rho < \frac{1}{2}$ have essentially the largest eigenvalues for their cardinality. This seems to indicate that at least the straightforward version of the approach of [19], as described above, does not lead to an improvement of the first JPL bound. The claim in [20] was derived from a *logarithmic Sobolev inequality* for highly concentrated functions on the boolean cube. We continue with a brief description of relevant notions.

Logarithmic Sobolev inequalities. Viewing both sides of (1) as functions of ϵ, and writing $L(\epsilon)$ for the LHS and $R(\epsilon)$ for the RHS, we have $L(0) = R(0) = \|f\|_2$, and $L(\epsilon) \leq R(\epsilon)$ for $0 \leq \epsilon \leq \frac{1}{2}$. Since both L and R are differentiable in ϵ this implies $L'(0) \leq R'(0)$. This inequality is the logarithmic Sobolev inequality ([3]) for the Hamming cube. We proceed to describe it in more detail. Recall that the Dirichlet form $\mathcal{E}(f,g)$ for functions f and g on the Hamming cube is defined by $\mathcal{E}(f,g) = \mathbb{E}_x \sum_{y \sim x} \left(f(x) - f(y)\right)\left(g(x) - g(y)\right)$. Here $y \sim x$ means that x and y differ in precisely one coordinate. The logarithmic Sobolev inequality then states that $\mathcal{E}(f,f) \geq 2 \ln 2 \cdot Ent(f^2)$. This inequality describes the behavior of the norm on the RHS of the hypercontractive inequality (1) as $\epsilon \to 0$ and, as such, can be viewed as a special case of (1). In point of fact, it was introduced in [3] as a way to prove (1) by (roughly speaking) integrating this inequality over the noise parameter (using the semigroup property of noise operators). Following [3], logarithmic Sobolev inequalities were shown to hold in many spaces of interest (see [21] for discussion and for many applications of these inequalities).

The logarithmic Sobolev inequality for highly concentrated functions in [20] (we will state this inequality explicitly in the discussion following Corollary 2 below) improves over the inequality $\mathcal{E}(f,f) \geq 2 \ln 2 \cdot Ent(f^2)$ similarly to the improvement to (1) provided by Theorem 1. However, deducing a tight hypercontractive inequality, such as Theorem 1, from the inequality in [20] by integration (following the approach of [3]) seems to be more challenging. Roughly speaking, the problem lies in the fact that the concentration of f might decrease very quickly under noise. With that, a family of logarithmic Sobolev inequalities, generalizing that of [20] was proved in [10]. Integrating these inequalities over noise leads to a family of hypercontractive inequalities which improve over (1) for highly concentrated functions and which are essentially tight in the vicinity of $\epsilon = 0$. These inequalities were used in [10] to prove a version of the *uncertainty principle* on $\{0,1\}^n$.

An uncertainty principle on $\{0,1\}^n$. We recall some basic notions in Fourier analysis on the Hamming cube (see [4]). For $\alpha \in \{0,1\}^n$, define the Walsh-Fourier character W_α on $\{0,1\}^n$ by setting $W_\alpha(y) = (-1)^{\sum \alpha_i y_i}$, for all $y \in \{0,1\}^n$. The *weight* of the character W_α is the Hamming weight $|\alpha|$ of α. The characters $\{W_\alpha\}_{\alpha \in \{0,1\}^n}$ form an orthonormal basis in the space of real-valued functions on $\{0,1\}^n$, under the inner product $\langle f, g \rangle =$

$\frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)g(x)$. The expansion $f = \sum_{\alpha \in \{0,1\}^n} \widehat{f}(\alpha) W_\alpha$ defines the Fourier transform $\widehat{f}$ of f. We also have the Parseval identity, $\|f\|_2^2 = \sum_{\alpha \in \{0,1\}^n} \widehat{f}^2(\alpha)$.

Uncertainty principle asserts that a function and its Fourier transform cannot be simultaneously narrowly concentrated. A well-known way (see, e.g., [22]) to state this for the Hamming cube is as follows. If f is a non-zero function on $\{0,1\}^n$ then $|supp(f)| \geq \frac{2^n}{|supp(\widehat{f})|}$. In [10], (see also the discussion following Theorem 1.10 in [9]) a different way to formalize this statement for the Hamming cube was presented. If f is a function on $\{0,1\}^n$ with $\frac{Ent_2\left(\frac{f}{\|f\|_1}\right)}{n} \geq 1 - H(\rho)$, then its Fourier transform $\widehat{f}$ cannot attain its ℓ_2 norm in a Hamming ball of radius much smaller than $\left(\frac{1}{2} - \sqrt{\rho(1-\rho)}\right) \cdot n$. This result was then used to establish some properties of binary linear codes.

Our results. We now pass to presenting our results which are relevant to the topics above. We first remark that the idea of using hypercontractivity to study binary codes was discussed already in [23]. In [24], the hypercontractive inequality (1) was used to obtain bounds on the distance components and other parameters of binary codes. We observe (a similar observation was made in [9]) that these bounds can be strengthened by replacing (1) by (stronger) inequalities of Theorem 1. We do not go into details.

Next, we consider some implications of Theorem 1, focussing on the behavior of the norm $\kappa = \kappa_{2,2}$ for values of the noise parameter ϵ in the vicinity of 0. Clearly, for any $0 \leq x \leq 1$ the function $\kappa_{2,2}(x, \epsilon)$ is 2 at $\epsilon = 0$. We prove the following technical claim.

Lemma 1. *Assume $0 < x < 1$. Let $\kappa(\epsilon) = \kappa_{2,2}(x, \epsilon)$.*

1.

$$\kappa'(0) = \frac{4}{\ln 2} \cdot \frac{\left(2\sqrt{H^{-1}(1-x)(1 - H^{-1}(1-x))} - 1\right)}{x}.$$

2. *Let $\epsilon \sim 0$ express the fact that ϵ is a sufficiently small absolute constant. Then for $\epsilon \sim 0$ holds $|\kappa'(\epsilon) - \kappa'(0)| \leq O(\epsilon)$, where the asymptotic notation hides absolute constants which may depend on x.*

We use the first part of this claim to rederive a slightly weaker (but sufficient for applications, see the dicussion following Corollary 4) version of the logarithmic Sobolev inequality from [20].

Corollary 2. *For any function f on $\{0,1\}^n$ holds*

$$\mathcal{E}(f,f) \geq \ell\left(\frac{Ent_2\left(\frac{f}{\|f\|_1}\right)}{n}\right) \cdot Ent\left(f^2\right),$$

where $\ell(x) = 2 \cdot \frac{1 - 2\sqrt{H^{-1}(1-x)\left(1 - H^{-1}(1-x)\right)}}{x}$ is a convex and increasing function on $[0,1]$, taking $[0,1]$ onto $[2\ln 2, 2]$.

We remark that in [20] (see also Theorem 6 in [10]) a somewhat stronger logarithmic Sobolev inequality $\mathcal{E}(f,f) \geq \ell\left(\frac{Ent\left(\frac{f^2}{\|f\|_2^2}\right)}{n}\right) \cdot Ent\left(f^2\right)$ was shown using a different approach. (It seems that it might be possible to recover this stronger inequality by differentiating a corresponding hypercontractive inequality at zero, if one considers a more general version of Theorem 1 which takes into account the ratio between ℓ_q and ℓ_p norms of f, for $q > p$, and in this case taking both q and p to be very close to 2. We omit the details.) Next, we use the second part of Lemma 1 to rederive the uncertainty principle from [10].

Corollary 3. *Let f be a non-zero function on $\{0,1\}^n$ such that $\frac{Ent_2\left(\frac{f}{\|f\|_1}\right)}{n} = 1 - H(\rho)$, for some $0 \le \rho < 1$. Let $0 \le \mu < \frac{1}{2} - \sqrt{\rho(1-\rho)}$. Then*

$$\sum_{|\alpha| \le \mu n} \widehat{f}^2(\alpha) \ \le \ 2^{-cn} \cdot \sum_{\alpha} \widehat{f}^2(\alpha),$$

where c is an absolute constant depending on ρ and μ.

Let us remark that it seems helpful to have an explicit hypercontractive inequality (given by Theorem 1) from which both Corollaries 2 and 3 can be derived as special cases.

The following two results are simple consequences of Corollaries 2 and 3, respectively. Recall that for a subset $D \subseteq \{0,1\}^n$, $\lambda(D)$ is the maximal eigenvalue of the adjacency matrix of the subgraph of the discrete cube induced by the vertices of D. Recall also that $R_{LP}(\delta)$ denotes the best possible upper bound on the asymptotic maximal rate $R(\delta)$ of a code with relative distance δ which is possible to obtain using the linear programming approach of [15].

Corollary 4.

- *Let D be a subset of $\{0,1\}^n$ of cardinality $|D| = 2^{H(\rho)n}$, for some $0 \le \rho \le 1$. Then*

$$\lambda(D) \ \le \ 2\sqrt{\rho(1-\rho)} \cdot n.$$

 This is almost tight if D is a Hamming ball of exponentially small cardinality.
- *For any $0 \le \delta \le \frac{1}{2}$ holds*

$$R_{LP}(\delta) \ \ge \ \frac{1 - H(\delta) + H\left(1/2 - \sqrt{\delta(1-\delta)}\right)}{2}.$$

Some comments.

– As discussed above, the first of the these claims answers the question of [19] and shows that a certain approach to bound binary codes does not lead to an improvement of the first JPL bound. The second claim shows that the best possible bound obtainable via the linear programming approach of [15] is not better than the arithmetic average of the Gilbert-Varshamov bound and the first JPL bound. Observe that the first claim is a consequence of the logarithmic Sobolev inequality in Corollary 2, and hence of the behavior of the norm $\kappa_{2,2}$ in Theorem 1 as $\epsilon \to 0$. The second claim is a consequence of the uncertainty principle in Corollary 3, and hence of the behavior of the norm $\kappa_{2,2}$ in Theorem 1 as $\epsilon \sim 0$. We find these connections between notions to be rather intriguing.
– As we have mentioned, the first of the claims recovers a result of [20], where it was also derived from the appropriate logarithmic Sobolev inequality. (Apart from this claim being a simple corollary of Theorem 1, an additional reason for stating it here is that it has only appeared in the unpublished arXiv preprint [20].) The second claim of recovers a result of [17].

Finally we present a result of a somewhat different nature. The question of the maximal possible ratio $\frac{\|f\|_2}{\|f\|_1}$ for a polynomial f of degree s on $\{0,1\}^n$ is considered in analysis [25,26] in connection with a conjecture of Pelczynski. The following claim is a simple consequence of Corollary 2.

Corollary 5. *Let $0 \le s \le \frac{n}{2}$ and let f be a polynomial of degree s on $\{0,1\}^n$ (that is, f a restriction of a degree s polynomial on $\mathbb{R}^n$ to $\{0,1\}^n$). Then, writing σ for $\frac{s}{n}$,*

$$\frac{1}{n} \log_2\left(\frac{\|f\|_2}{\|f\|_1}\right) \ \le \ \frac{1 - H\left(\frac{1}{2} - \sqrt{\sigma(1-\sigma)}\right)}{2}.$$

We remark that this improves the estimate of [25] for $0.3.. \leq \frac{s}{n} < \frac{1}{2}$.

1.3. Related Work

In [10], it was shown that if $\frac{\|f\|_p}{\|f\|_1} \geq 2^{\rho n}$, for some $p \geq 1$ and $\rho \geq 0$, then $\|f\|_p \geq \|f_\epsilon\|_{1 + \frac{p-1}{(1-2\epsilon)^2} + \Delta(p, \rho, \epsilon)}$, where $\Delta(p, \rho, \epsilon) > 0$ for all $p > 1, \epsilon, \rho > 0$ (cf. with (1), which can be restated as $\|f\|_p \geq \|f_\epsilon\|_{1 + \frac{p-1}{(1-2\epsilon)^2}}$, for $p = 1 + (1 - 2\epsilon)^2 (q - 1)$). The function $\Delta(p, \rho, \epsilon)$ is "semi-explicit", in the following sense: it is an explicit function of the (unique) solution of a certain explicit differential equation.

In [11], it was shown, using a different approach, that (restating the result in the notation of this paper) $\|f_\epsilon\|_2 \leq \|f\|_q$, where q is determined by $F_{f,\epsilon}(q) = q$ (in the notation of the last comment above to Theorem 1). As we have observed, this is the best possible value for κ in Theorem 1, but it might not be easy to determine explicitly in practice (compare with Corollary 1).

In [27], Mrs. Gerber type inequalities for Rényi divergence and arbitrary distributions on Polish spaces were proved, using a different approach. The results in [27] apply in higher generality, but they seem to be somewhat less explicit than these in Proposition 1.

This paper is organized as follows. We prove Proposition 1 in Section 2 and Theorem 1 in Section 3. We prove the remaining claims, including some technical lemmas and claims made above in the comments to the main results, in Section 4.

2. Proof of Proposition 1

We first prove (4) and then show it to be tight. We prove (4) in two steps, using Theorem 2 to reduce it to a claim about properties of the function ϕ_ϵ, and then proving that claim.

