Lecture 4

Applications of Harmonic Analysis

February 4, 2005
Lecturer: Nati Linial
Notes: Matthew Cary

4.1 Useful Facts

Most of our applications of harmonic analysis to computer science will involve only Parseval's identity.

Theorem 4.1 (Parseval's Identity).

$$\|f\|_2 = \|\hat{f}\|_2$$

Corollary 4.2.

$$\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle.$$

Proof. Note that $\langle f+g, f+g\rangle = \|f+g\|_2 = \|\widehat{f+g}\|_2 = \|\hat{f}+\hat{g}\|_2$. Now as $\langle f+g, f+g\rangle = \|f\|_2^2 + \|g\|_2^2 + 2\langle f, g\rangle$, and similarly $\|\hat{f}+\hat{g}\|_2^2 = \|\hat{f}\|_2^2 + \|\hat{g}\|_2^2 + 2\langle \hat{f}, \hat{g}\rangle$, applying Parseval to $\|f\|_2$ and $\|g\|_2$ and equating finishes the proof. □

The other basic identity is the following.

Lemma 4.3.

$$\widehat{f * g} = \hat{f} \cdot \hat{g}$$

Proof. We will show this for the unit circle $\mathbb{T}$, but one should note that it is true more generally. Recall that by definition $h = f * g$ means that

$$h(t) = \frac{1}{2\pi}\int_{\mathbb{T}} f(s)g(t-s)ds.$$

Now to calculate $\widehat{f * g}$ we manipulate $\hat{h}$.

$$\begin{aligned}
\hat{h}(r) &= \frac{1}{2\pi}\int_{\mathbb{T}} h(x)e^{-irx}dx \\
&= \frac{1}{4\pi^2}\iint_{\mathbb{T}^2} f(s)g(x-s)e^{-irx}ds\,dx \\
&= \frac{1}{4\pi^2}\iint_{\mathbb{T}^2} f(s)g(x-s)e^{-irs}e^{-ir(x-s)}dx\,ds
\end{aligned}$$

using $e^{-irx} = e^{-irs}e^{-ri(x-s)}$ and interchanging the order of integration. Then by taking $u = x - s$ we have

$$\begin{aligned}
&= \frac{1}{4\pi^2}\iint_{\mathbb{T}^2} f(s)g(u)e^{-irs}e^{-iru}du\,ds \\
&= \frac{1}{4\pi^2}\int_{\mathbb{T}} f(s)e^{-irs}ds \int_{\mathbb{T}} g(u)e^{-iru}du \\
&= \left(\frac{1}{2\pi}\int_{\mathbb{T}} f(s)e^{-irs}ds\right)\left(\frac{1}{2\pi}\int_{\mathbb{T}} g(u)ds\right) \\
&= \hat{f}\cdot\hat{g}.
\end{aligned}$$

□

4.2 Hurwitz's Proof of the Isoperimetric Inequality

Recall from last lecture that the isoperimetric problem is to show that a circle encloses the largest area for all curves of a fixed length. Formally, if L is the length of a curve and A the area enclosed, then we want to show that $L^2 - 4\pi A \geq 0$ with equality if and only if the curve is a circle. We will prove the following stronger theorem.

Theorem 4.4. *Let $(x, y) : \mathbb{T} \to \mathbb{R}^2$ be an anticlockwise arc length parametrization of a non self-intersecting curve Γ of length L enclosing an area A. If $x, y \in C^1$, then*

$$L^2 - 4\pi A = 2\pi^2\Big(\sum_{n\neq 0} |n\hat{x}(n) - i\hat{y}(n)|^2 + |n\hat{y}(n) + i\hat{x}(n)|^2 + (n^2-1)\big(|\hat{x}(n)|^2 + |\hat{y}(n)|^2\big)\Big).$$

In particular, $L^2 \geq 4\pi A$ with equality if and only if Γ is a circle.

