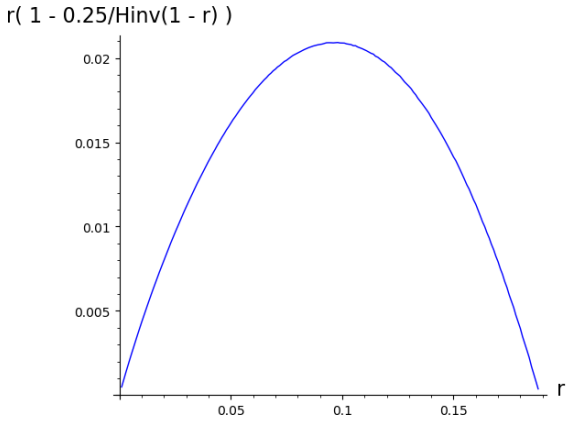
Class 7 Exercises

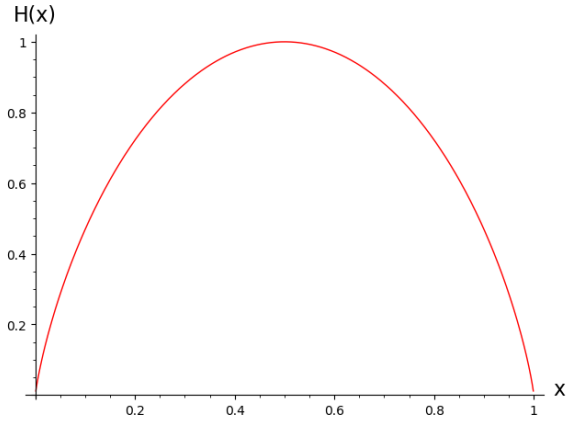
CS250/EE387, Winter 2025

1. What is (approximately) the rate of the best explicit binary code that we know exists with distance 1/4? Compare this to (approximately) the rate of the best non-explicit binary code that we know exists with distance 1/4.

   If it helps, here is a plot of the function $\phi(r) = r \cdot \left(1 - \frac{1/4}{H^{-1}(1-r)}\right)$:



   And here is a plot of the binary entropy function:



Solution

The best explicit binary code that we know of lies on the Zyablov bound, which is conveniently $R = \max_r \phi(r)$. Eyeballing the plot of $\phi(r)$, it looks like this is about 0.02 or a bit bigger.

The best non-explicit binary code that we know of lies on the GV bound, or $R = 1 - H(\delta)$. Again eyeballing the plot of $H(x)$, we see that $H(1/4) \approx 0.8$, so $1 - H(1/4) \approx 0.2$. So this is quite a bit bigger than the Zyablov bound.

2. With the Zyablov bound, we saw (finally) a construction of an "explicit" asymptotically good binary code using concatenated codes. However, the definition of "explicit" may not have been super-satisfying. In this exercise we'll see a construction with a more satisfying definition of "explicit."

   (a) Explain why the construction from the videos might not be super satisfying in terms of explicitness. (Recall that "explicit" meant that "there exists a polynomial-time algorithm to find it." What could we hope for instead?)

   Solution

   Our definition still involved exhaustive search for the generator matrix of a good inner code. A more satisfying definition (sometimes called "strongly explicit") might be for a description that efficiently tells us a single entry of this matrix, say in time polylog(n).

   For example, Reed-Solomon codes are "strongly explicit", because I can tell you that the i, j entry of the generator matrix is α_i^j really quickly. (Perhaps I need to take $O(\log n)$ time to compute α_i^j).

   (b) Fix a parameter k. Let $Q = 2^k$. We can identify the finite field $\mathbb{F}_Q$ with the vector space $\mathbb{F}_2^k$. These aren't the same thing, but it turns out that they have the same additive structure. That is, there is some map $\varphi : \mathbb{F}_2^k \to \mathbb{F}_Q$ so that

   $$\varphi(\mathbf{0}) = 0 \qquad \text{and} \qquad \varphi(\mathbf{x} + \mathbf{y}) = \varphi(\mathbf{x}) + \varphi(\mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{F}_2^k.$$

   For $\alpha \in \mathbb{F}_Q$, with $\alpha \neq 0$, consider the linear code $\mathcal{C}_\alpha \subset \mathbb{F}_2^{2k}$ of dimension k and length $2k$ over $\mathbb{F}_2$ given by the encoding map

   $$\mathbf{x} \mapsto \mathbf{x} \circ \varphi^{-1}(\varphi(\mathbf{x}) \cdot \alpha),$$

   where $\circ$ denotes concatenation.