We start with the first step. It follows closely the proof of Theorem 1.8 in [9], and hence will be presented rather briefly, and not in a self-contained manner. Let f be a function on $\{0, 1\}^n$, for which we want to show (4). Recall that, by assumption, $\mathbb{E} f = 1$. This means that $\|f\|_\infty \leq 2^n$, and that the points at which $f < 2^{-n}$, say, contribute little to both sides ot (4), so we may ignore them for the sake of the discussion (that is, we may and will assume that f vanishes on these points). All the remaining points can be partitioned into $O(n)$ level sets $A_1, ... A_r$ such that f varies by a factor of 2 at most in each level set. Let $\alpha_i = \frac{1}{n} \log_2(|A_i|)$, and let $\nu_i = \frac{1}{n} \log_2(v_i)$, where v_i is the minimal value of f on A_i. Then, as shown in the proof of Theorem 1.8 in [9], up to an additive error term of $O\left(\frac{\log(n)}{n}\right)$, we have,

$$\frac{Ent_2(f_\epsilon)}{n} = \frac{1}{n} \log_2 \|f_\epsilon\|_2^2 \leq 2 \cdot \max_{1 \leq i \leq r} \left\{ \phi_\epsilon(\alpha_i) + \nu_i \right\}.$$

The negligible error here contributes towards a negligible error in (4), which can then be removed by a tensorization argument, so we will ignore it from now on.

Let $N = \frac{1}{n} \log_2(\|f\|_q)$. Note that $N = \frac{q-1}{q} \cdot \frac{Ent_q(f)}{n}$. Hence, in particular, $N \leq \frac{q-1}{q}$. Note also that for any $1 \leq i \leq r$ holds $\alpha_i + \nu_i \leq 1$ (since $\mathbb{E} f = 1$) and $\frac{\alpha_i - 1}{q} + \nu_i \leq N$ (since $\frac{|A_i|}{2^n} 2^{q\nu_i n} \leq \frac{1}{2^n} \sum_{x \in A_i} f^q(x) \leq \|f\|_q^q$). We also have $0 \leq \alpha_i \leq 1$ and $-1 \leq \nu_i \leq 1$. This discussion leads to the definition of the following two subsets of $\mathbb{R}^2$, which will play an important role in the proof of Theorem 1 as well. (We remark that the relevance of the set Ω in the following definition is not immediately obvious. It will be made clear in the following arguments.)

Definition 2. *Let $q > 1$ and $0 < N \leq \frac{q-1}{q}$. Let $\Omega_0 \subseteq \mathbb{R}^2$ be defined by*

$$\Omega_0 = \left\{ (\alpha, \nu) : 0 \leq \alpha \leq 1, \ -1 \leq \nu \leq 1, \ \alpha + \nu \leq 1, \ \frac{\alpha - 1}{q} + \nu \leq N \right\}.$$

Let $\Omega \subseteq \Omega_0$ be the set of all pairs $(\alpha, \nu) \in \Omega_0$ with $\nu \geq 0$.

By the preceding discussion, (4) will follow from the following claim.

Lemma 2. *For all $0 \leq \epsilon \leq \frac{1}{2}$ holds*

$$\max_{(\alpha,\nu)\in\Omega_0} \left\{ \phi_\epsilon(\alpha) + \nu \right\} = \frac{1}{2} \cdot \psi_{2,q}\left(\frac{qN}{q-1}, \epsilon \right),$$

where $\psi_{2,q}$ is defined in Definition 1.

Before proving Lemma 2, we collect the relevant properties of the function ϕ_ϵ in the following lemma.

Lemma 3. *Let $0 < \epsilon < \frac{1}{2}$. Let $q_0 = q_0(\epsilon) = 1 + (1-2\epsilon)^2$. The function ϕ_ϵ has the following properties.*

1. *$\phi_\epsilon(\alpha)$ is strictly concave and increasing from $\phi_\epsilon(0) = -\frac{\log_2\left(\frac{4}{q_0}\right)}{2}$ to 0 on $0 \leq \alpha \leq 1$.*
2. *$\phi_\epsilon'(0) = 1$, $\phi_\epsilon'(1) = \frac{1}{q_0}$.*
3. *$\frac{\alpha-1}{\phi_\epsilon(\alpha)}$ is strictly increasing in α, going up to q_0, as $\alpha \to 1$.*
4. *The function $g(\alpha) = 1 - \alpha - \frac{\alpha}{1-\alpha} \cdot \phi_\epsilon(\alpha)$ is strictly decreasing on $[0,1]$. Moreover, $g(0) = 1$ and $g(1) = \frac{1}{q_0}$.*

This lemma will be proved in Section 4. For now we assume its correctness, and proceed with the proof of Lemma 2.

Proof. Our first observation is that the maximum of $\phi_\epsilon(\alpha) + \nu$ on Ω_0 is located in Ω, since for any point $(\alpha, \nu) \in \Omega_0$ with $\nu < 0$, the point $(\alpha, 0)$ is in Ω. So we may and will replace Ω_0 with Ω in the following argument.

Since ϕ_ϵ is increasing, any local maximum of $\phi_\epsilon(\alpha) + \nu$ is located on the upper boundary of Ω, that is on the piecewise linear curve which starts as the straight line $\frac{\alpha}{q} + \nu = N + \frac{1}{q}$, for $0 \leq \alpha \leq 1 - \frac{qN}{q-1}$ and continues as the straight line $\alpha + \nu = 1$ for $1 - \frac{qN}{q-1} \leq \alpha \leq 1$.

Note that, since $\phi_\epsilon' < 1$ for $\alpha > 0$, the function $\phi_\epsilon(\alpha) + \nu$ decreases (as a function of α) on the line $\alpha + \nu = 1$ for $1 - \frac{qN}{q-1} \leq \alpha \leq 1$. Next, let $h(\alpha) = \phi_\epsilon(\alpha) - \frac{\alpha}{q} + \left(N + \frac{1}{q}\right)$. The function h describes the restriction of $\phi_\epsilon(\alpha) + \nu$ to the line $\frac{\alpha}{q} + \nu = N + \frac{1}{q}$, and we are interested on the maximum of h on the interval $I = \left\{ 0 \leq \alpha \leq 1 - \frac{qN}{q-1} \right\}$. We have $h'(\alpha) = \phi_\epsilon'(\alpha) - \frac{1}{q}$. By Lemma 3, the function h is concave, and hence there are two possible cases:

-   $\phi_\epsilon'\left(1 - \frac{qN}{q-1}\right) \geq \frac{1}{q}$. In this case h is increasing on I and we get

$$\max_{(\alpha,\nu)\in\Omega} \left\{ \phi_\epsilon(\alpha) + \nu \right\} = \max_{\alpha \in I} \{h(\alpha)\} = h\left(1 - \frac{qN}{q-1} \right) =$$

$$\phi_\epsilon\left(1 - \frac{qN}{q-1} \right) + \frac{qN}{q-1} = \frac{1}{2} \cdot \psi_{2,q}\left(\frac{qN}{q-1}, \epsilon \right).$$

The last equality follows from the definition of $\psi_{2,q}$ in this case.

- $\phi'_\epsilon\left(1 - \frac{qN}{q-1}\right) < \frac{1}{q}$. Note that, by Lemma 3, $1 = \phi'_\epsilon(0) > \frac{1}{q}$. Hence, in this case the maximum of h on I is located at the unique zero of its derivative, that is at the point α_0 such that $\phi'_\epsilon(\alpha_0) = \frac{1}{q}$. Using the definition of $\psi_{2,q}$ in this case, we get

$$\max_{(\alpha,\nu)\in\Omega} \left\{\phi_\epsilon(\alpha) + \nu\right\} = h(\alpha_0) = N + \left(\phi_\epsilon(\alpha_0) + \frac{1 - \alpha_0}{q}\right) = \frac{1}{2} \cdot \psi_{2,q}\left(\frac{qN}{q-1}, \epsilon\right).$$

□

This concludes the proof of (4). The fact that $\psi_{2,q}(x, \epsilon)$ is strictly increasing and concave in its first argument is an easy implication of Lemma 3.

We pass to showing the tightness of (4). Let $0 < \epsilon < \frac{1}{2}$ and $0 < x < 1$. Set $N = \frac{q-1}{q} \cdot x$. Let Ω be the domain defined in Definition 2, and let (α^*, ν^*) be the maximum point of $\phi_\epsilon(\alpha) + \nu$ on Ω (note that the discussion above determines this point uniquely). We proceed to define the function f. Let n be sufficiently large. For $y \in \{0,1\}^n$, let $|y|$ denotes the Hamming weight of y, that is the number of 1-coordinates in y. Let $r = \lfloor H^{-1}(\alpha^*) \cdot n \rfloor$. Let $S = \{y \in \{0,1\}^n, |y| = r\}$ be the Hamming sphere around zero of radius r in $\{0,1\}^n$. Now there are two cases to consider.

- If $\phi'_\epsilon(1 - x) < \frac{1}{q}$, then by the discussion above, the point (α^*, ν^*) lies on the line $\frac{\alpha}{q} + \nu = N + \frac{1}{q}$, but not on the line $\alpha + \nu = 1$. Observe that $2^{\alpha^* n - o(n)} \leq |S| \leq 2^{\alpha^* n}$ (the first estimate follows from the Stirling formula, for the second estimate see, e.g., Theorem 1.4.5. in [28]). As the first attempt, let $g = 2^{\nu^* n} \cdot 1_S$. Then $N - o(n) \leq \frac{\alpha^* - 1}{q} + \nu^* - o(n) \leq \frac{1}{n} \log_2 \|g\|_q \leq \frac{\alpha^* - 1}{q} + \nu^* = N$. That is, $x - o_n(1) \leq \frac{Ent_q(g)}{n} \leq x$. However, $\mathbb{E}\, g$ is exponentially small. To correct that, we define f to be $v = 2^{(\nu^* - \delta) \cdot n}$ on S, and $\frac{2^n - |S|v}{2^n - |S|}$ on the complement of S. Then $\mathbb{E}\, f = 1$. We choose δ to be as small as possible, while ensuring that $\frac{Ent_q(f)}{n} \leq x$. Since the contribution of the constant-1 function to $\|f\|_q$ is exponentially small w.r.t. $\|f\|_q$, we can choose $\delta = o_n(1)$. We now have $\mathbb{E}\, f = 1$, $\frac{Ent_q(f)}{n} \leq x$, and

$$\frac{Ent_2(f_\epsilon)}{n} = \frac{1}{n} \log_2 \|f_\epsilon\|_2^2 = \frac{1}{n} \log_2 \langle f_{2\epsilon(1-\epsilon)}, f \rangle \geq$$

$$2 \cdot (\phi_\epsilon(\alpha^*) + \nu^*) - o_n(1) \geq \psi_{2,q}(x, \epsilon) - o_n(1).$$

Here the second equality follows from the semigroup property of the noise operator: $T_\epsilon \circ T_\epsilon = T_{2\epsilon(1-\epsilon)}$. The first inequality follows from the tightness part of Theorem 2 and the definition of ϕ_ϵ. The second inequality follows from Lemma 2.
The tightness of (4) in this case now follows, taking into account the fact that $\psi_{2,q}$ is strictly increasing.

- If $\phi'_\epsilon(1 - x) \geq \frac{1}{q}$, the point (α^*, ν^*) lies on the intersection of the lines $\frac{\alpha}{q} + \nu = N + \frac{1}{q}$, and $\alpha + \nu = 1$. Hence the function $g = 2^{\nu^* n} \cdot 1_S$ has both $x - o_n(1) \leq \frac{Ent_q(g)}{n} \leq x$, and $1 - o_n(1) \leq \mathbb{E}\, g \leq 1$. It is easy to see that g can be corrected as in the preceding case, by decreasing it slightly on S and adding a constant component, to obtain a function f with expectation 1 and $Ent_q(f) \leq x$, and with $\frac{Ent_2(f_\epsilon)}{n} \geq \psi_{2,q}(x, \epsilon) - o_n(1)$, proving the tightness of (4) in this case as well. We omit the details.

This completes the proof of Proposition 1. □

3. Proof of Theorem 1

The high-level outline of the argument in this proof is similar to that of Proposition 1. We start with proving (5), doing this in two steps. In the first step Theorem 2 is used to reduce (5) to a claim about properties of the function ϕ_ϵ. That claim is proved in the second step.

We will give only a brief description of the first step since, similarly to the first step in the proof of Proposition 1, it follows closely the proof of Theorem 1.8 in [9]. Let f be a function on $\{0,1\}^n$, for which we may and will assume that $f \geq 2^{-n}$ and that $\mathbb{E} f = \|f\|_1 = 1$. There are $O(n)$ real numbers $0 \leq \alpha_1, ..., \alpha_r \leq 1$ and $-1 \leq \nu_1, ..., \nu_r \leq 1$, such that, up to a negligible error, which may be removed by tensorization, we have

$$\frac{1}{n} \log_2 \|f_\epsilon\|_2 \leq \max_{1 \leq i \leq r} \left\{ \phi_\epsilon(\alpha_i) + \nu_i \right\} \quad \text{and} \quad \frac{1}{n} \log_2 \|f\|_q = \max_{1 \leq i \leq r} \left\{ \frac{\alpha_i - 1}{q} + \nu_i \right\}.$$

Hence (5) reduces to claim (6) in the following proposition.

Proposition 2. *Let $q > 1$ and $0 \leq \alpha_1, ..., \alpha_r \leq 1$, $-1 \leq \nu_1, ..., \nu_r \leq 1$ with $\max_{1 \leq i \leq r} \left\{ (\alpha_i - 1) + \nu_i \right\} = 0$. Let $N = \max_{1 \leq i \leq r} \left\{ \frac{\alpha_i - 1}{q} + \nu_i \right\}$. Then for any $0 \leq \epsilon \leq \frac{1}{2}$ holds*

$$\max_{1 \leq i \leq r} \left\{ \phi_\epsilon(\alpha_i) + \nu_i \right\} \leq \max_{1 \leq i \leq r} \left\{ \frac{\alpha_i - 1}{\kappa} + \nu_i \right\}, \tag{6}$$

where $\kappa = \kappa_{2,q}\left(\frac{qN}{q-1}, \epsilon \right)$ is defined in Definition 1 (it is easy to see that $0 \leq N \leq \frac{q-1}{q}$, and hence κ is well defined).