We will not define "arc length parameterization" formally, only remark that intuitively it means that if one views the parameterization as describing the motion of a particle in the plane, then an arc length parameterization is one so that the speed of the particle is constant. In our context, where we view time as the unit circle $\mathbb{T}$ of circumference 2π, we have that $(\dot{x})^2 + (\dot{y})^2$ is a constant so that the total distance covered is $\left(\frac{L}{2\pi}\right)^2$.

Proof. First we use our identity about the parameterization to relate the length to the transform of the parameterization.

$$\begin{aligned}
\left(\frac{L}{2\pi}\right)^2 &= \frac{1}{2\pi}\int_{\mathbb{T}} \left(\big(\dot{x}(s)\big)^2 + \big(\dot{y}(s)\big)^2\right) ds \\
&= \|\hat{\dot{x}}\|_2^2 + \|\hat{\dot{y}}\|_2^2 && \text{by Parseval's} \\
&= \sum_{-\infty}^{\infty} |in\hat{x}(n)|^2 + |in\hat{y}(n)|^2 && \text{by Fourier differentiation identities} \\
&= \sum_{-\infty}^{\infty} -n^2\big(|\hat{x}(n)|^2 + |\hat{y}(n)|^2\big) && (4.1)
\end{aligned}$$

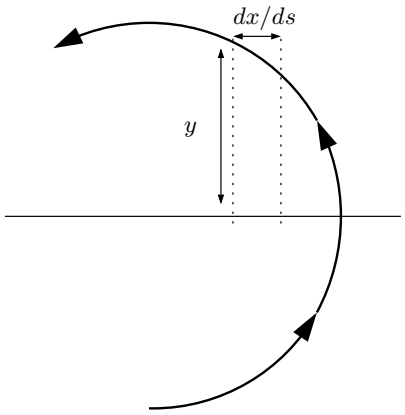

Figure 4.1: Computing the area enclosed by a curve

Now we compute the area. As the curve is anticlockwise,

$$A = -\int y \frac{dx}{ds} ds,$$

where the negative sign comes from the fact that the curve is anticlockwise. See Figure 4.1. This area integral looks like an inner product, so we write

$$\frac{A}{2\pi} = -\langle y, \dot{x} \rangle = -\langle \hat{y}, \hat{\dot{x}} \rangle.$$

By symmetry, considering the area integral from the other direction, we also have that

$$\frac{A}{2\pi} = \langle \hat{x}, \hat{\dot{y}} \rangle,$$

note there is no negative sign in this expression. Hence by adding we have that

$$\begin{aligned} \frac{A}{\pi} &= \langle \hat{x}, \hat{\dot{y}} \rangle - \langle \hat{y}, \hat{\dot{x}} \rangle \\ &= \sum_{-\infty}^{\infty} in\big(\hat{x}(n)^* \hat{y}(n) - \hat{x}(n)\hat{y}(n)^*\big), \end{aligned} \tag{4.2}$$

using the Fourier differentiation identities and using the notation a^* for the complex conjugate of a. Now subtract (4.1) and (4.2) to prove the theorem.

To see why the right hand side is zero if and only if Γ is a circle, consider when the right hand side vanishes. As it is a sum of many squares, $\hat{x}$ and $\hat{y}$ must vanish for all $n \neq 0$ or ± 1. Looking carefully at what those terms mean shows that they vanish if and only if Γ is a circle. □

4.3 Harmonic Analysis on the Cube for Coding Theory

The theory of error correcting codes is broad and has numerous practical applications. We will look at the asymptotic theory of block coding, which like many problems in coding theory is well-known, has a long

history and is still not well understood. The Boolean or Hamming cube $\{0,1\}^n$ is the set of all n-bit strings over $\{0,1\}$. The usual distance on $\{0,1\}^n$ is the *Hamming distance* $d_H(x,y)$, defined over $x, y \in \{0,1\}^n$ by the number of positions where x and y are not equal: $d_H(x,y) = |\{i : x_i \neq y_i\}|$. A code $\mathcal{C}$ is a subset of $\{0,1\}^n$. The *minimum distance* of $\mathcal{C}$ is the minimum distance between any two elements of $\mathcal{C}$:

$$\mathrm{dist}(\mathcal{C}) = \min\{d_H(x,y) : x, y \in \mathcal{C}\}.$$

The asymptotic question is to estimate the size of the largest code for any given minimum distance,

$$A(d,n) = \max\{|\mathcal{C}|, \mathcal{C} \subset \{0,1\}^n, \mathrm{dist}(\mathcal{C}) \geq d\},$$

as $n \to \infty$. The problem is easier if we restrict the parameter space by fixing d to be a constant fraction of the bit-length n, that is, consider $A(\delta n, n)$. Simple constructions show for $1/2 > \delta > 0$ that $A(\delta n, n)$ is exponential in n, so the interesting quantity will be the bit-rate of the code. Accordingly, we define the *rate* of a code as $\frac{1}{n}\log|\mathcal{C}|$, and then define the asymptotic rate limit as

$$R(\delta) = \limsup_{n\to\infty} \left\{ \frac{1}{n}\log|\mathcal{C}| : \mathcal{C} \subset \{0,1\}^n, \mathrm{dist}(\mathcal{C}) \geq \delta n \right\}.$$

It is a sign of our poor knowledge of the area that we do not even know if in the above we can replace the $\limsup$ by lim, i.e., if the limit exists. If $|\mathcal{C}| = 2^k$, we may think of the code as mapping k-bit strings into n-bit strings which are then communicated over a channel. The rate is then the ratio k/n, and measures the efficiency with which we utilize the channel.

A code is *linear* if $\mathcal{C}$ is a linear subspace of $\{0,1\}^n$, viewed as a vector space over $\mathbf{GF}(2)$. In a linear code, if the minimum distance is realized by two codewords x and y, then $x - y$ is a codeword whose (Hamming) length equals the minimum distance. Hence for linear codes we have that

$$\mathrm{dist}(\mathcal{C}) = \min\big\{|w| : w \in \mathcal{C} \setminus \{0\}\big\}.$$

Here we use $|\cdot|$ to indicate the *Hamming weight* of a codeword, the number of nonzero positions. Note that this is equal to several other, common norms on $\mathbf{GF}(2)$.

A useful entity is the *orthogonal code* to a given code. If $\mathcal{C}$ a linear code, we define

$$\mathcal{C}^\perp = \{y : \forall x \in \mathcal{C}, \langle x, y\rangle = 0\},$$

where we compute the inner product $\langle\cdot,\cdot\rangle$ over $\mathbf{GF}(2)$, that is, $\langle x, y\rangle = \sum_{i=1}^n x_i y_i \pmod 2$.

4.3.1 Distance Distributions and the MacWilliams Identities

Our first concrete study of codes will be into the *distance distribution*, which are the probabilities

$$\Pr[|x - y| = k : x, y \text{ chosen randomly from } \mathcal{C}]$$

for $0 \leq k \leq n$. If $\mathcal{C}$ is linear, our discussion above shows that the question of distance distribution is identical to the weight distribution of a code, the probabilities that a randomly selected codeword has a specified weight.

The MacWilliams Identities are important identities about this distribution that are easily derived using Parseval's Identity. Let $f = 1_{\mathcal{C}}$, the indicator function for the code. We first need the following lemma.

Lemma 4.5.

$$\hat{f} = \frac{|\mathcal{C}|}{2^n} 1_{\mathcal{C}^\perp}$$

Proof.

$$\begin{aligned}\hat{f}(u) &= \frac{1}{2^n} \sum_v f(v)\chi_v(u) \\ &= \frac{1}{2^n} \sum_v f(v)(-1)^{\langle u,v\rangle} \\ &= \frac{1}{2^n} \sum_{v\in\mathcal{C}} (-1)^{\langle u,v\rangle}\end{aligned}$$