   There is no question for this part, just make sure you understand the construction. Notice that there are $2^k - 1$ different codes C_α that you can create like this.

   (c) Let $\varepsilon > 0$, and let $\eta = 2^{-2\varepsilon k}$. Show that at least a $1 - \eta$ fraction of the α's in $\mathbb{F}_2$ have

   $$\text{dist}(\mathcal{C}_\alpha) \geq H_2^{-1}(1/2 - \varepsilon).$$

   Hint. Consider a fixed nonzero $\mathbf{y} \in \mathbb{F}_2^{2k}$ with weight less than $2k \cdot H_2^{-1}(1/2 - \varepsilon)$. How many such vectors $\mathbf{y}$ are there? How many different codes $\mathcal{C}_\alpha$ can each vector $\mathbf{y}$ lie in?

   Solution

   First, observe that any nonzero $\mathbf{y} \in \mathbb{F}_2^{2k}$ can lie in at most one $\mathcal{C}_\alpha$. Indeed, suppose that $\mathbf{y} = (\mathbf{x}, \alpha\mathbf{x}) = (\mathbf{x}', \beta\mathbf{x}')$. Then, since $\mathbf{y}$ is nonzero, we must have $\mathbf{x}$ nonzero, hence $\mathbf{x} = \mathbf{x}' \neq 0$, and thus $\alpha\mathbf{x} = \beta\mathbf{x}$, which implies that $\alpha = \beta$. (Above, we are abusing notation and using $\alpha\mathbf{x}$ to mean $\alpha\varphi(\mathbf{x})$).

   Now, let's count the number of nonzero $\mathbf{y}$'s with weight less than $\rho := H_2^{-1}(1/2 - \varepsilon)$. There are at most

   $$\text{Vol}_2(\rho, 2k) - 1 \leq 2^{2kH_2(\rho)} - 1 = 2^{2k(1/2-\varepsilon)} - 1 = 2^k\eta - 1$$

   such vectors (where the minus 1 is just discounting the all-zero vector). Thus, at most $2^k\eta - 1$

values of α can have $\text{dist}(\mathcal{C}_\alpha) \leq \rho$. Therefore, the fraction of "good" $\mathcal{C}_\alpha$'s is

$$1 - \frac{2^k\eta - 1}{2^k - 1} \geq 1 - \eta,$$

as desired.
(Note: Above I have been a bit pedantic to make sure that removing the all-zero vector is okay with the desired inequality. But even if it weren't, it would be a negligable difference, so it's fine to ignore it for the answer to this exercise).

(d) Let $Q = 2^k$ as before. Consider the code $\mathcal{C}$ (called the *Justesen code*) formed as follows:

- Let $\mathcal{RS}$ be a Reed-Solomon code over $\mathbb{F}_Q$ with length $N = Q - 1$, evaluation points $\mathbb{F}_Q \setminus \{0\}$, and rate R. Thus, each symbol of $\mathcal{RS}$ is associated with an element $\alpha \in \mathbb{F}_Q \setminus \{0\}$.
- For each $\alpha \in \mathbb{F}_Q \setminus \{0\}$, let $\mathcal{C}_\alpha$ be as in the previous part.
- Consider the "concenated-like" code where the inner code is *different* for each symbol: we concatenate the α'th symbol of a codeword in $\mathcal{RS}$ with the inner code $\mathcal{C}_\alpha$.

Show that the rate of this code is at least $R/2$, and that the distance is at least $(1 - R - \eta) \cdot H_2^{-1}(1/2 - \varepsilon)$.

Solution

To see the rate, notice that there are $Q^{RN} = 2^{kRN}$ codewords, and that the length (in bits) of the output is $N \cdot 2k$. So the rate is

$$\log_2(Q^{RN})/(2Nk) = (kRN)/(2Nk) = \frac{R}{2}.$$

For the distance, observe that this code is $\mathbb{F}_2$-linear, so it suffices to show that any nonzero codeword has relative weight at least $(1 - R - \eta) \cdot H_2^{-1}(1/2 - \varepsilon)$. First, any nonzero RS codeword has weight at least $(1 - R)N$, by the distance of RS codes. At most ηN of the symbols are concatenated with a "bad" inner code, i.e., one that has distance less than $H_2^{-1}(1/2 - \varepsilon)$. Thus, at least $(1 - R - \eta)N$ of the inner codewords have weight at least $H_2^{-1}(1/2 - \varepsilon) \cdot 2k$, so any nonzero codeword has weight at least

$$(1 - R - \eta)H_2^{-1}(1/2 - \varepsilon) \cdot 2Nk,$$

which is what we wanted to show.