Moreover, this is tight, in the following sense. For any $0 < N < \frac{q-1}{q}$ and $0 < \epsilon < \frac{1}{2}$, and for any $\tilde{\kappa} < \kappa_{2,q}(x, \epsilon)$, there exist $0 \leq \alpha_1, \alpha_2 \leq 1$ and $-1 \leq \nu_1, \nu_2 \leq 1$ such that $\max_{1 \leq i \leq 2} \left\{ (\alpha_i - 1) + \nu_i \right\} = 0$, $\max_{1 \leq i \leq 2} \left\{ \frac{\alpha_i - 1}{q} + \nu_i \right\} = N$, and $\max_{1 \leq i \leq 2} \left\{ \phi_\epsilon(\alpha_i) + \nu_i \right\} > \max_{1 \leq i \leq r} \left\{ \frac{\alpha_i - 1}{\tilde{\kappa}} + \nu_i \right\}$.

Proof of Proposition 2. We start with verifying simple boundary cases. First, we observe that $\phi_0(x) = \frac{x-1}{2}$ (Lemma 9) and that $\phi_{\frac{1}{2}}(x) = x - 1$ (see the relevant discussion in the proof of Corollary 1). In addition, it is easy to see that $\kappa_{2,q}\left(x, \frac{1}{2} \right) = 1$ for all $q \geq 1$ and $0 \leq x \leq 1$; and (bearing in mind that $\phi_0(x) = \frac{x-1}{2}$) that $\kappa_{2,q}(x, 0) = 2$ for all $q \geq 1$ and $0 \leq x \leq 1$. Therefore (6) is an identity for $\epsilon = 0$ and $\epsilon = \frac{1}{2}$. Hence we may and will assume from now on that $0 < \epsilon < \frac{1}{2}$.

Let $q_0 = 1 + (1 - 2\epsilon)^2$. We proceed to consider the (simple) cases $N = 0$ or $N + \frac{1}{q} \leq \frac{1}{q_0}$. Note that in these cases we have $\kappa = \kappa_{2,q}\left(\frac{qN}{q-1}, \epsilon \right) = q_0$. Next, observe that, by the first and the second claims of Lemma 3, for any $0 \leq \alpha \leq 1$ holds $\phi_\epsilon(\alpha) \leq \frac{\alpha-1}{q_0} = \frac{\alpha-1}{\kappa}$ and hence (6) holds trivially in these cases.

We continue to prove (6), assuming from now on that $N > 0$ and that $N + \frac{1}{q} > \frac{1}{q_0}$. Let $\Omega \subseteq \mathbb{R}^2$ be the set defined in Definition 2. We now define a family of continuous functions on Ω, which will play an important role in the following argument. Let (α_1, ν_1) be a point in Ω with $\frac{\alpha_1 - 1}{q} + \nu_1 = N$. Define a function $f = f_{\alpha_1, \nu_1}$ on Ω as follows. For $(\alpha, \nu) \in \Omega$ with $\alpha < 1$ let $f(\alpha, \nu)$ be the value of κ for which $\phi_\epsilon(\alpha) + \nu = \max\left\{ \frac{\alpha_1 - 1}{\kappa} + \nu_1, \frac{\alpha-1}{\kappa} + \nu \right\}$. In addition, let $f(1, 0) = \frac{1-\alpha_1}{\nu_1}$.

Lemma 4. *For any choice of (α_1, ν_1) as above the function f_{α_1, ν_1} is well-defined and continuous on Ω.*

Let $M(\alpha_1, \nu_1) = \max_\Omega f_{\alpha_1, \nu_1}$. The inequality (6) will follow from the next main technical claim, describing the behavior of $M(\alpha_1, \nu_1)$, as a function of α_1 and ν_1. Before stating this claim, let us make some preliminary comments. Note that the points $\left(1 - \frac{qN}{q-1}, \frac{qN}{q-1} \right)$

and $\left(0, N + \frac{1}{q}\right)$ are possible choices for (α_1, ν_1). Note also that α_0 in the third part of the claim is well-defined, by the fourth claim of Lemma 3.

Proposition 3.

1.

$$M\left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right) = \frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)}.$$

2. If $\dfrac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)} \geq q$, then for any choice of (α_1, ν_1) holds

$$M(\alpha_1, \nu_1) \leq M\left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right).$$

3. If $\dfrac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)} \leq q$, then for any choice of (α_1, ν_1) holds

$$M\left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right) \leq M(\alpha_1, \nu_1) \leq M\left(0, N + \frac{1}{q}\right) = \frac{\alpha_0 - 1}{\phi_\epsilon(\alpha_0)},$$

where α_0 is determined by $1 - \alpha_0 - \dfrac{\alpha_0 \phi_\epsilon(\alpha_0)}{1 - \alpha_0} = N + \dfrac{1}{q}$.

We will prove Lemma 4 and Proposition 3 in Sections 3.1 and 3.2. For now we assume their validity and complete the proof of Proposition 2.

We first prove (6). Note that if $x = \frac{qN}{q-1}$ then in the definition of $\kappa_{2,q}(x, \epsilon)$ we have $y = \frac{q-1}{q} \cdot x + \frac{1}{q} = N + \frac{1}{q}$. Recall also that we may assume that $N > 0$ and that $y = N + \frac{1}{q} > \frac{1}{q_0}$.

By assumption $\alpha_i + \nu_i \leq 1$, and $\frac{\alpha_i - 1}{q} + \nu_i \leq N$ for all $1 \leq i \leq r$. Moreover there is an index $1 \leq i \leq r$ for which $\frac{\alpha_i - 1}{q} + \nu_i = N$. Assume, w.l.o.g., that $i = 1$. We apply Proposition 3 to the function f_{α_1, ν_1}. Observe that the claim of the proposition together with the definition of κ imply $M(\alpha_1, \nu_1) \leq \kappa$. By the definition of f_{α_1, ν_1}, this means that for any point $(\alpha, \nu) \in \Omega$ holds $\phi_\epsilon(\alpha) + \nu \leq \max\left\{\frac{\alpha_1 - 1}{\kappa} + \nu_1, \frac{\alpha - 1}{\kappa} + \nu\right\}$. We now claim that this inequality holds for all the points (α_i, ν_i), $1 \leq i \leq r$, which will immediately imply (6). In fact, points (α_i, ν_i) with $0 \leq \nu_i \leq 1$ lie in Ω and hence the inequality holds for these points. Furthermore, if $\nu_i < 0$ for some $1 \leq i \leq r$, then the point $(\alpha_i, 0)$ lies in Ω, and hence $\phi_\epsilon(\alpha_i) \leq \max\left\{\frac{\alpha_1 - 1}{q} + \nu_1, \frac{\alpha_i - 1}{q}\right\}$. However, then $\phi_\epsilon(\alpha_i) + \nu_i \leq \max\left\{\frac{\alpha_1 - 1}{q} + \nu_1, \frac{\alpha_i - 1}{q} + \nu_i\right\}$, proving the inequality in this case as well.

We pass to proving the tightness of (6), starting with the case $N + \frac{1}{q} \leq \frac{1}{q_0}$. In this case, by definition, $\kappa = q_0$. Let $\tilde{\kappa} < \kappa$ be given. Observe that since, by assumption, $N > 0$, we have $q > q_0$. Set $\alpha_1 = \frac{\frac{1}{q_0} - \frac{1}{q} - N}{\frac{1}{q_0} - \frac{1}{q}}$. Set $\nu_1 = \frac{1 - \alpha_1}{q_0}$. Let $\delta > 0$ be sufficiently small (depending on N and $\tilde{\kappa}$). Set $\alpha_2 = 1 - \delta$ and $\nu_2 = \delta$. It is easy to see that α_1, α_2 and ν_1, ν_2 satisfy the required constraints. We claim that $\phi_\epsilon(\alpha_2) + \nu_2 > \max_{1 \leq i \leq 2}\left\{\frac{\alpha_i - 1}{\tilde{\kappa}} + \nu_i\right\}$. In fact, for a sufficiently small δ we have, using the second claim of Lemma 3 (and observing that ϕ'_ϵ is continuous), that

$$\phi_\epsilon(\alpha_2) + \nu_2 = \phi_\epsilon(1 - \delta) + \delta \approx -\frac{\delta}{q_0} + \delta > -\frac{\delta}{\tilde{\kappa}} + \delta = \frac{\alpha_2 - 1}{\tilde{\kappa}} + \nu_2,$$

and

$$\phi_\epsilon(\alpha_2) + \nu_2 \approx -\frac{\delta}{q_0} + \delta > 0 \geq \frac{\alpha_1 - 1}{\tilde{\kappa}} + \frac{1 - \alpha_1}{q_0} = \frac{\alpha_1 - 1}{\tilde{\kappa}} + \nu_1.$$

We pass to the case $N + \frac{1}{q} > \frac{1}{q_0}$ and $\frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1-\frac{qN}{q-1}\right)} \geq q$. In this case $\kappa = \frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1-\frac{qN}{q-1}\right)}$. Set $\alpha_1 = \alpha_2 = 1 - \frac{qN}{q-1}$ and $\nu_1 = \nu_2 = \frac{qN}{q-1}$. It is easy to see that α_1, α_2 and ν_1, ν_2 satisfy the required constraints. It is also easy to see that for any $\tilde{\kappa} < \kappa$ holds

$$\phi_\epsilon(\alpha_1) + \nu_1 = \frac{\alpha_1 - 1}{\kappa} + \nu_1 > \frac{\alpha_1 - 1}{\tilde{\kappa}} + \nu_1.$$

It remains to deal with the case $N + \frac{1}{q} > \frac{1}{q_0}$ and $\frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1-\frac{qN}{q-1}\right)} < q$. Let α_0 be determined by $1 - \alpha_0 - \frac{\alpha_0\phi_\epsilon(\alpha_0)}{1-\alpha_0} = N + \frac{1}{q}$. Then $\kappa = \frac{\alpha_0-1}{\phi_\epsilon(\alpha_0)}$. Set $\alpha_1 = 0$ and $\nu_1 = N + \frac{1}{q}$. Set $\alpha_2 = \alpha_0$ and $\nu_2 = 1 - \alpha_0$. It is easy to see that in this case the function $1 - \alpha - \frac{\alpha\phi_\epsilon(\alpha)}{1-\alpha}$ is larger than $N + \frac{1}{q}$ at $\alpha = 1 - \frac{qN}{q-1}$, and hence the fourth claim of Lemma 3 implies that $\alpha_2 = \alpha_0 > 1 - \frac{qN}{q-1}$. Using this, it is easy to see that α_1, α_2 and ν_1, ν_2 satisfy the required constraints. Furthermore, note that $\alpha_2 < 1$ (again, using the fourth claim of Lemma 3). It is also easy to verify, using the definition of α_0, that

$$\phi_\epsilon(\alpha_2) + \nu_2 = \frac{\alpha_1 - 1}{\kappa} + \nu_1 = \frac{\alpha_2 - 1}{\kappa} + \nu_2,$$

which implies that for any $\tilde{\kappa} < \kappa$ holds $\phi_\epsilon(\alpha_2) + \nu_2 > \max_{1 \leq i \leq 2}\left\{\frac{\alpha_i-1}{\tilde{\kappa}} + \nu_i\right\}$. This completes the proof of Proposition 2. □

We now prove Lemma 4 and Proposition 3. Recall that we may assume $N > 0$ and $N + \frac{1}{q} > \frac{1}{q_0}$.

3.1. Proof of Lemma 4

Let (α_1, ν_1) be a point in Ω with $\frac{\alpha_1-1}{q} + \nu_1 = N$. We start with some simple but useful observations about α_1 and ν_1.

Lemma 5.
1. $\alpha_1 \leq 1 - \frac{qN}{q-1}$ and $\nu_1 \geq \frac{qN}{q-1}$.
2. $\frac{1-\alpha_1}{\nu_1} < q_0$.

Proof. The first claim of the lemma is an easy consequence of the inequalities $\frac{\alpha_1-1}{q} + \nu_1 = N$ and $\alpha_1 + \nu_1 \leq 1$. We omit the details.

We pass to the second claim of the lemma, distinguishing two cases, $q \leq q_0$ and $q > q_0$. If $q \leq q_0$, then $\nu_1 = N + \frac{1-\alpha_1}{q} > \frac{1-\alpha_1}{q} \geq \frac{1-\alpha_1}{q_0}$. If $q > q_0$, we use the fact that $N + \frac{1}{q} > \frac{1}{q_0}$ to obtain $\frac{1-\alpha_1}{q_0} < (1-\alpha_1)\left(N + \frac{1}{q}\right) = (1-\alpha_1)\left(\frac{\alpha_1}{q} + \nu_1\right)$. Viewing the last expression as a function of α_1, it is easy to see that it equals ν_1 at $\alpha_1 = 0$ and that it decreases in α_1. Hence $\nu_1 \geq (1-\alpha_1)\left(\frac{\alpha_1}{q} + \nu_1\right) > \frac{1-\alpha_1}{q_0}$, completing the argument in this case as well. □

We now show that the function $f = f_{\alpha_1,\nu_1}$ is well-defined and that its values lie in the interval $(0, q_0)$. By Lemma 5, $\alpha_1 < 1$ and $0 < f(1,0) = \frac{1-\alpha_1}{\nu_1} < q_0$. Let now $\alpha < 1$. In this case the function $g(\kappa) = \max\left\{\frac{\alpha_1-1}{\kappa} + \nu_1, \frac{\alpha-1}{\kappa} + \nu\right\}$ is a strictly increasing continuous function of κ, which is $-\infty$ at $\kappa = 0$. Furthermore, by Lemma 3, $\phi_\epsilon(\alpha) < \frac{\alpha-1}{q_0}$, implying that $g(q_0) > \phi_\epsilon(\alpha) + \nu$. Hence, by the intermediate value theorem, there exists a unique $0 < \kappa < q_0$ for which $\phi_\epsilon(\alpha) + \nu = \max\left\{\frac{\alpha_1-1}{\kappa} + \nu_1, \frac{\alpha-1}{\kappa} + \nu\right\}$.