If $u \in \mathcal{C}^\perp$, then $\langle u, v\rangle = 0$ for all $v \in \mathcal{C}$, so that $\hat{f}(u) = |\mathcal{C}|/2^n$. Suppose otherwise, so that $\sum_{\mathcal{C}}(-1)^{\langle u,v\rangle} = |\mathcal{C}_0| - |\mathcal{C}_1|$, where $\mathcal{C}_0$ are the codewords of $\mathcal{C}$ that are perpendicular to u, and $\mathcal{C}_1 = \mathcal{C} \setminus \mathcal{C}_1$. As $u \notin \mathcal{C}^\perp$, $\mathcal{C}_1$ is nonempty. Pick an arbitrary w in $\mathcal{C}_1$. Then, any $y \in \mathcal{C}_1 \setminus \{w\}$ corresponds to a unique $x \in \mathcal{C}_0$, namely $w + y$. Similarly, any $x \in \mathcal{C}_0 \setminus \{0\}$ corresponds to $w + x \in \mathcal{C}_1 \setminus w$. As $w \in \mathcal{C}_1$ corresponds to $0 \in \mathcal{C}_0$, we have that $|\mathcal{C}_0| = |\mathcal{C}_1|$. Hence $\sum_{\mathcal{C}}(-1)^{\langle e_1,v\rangle} = 0$, so that

$$\hat{f}(u) = \begin{cases} |\mathcal{C}|/2^n & \text{if } u \in \mathcal{C}^\perp \\ 0 & \text{otherwise} \end{cases}$$

which proves the lemma. □

We now define the *weight enumerator* of a code to be

$$P_{\mathcal{C}}(x, y) = \sum_{w\in\mathcal{C}} x^{|w|} y^{n-|w|}.$$

The MacWilliams Identity connects the weight enumerators of $\mathcal{C}$ and $\mathcal{C}^\perp$ for linear codes.

Theorem 4.6 (The MacWilliams Identity).

$$P_{\mathcal{C}}(x, y) = |\mathcal{C}| P_{\mathcal{C}^\perp}(y - x, y + x)$$

Proof. Harmonic analysis provides a nice proof of the identity by viewing it as an inner product. Define $f = 1_{\mathcal{C}}$ and $g(w) = x^{|w|} y^{n-|w|}$. Then, using Parseval's,

$$P_{\mathcal{C}}(x, y) = 2^n \langle f, g\rangle = 2^n \langle \hat{f}, \hat{g}\rangle.$$

$\hat{f}$ has already been computed in Lemma 4.5, so we turn our attention to $\hat{g}$.

$$\begin{aligned}\hat{g}(u) &= \frac{1}{2^n} \sum_v g(v)(-1)^{\langle u,v\rangle} \\ &= \frac{1}{2^n} \sum_v x^{|v|} y^{n-|v|} (-1)^{\langle u,v\rangle}\end{aligned}$$

Let u have k ones and $n-k$ zeros. For a given v, let s be the number of ones of v that coincide with those of u, and let t be the number ones of v coinciding with the zeros of u. Then we rewrite the sum as

$$\begin{aligned}
&= \frac{1}{2^n} \sum_{s,t,k} \binom{k}{s} \binom{n-k}{t} x^{s+t} y^{n-s-t} (-1)^s \\
&= \frac{y^n}{2^n} \sum_{s} \binom{k}{s} \left(-\frac{x}{y}\right)^s \sum_{t} \binom{n-k}{t} \left(\frac{x}{y}\right)^t \\
&= \frac{y^n}{2^n} \left(1 - \frac{x}{y}\right)^k \left(1 + \frac{x}{y}\right)^{n-k} \\
&= \frac{1}{2^n} (y-x)^k (y+x)^{n-k} \\
&= \frac{1}{2^n} (y-x)^{|u|} (y+x)^{n-|u|}
\end{aligned}$$