(e) Does the Justesen code meet your definition of "satisfyingly explicit" from part (a)? Why or why not?

Note: If you haven't seen the correspondence between $\mathbb{F}_{2^k}$ and $\mathbb{F}_2^k$ (as per the map φ above) before, it may be tricky to answer this question precisely. In that case, just try to think about whether it seems plausible and then move on.

Solution

It meets my definition. The generator matrix $G \in \mathbb{F}_q^{2kN \times Rkn}$ is given by $G_1 \cdot G_2$, where $G_2 \in \mathbb{F}_2^{kN \times kRN}$ represents the generator matrix for the RS code and $G_1 \in \mathbb{F}_2^{2kN \times kN}$ does the concatenation.
In more detail, G_2 is the block matrix where each symbol $\beta \in \mathbb{F}_Q$ of the RS code's generator matrix in $\mathbb{F}_Q^{N \times RN}$ is replaced by a $k \times k$ binary matrix that represents multiplication by β in $\mathbb{F}_2^k$.
The matrix G_1 is block-diagonal, and the $2k \times k$ block corresponding to $\alpha \in \mathbb{F}_Q$ is again a block matrix where the top part is the identity and the bottom part is the $k \times k$ matrix

representation of α.
Thus, $G = G_1 \cdot G_2 \in \mathbb{F}_2^{2kN \times kRN}$ is a block matrix consisting of $2k \times k$ blocks, and the block corresponding to $\alpha \in \mathbb{F}_Q$ and $i \in [RN]$ is given by

$$\begin{bmatrix} I \\ \hline\hline M_\alpha \end{bmatrix} \cdot \left[M_\alpha^i\right] = \begin{bmatrix} M_\alpha \\ \hline\hline M_\alpha^{i+1} \end{bmatrix},$$

where $M_\alpha \in \mathbb{F}_2^{k \times k}$ is the matrix that represents multiplication by α. To compute this, we just need to be able to compute powers α^{i+1} and then to turn it into binary. Both of these can be done in time $\text{poly}(k)$, and k is logarithmic in the block length of the code.
(Note: Since we haven't gone into detail about how $\mathbb{F}_Q$ works as a vector space over $\mathbb{F}_2$ in this class, you didn't need to write out exactly how to compute each entry of the generator matrix, just convince yourself that it's plausible).

(f) Does this construction give you an asymptotically good code of distance, say, 1/4 (c.f. Exercise 1)? How about for distance 1/20? How does this compare to the Zyablov/GV bound? (Better or worse?) (**Note:** for distance 1/20, the Zyablov bound is about 0.27. The binary entropy of 1/20 is about 0.29).

Hint. Since the question is about asymptotics, you can set $\varepsilon, \eta = 0$ in your answer to (d). (The reason is that we can choose $\varepsilon > 0$ to be arbitrarily small; and then suppose that k is big enough compared to $1/\varepsilon$ that $\eta = 2^{-2k\varepsilon}$ is also arbitrarily small.)

Solution

Letting R^* be the rate of the final code, we see that the distance δ can approach

$$\delta = (1 - R)H_2^{-1}(1/2) = (1 - 2R^*)H_2^{-1}(1/2).$$

(if we let $\varepsilon, \eta \to 0$). Solving for R^* in terms of δ, we get

$$R^* = \frac{1}{2} - \frac{\delta}{2H_2^{-1}(1/2)}.$$

Looking at the plot from Exercise 1, $H_2^{-1}(1/2) \approx 0.1$, so the best rate we can get with this construction and distance δ is

$$R^* \approx \frac{1}{2} - \frac{\delta}{2 \times 0.1} = \frac{1}{2} - 5\delta.$$

For $\delta = 1/4$, this is $1/2 - 5/4 < 0$, so we don't get anything useful here. Since we *did* get something useful for $\delta = 1/4$ from the Zyablov bound, this guarantee is substantially worse than the Zyablov bound. However, if δ is small enough, say $\delta = 1/20$, we'd get

$$R^* \approx \frac{1}{2} - \frac{5}{20} = \frac{5}{20} = 0.25,$$

which at least gives us a satisfyingly explicit asymptotically good code. It's a bit worse than the Zyablov bound (0.27) and definitely not as good as the GV bound (0.71), but not bad!

(g) **(Bonus).** How would you modify the construction above to achieve a better trade-off? Can you match the Zyablov bound this way?