Next, we argue that f is continuous on Ω. Let $(\alpha, \nu) \in \Omega$. If $\alpha < 1$, then there exists a compact neighborhood of (α, ν) in which both one-sided derivatives of $g(\kappa)$ are positive

and bounded. This, together with the fact that $\phi_\epsilon(\alpha) + \nu$ is continuous, implies that f is continuous at (α, ν).

It remains to argue that f is continuous at $(1,0)$. Let O be a sufficiently small neighbourhood of $(1,0)$ in Ω. Let $(\alpha, \nu) \in O$, with $\alpha < 1$. Then $\phi_\epsilon(\alpha) + \nu$ is close to $\phi_\epsilon(1) + 0 = 0$. We would like to claim that $f(\alpha, \nu)$ is close to $f(1,0) = \frac{1 - \alpha_1}{\nu_1}$. In fact, assume towards contradiction that $f(\alpha, \nu)$ is significantly larger than $\frac{1 - \alpha_1}{\nu_1}$. In this case $\phi_\epsilon(\alpha) + \nu = \max\left\{ \frac{\alpha_1 - 1}{f(\alpha, \nu)} + \nu_1, \frac{\alpha - 1}{f(\alpha, \nu)} + \nu \right\} \geq \frac{\alpha_1 - 1}{f(\alpha, \nu)} + \nu_1$ is significantly larger than 0 (taking into account that $\alpha_1 < 1$), reaching contradiction. On the other hand, assume that $f(\alpha, \nu)$ is significantly smaller than $\frac{1 - \alpha_1}{\nu_1}$, and hence significantly smaller than q_0 (by the second claim of Lemma 5). Recall that $\phi_\epsilon(1) = 0$ and that $\phi_\epsilon'(1) = \frac{1}{q_0}$. Hence $\phi_\epsilon(\alpha) = \frac{\alpha - 1}{q_0} + O\left((1 - \alpha)^2\right) > \frac{\alpha - 1}{f(\alpha, \nu)}$. This means that $\phi_\epsilon(\alpha) + \nu = \frac{\alpha_1 - 1}{f(\alpha, \nu)} + \nu_1$, which is significantly smaller than 0, again reaching contradiction. This completes the proof of Lemma 4.

We collect some useful properties of $f = f_{\alpha_1, \nu_1}$ in the following claim.

Corollary 6. *1.   For any $(\alpha, \nu) \in \Omega$ holds $\phi_\epsilon(\alpha) + \nu = \max\left\{ \frac{\alpha_1 - 1}{f(\alpha, \nu)} + \nu_1, \frac{\alpha - 1}{f(\alpha, \nu)} + \nu \right\}$.*

2. $0 < f \leq M(\alpha_1, \nu_1) < q_0$ on Ω.

3.   For any $(\alpha, \nu) \in \Omega$ holds $f(\alpha, \nu) \leq \frac{\alpha - 1}{\phi_\epsilon(\alpha)}$. (If $\alpha = 1$ we replace the RHS of this inequality with q_0.)

Proof. The first two claims follow immediately from the preceding discussion and from the continuity of f. For the third claim, recall that

$$\phi_\epsilon(\alpha) + \nu = \max\left\{ \frac{\alpha_1 - 1}{f(\alpha, \nu)} + \nu_1, \frac{\alpha - 1}{f(\alpha, \nu)} + \nu \right\} \geq \frac{\alpha - 1}{f(\alpha, \nu)} + \nu$$

$\square$

3.2. Proof of Proposition 3

Let (α_1, ν_1) be given, let $f = f_{\alpha_1, \nu_1}$, and let $M = M(\alpha_1, \nu_1) = \max_\Omega f$. Let (α^*, ν^*) be a maximum point of f. Then $f(\alpha^*, \nu^*) = M$ and hence $\phi_\epsilon(\alpha^*) + \nu^* = \max\left\{ \frac{\alpha_1 - 1}{M} + \nu_1, \frac{\alpha^* - 1}{M} + \nu \right\}$. Clearly either $\frac{\alpha_1 - 1}{M} + \nu_1 \neq \frac{\alpha^* - 1}{M} + \nu^*$ or $\frac{\alpha_1 - 1}{M} + \nu_1 = \frac{\alpha^* - 1}{M} + \nu^*$. In the first case we say that (α^*, ν^*) is a maximum point of the *first type*, and otherwise it is a maximum point of the *second type*.

The following two claims constitute the main steps of the proof of Proposition 3. They describe the respective behavior of maxima points of the first and the second type.

Lemma 6. *Let (α^*, ν^*) be a maximum point of f of the first type. Then the following two claims hold.*

- $\frac{\alpha_1 - 1}{f(\alpha^*, \nu^*)} + \nu_1 > \frac{\alpha^* - 1}{f(\alpha^*, \nu^*)} + \nu^*$.

- $\alpha^* \leq 1 - \frac{qN}{q-1}$.

Lemma 7. *If $(\alpha_1, \nu_1) = \left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right)$, then $\left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right)$ is the unique maximum point of f. This is a maximum point of the second type.*

If $(\alpha_1, \nu_1) \neq \left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right)$, then there are two possible cases.

- $\frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)} \geq q$. Let (α^*, ν^*) be a maximum point of f of the second type in this case. Then $\alpha^* \leq 1 - \frac{qN}{q-1}$.

- $\dfrac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1-\frac{qN}{q-1}\right)} < q$. In this case f has a unique maximum point (α^*, ν^*). This point is of the second type. Furthermore, $\alpha^* > 1 - \frac{qN}{q-1}$, and it is uniquely determined by the following identity:

$$\frac{\alpha^* - 1}{\phi_\epsilon(\alpha^*)} = \frac{\alpha^* - \alpha_1}{\alpha^* - (1 - \nu_1)}.$$

Lemmas 6 and 7 will be proved in Section 3.3. At this point we prove Proposition 3 assuming these lemmas hold.

We start with the first claim of Proposition 3. Let $\alpha_1 = 1 - \frac{qN}{q-1}$ and $\nu_1 = \frac{qN}{q-1}$. Let $f = f_{\alpha_1, \nu_1}$. By the first claim of Lemma 7, we have

$$M(\alpha_1, \nu_1) = f(\alpha_1, \nu_1) = \frac{\alpha_1 - 1}{\phi_\epsilon(\alpha_1, \nu_1)} = \frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)}.$$

We pass to the second claim of the proposition. Assume that $\dfrac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1-\frac{qN}{q-1}\right)} \geq q$. Let $f = f_{\alpha_1, \nu_1}$, for some α_1 and ν_1. Let (α^*, ν^*) be a maximum point of f. Then Lemmas 6 and 7 imply that $\alpha^* \leq 1 - \frac{qN}{q-1}$. Hence

$$M(\alpha_1, \nu_1) = f(\alpha^*, \nu^*) \leq \frac{\alpha^* - 1}{\phi_\epsilon(\alpha^*)} \leq \frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)} = M\left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right).$$

Here in the second step we have used the third claim of Corollary 6, in the third step the third claim of Lemma 3 and in the fourth step the first claim of the proposition.

We pass to the third claim of the proposition. Assume that $\dfrac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1-\frac{qN}{q-1}\right)} < q$. Let $f = f_{\alpha_1, \nu_1}$, for some α_1 and ν_1. Then, by Lemma 7, f has a unique maximum point (α^*, ν^*). This means that α^* is determined by α_1 and ν_1, and furthermore, since $\nu_1 = N + \frac{1-\alpha_1}{q}$, α^* is a function of α_1. We will show the following claim below.

Lemma 8. *If $(\alpha_1, \nu_1) \neq \left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right)$ and $\dfrac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1-\frac{qN}{q-1}\right)} < q$, then α^* is a decreasing function of α_1.*

Assume Lemma 8 to hold. We have

$$M(\alpha_1, \nu_1) = f(\alpha^*, \nu^*) = \frac{\alpha^*(\alpha_1) - 1}{\phi_\epsilon(\alpha^*(\alpha_1))} \leq \frac{\alpha^*(0) - 1}{\phi_\epsilon(\alpha^*(0))} = M\left(0, N + \frac{1}{q}\right).$$

The second step uses the fact that (α^*, ν^*) is a maximum point of the second type, and hence $f(\alpha^*, \nu^*) = \frac{\alpha^* - 1}{\phi(\alpha^*)}$. The third step uses Lemma 8 and the third claim of Lemma 3, and the fourth step the fact that $\alpha_1 = 0$ implies $\nu_1 = N + \frac{1}{q}$.

Next, by Lemma 7, $\alpha = \alpha^*(0)$ is determined by the identity $\dfrac{\alpha - 1}{\phi_\epsilon(\alpha)} = \dfrac{\alpha}{\alpha - \left(\frac{q-1}{q} - N\right)}$ which, after rearranging, gives $1 - \alpha - \frac{\alpha \phi_\epsilon(\alpha)}{1 - \alpha} = N + \frac{1}{q}$. Hence, by the fourth claim of Lemma 3, $\alpha^*(0) = \alpha_0$ and $M\left(0, N + \frac{1}{q}\right) = \frac{\alpha_0 - 1}{\phi_\epsilon(\alpha_0)}$.

To conclude the proof of the third claim of the proposition, observe that since $\alpha^* > 1 - \frac{qN}{q-1}$, we have

$$M(\alpha_1, \nu_1) \;=\; f(\alpha^*, \nu^*) \;=\; \frac{\alpha^* - 1}{\phi_\epsilon(\alpha^*)} \;>\; \frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)},$$

where the last inequality is by the third claim of Lemma 3. This completes the proof of Proposition 3.

It remains to prove Lemmas 6–8.

3.3. Proofs of the Remaining Lemmas

Proof of Lemma 6. We start with the first claim of the lemma. Assume towards contradiction that $\frac{\alpha_1 - 1}{f(\alpha^*, \nu^*)} + \nu_1 < \frac{\alpha^* - 1}{f(\alpha^*, \nu^*)} + \nu^*$. Since f is a positive continuous function on Ω, there is a neighborhood O of (α^*, ν^*) in Ω on which $\frac{\alpha_1 - 1}{f(\alpha, \nu)} + \nu_1 < \frac{\alpha - 1}{f(\alpha, \nu)} + \nu$. This means that any point $(\alpha, \nu) \in O$ satisfies $\phi_\epsilon(\alpha) + \nu = \frac{\alpha - 1}{f(\alpha, \nu)} + \nu$, and hence $f(\alpha, \nu) = \frac{\alpha - 1}{\phi_\epsilon(\alpha)}$. Since $f(\alpha^*, \nu^*) \geq f(\alpha, \nu)$, this implies that $\frac{\alpha^* - 1}{\phi_\epsilon(\alpha^*)} \geq \frac{\alpha - 1}{\phi_\epsilon(\alpha)}$, and hence, by the third claim of Lemma 3, that $\alpha^* \geq \alpha$. It follows that α^* has to be 1, and hence $(\alpha^*, \nu^*) = (1, 0)$. However, in this case $\frac{\alpha_1 - 1}{f(\alpha^*, \nu^*)} + \nu_1 = \frac{\alpha^* - 1}{f(\alpha^*, \nu^*)} + \nu^* = 0$, reaching contradiction.

We pass to the second claim of the lemma. By the first claim $\frac{\alpha_1 - 1}{f(\alpha^*, \nu^*)} + \nu_1 > \frac{\alpha^* - 1}{f(\alpha^*, \nu^*)} + \nu^*$. We claim that this implies that (α^*, ν^*) is a local maximum of $\phi_\epsilon(\alpha) + \nu$. In fact, arguing as above, there is a neighborhood O of (α^*, ν^*) on which $\frac{\alpha_1 - 1}{f(\alpha, \nu)} + \nu_1 > \frac{\alpha - 1}{f(\alpha, \nu)} + \nu$. This means that for any point $(\alpha, \nu) \in O$ we have $\phi_\epsilon(\alpha) + \nu = \frac{\alpha_1 - 1}{f(\alpha, \nu)} + \nu_1$. Since $f(\alpha^*, \nu^*) \geq f(\alpha, \nu)$, this implies that $\phi_\epsilon(\alpha) + \nu \leq \phi(\alpha^*) + \nu^*$. To complete the proof, recall that any local maximum (α, ν) of $\phi(\alpha) + \nu$ has $\alpha \leq 1 - \frac{qN}{q-1}$ (as shown in the proof of Proposition 1). □

Proof of Lemma 7. Let (α^*, ν^*) be a maximum point of f of the second type. The first observation is that (α^*, ν^*) has to lie on the upper boundary of Ω. In fact, assume not. Then for a sufficiently small $\tau > 0$ the point $(\alpha, \nu) = (\alpha^*, \nu^* + \tau)$ is in Ω. Since $f(\alpha, \nu) \leq f(\alpha^*, \nu^*)$, we have $\phi_\epsilon(\alpha) + \nu > \phi_\epsilon(\alpha^*) + \nu^* = \frac{\alpha_1 - 1}{f(\alpha^*, \nu^*)} + \nu_1 \geq \frac{\alpha_1 - 1}{f(\alpha, \nu)} + \nu_1$. Hence $f(\alpha, \nu)$ is determined by the equality $\phi_\epsilon(\alpha) + \nu = \frac{\alpha - 1}{f(\alpha, \nu)} + \nu$, which implies $f(\alpha, \nu) = f(\alpha^*, \nu^*) = \frac{\alpha^* - 1}{\phi_\epsilon(\alpha^*)}$. Hence (α, ν) is a point of maximum of f of the first type with $\frac{\alpha_1 - 1}{f(\alpha, \nu)} + \nu_1 < \frac{\alpha - 1}{f(\alpha, \nu)} + \nu$. This, however, contradicts the first claim of Lemma 6.