Now, as $\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle = 2^{-n} P_{\mathcal{C}}(x, y)$, we plug in our expressions for $\hat{f}$ and $\hat{g}$ to get

$$\begin{aligned}
2^{-n} P_{\mathcal{C}}(x, y) &= \frac{|\mathcal{C}|}{2^n} \sum_{w \in \mathcal{C}^\perp} \frac{1}{2^n} (y-x)^{|w|} (y+x)^{n-|w|} \\
&= \frac{|\mathcal{C}|}{2^n} P_{\mathcal{C}^\perp}(y-x, y+x),
\end{aligned}$$

which implies

$$P_{\mathcal{C}} = |\mathcal{C}| P_{\mathcal{C}^\perp}(y-x, y+x).$$

□

4.3.2 Upper and Lower Bounds on the Rate of Codes

We now turn our attention to upper and lower bounds for codes. We remind any complexity theorists reading these notes that the senses of "upper bound" and "lower bound" are reversed from their usage in complexity theory. Namely, a lower bound on $R(\delta)$ shows that good codes exist, and an upper bound shows that superb codes don't.

In the remainder of this lecture we show several simple upper and lower bounds, an then set the stage for the essentially strongest known upper bound on the rate of codes, the MacEleiece, Rumsey, Rodemich and Welsh (MRRW) upper bound. This is also referred to as the JPL bound, after the lab the authors worked in, or the linear programming (LP) bound, after its proof method.

Our first bound is a lower bound. Recall the binary entropy function

$$H(x) = -x \log x - (1-x) \log(1-x).$$

Theorem 4.7 (Gilbert-Varshamov Bound).

$$R(\delta) \geq 1 - H(\delta),$$

and there exists a linear code satisfying the bound.

Proof. We will sequentially pick codewords where each new point avoids all δn-spheres around previously selected points. The resulting code $\mathcal{C}$ will satisfy

$$|\mathcal{C}| \geq \frac{2^n}{\text{vol(sphere of radius } \delta n)} = \frac{2^n}{\sum_{j=0}^{\delta n} \binom{n}{j}}.$$

Now, note that $\log \binom{n}{\alpha n}/n \to H(\alpha)$ as $n \to \infty$, so that $2^n / \sum^{\delta n} \binom{n}{j} \sim 2^{n(1-H(\delta))}$, and take logs to prove the first part of the theorem.

We now show that there's a linear code satisfying this rate bound. This proof is different than the one given in class, as I couldn't get that to work out. The presentation is taken from Trevison's survey of coding theory for computational complexity. We can describe a linear k-dimensional code $\mathcal{C}_A$ by a $k \times n$ 0-1 matrix A by $\mathcal{C}_A = \{Ax : x \in \{0,1\}^k\}$. We'll show that if $k/n \leq 1 - H(\delta)$, with positive probability $\text{dist}(\mathcal{C}_A) \geq \delta n$. As the code is linear, it suffices to show that the weight of all nonzero codewords is at least δn. As for a given $x \in \{0,1\}^k$, Ax is uniformly distributed over $\{0,1\}^n$, we have

$$\Pr[|Ax| < \delta n] = 2^{-n} \sum_{i=0}^{\delta n - 1} \binom{n}{i} \leq 2^{-n} 2^{nH(\delta)+o(n)},$$

using our approximation to the binomial sum. Now we take a union bound over all 2^k choices for x to get

$$\Pr[\exists x \neq 0 : Ax < d] \leq 2^k \cdot 2^{-n} \cdot 2^{nH(\delta)+o(n)} = 2^{k+n(H(\delta)-1)+o(1)} < 1$$

by our choice of $k \leq n(1 - H(\delta))$. □

We now turn to upper bounds on $R(\delta)$.

Theorem 4.8 (Sphere-Packing Bound).

$$R(\delta) \leq 1 - H(\delta/2)$$

Proof. The theorem follows from noting that balls of radius $\delta n/2$ around codewords must be disjoint, and applying the approximations used above for the volume of spheres in the cube. □

We note in Figure 4.2 that the sphere-packing bound is far from the GV bound. In particular, that GV bound reaches zero at $\delta = 1/2$, while the sphere-packing bound is positive until $\delta = 1$. However, we have the following simple claim.

Claim 4.1. $R(\delta) = 0$ for $\delta > 1/2$.