Solution

(Sketch). One way to modify it is to improve the rate of the inner code by keeping just the first s bits of $\varphi^{-1}(\varphi(\mathbf{x}) \cdot \alpha)$ in the definition of C_α. It turns out that most codes in this family have distance at least $H^{-1}(s/(s+k) - \varepsilon)$. By doing this, we can get a final code with rate

$$R^* = \max_{1/2 < r < 1 - H_2(\delta)} r \cdot \left(1 - \frac{\delta}{H^{-1}(1-r)}\right),$$

which looks a lot like the Zyablov bound except for the extra constraint on r. It turns out that this coincides with the Zyablov bound for rates larger than $1/3$ or so.

(h) **(Bonus 2).** Does the algorithm that we saw for decoding concatenated codes work for our Justesen code?

Solution

Essentially, yes. In the analysis, we just need to ignore the η fraction of inner codes that don't have good distance, and this doesn't really affect anything.

Stuff below here is super-bonus, we definitely won't get to it in class. The lecture notes promised that "in class" we'd see a construction of a code near the GV bound in time $2^{O(n)}$ (rather than $2^{O(n^2)}$)...but then we decided that the above would be more fun. If you want to know how to construct such a code, work through the steps below!

3. (**Deterministic codes on the GV bound in time $2^{O(n)}$.**) In the lecture videos/notes, we said that it is open (in most parameter regimes) to find an efficient deterministic construction of an asymptotically good code near the GV bound, even though "most" linear codes lie close to this bound.

   (0) Fix some constant δ and let $\varepsilon > 0$. Describe a straightforward algorithm that finds an asymptotically good code of length n near the GV bound (that is, with dimension at least $(1 - H(\delta) - \varepsilon)n$ and distance at least δ) in time $2^{O(n^2)}$. In the big-Oh notation, we treat ε, δ as constants and n as growing.

Solution

We exhaust over all generator matrices $G \in \mathbb{F}_2^{k \times n}$, where $k = n(1 - H_2(\delta) - \varepsilon)$. The GV bound tells us that at least one such matrix has good distance. To test distance, we look at all 2^k codewords and see what the lowest non-zero weight is. This all takes time

$$O(2^{kn} \cdot 2^k \cdot n) = O(2^{kn + k + log_2(n)}) = 2^{O(n^2)}.$$

However, we can do a little bit better than $2^{O(n^2)}$ (and this may be useful soon for using concatenated codes to get an explicit construction of an asymptotically good code...).

In this exercise we'll prove the following theorem:

Theorem 1. *There is a deterministic algorithm that finds a binary linear code $\mathcal{C} \subset \mathbb{F}_2^n$ with rate $R = 1 - H_2(\delta) - o(1)$ in time $2^{O(n)}$.*

Fix constants $\varepsilon, \delta > 0$. Suppose that $k \geq 1 - H_2(\delta) - \varepsilon$. In the following, we will come up with a distribution $\mathcal{D}$ on matrices $G \in \mathbb{F}_2^{n \times k}$ so that

(i) A matrix $G \sim \mathcal{D}$ can be sampled using $O(n)$ bits of randomness.

(ii) A matrix $G \sim \mathcal{D}$ is full rank with probability at least $2/3$.

(iii) Let $\mathcal{C} \subseteq \mathbb{F}_2^n$ be a (random) code with generator matrix $G \sim \mathcal{D}$; then $\mathcal{C}$ has distance at least δ with probability at least $2/3$.

(a) Explain why coming up with such a distribution $\mathcal{D}$ would prove the theorem.

Solution

Suppose that there is such a $\mathcal{D}$, and say we sample from it via a function $G : 0, 1^r \to \mathbb{F}_2^{n \times k}$, so that $G(\omega)$ is the matrix that we'd draw with random seed ω. By (i), we may take $r = O(n)$. By (ii) and (iii) and a union bound, with probability at least $1/3$ G is both full rank and corresponds to a code with distance at least δ. In particular, there exists some ω so that $G(\omega)$ has these properties. Thus, we may enumerate over all possible seeds ω in time $2^{O(n)}$ and we will encounter such a G. To find such a matrix, we need to test each $G(\omega)$: this takes time $\text{poly}(n)$ (to check the rank) plus time $2^{O(k)}$ (to check the distance; we just enumerate over all codewords and check their weight). So the total time is $2^{O(n)}$.
Once we have found such a G, it is the generator matrix for a code with dimension k (using (ii)) and distance at least δ (using (iii)), as desired.