Recall that the upper boundary of Ω is a piecewise linear curve which starts as the straight line $\frac{\alpha}{q} + \nu = N + \frac{1}{q}$, for $0 \leq \alpha \leq 1 - \frac{qN}{q-1}$ and continues as the straight line $\alpha + \nu = 1$ for $1 - \frac{qN}{q-1} \leq \alpha \leq 1$. Hence there are two cases to consider: In the first case $\alpha^* \leq 1 - \frac{qN}{q-1}$ and $\frac{\alpha^*}{q} + \nu^* = N + \frac{1}{q}$. In the second case $1 - \frac{qN}{q-1} < \alpha^* \leq 1$ and $\alpha^* + \nu^* = 1$. Assume that the second case holds. Then (α^*, ν^*) satisfies

1. $\frac{\alpha_1 - 1}{f(\alpha^*, \nu^*)} + \nu_1 = \frac{\alpha^* - 1}{f(\alpha^*, \nu^*)} + \nu^* = \phi_\epsilon(\alpha^*) + \nu^*$.
2. $1 - \frac{qN}{q-1} < \alpha^* \leq 1$ and $\alpha^* + \nu^* = 1$.

In particular, $f(\alpha^*, \nu^*) = \frac{\alpha^* - 1}{\phi_\epsilon(\alpha^*)} = \frac{\alpha^* - \alpha_1}{\alpha^* - (1 - \nu_1)}$. Consider the following two functions of α: $g_1(\alpha) = \frac{\alpha - 1}{\phi_\epsilon(\alpha)}$ and $g_2(\alpha) = \frac{\alpha - \alpha_1}{\alpha - (1 - \nu_1)}$, for $\alpha > 1 - \frac{qN}{q-1}$. Note that g_2 is well-defined since, by Lemma 5, $\nu_1 \geq \frac{qN}{q-1}$. By the third claim of Lemma 3, g_1 is strictly increasing. On the other hand, $g_2(\alpha) = 1 + \frac{1 - \alpha_1 - \nu_1}{\alpha - (1 - \nu_1)}$ is non-increasing. Note also that $g_1(1) = q_0$ (more precisely, $\lim_{\alpha \to 1} g_1(\alpha) = q_0$) and, by Lemma 5, $g_2(1) = \frac{1 - \alpha_1}{\nu_1} < q_0$. This means that g_1 and g_2 coincide at a (unique) point $1 - \frac{qN}{q-1} < \alpha < 1$ iff $g_1\left(1 - \frac{qN}{q-1}\right) < g_2\left(1 - \frac{qN}{q-1}\right)$.

Observe that if $(\alpha_1, \nu_1) = \left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right)$ then g_2 is the constant 1-function. Furthermore, by the first and the third claims of Lemma 3, $g_1\left(1 - \frac{qN}{q-1}\right) \geq g_1(0) = \frac{2}{\log_2(4/q_0)} \geq 1$, and hence in this case g_1 and g_2 cannot coincide for $\alpha > 1 - \frac{qN}{q-1}$. If $(\alpha_1, \nu_1) \neq \left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right)$ then it is easy to see (recall that $\frac{\alpha_1}{q} + \nu_1 = N + \frac{1}{q}$) that $g_2\left(1 - \frac{qN}{q-1}\right) = q$, and hence the two functions have a unique intersection at some $\alpha > 1 - \frac{qN}{q-1}$ iff $g_1\left(1 - \frac{qN}{q-1}\right) = \frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)}$ is smaller than q.

To recap, the second case can hold only provided $(\alpha_1, \nu_1) \neq \left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right)$ and $\frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)} < q$. Furthermore, if it holds then $1 - \frac{qN}{q-1} < \alpha^* < 1$ is uniquely determined by the equality $g_1(\alpha^*) = g_2(\alpha^*)$.

We can now complete the proof of the lemma. First, let $(\alpha_1, \nu_1) = \left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right)$. By the preceding discussion, in this case a maximum point (α^*, ν^*) of f of the second type has to have $\alpha^* \leq \alpha_1$. Moreover, taking into account Lemma 6, this is true for any maximum point of f. By the third claim of Corollary 6, this means that $M(\alpha_1, \nu_1) \leq \frac{\alpha_1 - 1}{\phi_\epsilon(\alpha_1)} = f(\alpha_1, \nu_1)$. Hence (α_1, ν_1) is a maximum point of f. It is trivially a maximum point of the second type. To see that it is a unique maximum point, note that for any point (α, ν) on the upper boundary of Ω, if $\alpha = \alpha_1$, then necessarily $\nu = \nu_1$. So, for any other putative maximum point (α, ν), we would have $\alpha < \alpha_1$ and hence, by the third claims of Lemma 3 and the third claim of Corollary 6, $f(\alpha, \nu) \leq \frac{\alpha - 1}{\phi_\epsilon(\alpha)} < \frac{\alpha_1 - 1}{\phi_\epsilon(\alpha_1)} = f(\alpha_1, \nu_1)$. This proves the first claim of the lemma.

Assume now that $(\alpha_1, \nu_1) \neq \left(1 - \frac{qN}{q-1}, \frac{qN}{q-1}\right)$. Let (α^*, ν^*) be a maximum point of f of the second type. If $g_1\left(1 - \frac{qN}{q-1}\right) = \frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)} \geq q$, then the preceding discussion implies that $\alpha^* \leq 1 - \frac{qN}{q-1}$, proving the second claim of the lemma.

If $\frac{-\frac{qN}{q-1}}{\phi_\epsilon\left(1 - \frac{qN}{q-1}\right)} < q$, let α be the unique solution for $g_1(\alpha) = g_2(\alpha)$ on $1 - \frac{qN}{q-1} < \alpha < 1$. Set $\alpha^* = \alpha$ and $\nu^* = 1 - \alpha$. We claim that (α^*, ν^*) is the unique maximum point of f (note that by Lemma 6 it would necessarily be of the second type). In fact, let us first verify that $\frac{\alpha_1 - 1}{\kappa} + \nu_1 = \frac{\alpha^* - 1}{\kappa} + \nu^* = \phi_\epsilon(\alpha^*) + \nu^*$, for $\kappa = \frac{\alpha^* - 1}{\phi_\epsilon(\alpha^*)}$. The second equality is immediate, by the definition of κ. The first equality is equivalent to $\kappa = \frac{\alpha^* - \alpha_1}{\alpha^* - (1 - \nu_1)}$, which follows from the definitions of α^* and κ. Hence $f(\alpha^*, \nu^*) = \kappa = \frac{\alpha^* - 1}{\phi_\epsilon(\alpha^*)}$. For any other putative maximum point (α, ν), we would have, by the preceding discussion, that $\alpha \leq 1 - \frac{qN}{q-1} < \alpha^*$ and hence, as above, $f(\alpha, \nu) \leq \frac{\alpha - 1}{\phi_\epsilon(\alpha)} < f(\alpha^*, \nu^*)$. This proves the third claim of the lemma. □

Proof of Lemma 8. In the assumptions of the lemma, α^* is the unique solution on $\left(1 - \frac{qN}{q-1}, 1\right)$ of the identity

$$\frac{\alpha^* - 1}{\phi_\epsilon(\alpha^*)} = \frac{\alpha^* - \alpha_1}{\alpha^* - (1 - \nu_1)}.$$

Here the LHS is a strictly increasing and the RHS a strictly decreasing (since by assumption $\alpha_1 \neq 1 - \frac{qN}{q-1}$, and hence $\alpha_1 + \nu_1 < 1$) functions of α^*. It follows that to prove the claim of the lemma it suffices to show that for a fixed $\alpha^* > 1 - \frac{qN}{q-1}$ the RHS is a decreasing function of α_1 (keeping in mind that $\nu_1 = -\frac{\alpha_1}{q} + \left(N + \frac{1}{q}\right)$). However, this is easily verifiable by a direct differentiation of the RHS w.r.t. α_1. □

This completes the proof of Proposition 2 and of (5). We proceed to complete the proof of Theorem 1. The tightness of (5) follows from the tightness of (6), similarly to the way the tightness of (4) was shown in the proof of Proposition 1. We omit the details.

It remains to consider the properties of the function $\kappa_{2,q}$. We first remark that it is easy to see, using the properties of the function ϕ_ϵ given in Lemma 3, that $\kappa_{2,q}$ is a continuous function of its first variable (we omit the details). In particular, we can replace strict inequalities with non-strict ones in the definition of $\kappa_{2,q}$ in Definition 1. Now there are two cases to consider.

- $q \geq q_0$. In this case, by the third claim of Lemma 3, $-\frac{x}{\phi_\epsilon(1-x)}$ is never larger than q, and hence

$$\kappa_{2,q}(x,\epsilon) = \begin{cases} q_0 & \text{if } y \leq \frac{1}{q_0} \\ \frac{\alpha_0 - 1}{\phi_\epsilon(\alpha_0)} & \text{if } y \geq \frac{1}{q_0} \end{cases}$$

  Here $y = \frac{q-1}{q} \cdot x + \frac{1}{q}$, $q_0 = 1 + (1-2\epsilon)^2$, and α_0 is determined by $1 - \alpha_0 - \frac{\alpha_0 \phi_\epsilon(\alpha_0)}{1-\alpha_0} = y$. Note that α_0 is well-defined, by the fourth claim of Lemma 3. The fact that $\kappa_{2,q}$ is decreasing in x follows from combining the third and the fourth claims of Lemma 3. In fact, $\kappa_{2,q}$ is a constant-$(1 + (1-2\epsilon)^2)$ function for $0 \leq x \leq \frac{q-q_0}{(q-1)q_0}$, and it is strictly decreasing for larger x.

- $q < q_0$. In this case y is always greater than $\frac{1}{q_0}$ and we have that

$$\kappa_{2,q}(x,\epsilon) = \begin{cases} -\frac{x}{\phi_\epsilon(1-x)} & \text{if } -\frac{x}{\phi_\epsilon(1-x)} \geq q \\ \frac{\alpha_0 - 1}{\phi_\epsilon(\alpha_0)} & \text{if } -\frac{x}{\phi_\epsilon(1-x)} \leq q \end{cases}$$

  It suffices to show that $\kappa_{2,q}$ is decreasing on both relevant subintervals of $[0,1]$, and this again follows from the third and the fourth claims of Lemma 3. In this case $\kappa_{2,q}$ is strictly decreasing on $[0,1]$.

This completes the proof of Theorem 1. □

4. Remaining Proofs

Proof of Lemma 3. The strict concavity of ϕ_ϵ and the bounds on its derivative were shown in [9], Lemma 2.13 (note that $\phi_\epsilon(x) = \frac{1}{2}\tilde{\phi}(x, 2\epsilon(1-\epsilon))$ in terms of [9]). The value of ϕ_ϵ at the endpoints of the interval $[0,1]$ are directly computable.

We pass to the third claim of the lemma. Taking the derivative and rearranging, it suffices to prove that for any $\alpha \in (0,1)$ holds $\phi_\epsilon(\alpha) > (\alpha - 1)\phi'_\epsilon(\alpha)$. This follows immediately from the strict concavity of ϕ_ϵ and the fact that $\phi_\epsilon(1) = 0$.

We pass to the last claim of the lemma. Taking the derivative and rearranging, it suffices to prove that for any $\alpha \in (0,1)$ holds

$$(1 - \alpha)\big(\alpha\phi'_\epsilon(\alpha) + (1 - \alpha)\big) > -\phi_\epsilon(\alpha).$$

Since $(1 - \alpha) \cdot \phi'_\epsilon(\alpha) > -\phi_\epsilon(\alpha)$, it suffices to show that $\alpha\phi'_\epsilon(\alpha) + (1 - \alpha) \geq \phi'_\epsilon(\alpha)$, and this follows from the first two claims of the lemma. The values of the function g at the endpoints are directly computable. □

Proof of Lemma 1. We start with a technical lemma which deals with the behavior of the function $\phi_\epsilon(x)$ in the vicinity of $\epsilon = 0$. We write $\epsilon \sim 0$ as a shorthand for "ϵ close to 0". We again use the fact that $\phi_\epsilon(x) = \Phi(x, 2\epsilon(1-\epsilon)) = \frac{1}{2}\tilde{\phi}(x, 2\epsilon(1 - \epsilon))$, where the function $\tilde{\phi}$ was defined and studied in [9]. In the calculations below $\phi(x, \epsilon)$ is written instead of $\phi_\epsilon(x)$, for notational convenience.

