A Fourier analytic proof of the q-ary LP bound *

Sagnik Bhattacharya
Final Year Undergraduate
Department of Electrical Engineering, IIT Kanpur
Kanpur, India
sagnikb@iitk.ac.in

Advisor: Dr Adrish Banerjee
Associate Professor
Department of Electrical Engineering, IIT Kanpur
Kanpur, India
adrish@iitk.ac.in

November 2018

Abstract

We give a Fourier analytic proof of the q-ary linear programming bound on codes for the Hamming metric, extending the technique used to prove the same for the binary case by [NS07]. This method bypasses the Delsarte-MacWilliams inequality approach and it might be possible to generalise this to more general metric like the Lee metric, which is a hard problem by the previously known techniques.

1 Introduction

1.1 Context and Previously known results

A major focus of information and coding theory involves studying the limits of reliable communication over a channel. When the noise is random (the Shannon model), there is a huge body of work starting with Shannon's original paper [Sha48]. In this case, we have a precise characterisation of the capacity of the channel, with any rate below capacity being achievable with probability of error going to zero asymptotically, and conversely, any rate above capacity implying that probability of error is always bounded away from zero. We also know the rate at which the error goes to zero, that is, the error exponents, computationally efficient codes coming close to the capacity for many channels etc.

*Report for EE491A - Undergradute Project (UGP) in the Department of Electrical Engineering, Indian Institute of Technology, Kanpur (IITK)

The problem becomes much harder for adversarial noise (the Hamming model), which is also referred to as zero-error information theory. Even for binary input, binary output channels the best known upper (the second linear programming bound [MR77]) and lower (the Gilbert-Varshamov bound [Gil52]) bounds on the achievable rates don't match. Also, using bounds like the Elias-Bassalygo bound [Ber15][Rot06], we know that rate equal to the random noise capacity is not possible in the adversarial noise setting.

In the adversarial setting, there can be various constraints on the adversary that give rise to various underlying metrics that dictate the capacity of the channel. In the most general case, such a metric can simply be represented by a $q \times q$ matrix where the i, j^{th} entry gives the cost for the adversary to flip the i^{th} symbol to the j^{th} symbol, assuming that the input and output to the channel are both q-ary. We say that a channel is matched to a particular metric if nearest neighbour decoding according to the metric gives the same result as the most probable decoding according to the channel model[Rot06]. By far the simplest metric to analyse is the Hamming metric, which assigns a distance of one to a non-match, but there are other metrics as well, like the Lee metric [CW71] or the Manhattan/taxicab metric [MD12]. The Hamming metric is also the most well-studied metric, and several upper bounds are known for both the binary and the q-ary case with this metric - for example the Hamming bound [Rot06], the Plotkin bound [Rot06], the Elias-Bassalygo bound [Rot06] and the LP (or MRRW) bound [Rot06]. For other metrics not all of these analogous bounds are known - for example, only an analogous Plotkin bound is known for the Lee metric [CW71]. Proving bounds for very general metrics is hard and an attempt to generalise these bounds to slightly more structured metrics like the Lee metric is also an non-trivial problem.

1.2 The Linear Programming Bound

The linear programming bound for binary codes and the Hamming metric was first given in [MR77] using linear programming techniques. Their proof critically uses the linear program for code design introduced by Delsarte in [Del75], and the Delsarte-MacWilliams inequalities. The Delsarte-MacWilliams involve expressing quantities in terms of Krawtchouk polynomials, and it is not clear how this approach can be generalised to non-Hamming metrics - this approach is certainly not easy.

The main appeal of the much more recent Fourier analytic approach to this problem, introduced in [NS07], is that it bypasses the Krawtchouk polynomials in favour of Fourier analytic ideas, that allows a much simpler proof of the LP bound without resorting to linear programming at all. We hope that this method will generalise to non-Hamming metrics.

1.3 Our contribution

In this paper, we generalise the binary Fourier analysis used in [NS07] to give a Fourier analytic proof of the q-ary LP bound. This is a first step towards the ultimate goal of treating more general metrics, because a metric like the Lee metric only becomes interesting for $q > 3$. If we want to handle such cases, we should be able to handle the simpler q-ary Hamming metric.

1.4 Structure of the paper

In Section 2, we give the necessary definitions, state two necessary theorems without proof and show how the main result follows using the two theorems. In the next two sections, we prove the two theorems.