Proof. We will show the stronger statement that if $|\mathcal{C}|$ is substantial then not only is it impossible for $d_H(x,y) > \delta n$ for all $x, y \in \mathcal{C}$, but even the average of all $x, y \in \mathcal{C}$ will be at most $n/2$. This average distance is

$$\frac{1}{\binom{|\mathcal{C}|}{2}} \sum_{\mathcal{C} \times \mathcal{C}} d(x,y),$$

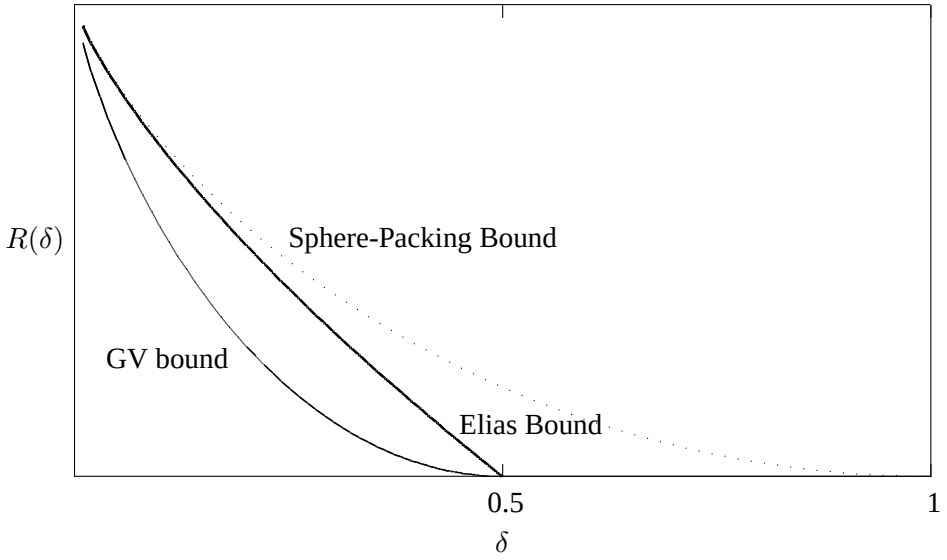

Figure 4.2: The GV bound contrasted with the Sphere-Packing Bound

and we will expand the distance $d(x, y) = |\{i : x_i \neq y_i\}|$. Reversing the order of summation,

$$\begin{aligned}\text{Average distance} &= \frac{1}{\binom{|\mathcal{C}|}{2}} \sum_i \sum_{x,y} 1_{x_i \neq y_i} \\ &= \frac{1}{\binom{|\mathcal{C}|}{2}} \sum_i z_i(|\mathcal{C}| - z_i),\end{aligned}$$

where z_i is the number of zeros in the i^{th} position of all the codewords of $\mathcal{C}$.

$$\begin{aligned}&\leq \frac{1}{\binom{|\mathcal{C}|}{2}} \sum \frac{|\mathcal{C}|^2}{4} \\ &\leq \frac{1}{2} n \cdot \frac{|\mathcal{C}|}{|\mathcal{C}| - 1}.\end{aligned}$$

So unless $\mathcal{C}$ is very small, the average distance is essentially $n/2$. □

Our next upper bound improves on the sphere packing bounds, at least achieving $R(\delta) = 0$ for $\delta > 1/2$. It still leaves a substantial gap with the GV bound.

Theorem 4.9 (Elias Bound).

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

Proof. The proof begins by considering the calculation of average distance from the previous theorem. It follows from Jensen's inequality that if the average weight of the vectors in $\mathcal{C}$ is αn, then the maximum of $\sum z_i(|\mathcal{C}| - z_i)$ is obtained if for all i, $z_i = (1 - \alpha)\mathcal{C}$ for some α. We sketch the argument for those not

familiar with Jensen's inequality. The inequality states that if f is convex, then for $x_1, \ldots, x_n$, $\frac{1}{n}\sum f(x_i) \leq f(\sum x_i/n)$, with equality if and only if $x_1 = \cdots = x_n$. For our case, the function $f(x) = x(|\mathcal{C}| - x)$ is easily verified convex and so its maximum over a set of z_i is achieved when the z_i are all equal. This makes the average distance in $\mathcal{C}$ at most $2\alpha(1-\alpha)n$.