Lemma 9. *Let $0 < t < 1$. Then*

1.

$$\phi(t,0) = \frac{t-1}{2} \quad \text{and for } \epsilon \sim 0 \text{ holds } \left|\phi(t,\epsilon) - \frac{t-1}{2}\right| \leq O(\epsilon).$$

2.

$$\frac{\partial \phi}{\partial \epsilon}(t,0) = \frac{2\sqrt{H^{-1}(t)(1 - H^{-1}(t))} - 1}{\ln(2)} \quad \text{and for } \epsilon \sim 0 \text{ holds}$$

$$\left| \frac{\partial \phi}{\partial \epsilon}(t,\epsilon) - \frac{2\sqrt{H^{-1}(t)(1 - H^{-1}(t))} - 1}{\ln(2)} \right| \leq O(\epsilon).$$

3.

$$\frac{\partial \phi}{\partial t}(t,0) = \frac{1}{2} \quad \text{and for } \epsilon \sim 0 \text{ holds} \quad \left| \frac{\partial \phi}{\partial t}(t,\epsilon) - \frac{1}{2} \right| \leq O(\epsilon).$$

Proof of Lemma 9. *Notation.* Here and below we write $a \pm \epsilon$ as a shorthand for the interval $[a - \epsilon, a + \epsilon]$.
Recall that

$$\tilde{\phi}(t,\epsilon) = t - 1 + \sigma H\left(\frac{z}{\sigma}\right) + (1 - \sigma)H\left(\frac{z}{1 - \sigma}\right) + 2z \log_2(\epsilon) + (1 - 2z) \log_2(1 - \epsilon),$$

where $\sigma = H^{-1}(t)$ and $z = z(t,\epsilon) = \frac{-\epsilon^2 + \epsilon\sqrt{\epsilon^2 + 4(1 - 2\epsilon)\sigma(1-\sigma)}}{2(1 - 2\epsilon)}$.

The fact that $\phi(t,0) = \frac{1}{2}\tilde{\phi}(t,0) = \frac{t-1}{2}$ is verified by inspection, observing that $z(t,0) = 0$ for any t. Note also that, by assumption, $\sigma > 0$, and hence $z(t,\epsilon) \in \sqrt{\sigma(1-\sigma)} \cdot \epsilon \pm O(\epsilon^2)$ for a sufficiently small ϵ.

Using (as in the proof of Lemma 2.13 in [9]) the fact that for $\epsilon > 0$ holds $\frac{(\sigma - z)(1 - \sigma - z)}{z^2} = \frac{(1-\epsilon)^2}{\epsilon^2}$, and writing $\delta = 2\epsilon(1 - \epsilon)$, we have that

$$\frac{\partial \phi(t,\epsilon)}{\partial \epsilon} = \frac{1}{2} \cdot \frac{\partial \tilde{\phi}(t,\delta)}{\partial \epsilon} = \frac{1 - 2\epsilon}{\ln(2)} \cdot \frac{2z - \delta}{\delta(1 - \delta)}.$$

Hence for $\epsilon \sim 0$ we have $\frac{\partial \phi(t,\epsilon)}{\partial \epsilon} \in \frac{2\sqrt{\sigma(1-\sigma)} - 1}{\ln(2)} \pm O(\delta)$, or equivalently $\frac{\partial \phi(t,\epsilon)}{\partial \epsilon} \in \frac{2\sqrt{\sigma(1-\sigma)} - 1}{\ln(2)} \pm O(\epsilon)$.
In particular,

$$\frac{\partial \phi(t,\epsilon)}{\partial \epsilon}\Big|_{\epsilon=0} = \lim_{\epsilon \to 0} \frac{\partial \phi(t,\epsilon)}{\partial \epsilon} = \frac{2\sqrt{\sigma(1-\sigma)} - 1}{\ln(2)} = \frac{2\sqrt{H^{-1}(t)(1 - H^{-1}(t))} - 1}{\ln(2)}.$$

This proves both the first and the second claims of the lemma.

We pass to the third claim of the lemma. As shown in the proof of Lemma 2.13 in KS2 we have $\frac{\partial \tilde{\phi}}{\partial t}(t,\epsilon) = \frac{\ln\left(\frac{1-\sigma-z}{\sigma-z}\right)}{\ln\left(\frac{1-\sigma}{\sigma}\right)}$. Hence

$$\frac{\partial \phi(t,\epsilon)}{\partial t} = \frac{1}{2} \cdot \frac{\partial \tilde{\phi}(t,\delta)}{\partial t} = \frac{1}{2} \cdot \frac{\ln\left(\frac{1-\sigma-z}{\sigma-z}\right)}{\ln\left(\frac{1-\sigma}{\sigma}\right)},$$

where $z = z(t,\delta)$. Recall for any $0 < t < 1$ we have $z(t,0) = 0$ and in addition for $\delta \sim 0$ we have $z(t,\delta) \in \sqrt{\sigma(1-\sigma)} \cdot \delta \pm O(\delta^2)$. The third claim of the lemma now follows by inspection. This completes the proof of Lemma 9. $\square$

We proceed with the proof of Lemma 1. First, consider the definition of $\kappa = \kappa_{2,2}$. For ϵ sufficiently close to zero, we have that $\frac{x+1}{2} > \frac{1}{q_0}$ (recall that $q_0 = 1 + (1 - 2\epsilon)^2$) and hence

$\kappa = \frac{\alpha-1}{\phi(\alpha,\epsilon)}$, where $\alpha = \alpha(\epsilon)$ is determined by $1 - \alpha + \frac{\alpha\phi(\alpha,\epsilon)}{\alpha-1} = \frac{x+1}{2}$. Taking the derivative w.r.t. ϵ in the definition of α and rearranging gives

$$\alpha'(\epsilon) = -\frac{\alpha(\alpha-1)\frac{\partial\phi}{\partial\epsilon}(\alpha,\epsilon)}{\alpha(\alpha-1)\frac{\partial\phi}{\partial\alpha}(\alpha,\epsilon) - \phi(\alpha,\epsilon) - (\alpha-1)^2}.$$

Using the first claim of Lemma 9, it is easy to see that $\alpha(0) = 1 - x$. Hence, using all claims of Lemma 9, we have that

$$\alpha'(0) = -\frac{2}{\ln 2} \cdot \frac{\alpha(0)\left(2\sqrt{H^{-1}(\alpha(0))(1 - H^{-1}(\alpha(0)))} - 1\right)}{1 - \alpha(0)} =$$

$$-\frac{2}{\ln 2} \cdot \frac{(1-x)\left(2\sqrt{H^{-1}(1-x)(1 - H^{-1}(1-x))} - 1\right)}{x}$$

Next, we compute κ and κ' at 0. Note that by the definition of κ, we have $1 - \alpha + \frac{\alpha}{\kappa} = \frac{x+1}{2}$. Hence, $\kappa = \frac{\alpha}{\frac{x+1}{2}+\alpha-1}$ and $\kappa' = \frac{\kappa(1-\kappa)\alpha'}{\alpha}$. In particular, $\kappa(0) = 2$ and

$$\kappa'(0) = = -\frac{2\alpha'(0)}{\alpha(0)} = \frac{4}{\ln 2} \cdot \frac{\left(2\sqrt{H^{-1}(1-x)(1 - H^{-1}(1-x))} - 1\right)}{x},$$

proving the first claim of the proposition.

Let now $\epsilon \sim 0$. We start with estimating $\alpha(\epsilon)$ and $\kappa(\epsilon)$. From the identity $1 - \alpha + \frac{\alpha\phi(\alpha,\epsilon)}{\alpha-1} = \frac{x+1}{2}$, using the monotonicity of the LHS in α (by Lemma 2.3) and Lemma 9, it is easy to see that $\alpha(\epsilon) \in 1 - x \pm O(\epsilon)$. From this, and from the identity $1 - \alpha + \frac{\alpha}{\kappa} = \frac{x+1}{2}$, we get $\kappa(\epsilon) = \frac{\alpha(\epsilon)}{\frac{x+1}{2}+\alpha(\epsilon)-1} \in 2 \pm O(\epsilon)$.

Proceeding in a similar vein, using the above expression for α', we get that

$$\alpha'(\epsilon) \in \frac{2}{\ln(2)} \cdot \frac{1-x}{x} \cdot \left(1 - 2\sqrt{H^{-1}(1-x)(1 - H^{-1}(1-x))}\right) \pm O(\epsilon),$$

and

$$\kappa'(\epsilon) = \frac{\kappa(\epsilon)(1-\kappa(\epsilon))\alpha'(\epsilon)}{\alpha(\epsilon)} \in = -\frac{2\alpha'(\epsilon)}{\alpha(\epsilon)} \subseteq$$

$$\frac{4}{\ln 2} \cdot \frac{\left(2\sqrt{H^{-1}(1-x)(1 - H^{-1}(1-x))} - 1\right)}{x} \pm O(\epsilon),$$

completing the proof of the lemma. □

Proof of Corollary 2. Let $q = 2$ and $\kappa = \kappa_{2,2}$ (see the second claim of Corollary 1 for a more explicit statement of Theorem 1 in this case). Viewing both sides of (5) as functions of ϵ, and writing $L(\epsilon)$ for the LHS and $R(\epsilon)$ for the RHS, we have $L(0) = R(0) = \|f\|_2$, and $L(\epsilon) \le R(\epsilon)$ for $0 \le \epsilon \le \frac{1}{2}$. It is easy to see that both L and R are differentiable, and we may deduce that $L'(0) \le R'(0)$. Computing the derivatives (see, e.g., [3]) gives

$$L'(0) = -\frac{1}{2} \cdot \frac{\mathcal{E}(f,f)}{\|f\|_2} \quad \text{and} \quad R'(0) = \frac{\ln(2)\kappa'(0)}{4} \cdot \frac{Ent(f^2)}{\|f\|_2},$$

where we write $\kappa'(0)$ for $\frac{\partial\kappa}{\partial\epsilon}|_{\epsilon=0}$. Hence $L'(0) \le R'(0)$ is equivalent to

$$\mathcal{E}(f,f) \ge -\frac{\ln(2)\kappa'(0)}{2} \cdot Ent\left(f^2\right). \tag{7}$$

The claim of the corollary now follows from the first claim of Lemma 1. It only remains to add that the fact that $\ell(\cdot)$ a convex and increasing function on $[0,1]$, taking $[0,1]$ onto $[2\ln 2, 2]$ was proved in [20]. $\square$

Proof of Corollary 3. Let us point out that our argument follows along the same lines as the proof of the same result in [10]. We do believe that the argument here is worth presenting in full, since it seems to be somewhat more explicit and easier to parse.

We use the simple fact (see, e.g., [4]) that for any $0 \leq \epsilon \leq \frac{1}{2}$ and for any $\alpha \in \{0,1\}^n$ holds $\widehat{f_\epsilon}(\alpha) = (1-2\epsilon)^{|\alpha|}\widehat{f}(\alpha)$. Hence, using Parseval's identity in the first step below, we have

$$\|f_\epsilon\|_2^2 = \sum_{\alpha\in\{0,1\}^n} (1-2\epsilon)^{2|\alpha|}\widehat{f}^2(\alpha) \geq (1-2\epsilon)^{2\mu n} \cdot \sum_{|\alpha|\leq\mu n} \widehat{f}^2(\alpha).$$

Since this holds for any $0 \leq \epsilon \leq \frac{1}{2}$, we deduce that

$$\sum_{|\alpha|\leq\mu n} \widehat{f}^2(\alpha) \leq \min_{0\leq\epsilon\leq\frac{1}{2}} \frac{\|f_\epsilon\|_2^2}{(1-2\epsilon)^{2\mu n}} \leq \min_{0\leq\epsilon\leq\frac{1}{2}} \frac{\|f\|_\kappa^2}{(1-2\epsilon)^{2\mu n}},$$

where we have used Theorem 1 with $q=2$ in the second step, and $\kappa = \kappa(\epsilon) = \kappa_{2,2}\left(\frac{Ent_2\left(\frac{f}{\|f\|_1}\right)}{n}, \epsilon\right)$.

Let $F(\epsilon) = \frac{1}{n}\log_2\left(\frac{\|f\|_\kappa^2}{(1-2\epsilon)^{2\mu n}}\right) = \frac{1}{n}\log_2(\|f\|_\kappa^2) - 2\mu\log_2(1-2\epsilon)$. Since $\kappa(0) = 2$, we have $F(0) = \frac{1}{n}\log_2(\|f\|_2^2)$. Hence the claim of the corollary is equivalent to the claim that $\min_{0\leq\epsilon\leq\frac{1}{2}} F(\epsilon)$ is negative and bounded away from $F(0)$ by some absolute constant. To show this, it suffices to show that $F'(\epsilon)$ is negative and bounded away from 0 by an absolute constant for ϵ in a constant length interval $[0, \epsilon_0]$.

Recall that for any nonnegative non-zero function g on $\{0,1\}^n$ holds $\frac{Ent(g^2)}{\mathbb{E} g^2} \geq \log_2\left(\frac{\mathbb{E} g^2}{\mathbb{E}^2 g}\right) = Ent_2\left(\frac{g}{\|g\|_1}\right)$ (see, e.g., [10]). Recall also that $\frac{\partial}{\partial\epsilon}\log_2\left(\|f\|_{\kappa(\epsilon)}\right) = \frac{\kappa'}{\kappa^2} \cdot \frac{Ent(|f|^\kappa)}{\|f\|_\kappa^\kappa}$. Hence, recalling that, by Lemma 1, $\kappa' < 0$ in the vicinity of 0, we have

$$F'(\epsilon) = 2\frac{\kappa'}{\kappa^2} \cdot \frac{1}{n}\frac{Ent(|f|^\kappa)}{\|f\|_\kappa^\kappa} + \frac{4}{\ln(2)} \cdot \frac{\mu}{1-2\epsilon} \leq 2\frac{\kappa'}{\kappa^2} \cdot \frac{1}{n}\log_2\left(\frac{\mathbb{E}(|f|^\kappa)}{\mathbb{E}^2|f|^{\kappa/2}}\right) + \frac{4}{\ln(2)} \cdot \frac{\mu}{1-2\epsilon}.$$

Let $x = \frac{Ent_2\left(\frac{f}{\|f\|_1}\right)}{n} = 1 - H(\rho)$. Recalling again that $\kappa(0) = 2$ and applying the first claim of Lemma 1 we get

$$F'(0) \leq \frac{\kappa'(0)}{2} \cdot x + \frac{4\mu}{\ln(2)} = \frac{4}{\ln(2)} \cdot \left(\mu - \left(\frac{1}{2} - \sqrt{\rho(1-\rho)}\right)\right) < 0.$$

It now suffices to show that for sufficiently small ϵ we have $F'(\epsilon) \leq F'(0) + O(\epsilon)$. Taking the second claim of Lemma 1 into account, it is enough to show that $\frac{1}{n}\log_2\left(\frac{\mathbb{E}(|f|^\kappa)}{\mathbb{E}^2|f|^{\kappa/2}}\right) \geq x - O(\epsilon)$. Let $G(\epsilon) = \frac{1}{n}\log_2\left(\frac{\mathbb{E}(|f|^\kappa)}{\mathbb{E}^2|f|^{\kappa/2}}\right)$. Then $G(0) = x$ and it suffices to show that $|G'|$ is bounded by an absolute constant. A simple calculation gives that

$$G' = \frac{\kappa'}{\kappa} \cdot \left(\frac{1}{n}\frac{Ent(|f|^\kappa)}{\mathbb{E}|f|^\kappa} - \frac{2}{n}\frac{Ent\left(|f|^{\kappa/2}\right)}{\mathbb{E}|f|^{\kappa/2}} + G\right).$$

The RHS in the last expression is bounded by a constant, since for any nonnegative non-zero function g on $\{0,1\}^n$ both $\frac{Ent(g)}{\mathbb{E} g}$ and $\log_2\left(\frac{\mathbb{E} g^2}{\mathbb{E}^2(g)}\right)$ are bounded by n. $\square$

Proof of Corollary 4.