2 Definitions, Lemmas and Main Results

Definition 1 (The Adjacency Matrix of the Hamming cube). *Let the elements of $[q]^n$ be the vertices of a graph. Two vertices of the graph have an edge between them iff the Hamming distance between the corresponding n-length vectors is one. For the graph obtained, A is the adjacency matrix.*

Definition 2 (The maximum eigenvalue of a subset of the Hamming cube). *For $\mathcal{B} \subseteq [q]^n$, we define the maximum eigenvalue of $\mathcal{B}$ as*

$$\lambda_{\mathcal{B}} = \max\left\{\frac{\langle Af, f\rangle}{\langle f, f\rangle}; f : [q]^n \to \mathbb{R}, supp(f) \subseteq \mathcal{B}\right\}$$

Theorem 1. *Let $\mathcal{C}$ be a code that is closed under inverses. Let it have minimal distance d and block length n. Let $\mathcal{B}$ be a subset of $\mathbb{F}_q^n$ with $\lambda_{\mathcal{B}} \geq (q-1)n - qd + 1$. Then,*

$$|\mathcal{C}| \leq n|\mathcal{B}|$$

The previous theorem gives a bound on the size of the code in terms of the size of a subset of the Hamming cube which has its maximum eigenvalue greater than a certain quantity. The next theorem guarantees the existence of a subset of the Hamming cube with a given maximum eigenvalue. We will use both of these to prove the required result.

Theorem 2. *Let $\mathcal{B}(r)$ be a Hamming ball of radius r. Then,*

$$\lambda_{\mathcal{B}(r)} \geq 2\sqrt{q-1}(\sqrt{r(n-r)} - o(n)) + (q-2)(r - o(n))$$

Theorem 3 (The Linear Programming Bound). *A code with relative distance δ has rate upper bounded as follows*

$$R(\delta) \leq H_q\left(\frac{1}{q}\left(q - 1 - (q-2)\delta - 2\sqrt{(q-1)\delta(1-\delta)}\right)\right)$$

Proof. First, we want to choose a value of r in theorem 1 such that the required condition in theorem 2 is satisfied. We first need to solve the equation

$$2\sqrt{q-1}\left(\sqrt{r(n-r)}\right) + (q-2)r = (q-1)n - qd$$

Simplifying, we get,

$$\implies 2\sqrt{q-1}\left(\sqrt{r(n-r)}\right) = (q-1)n - qd - (q-2)r$$

$$\implies 4(q-1)r(n-r) = ((q-1)n - qd)^2 + (q-2)^2r^2 - 2(q-2)((q-1)n - qd)r$$

$$\implies q^2r^2 - 2q\left[(q-1)n - (q-2)d\right] + ((q-1)n - qd)^2 = 0$$

$$\implies r^2 - 2\left[\frac{q-1}{q}n - \frac{q-2}{q}d\right] + \left(\frac{q-1}{q}n - d\right)^2 = 0$$

Which implies

$$r = \frac{2\left[\frac{q-1}{q}n - \frac{q-2}{q}d\right] \pm \sqrt{4\left(\frac{q-1}{q}n - \frac{q-2}{q}d\right)^2 - 4\left(\frac{q-1}{q}n - d\right)^2}}{2}$$
$$= \frac{q-1}{q}n - \frac{q-2}{q} \pm \frac{2}{q}\sqrt{(q-1)d(n-d)}$$

Therefore, if we choose

$$r = \frac{q-1}{q}n - \frac{q-2}{q} - \frac{2}{q}\sqrt{(q-1)d(n-d)} + o(n)$$

then the condition in theorem 1

$$\lambda_{\mathcal{B}} \geq (q-1)n - qd + 1$$

is satisfied. It is well known [Rud] that the size of a Hamming ball is well-approximated by the q-ary entropy function, that is,

Lemma 4 (Size of hamming ball)**.** *For an integer $q \geq 2$ and $p \in [0, 1 - \frac{1}{q}]$,*

$$q^{h_q(p)n - o(n)} \leq Vol_q(n, pn) \leq q^{h_q(p)n}$$

where pn is the radius of the ball.

We use theorem 1 by using the above estimate of the size of the set $\mathcal{B}$ which we are taking to be a Hamming ball, with the same radius as the value of r obtained above to get that

$$|\mathcal{C}| \leq 2^{nH_q\left(\frac{1}{q}\left(q-1-(q-2)\delta - 2\sqrt{(q-1)\delta(1-\delta)}\right) + o(1)\right)}$$

Taking log, this implies

$$R(\delta) \leq H_q\left(\frac{1}{q}\left(q - 1 - (q-2)\delta - 2\sqrt{(q-1)\delta(1-\delta)}\right)\right)$$

which proves the result. □

Observe that for $\delta = 1 - \frac{1}{q}$, $R(\delta) = 0$. This is the same threshold above which the Plotkin bound guarantees asymptotically zero rate.