With this calculation in mind, chose a spherical shell S in $\{0,1\}^n$ centered at some x_0 with radius r such that

$$|S \cap \mathcal{C}| \geq |C| \cdot \frac{|S|}{2^n}.$$

Such a shell exists as the right hand side of the inequality is the expected intersection size if the sphere is chosen randomly. Set $r = pn$ so that $|S| \approx 2^{nH(p)}$, which means

$$|S \cap \mathcal{C}| \geq \frac{|\mathcal{C}|}{2^{n\left(1-H(p)\right)}}.$$

Now apply the argument above on $x_0 + \mathcal{C} \cap S$. It follows from our discussion that we actually have a p fraction of ones in each row, so if $\delta > 2p(1-p)$, the $|S \cap \mathcal{C}|$ is subexponential, but this is equal to $|\mathcal{C}|2^{-n\left(1-H(p)\right)}$, implying

$$\frac{1}{n}\log|\mathcal{C}| < 1 - H(p).$$

Let us rewrite our condition $\delta > 2p(1-p)$ as follows:

$$1 - 2p \geq \sqrt{1-2\delta} \Rightarrow p = \frac{1-\sqrt{1-2\delta}}{2}.$$

This is the critical value of p—when p is below this the code is insubstantially small. □

Figure 4.2 shows how the Elias bound improves the sphere-packing bound to something reasonable. The gap between it and the GV bound is still large, however.

4.4 Aside: Erdoös-Ko-Rado Theorem

The proof of the Elias bound that we just saw is based on the following clever idea: we investigate and unknown object (the code $\mathcal{C}$) by itersecting it with random elements of a cleverly chosen set (the sphere). This method of "a randomly chosen fish-net" is also the basis for the following beautiful proof, due to Katona, of the Erdös-Ko-Rado theorem.

Definition 4.1. An *intersecting family* is a family $\mathcal{F}$ of k-sets in $1 \ldots n$ (compactly, $\mathcal{F} \subseteq \binom{[n]}{k}$), with $2k \leq n$, such that for any $A, B \in \mathcal{F}$, $A \cap B \neq \emptyset$.

Informally, an intersecting family is a collection of sets which are pairwise intersecting. One way to construct such a set is to pick a common point of intersection, and then choose all possible $(k-1)$-sets to fill out the sets. The Erdoös-Ko-Rado Theorem says that this easy construction is the best possible.

Theorem 4.10 (Erdös-Ko-Rado). *If $\mathcal{F} \subseteq \binom{[n]}{k}$ is an intersecting family with $2k \leq n$, then*

$$\mathcal{F} \leq \binom{n-1}{k-1}.$$

Proof (Katona). Given an intersecting family $\mathcal{F}$, arrange $1 \dots n$ in a random permutation π along a circle, and count the number of sets $A \in \mathcal{F}$ such that A appears as an arc in π. This will be our random fish-net.

There are $(n-1)!$ cyclic permutations—that is, $n!$ permutations, divided by n as rotations of the circle are identical. There are $k!$ ways for a given k-set to be arranged, and $(n-k)!$ ways of the other elements not interfering with that arc, so that the set appears consecutively on the circle. Hence the probability that a given k-set appears as an arc is

$$\frac{k!(n-k)!}{(n-1)!} = \frac{n}{\binom{n}{k}},$$

which by linearity of expectation implies

$$E\begin{bmatrix}\text{\# arcs belonging} \\ \text{to } \mathcal{F}\end{bmatrix} = \frac{n|\mathcal{F}|}{\binom{n}{k}}.$$

Now, as $2k \leq n$, at most k member of an intersecting family can appear as arcs on the circle, otherwise two of the arcs wouldn't intersect. Hence

$$\frac{n|\mathcal{F}|}{\binom{n}{k}} \leq k$$

implying

$$|\mathcal{F}| \leq \frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1}$$

□