(b) Define a distribution $\mathcal{D}$ on matrices G by letting G be a random *Toeplitz matrix*. That is, G is of the form

$$G = \begin{pmatrix} X_0 & X_1 & \cdots & X_{k-1} \\ X_k & X_0 & X_1 & \ddots \\ X_{k+1} & X_k & X_0 & \ddots \\ \vdots & & & \\ X_{n+k-2} & & & \end{pmatrix}.$$

where $X_0, \ldots, X_{n+k-2}$ are i.i.d. uniform random variables in $\mathbb{F}_2$.
(There is no question here, just understand the distribution).

Solution

Understood!

(c) Show that, for any $x \in \mathbb{F}_2^k$, Gx is uniformly distributed when $G \sim \mathcal{D}$.

Solution

Fix an $x \in \mathbb{F}_2^k$, and suppose that x has support $I \subseteq \{0, \ldots, k-1\}$ and let i^* be the smallest element of I. Let $J \subset \{0, \ldots, k+n-2\}$ be

$$J = \{j \,:\, X_j \text{ appears in } G[:, i^*]\}.$$

Let φ denote an assignment of the variables X_j for $j \notin J$: that is,

$$\varphi : \{0, \ldots, n+k-2\} \setminus J \to \mathbb{F}_2,$$

and let $G|_\varphi$ denote the conditional random variable where these variables X_j for $j \notin J$ have been fixed to $X_j = \varphi(j)$. (But the X_j for $j \in J$ are still random).

Claim 2. *Fix any such φ and any $y \in \mathbb{F}_2^n$. We claim that there is exactly one realization $\psi : J \to \mathbb{F}_2$ so that*

$$G|_{\varphi,\psi} x = y.$$

, *aka*

$$\sum_{i \in I} G|_{\varphi,\psi}[:, i] = y. \tag{1}$$

Proof. To see this, consider trying to solve (1) for the assignment $\psi : J \to \mathbb{F}_2$, going variable-by-variable, where the variables in J are ordered from top-to-bottom of column $G[:, i^*]$. The first variable, in the spot $G[0, i^*]$, is determined by (1), because the variables in $G[0, i^* + 1 :]$ are already fixed by φ; this follows from the Toeplitz structure of G, which implies that $G[0, i^* + 1 :] \cap J = \emptyset$.
Next, the second variable, in the spot $G[1, i^*]$, was free (because we hadn't yet encountered it, but is now fixed by (1), by the same logic; there are no free variables in any of relevant positions in row 1.
We can continue in this way to uniquely determine the assignment ψ. □

Now, the statement in this part follows from the claim. Indeed, for any fixed y, let's count the number of instantiations of G so that $Gx = y$. By the claim, this is precisely the number of ways to assign boolean variables to a set of size $(n + k - 1) - n = k - 1$, which is 2^{k-1}. Thus, when G is drawn from the distribution described above, the probability of obtaining any fixed y is the number of ways to obtain y, divided by the number of instantiations of G, which is

$$\mathbb{P}\{Gx = y\} = \frac{2^{k-1}}{2^{n+k-1}} = \frac{1}{2^n},$$

and this proves this part.

(d) Show that (i), (ii), and (iii) are satisfied by $\mathcal{D}$.

Solution

(i) The number of random bits we need to draw $G \sim \mathcal{D}$ is $O(n)$, since we just need to pick bits for the first row and first column of G.

(ii) To see that the matrix is full-rank with high probability, we can use the previous part. If G is *not* full-rank, then there is some kernel vector $x \neq 0$ so that $Gx = 0$. For any fixed nonzero x,

$$\mathbb{P}\{Gx = 0\} = 2^{-n}$$

by the previous part. Thus, by a union bound over all 2^k such x,

$$\Pr[\exists x \neq 0, Gx = 0] \leq 2^{k-n}$$

which is tiny, way smaller than $1/3$.

(iii) The fact that G has good distance with good probability follows exactly as the proof of the Gilbert-Varshamov bound from the videos/notes (Class 3?). Since Gx is uniform for all x, the probability that $\mathrm{wt}(Gx) < \delta n$ is at most

$$\mathbb{P}\{\mathrm{wt}(Gx) < \delta n\} \leq 2^{n(H_2(\delta)-1)},$$

and so by a union bound over all 2^k possible x's, the probability that any codeword in the code $\mathcal{C}$ with generator matrix G is at most

$$\mathbb{P}\{\mathrm{dist}(\mathcal{C}) < \delta n\} \leq 2^k 2^{n(H_2(\delta)-1)},$$

and so if $k < n(1 - H_2(\delta) - \varepsilon)$ for any $\varepsilon > 0$, this probability is less than $1/3$.