3 Proof of theorem 1

In this section, we prove theorem 1. All required results from the Fourier analysis of finite groups can be found in appendix A.

We define the following function which allows us to express $A * f$ as a convolution.

Definition 3. *Define the function $L(x)$ as*

$$L(x) = \begin{cases} q^n & \text{if } |x| = 1 \\ 0 & \text{otherwise} \end{cases}$$

Lemma 5. *For a function on $[q]^n$, $Af = f * L$ where A is the adjacency matrix defined earlier.*

Proof. We start from the RHS. Fixing a vertex x, the $Af(x)$ is precisely the sum

$$\sum_{\substack{z\in[q]^n\\ d(x,z)=1}} f(z)$$

Now, consider the RHS.

$$\begin{aligned} f * L = \mathbb{E}_y[L(y)f(x-y)] &= \frac{1}{q^n}\sum_{y\in q^n} L(y)f(x-y) \\ &= \frac{1}{q^n}\sum_{|y|=1} q^n f(x-y) \\ &= \sum_{|y|=1} f(x-y) \end{aligned}$$

The final expression sums the function values at all those points which are at a distance 1 from the point x, and is thus precisely the same sum as obtained from the LHS. □

Lemma 6 (Fourier transform of the L function)**.** $\widehat{L}(S) = n(q-1) - q|S|$

Proof.

$$\widehat{L}(S) = \frac{1}{q^n}\sum_{x\in[q]^n} L(x)\bar{\chi}_S(x) = \sum_{\substack{x\in[q]^n\\ |x|=1}} \bar{\chi}_S(x)$$

We know (for example, from [Con], lemma 3.10) that each character for the group $\mathbb{F}_q^n$ can be expressed as a product of the characters for the group $\mathbb{F}_q$, and the vector S gives us exactly which character of $\mathbb{F}_q$ is present at a particular coordinate of the overall character χ_S. Now, if S is zero at some coordinate j, ie, $S_j = 0$, it means that the character at that coordinate is the trivial character, and such a character contributes $(q-1)$ to the sum (it contributes 1 whenever $x_j \neq 0$, and there are $(q-1)$ such choices.

If $S_j \neq 0$, then the character at that point is non-trivial. We need to evaluate the quantity

$$\sum_{\substack{g\in\mathbb{F}_q\\ g\neq 0}} \chi(g)$$

We know from theorem A.1 that for a non-trivial character χ, $\sum_{g\in\mathbb{F}_q}\chi(g) = 0$ and that $\chi(0) = 1\ \forall\chi$. Therefore, we get that for these coordinates the contribution to the sum is -1. Therefore, the net contribution is

$$\sum_{\substack{g\in\mathbb{F}_q\\ g\neq 0}} \chi(g) = (n-|S|)(q-1) - |S| = n(q-1) - q|S|$$

Note that when we substitute $q = 2$ in this equation, we recover the bound for the binary case given in [NS07]. □

We now have all that we need to prove the theorem, which we restate for convenience.

Theorem 7 (theorem 1 restated)**.** *Let $\mathcal{C}$ be a code that is closed under inverses. Let it have minimal distance d and block length n. Let $\mathcal{B}$ be a subset of $\mathbb{F}_q^n$ with $\lambda_{\mathcal{B}} \geq (q-1)n - qd + 1$. Then,*

$$|\mathcal{C}| \leq n|\mathcal{B}|$$

Proof. Let ϕ be a real function on $[q]^n$ such that

$$(\widehat{\phi})^2 = \mathbb{1}_{\mathcal{C}} * \mathbb{1}_{\mathcal{C}}$$

We then have, by Parseval's theorem (corollary A.4)

$$\widehat{\phi * \phi} = \mathbb{1}_{\mathcal{C}} * \mathbb{1}_{\mathcal{C}}$$

We also have, by Fourier duality (lemma A.6)

$$\phi * \phi = q^n \widehat{\mathbb{1}_{\mathcal{C}} * \mathbb{1}_{\mathcal{C}}} = q^n \widehat{\mathbb{1}_{\mathcal{C}}}^2$$