The first claim of the corollary

Let $D \subseteq \{0,1\}^n$, $|D| = 2^{H(\rho)n}$. Let M_D be the adjacency matrix of the subgraph of the discrete cube induced by the vertices of D. Let $\lambda(D)$ be the maximal eigenvalue of M_D. Let f be a maximal eigenvector of M_D. We view f as a function on D and extend it to a function on $\{0,1\}^n$ by defining it to be zero outside D. Let A be the adjacency matrix of $\{0,1\}^n$. Then $\lambda(D) = \frac{\langle f, M_D f \rangle}{\langle f, f \rangle} = \frac{\langle f, Af \rangle}{\langle f, f \rangle}$. Note also that since f is supported on D we have

$$\mathbb{E}^2 |f| = \left(\langle f, \operatorname{sign}(f) \cdot 1_D \rangle \right)^2 \leq \mathbb{E} f^2 \cdot \mathbb{E} \left(\operatorname{sign}(f) \cdot 1_D \right)^2 = \mathbb{E} f^2 \cdot \frac{|D|}{2^n} = \mathbb{E} f^2 \cdot 2^{(H(\rho)-1)n}.$$

It follows that $\frac{Ent_2\left(\frac{f}{\|f\|_1} \right)}{n} \geq 1 - H(\rho)$.

Next, it is easy to check that for any function g on $\{0,1\}^n$ holds $\mathcal{E}(g,g) = 2\langle g, (nI - A)g \rangle$, where I is the $2^n \times 2^n$ identity matrix. Hence, using Corollary 2 and the fact that $\frac{Ent(f^2)}{\mathbb{E} f^2} \geq \log_2\left(\frac{\mathbb{E} f^2}{\mathbb{E}^2 |f|} \right) = Ent_2\left(\frac{f}{\|f\|_1} \right)$, we have, writing x for $\frac{1}{n} Ent_2\left(\frac{f}{\|f\|_1} \right)$,

$$\lambda(D) = \frac{\langle f, Af \rangle}{\langle f, f \rangle} = n - \frac{1}{2} \frac{\mathcal{E}(f,f)}{\langle f, f \rangle} \leq n - \frac{1}{2} \frac{\ell(x) \cdot Ent(f^2)}{\mathbb{E} f^2} \leq$$

$$n - \frac{n}{2} x \ell(x) \leq n \left(1 - \frac{1}{2}(1 - H(\rho)\ell(1 - H(\rho)) \right) = 2\sqrt{\rho(1-\rho)} \cdot n.$$

This is almost tight if we set $r = \lceil \rho n \rceil$ and take $D = \{x \in \{0,1\}^n : |x| \leq r\}$ to be the Hamming ball of radius r around 0. In fact, recall that $|D| \approx 2^{H(\rho)n}$ (see, e.g., [19]) and, as shown in [19], $\lambda(D) \geq 2\sqrt{\rho(1-\rho)} \cdot n - o(n)$. $\square$

The second claim of the corollary

Let $0 < \delta < \frac{1}{2}$. Let $d = \lfloor \delta n \rfloor$, and let f be a feasible solution of the dual linear program of [15] with parameters n and d. Then, as observed by [29] f can be viewed as a function on $\{0,1\}^n$ with the following properties:

- f is symmetric, that is $f(x)$ depends only on $|x|$.
- $f(x) \leq 0$ for $|x| \geq d$.
- $\widehat{f} \geq 0$ and $\widehat{f}(0) = 1$.
- $f(0) \leq 2^{R_{LP}(\delta) \cdot n + o(n)}$.

To prove the claim, we will show that any function f with the first three of these properties satisfies $\frac{1}{n} \log_2(f(0)) \geq \frac{1 - H(\delta) + H\left(\frac{1}{2} - \sqrt{\delta(1-\delta)} \right)}{2} - o_n(1)$.

Notation: We write $\|g\|_{q,\mathcal{F}}$ for $\left(\sum_{\alpha \in \{0,1\}^n} |g(\alpha)|^q \right)^{1/q}$. Note that Parseval's identity states $\|f\|_2 = \|\widehat{f}\|_{2,\mathcal{F}}$. We write $\approx$, $\lesssim$, and $\gtrsim$ to denote equality or inequality which hold up to lower order terms. To give an example, recall that for $0 < \rho \leq \frac{1}{2}$ the cardinalities of the Hamming ball $\{x \in \{0,1\}^n : |x| \leq r\}$ and the Hamming sphere $\{x \in \{0,1\}^n : |x| = r\}$ are $2^{H(\rho)n}$, up to lower order terms. We write this as $\frac{1}{n} \log_2(|\{x \in \{0,1\}^n : |x| \leq r\}|) \approx H(\rho)$.

We start with some preliminary observations. First, we need some simple and well-known facts from Fourier analysis on $\{0,1\}^n$. If f is symmetric, then so is $\widehat{f}$. Next, $\widehat{f}(0) = \mathbb{E} f \leq \|f\|_1$. Furthermore, finally, using the fact that in our case $\widehat{f} \geq 0$, $f(0) = \sum_{\alpha \in \{0,1\}^n} \widehat{f}(\alpha) = \|\widehat{f}\|_{1,\mathcal{F}}$.

Next, we claim that if f is symmetric and if, for some $0 \leq i \leq n$ holds $\frac{1}{2^n}\binom{n}{i}|f(i)| \geq \Omega\left(\frac{1}{n} \right) \cdot \|f\|_1$ then $\frac{\|f\|_2}{\|f\|_1} \geq \Omega\left(\frac{1}{n} \right) \cdot \sqrt{\frac{2^n}{\binom{n}{i}}}$. In fact, we will have

$$\|f\|_2^2 \geq \frac{1}{2^n}\binom{n}{i} f^2(i) \geq \Omega\left(\frac{1}{n^2} \right) \cdot \frac{1}{2^n}\binom{n}{i} \left(\frac{2^n}{\binom{n}{i}} \|f\|_1 \right)^2 = \Omega\left(\frac{1}{n^2} \right) \cdot \frac{2^n}{\binom{n}{i}} \|f\|_1^2.$$

Similarly, if for some $0 \leq j \leq n$ holds $\binom{n}{j}\widehat{f}^2(j) \geq \Omega\left(\frac{1}{n}\right) \cdot \|\widehat{f}\|_{2,\mathcal{F}}^2$ then $\frac{\|\widehat{f}\|_{1,\mathcal{F}}}{\|\widehat{f}\|_{2,\mathcal{F}}} \geq$ $\Omega\left(\frac{1}{n}\right) \cdot \sqrt{\binom{n}{j}}$.

Finally, we need a slight extension of Corollary 3. As stated, it shows that if f has a large second entropy, then $\widehat{f}$ cannot attain its ℓ_2 norm in a Hamming ball of small radius *around* 0. We claim, as was also observed in [10], that this holds more generally for Hamming balls with arbitrary centers in $\{0,1\}^n$. To see that, let $z \in \{0,1\}^n$, and define $g = f \cdot W_z$, where W_z is the corresponding Walsh-Fourier character. It is easy to see that for any $y \in \{0,1\}^n$ holds $\widehat{g}(y) = \widehat{f}(y+z)$, and hence g has the same first and second norms as f. Moreover, writing $B(z,r)$ for the Hamming ball of radius r around z, we have $\sum_{\alpha \in B(z,r)} \widehat{f}^2(\alpha) = \sum_{\beta \in B(0,r)} \widehat{g}^2(\beta)$.

We pass to the proof of the claim. Note that since $f(x) \leq 0$ for $|x| \geq d$ and since $\mathbb{E}\, f \geq 0$, there exists $0 \leq i \leq d-1$ such that $\frac{1}{2^n}\binom{n}{i}|f(i)| \geq \Omega\left(\frac{1}{n}\right) \cdot \|f\|_1$. Hence

$$\frac{1}{n}Ent_2\left(\frac{f}{\|f\|_1}\right) \;=\; \frac{1}{n}\log_2\left(\frac{\|f\|_2^2}{\|f\|_1^2}\right) \;\gtrsim\; 1 - H\left(\frac{i}{n}\right) \;\geq\; 1 - H(\delta).$$

By Corollary 3 this means that $\widehat{f}$ cannot attain its ℓ_2 norms inside Hamming balls or radii much smaller than $r(\delta) := \left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right) \cdot n$ around the all-0 and all-1 vectors. Hence there exists $r(\delta) - o(n) \leq j \leq r(\delta) + o(n)$ such that $\binom{n}{j}\widehat{f}^2(j) \geq \Omega\left(\frac{1}{n}\right) \cdot \|\widehat{f}\|_{2,\mathcal{F}}^2$. It follows that

$$\frac{1}{n}\log_2\left(\frac{\|\widehat{f}\|_{1,\mathcal{F}}}{\|\widehat{f}\|_{2,\mathcal{F}}}\right) \;\gtrsim\; \frac{H\left(\frac{j}{n}\right)}{2} \;\gtrsim\; \frac{H\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right)}{2}.$$

We can now complete the proof of the second claim of the corollary. We have

$$0 \;=\; \frac{1}{n}\log_2\left(\widehat{f}(0)\right) \;\leq\; \frac{1}{n}\log_2(\|f\|_1) \;\lesssim\; \frac{1}{n}\log_2(\|f\|_2) - \frac{1-H(\delta)}{2} \;=$$

$$\frac{1}{n}\log_2\left(\|\widehat{f}\|_{2,\mathcal{F}}\right) - \frac{1-H(\delta)}{2} \;\lesssim\; \frac{1}{n}\log_2\left(\|\widehat{f}\|_{1,\mathcal{F}}\right) - \frac{1-H(\delta) + H\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right)}{2} \;=$$

$$\frac{1}{n}\log_2(f(0)) - \frac{1-H(\delta) + H\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right)}{2}.$$

$\square$

Proof of Corollary 5. Let $0 \leq s \leq n/2$ and let f be a polynomial of degree s on $\{0,1\}^n$. We need two simple and well-known facts from Fourier analysis on $\{0,1\}^n$. First, that the Fourier expansion of f is supported on characters of weight at most s; and second, that for any function g on $\{0,1\}^n$ holds $\mathcal{E}(g,g) = 4\sum_{\alpha \in \{0,1\}^n} |\alpha|\widehat{g}^2(\alpha)$. Combining these two facts implies that

$$\mathcal{E}(f,f) \;=\; 4\sum_{\alpha \in \{0,1\}^n} |\alpha|\widehat{f}^2(\alpha) \;=\; 4\sum_{|\alpha| \leq s} |\alpha|\widehat{f}^2(\alpha) \;\leq\; 4s \cdot \sum_{|\alpha| \leq s} \widehat{f}^2(\alpha) \;=\; 4s \cdot \mathbb{E}\, f^2,$$

where in the last step we used Parseval's identity.

Write σ for s/n and x for $\frac{1}{n}Ent_2\left(\frac{f}{\|f\|_1}\right)$. We have, using Corollary 2,

$$4\sigma n \;=\; 4s \;\geq\; \frac{\mathcal{E}(f,f)}{\mathbb{E}\, f^2} \;\geq\; \ell(x) \cdot \frac{Ent\left(f^2\right)}{\mathbb{E}\, f^2} \;\geq\; nx\ell(x) \;=$$

$$n \cdot \left(2 - 4\sqrt{H^{-1}(1-x)(1-H^{-1}(1-x))}\right).$$

Rearranging and simplifying, this is equivalent to

$$\frac{1}{n}\log_2\left(\frac{\|f\|_2}{\|f\|_1}\right) = \frac{x}{2} \le \frac{1 - H\left(\frac{1}{2} - \sqrt{\sigma(1-\sigma)}\right)}{2},$$

completing the proof. $\square$

Proof of Corollary 1. We start with the first claim of the corollary. First consider the case $\epsilon = \frac{1}{2}$. It is easy to see that $\phi_{\frac{1}{2}}(x) = x - 1$ (note that in the definition of $\Phi(x, \epsilon)$ we have $y\left(x, \frac{1}{2}\right) = \lim_{\epsilon \to \frac{1}{2}} y(x, \epsilon) = H^{-1}(x)\left(1 - H^{-1}(x)\right)$) and hence in this case the value of κ given by the claim is 1 (as it should be).