Note that $\phi * \phi \geq 0$ because of the previous equation and that under the assumption that the code is closed under inverses,

$$\frac{\mathbb{E}\phi^2}{\mathbb{E}^2\phi} = \frac{(\phi * \phi)(0)}{\widehat{\phi * \phi}(0)} = \frac{\frac{|\mathcal{C}|^2}{q^n}}{\frac{|\mathcal{C}|}{q^n}} = |\mathcal{C}|$$

We now consider the adjacency matrix of the subgraph of $[q]^n$ restricted to the vertices of $\mathcal{B}$. Let $f_{\mathcal{B}}$ be an eigenfunction of this matrix with maximum eigenvalue $\lambda_{\mathcal{B}}$. Thus, by definition, $f_{\mathcal{B}}$ is supported on $\mathcal{B}$ and

$$\lambda_{\mathcal{B}} = \frac{\langle Af_{\mathcal{B}}, f_{\mathcal{B}}\rangle}{\langle f_{\mathcal{B}}, f_{\mathcal{B}}\rangle}$$

The matrix A and the function $f_{\mathcal{B}}$ are both non-negative. Now, $Af_{\mathcal{B}} = \lambda_{\mathcal{B}} f_{\mathcal{B}}$ on $\mathcal{B}$ and $Af_{\mathcal{B}} \geq \lambda_{\mathcal{B}} f_{\mathcal{B}}$ outside $\mathcal{B}$. Therefore,

$$Af_{\mathcal{B}} \geq \lambda_{\mathcal{B}} f_{\mathcal{B}}$$

For convenience, we let $\lambda = \lambda_{\mathcal{B}}$ and $f = f_{\mathcal{B}}$ in the rest of the proof.

Choose $F = \phi * f$. We first show that $\mathbb{E}F^2 \leq \mathbb{E}^2 F$. For this, we estimate the inner product $\langle AF, F\rangle$ in two ways.

$$\begin{aligned}
\langle AF, F\rangle &\overset{(a)}{=} \langle (\phi * f) * L, \phi * f\rangle \\
&\overset{(b)}{=} \langle \phi * \phi * f, f * L\rangle \\
&\overset{(c)}{=} \langle \phi * \phi * f, Af\rangle \\
&\overset{(d)}{\geq} \lambda\langle \phi * \phi * f, f\rangle \\
&\overset{(e)}{=} \lambda\langle \phi * f, \phi * f\rangle \\
&\overset{(f)}{=} \lambda\langle F, F\rangle \\
&\overset{(g)}{=} \lambda\mathbb{E}F^2
\end{aligned}$$

Here, (a) follows from the choice of F, lemma 5 and the commutativity of the convolution operator. (b) follows from lemma A.7. (c) again uses the fact that convolution is commutative and lemma 5. (d) uses the fact that $Af_{\mathcal{B}} \geq \lambda_{\mathcal{B}} f_{\mathcal{B}}$, which was shown earlier. (e) follows similarly to (b). (f) again follows from the choice of F, and (g) is clear from the form of the inner product. By the Plancherel formula (theorem A.3),

$$\begin{aligned}\langle AF, F\rangle =& \left\langle \widehat{AF}, \widehat{F}\right\rangle \\ \overset{(a)}{=}& \left\langle \widehat{L}\cdot \widehat{F}, \widehat{F}\right\rangle \\ \overset{(b)}{=}& \left\langle (n(q-1) - q|S|)\widehat{F}, \widehat{F}\right\rangle \\ \overset{(c)}{=}& \sum_S (n(q-1) - q|S|)|\widehat{F}(S)|^2\end{aligned}$$

Here, (a) follows by first using Plancherel's formula (theorem A.3), then 5 and finally taking the convolution to the Fourier domain. (b) follows from lemma 6, and (c) follows from the inner product in the Fourier domain.

Now,

$$F = \phi * f$$

Taking the fourier transform,

$$\begin{aligned}\widehat{F} =& \widehat{\phi}\cdot\widehat{f} \\ \implies |\widehat{F}|^2 =& \widehat{\phi}^2 \cdot |\widehat{f}|^2 = (\mathbb{1}_{\mathcal{C}} * \mathbb{1}_{\mathcal{C}})\cdot \hat{f}^2 \\ =& 0 \text{ for } 0 < |S| < d\end{aligned}$$

The second step follows because $\widehat{\phi}$ is real. The last step follows because the minimum distance of the code is d.Therefore we have,

$$\begin{aligned}\sum_S (n(q-1) - q|S|)\widehat{F}^2(S) =& n\widehat{F}^2(0) + \sum_{|S|\geq d}(n(q-1) - q|S|)\widehat{F}^2(S) \\ \overset{(a)}{\leq}& n\widehat{F}^2(0) + (n(q-1) - qd)\sum_{|S|\geq d}\widehat{F}^2(S) \\ \overset{(b)}{\leq}& n\widehat{F}^2(0) + (n(q-1) - qd)\sum_{S}\widehat{F}^2(S) \\ =& n\mathbb{E}^2[F] + (n(q-1) - qd)\mathbb{E}[F^2]\end{aligned}$$

(a) follows because of the observation immediately above. (b) follows because of the non-negativity of $\widehat{F}^2(S)$. The second term in the last expression follows from Parseval's theorem. Combining the two estimates, we get that

$$\lambda_{\mathcal{B}}\mathbb{E}[F^2] \leq n\mathbb{E}^2[F] + (n(q-1) - qd)\mathbb{E}[F^2]$$

From the value of $\lambda_{\mathcal{B}}$

$$n\mathbb{E}^2[F] \geq \mathbb{E}[F^2]$$

Now,

$$\mathbb{E}^2 F = \mathbb{E}^2(\phi * f) = \mathbb{E}^2\phi \cdot \mathbb{E}^2 f$$

We also have

$$\mathbb{E}F^2 = \langle F, F\rangle = \langle \phi * f, \phi * f\rangle$$

by Plancherel's theorem, theorem A.3

$$\begin{aligned} &= \langle \widehat{\phi * f}, \widehat{\phi * f}\rangle \\ &= \langle \widehat{\phi} \cdot \widehat{f}, \widehat{\phi} \cdot \widehat{f}\rangle \\ &\geq (\phi * \phi)(0)(f * f)(0) \\ &= \frac{1}{q^n}\mathbb{E}\phi^2 \cdot \mathbb{E}f^2 \end{aligned}$$

Now, let a function F be supported on a set U. Then, by Cauchy-Shwarz,

$$\mathbb{E}[F]^2 = \langle F, \mathbb{1}_U\rangle^2 \leq \mathbb{E}[F^2] \cdot \mathbb{E}[\mathbb{1}_U^2] = \frac{|U|}{q^n} \cdot \mathbb{E}[F^2]$$

This implies

$$|\mathcal{B}| \geq q^n \frac{\mathbb{E}^2 f}{\mathbb{E}f^2} \geq \frac{1}{n}\frac{\mathbb{E}\phi^2}{\mathbb{E}^2\phi} = \frac{1}{n}|\mathcal{C}|$$

This completes the proof of the theorem. □

4 Proof of theorem 2

We restate the theorem for convenience.

Theorem 8 (Restatement of theorem 2). *Let $\mathcal{B}(r)$ be a Hamming ball of radius r. Then,*

$$\lambda_{\mathcal{B}(r)} \geq 2\sqrt{q-1}(\sqrt{r(n-r)} - o(n)) + (q-2)(r - o(n))$$

Proof. We give an explicit construction of the function f supported on $\mathcal{B} = \mathcal{B}(r)$ with

$$\frac{\langle Af, f\rangle}{\langle f, f\rangle} \geq \lambda = 2\sqrt{q-1}(\sqrt{r(n-r)} - o(n)) + (q-2)(r - o(n))$$

$f \geq 0$ and $Af \geq \lambda f$. The function will be *strongly symmetric*, that is, it will be constant on all inputs with the same Hamming weight. The function is thus uniquely defined when we define it on all integers $0, 1, \ldots, n$.

Lemma 9. *For a strongly symmetric function g on $[q]^n$, we have*

$$Ag(i) = ig(i-1) + (q-2)ig(i) + (q-1)(n-i)g(i+1)$$

Proof. For the proof, it will be helpful to think of the function f as a q^n-long vector instead of thinking of it as being defined on the integers as earlier [1]. For a general function g on $[q]^n$, the quantity $Ag(x)$ can be written as

$$Ag(x) = \sum_{|y|=1} g(x+y)$$

That is, we are summing the values of the function g over all vectors at a (Hamming) distance of 1 from the vector under consideration. We can split this into three separate cases.