Assume now $\epsilon < \frac{1}{2}$. This implies that $q_0 = 1 + (1 - 2\epsilon)^2 > 1$. By the first claim of Lemma 3, this means that for any $0 \le x \le 1$ we have $\frac{-x}{\phi_\epsilon(1-x)} \ge -\frac{1}{\phi_\epsilon(0)} = \frac{2}{\log_2\left(\frac{4}{q_0}\right)} > 1$.

Hence, it is easy to see that for q sufficiently close to 1 the first and the third clauses in the definition of $\kappa_{2,q}$ in Definition 1 do not apply, and we have $\kappa_{2,q}(x, \epsilon) = \frac{-x}{\phi_\epsilon(1-x)}$. Theorem 1 then gives

$$\|f_\epsilon\|_2 \le \|f\|_\kappa, \quad \text{with} \quad \kappa = -\frac{\frac{Ent_q\left(\frac{f}{\|f\|_1}\right)}{n}}{\phi_\epsilon\left(1 - \frac{Ent_q\left(\frac{f}{\|f\|_1}\right)}{n}\right)}.$$

Taking $q \to 1$ and recalling that $Ent_q(\cdot) \to_{q \to 1} Ent(\cdot)$ completes the proof of the claim.

We pass to the second claim of the corollary. First consider the case $\epsilon = 0$. Note that in this case $q_0 = 2$. Furthermore, by the first claim of Lemma 9, $\phi_0(x) = \frac{x-1}{2}$, and hence the value of κ given by the claim is 2 (as expected).

Assume now $\epsilon > 0$. This implies that $q_0 < 2$, and hence, by the third claim of Lemma 3, for any $0 \le x \le 1$ we have $\frac{-x}{\phi_\epsilon(1-x)} \le q_0 < 2 = q$. Hence the second clause in the definition of $\kappa_{2,q}$ in Definition 1 does not apply. The remaining two clauses give the claim, as stated. $\square$

Proofs of Comments to Theorem 1

Some of the claims in these comments require a proof. These claims are restated and proved in the following lemma.

Lemma 10.

- *If $q \ge 2$ then for any $0 < \epsilon < \frac{1}{2}$ the function $\kappa_{2,q}(x, \epsilon)$ starts as a constant-$\left(1 + (1 - 2\epsilon)^2\right)$ function up to some $x = x(q, \epsilon) > 0$, and becomes strictly decreasing after that. For $1 < q < 2$ there is a value $0 < \epsilon(q) < \frac{1}{2}$, such that for all $\epsilon \le \epsilon(q)$ the function $\kappa_{2,q}(x, \epsilon)$ is strictly decreasing (in which case we say that $x(q, \epsilon) = 0$). However, $x(q, \epsilon) > 0$ for all $\epsilon > \epsilon(q)$. The function $\epsilon(q)$ decreases with q (in particular, $\epsilon(q) = 0$ for $g \ge 2$). The function $x(q, \epsilon)$ increases both in q and in ϵ.*
- *The function $\kappa_{2,1}(x, \epsilon) = -\frac{x}{\phi_\epsilon(1-x)}$ is strictly decreasing in its first argument for any $0 < \epsilon < \frac{1}{2}$. It satisfies $\kappa_{2,1}(0, \epsilon) = \lim_{x \to 0} \kappa_{2,1}(x, \epsilon) = 1 + (1 - 2\epsilon)^2$, for all $0 \le \epsilon \le \frac{1}{2}$.*
- *Let f be a non-constant function on $\{0, 1\}^n$. Let $0 < \epsilon < \frac{1}{2}$. Let $F(q) = F_{f,\epsilon}(q) = \kappa_{2,q}\left(Ent_q\left(\frac{f}{\|f\|_1}\right)/n, \epsilon\right)$. There is a unique value $1 < q(f, \epsilon) \le 1 + (1 - 2\epsilon)^2$ of q for which $F(q) = q$. Moreover, $q(f, \epsilon) = \min_{q \ge 1} F(q)$. Furthermore, $\lim_{\epsilon \to 0} q(f, \epsilon) = 2$ for any f.*

Proof. The first claim of the lemma follows from the properties of $\kappa_{2,q}$ as shown in the proof of Theorem 1. In particular, it is easy to see that for $q \le 2$ we have $\epsilon(q) = \frac{1 - \sqrt{q-1}}{2}$ and for $\epsilon \ge \epsilon(q)$ we have $x(q, \epsilon) = \frac{q - \left(1 + (1 - 2\epsilon)^2\right)}{(1 + (1 - 2\epsilon)^2) \cdot (q-1)}$. The claim that $\epsilon(q)$ decreases with q and that $x(q, \epsilon)$ increases in both q and ϵ follows by direct verification.

The second claim of the lemma follows immediately from the third claim of Lemma 3.

We pass to the third claim of the lemma. Note that the function $x(q) = Ent_q\left(\frac{f}{\|f\|_1}\right)/n$ is positive and strictly increasing in q. We need the following auxiliary claim.

Lemma 11. *The function* $y(q) = \frac{q-1}{q} \cdot x(q) + \frac{1}{q}$ *is strictly decreasing in* q.

Proof of Lemma 11. Assume w.l.o.g. that $f \geq 0$ and that $\mathbb{E} f = 1$. Let $P = \frac{f}{2^n}$ be a distribution on $\{0,1\}^n$. A simple calculation gives that

$$y(q) = 1 + \frac{1}{n} \cdot \log_2\left(\left(\sum_{a \in \{0,1\}^n} P(a)^q\right)^{\frac{1}{q}}\right),$$

which is strictly decreasing in q, by Hölder's inequality.　□

We proceed with the proof of of the third claim of Lemma 10. Let $q_0 = 1 + (1 - 2\epsilon)^2$. We claim, first, that F is strictly increasing on $q_0 \leq q < \infty$. In fact, for these values of q the second clause of Definition 1 does not apply (by the third claim of Lemma 3) and we have

$$\kappa_{2,q}(x,\epsilon) = \begin{cases} q_0 & \text{if} \quad y \leq \frac{1}{q_0} \\ \frac{\alpha_0 - 1}{\phi_\epsilon(\alpha_0)} & \text{if} \quad y > \frac{1}{q_0} \end{cases},$$

where $y = y(q)$ and α_0 is determined by $1 - \alpha_0 - \frac{\alpha_0 \phi_\epsilon(\alpha_0)}{1-\alpha_0} = y$. The claim now follows by combining Lemma 11, and the third and fourth claims of Lemma 3.

Next, we claim that there exists a unique value $1 \leq q = q^* \leq q_0$ for which $\frac{-x}{\phi_\epsilon(1-x)} = q$ (here $x = x(q)$). Moreover, F decreases for $1 \leq q \leq q^*$ and increases for $q \geq q^*$. Finally, $F(q^*) = q^*$. Observe that verifying these claims will essentially complete the proof of the third claim of Lemma 10 (apart from the fact that $\lim_{\epsilon \to 0} q(f,\epsilon) = 2$).

In fact, by the first and third claims of Lemma 3, and the fact that x is strictly increasing in q, the function $\frac{-x}{\phi_\epsilon(1-x)}$ is strictly decreasing in q, taking values between $\frac{2}{\log_2\left(\frac{4}{q_0}\right)}$ and q_0. This means that it has a unique intersection $q = q^*$ with the function q in $[1, q_0]$. Next, observe that by Definition 1 for $q \leq q_0$ we have

$$\kappa_{2,q}(x,\epsilon) = \begin{cases} -\frac{x}{\phi_\epsilon(1-x)} & \text{if} \quad -\frac{x}{\phi_\epsilon(1-x)} \geq q \\ \frac{\alpha_0 - 1}{\phi_\epsilon(\alpha_0)} & \text{if} \quad -\frac{x}{\phi_\epsilon(1-x)} \leq q \end{cases}$$

This means that for $q < q^*$ we have $F(q) = \kappa_{2,q}(x,\epsilon) = -\frac{x}{\phi_\epsilon(1-x)}$, which is decreasing in q, and for for $q > q^*$ we have $F(q) = \frac{\alpha_0 - 1}{\phi_\epsilon(\alpha_0)}$, which increases in q. Finally, for $q = q^*$, we have $F(q) = -\frac{x}{\phi_\epsilon(1-x)} = q$.

It remains to verify that $\lim_{\epsilon \to 0} q(f,\epsilon) = 2$. By the first claim of Lemma 9, $\phi_0(x) = \frac{x-1}{2}$. This means that for any $0 < x \leq 1$ we have $\lim_{\epsilon \to 0} \frac{-x}{\phi_\epsilon(1-x)} = 2$. The claim follows since, by the preceding discussion, $q = q(f,\epsilon) = \frac{-x(q)}{\phi_\epsilon(1-x(q))}$.　□

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Beckner, W. Inequalities in Fourier Analysis. *Ann. Math.* **1975**, *102*, 159–182. [CrossRef]
2. Bonami, A. Etude des coefficients Fourier des fonctions de $L_p(G)$. *Annales de l'Institut Fourier* **1970**, *20*, 335–402. [CrossRef]
3. Gross, L. Logarithmic Sobolev inequalities. *Am. J. Math.* **1975**, *97*, 1061–1083. [CrossRef]
4. O'Donnell, R. *Analysis of Boolean Functions*; Cambridge University Press: New York, NY, USA, 2014.
5. Samorodnitsky, A. An upper bound on ℓ_q norms of noisy functions. *IEEE Trans. Inf. Theory* **2020**, *66*, 742–748. [CrossRef]
6. Wyner, A.D.; Ziv, J. A theorem on the entropy of certain binary sequences and applications: Part I. *IEEE Trans. Inf. Theory* **1973**, *19*, 769–772. [CrossRef]
7. Cover, T.; Thomas, J. *Elements of Information Theory*; Wiley: Hoboken, NJ, USA, 2006.
8. Rényi, A. On the measures of entropy and information. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20–30 June 1961; Volume 1, pp. 547–561.
9. Kirshner, N.; Samorodnitsky, A. A moment ratio bound for polynomials and some extremal properties of Krawchouk polynomials and Hamming spheres. *IEEE Trans. Inf. Theory* **2021**, *67*, 3509–3541. [CrossRef]
10. Polyanskiy, Y.; Samorodnitsky, A. Improved log-Sobolev inequalities, hypercontractivity and uncertainty principle on the hypercube. *J. Funct. Anal.* **2019**, *277*, 108280. [CrossRef]
11. Yu, L.; Anantharam, V.; Chen, J. Graphs of joint types. In Proceedings of the ISIT, Melbourne, Australia, 12–20 July 2021.
12. Chandar, V.; Tchamkerten, A. Most informative quantization functions. In Proceedings of the ITA Workshop, San Diego, CA, USA, 9–14 February 2014.
13. MacWilliams, J.; Sloane, N.J.A. *The Theory of Error Correcting Codes*; Elsevier: Amsterdam, The Netherlands, 1977.
14. McEliece, R.J.; Rodemich, E.R.; Rumsey, H., Jr.; Welch, L.R. New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities. *IEEE Trans. Inf. Theory* **1977**, *23*, 157–166. [CrossRef]
15. Delsarte, P. An algebraic approach to the association schemes of coding theory. *Philips Res. Rep. Suppl.* **1973**, *10*, vi+–97.
16. Barg, A.; Jaffe, D.B. Numerical results on the asymptotic rate of binary codes. In *Codes and Association Schemes*; Barg, A., Litsyn, S., Eds.; American Mathematical Society: Providence, RI, USA, 2001.
17. Samorodnitsky, A. On the Optimum of Delsarte's Linear Program. *J. Comb. Theory Ser. A* **2001**, *96*, 261–287. [CrossRef]
18. Navon, M.; Samorodnitsky, A. On Delsarte's Linear Programming Bounds for Binary Codes. In Proceedings of the FOCS 2005, Pittsburgh, PA, USA, 23–25 October 2005; pp. 327–338.
19. Friedman, J.; Tillich, J.-P. Generalized Alon-Boppana theorems and error-correcting codes. *SIAM J. Discret. Math.* **2005**, *19*, 700–718. [CrossRef]
20. Samorodnitsky, A. A modified logarithmic Sobolev inequality for the hamming cube and some applications. *arXiv* **2008**, arXiv:0807.1679.
21. Boucheron, S.; Lugosi, G.; Massart, P. *Concentration Inequalities*; Oxford University Press: Oxford, UK, 2013.
22. Donoho, D.L.; Stark, P.B. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.* **1989**, *49*, 906–931. [CrossRef]
23. Kalai, G.; Linial, N. On the distance distribution of codes. *IEEE Trans. Inf. Theory* **1995**, *41*, 1467–1472. [CrossRef]
24. Ashikhmin, A.; Cohen, G.; Krivelevich, M.; Litsyn, S. Bounds on distance distributions in codes of known size. *IEEE Trans. Inf. Theory* **2005**, *51*, 250–258. [CrossRef]
25. Eskenazis, A.; Ivanisvili, P. Polynomial inequalities on the Hamming cube. *Prob. Theory Relat. Fields* **2020**, *178*, 235–287. [CrossRef]
26. Ivanisvili, P.; Tkocz, T. Comparison of moments of Rademacher Chaoses. *Ark. Mat.* **2019**, *57*, 121–128. [CrossRef]
27. Yu, L. Strong Brascamp-Lieb Inequalities. *arXiv* **2021**, arXiv:2102.06935.
28. van Lint, J.H. *Introduction to Coding Theory*, 3rd ed.; Graduate Texts in Mathematics; Springer: Berlin, Germany, 1999; Volume 86.
29. Kalai, G. (The Hebrew University of Jerusalem, Jerusalem, Israel); Linial, N. (The Hebrew University of Jerusalem, Jerusalem, Israel). Personal communication, 1994.