- In the first case, adding the y to x lowers the Hamming weight of x by 1. This can only happen if the vector y takes, at an index where x is non-zero, a value that when added to that entry of x makes that entry non-zero, and there is exactly one choice for that entry of the y vector for each non-zero x entry.We know that for a strongly symmetric function, the value depends only on the Hamming weight of the argument. Therefore, y's like this contribute $ig(i-1)$ to the sum.
- In the other extreme case, adding y to x increases the Hamming weight of x by one. This will happen if y takes a non-zero value at a point where x is zero. There are then $(q-1)$ choices for each non-zero entry of the x vector. Therefore, y's like this contribute $(q-1)(n-i)g(i+1)$ to the sum.
- In the intermediate case, adding y to x doesn't change the Hamming weight. This happens if, at a point where x is non-zero, y takes a non-zero value that *is not the value that makes the entry there zero.* Therefore, there are $(q-2)$ choices for each non-zero entry of the x vector that leads to such a situation. The contribution of this to the sum is $(q-2)ig(i)$.

Putting all of these contributions together, we get the required result. □

Let $M = \sqrt{n} = o(n)$ and define a function

$$f(i) = \begin{cases} \frac{1}{\sqrt{\binom{n}{i}(q-1)^i}} & \text{for } i \in [r-M, r] \\ 0 & \text{otherwise} \end{cases}$$

Lemma 10.

$$\frac{f(i)}{f(i+1)} = \sqrt{\frac{(n-i)(q-1)}{(i+1)}}$$

Using this to express $f(i-1)$ and $f(i+1)$ in terms of $f(i)$ and substituting in the recurrence relation obtained in lemma 9, we get that

$$Af(i) = \sqrt{(q-1)i(n-i+1)}f(i) + (q-2)if(i) + \sqrt{(q-1)(i+1)(n-i)}f(i)$$

for all values of i other than $i = r - M$ and $i = r$, since at these places $f(i-1)$ and $f(i+1)$ respectively are undefined. Still, by the construction of f we are assured that

[1] We can order the elements of $\mathbb{F}_q^n$ lexicographically, and the element of the vector corresponding to a particular point of $\mathbb{F}_q^n$ is the value of the function at that point.

it takes positive values at these points. In what follows, it will again be helpful to think of A as a matrix and f as a vector. Since f is strongly symmetric, for $\binom{n}{k}(q-1)^k$ many values of the input for which the function value is fixed for a give k, the contribution to the product $f^\dagger Af$ is $\left(\sqrt{(q-1)k(n-k+1)}+(q-2)k+\sqrt{(q-1)(k+1)(n-k)}\right)f(k)^2$. This implies that

$$\begin{aligned} f^\dagger Af &\geq \sum_{k=r-M+1}^{r-1}\binom{n}{k}(q-1)^k\left(\sqrt{(q-1)k(n-k+1)}+(q-2)k+\sqrt{(q-1)(k+1)(n-k)}\right)f(k)^2 \\ &= \sum_{k=r-M+1}^{r-1}\left(\sqrt{(q-1)k(n-k+1)}+(q-2)k+\sqrt{(q-1)(k+1)(n-k)}\right) \end{aligned}$$

Note that

$$\sqrt{k(n-k+1)},\sqrt{(n-k)(k+1)} \geq \sqrt{(r-M)(n-r)} \geq \sqrt{r(n-r)}-M = \sqrt{r(n-r)}-o(n)$$

and therefore,

$$f^\dagger Af \geq 2(M-1)\sqrt{(q-1)}\left(\sqrt{r(n-r)}-o(n)\right)+(q-2)\sum_{k=r-M+1}^{r-1}k$$

Now,

$$\begin{aligned} \sum_{k=r-M+1}^{r-1}k &= \frac{M-1}{2}(2r-M) = \frac{M-1}{2}(2r-o(n)) \\ &= (M-1)(r-o(n)) \end{aligned}$$

Plugging this back into the last equation, we therefore get

$$f^\dagger Af \geq 2(M-1)\sqrt{(q-1)}\left(\sqrt{r(n-r)}-o(n)\right)+(q-2)(M-1)(r-o(n))$$

Also,

$$\begin{aligned} \langle f,f\rangle &= \sum_{k=[r-M,r]}\sum_{wt(x)=k}\frac{1}{\binom{n}{i}(q-1)^i} \\ &= \sum_{k=[r-M,r]}\frac{\binom{n}{i}(q-1)^i}{\binom{n}{i}(q-1)^i} = M \end{aligned}$$

Therefore,

$$\frac{\langle Af,f\rangle}{\langle f,f\rangle} = 2\sqrt{q-1}\left(\sqrt{r(n-r)}-o(n)\right)+(q-2)(r+o(n))$$

which proves the theorem. □

5 Open Questions and Future Work

We needed to assume that the code is closed under inverses for the proof to work. One future line of investigation is whether this is necessary, or if we can show that in some domains of interest, this assumptions does not change the achievable rate. Also, a major motivation for investigating this proof technique is that we want to find analogous bounds for other metrics. Two immediate candidates are the Lee metric, investigated in [CW71], and the Manhattan metric. For the Lee metric, a Plotkin-type bound is known. We have already observed that for the Hamming metric, the Plotkin bound and the LP bound give the same zero-rate threshold. It is interesting to investigate whether this is true for the Lee metric too. LP-type bounds for more general metrics than any of these remain open problems.

Appendices

A Fourier Analysis on Finite Abelian Groups

In this section we record the basic definitions and theorems in the Fourier analysis of finite Abelian groups. We will confine our attention mostly to the group $\mathbb{F}_q^n$, but some initial results will be stated in more generality. Most of the contents of this appendix can be found in [Con]. The section makes this write-up self contained and includes only those theorems and definitions that are used in the main proof.

$\mathbb{F}_q^n$ is a finite Abelian group, and therefore its characters $\{\chi_S\}_{S\in\mathbb{F}_q^n}$ form a group (called the dual group) that is isomorphic to $\mathbb{F}_q^n$. Characters are homomorphisms from the group $\mathbb{F}_q^n$ to $\mathbb{C}^\times$, and they form an orthonormal basis for the space of real-valued functions $f : \mathbb{F}_q^n \to \mathbb{R}$ equipped with uniform probability distribution.

Theorem A.1. *If G is a finite abelian group,*

$$\sum_{g\in G} \chi(g) = \begin{cases} |G| & \text{if } \chi = \mathbb{1}_{\widehat{G}} \\ 0 & \text{otherwise} \end{cases}$$

Proof. Let

$$S = \sum_{g\in G} \chi(g)$$

If $\chi = \mathbb{1}_{\widehat{G}}$ then $S = |G|$. If $\chi \neq \mathbb{1}_G$, then there exists $g' \in G$ such that $\chi(g) \neq 0$. Then,

$$\chi(g')S = \sum_{g\in G} \chi(gg') = \sum_{g\in G} \chi(g) = S \implies S = 0$$

If $\chi = \mathbb{1}_G$, then the result follows immediately. □

Corollary A.2. *If χ_1 and χ_2 are characters of a finite abelian group G, then*

$$\sum_{g\in G} \chi_1(g)\bar{\chi_2}(g) = \begin{cases} |G| & \text{if } \chi_1 = \chi_2 \\ 0 & \text{otherwise} \end{cases}$$

Proof. Follows by taking $\chi = \chi_1 \bar{\chi}_2$ in the theorem. □

Definition 4 (Expected value of a function on $\mathbb{F}_q^n$)**.**

$$\mathbb{E}[f] = \frac{1}{q^n} \sum_{x \in \mathbb{F}_q^n} f(x)$$

Definition 5 (Inner product in function domain)**.** *Define the inner product of two functions f and g as*

$$\langle f, g \rangle = \mathbb{E}[f \cdot \bar{g}]$$

Definition 6 (Fourier Transform of a function)**.** *For $f : \mathbb{F}_q^n \to \mathbb{R}$, define the Fourier transform of f, $\widehat{f} : \mathbb{F}_q^n \to \mathbb{R}$ as*

$$\widehat{f}(S) = \langle f, \chi_S \rangle$$

This means that (using corollary A.2)

$$f(x) = \sum_S \widehat{f}(S) \chi_S$$

Definition 7 (Inner Product in Fourier domain)**.** *In the Fourier space, define the inner product as*

$$\langle \widehat{f}, \widehat{g} \rangle = \sum_S \widehat{f}(S) \bar{\widehat{g}}(S)$$

Theorem A.3 (Plancherel's Theorem)**.**

$$\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle$$

Proof. We start from the LHS.

$$\langle f, g \rangle = \left\langle \sum_S \widehat{f}(S) \chi_S, \sum_T \widehat{g}(T) \chi_T \right\rangle = \sum_{S,T} \widehat{f}(S) \widehat{g}(T) \langle \chi_S, \chi_T \rangle$$

From corollary A.2 and definition 5, we get

$$\langle f, g \rangle = \sum_S \widehat{f}(S) \widehat{g}(S) = \langle \widehat{f}, \widehat{g} \rangle$$

□

Corollary A.4 (Parseval's Formula)**.**

$$\langle f, f \rangle = \sum_S \widehat{f}(S)^2$$

Definition 8 (Convolution)**.** *The convolution of two functions f and g is defined as*

$$f * g(x) = \mathbb{E}_{y \in [q]^n} [f(y) g(x - y)]$$

Theorem A.5. *Convolution in function domain gives multiplication in Fourier domain*

$$\widehat{f * g}(S) = \widehat{f}(S) \widehat{g}(S)$$

Proof. □

Lemma A.6. *If $\widehat{f} = g$ then $f = q^n \widehat{g}$*

Proof.

$$\begin{aligned}
\widehat{f}(S) &= \frac{1}{q^n} \sum_{x \in [q]^n} f(x)\bar{\chi}_S(x) \\
\implies g(S) &= \frac{1}{q^n} \sum_{x \in [q]^n} f(x)\bar{\chi}_S(x) \\
\implies \widehat{g}(S') &= \frac{1}{q^n} \sum_S \left[\frac{1}{q^n} \sum_x f(x)\bar{\chi}_S(x) \right] \chi_{S'}(s)
\end{aligned}$$

Everything is finite so we can interchange the sums

$$\begin{aligned}
\implies \widehat{g}(S') &= \frac{1}{q^{2n}} \sum_x f(x) \sum_S \bar{\chi}_S(x)\chi_{S'}(s) \\
&= \frac{1}{q^n} f(S')
\end{aligned}$$

which proves the result. □

Lemma A.7 (Interchange Lemma)**.** *If A, B, C and D are functions from $\mathbb{F}_q^n$ to $\mathbb{R}$, and the Fourier coefficients of A and C are real, then*

$$\langle A * B, C * D \rangle = \langle B, A * C * D \rangle = \langle C * A * B, D \rangle$$

Essentially, this means that if the inner product is of the form shown, then functions with real fourier coefficients can be taken to the other side of the inner product without changing the value of the inner product.

Proof. We show one of the equalities. The other one follows similarly. Using Parseval's theorem, we can write,

$$\begin{aligned}
\langle A * B, C * D \rangle &= \langle \widehat{A * B}, \widehat{C * D} \rangle \\
&= \langle \widehat{A} \cdot \widehat{B}, \widehat{C} \cdot \widehat{D} \rangle
\end{aligned}$$

Using the definition of the inner product,

$$= \sum_S \widehat{A}(S)\widehat{B}(S)\bar{\widehat{C}}(S)\bar{\widehat{D}}(S)$$

By assumption, $\widehat{A}$ is real

$$\begin{aligned}
&= \sum_S \widehat{B}(S)\widehat{A}\bar{(}S)\widehat{C}\bar{(}S)\widehat{D}\bar{(}S) \\
&= \langle \widehat{B}, \widehat{A} \cdot \widehat{C} \cdot \widehat{D} \rangle \\
&= \langle \widehat{B}, A \widehat{* C *} D \rangle
\end{aligned}$$

Using Plancherel's theorem again,

$$= \langle B, A * C * D \rangle$$

□

References

[Ber15] Elwyn R. Berlekamp. *Algebraic Coding Theory - Revised Edition.* World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2015.

[Con] Keith Conrad. Expository papers - characters of finite abelian groups. `http://www.math.uconn.edu/~kconrad/blurbs/grouptheory/charthy.pdf`. Accessed: 2018-10-10.

[CW71] J. Chung-Yaw Chiang and Jack K. Wolf. On channels and codes for the lee metric. *Information and Control*, 19(2):159 – 173, 1971.

[Del75] P. Delsarte. The association schemes of coding theory. In M. Hall and J. H. van Lint, editors, *Combinatorics*, pages 143–161, Dordrecht, 1975. Springer Netherlands.

[Gil52] E. N. Gilbert. A comparison of signalling alphabets. *The Bell System Technical Journal*, 31(3):504–522, May 1952.

[MD12] Krassimir Markov (eds.) Michel Deza, Michel Petitjean. Mathematics of distances and applications. 2012.

[MR77] Rumsey H. ; Welch L. McEliece R., Rodemich E. New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities. *IEEE Transactions on Information Theory*, March 1977.

[NS07] M. Navon and A. Samorodnitsky. Linear programming bounds for codes via a covering argument. *ArXiv Mathematics e-prints*, February 2007.

[Rot06] Ron Roth. *Introduction to Coding Theory.* Cambridge University Press, New York, NY, USA, 2006.

[Rud] Atri Rudra. Error correcting codes: Combinatorics, algorithms and applications - lecture notes - fall 2007. `https://cse.buffalo.edu/faculty/atri/courses/coding-theory/lectures/lect9.pdf`. Accessed: 2018-10-10.

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.