Class 9 Exercises

CS250/EE387, Winter 2025

Today we'll show (from scratch, without information theory) that the capacity of the binary erasure channel (with erasure probability p) is $1-p$: that is, there exist codes of rate $1-p-\varepsilon$ that can communicate reliably over BEC_p, but any code of rate $1-p+\varepsilon$ cannot.

1. Suppose that G is a random matrix in $\mathbb{F}_2^{n\times k}$. Show that the probability that G has rank less than k is at most 2^{k-n}.

2. Consider the linear code given by the encoding map $x \mapsto Gx$. Define a decoder $D : \{0,1,\perp\}^n \to \{0,1\}^k \cup$ `FAIL` for this code to be the decoding algorithm that, on input $y \in \{0,1,\perp\}^n$, returns x if Gx agrees with y on all of the un-erased symbols if such an x exists and is unique; otherwise it returns `FAIL`.

   Let $\varepsilon > 0$ and suppose that $k \le (1-p-\varepsilon)n$. Let $G \in \mathbb{F}_2^{n\times k}$ be a random binary matrix. In the following parts, you will show that, for all $x \in \mathbb{F}_2^k$, there is some constant γ so that

   $$\mathbb{E}_G\left[\mathbb{P}_{BEC_p}\left\{D(BEC_p(Gx)) \neq x \,|\, G\right\}\right] \le 2^{-\gamma n},$$

   where the randomness in the $\mathbb{E}$ is over the choice of the matrix G, and the randomness in the $\mathbb{P}$ is over the channel BEC_p.

   (a) Let $J \subset [n]$ be the set of erased positions. Explain why the decoder succeeds if and only if the rank of $G|_{\bar{J}}$ is k, where $G|_{\bar{J}}$ is the restriction of G to the rows *not* in J.

   (b) Explain why
   $$\mathbb{P}\left\{\text{rank}(G|_{\bar{J}}) < k\right\} \le \min\{1, 2^{k-(n-|J|)}\},$$
   and conclude that, for any $x \in \mathbb{F}_2^k$,
   $$\mathbb{P}_G\left\{D(BEC_p(Gx)) \neq x|J\right\} \le \min\{1, 2^{k-(n-|J|)}\},$$
   where above the conditional probability means "the probability that the decoder fails, given that J is the set of erased positions."

   (c) We are later going to want to take the expectation over J (that is, over the set of positions erased by the BSC_p). When we do that in the conclusion above, we get:
   $$\mathbb{E}_J\mathbb{P}_G\left\{D(BEC_p(Gx)) \neq x|J\right\} \le \mathbb{E}_J \min\{1, 2^{k-(n-|J|)}\},$$
   so we want to bound the right hand side, which we'll do in this part.

   Show that
   $$\mathbb{E}_J \min\left\{1, 2^{k+|J|-n}\right\} \le \mathbb{P}_J\left\{|J| \ge n(p+\varepsilon/2)\right\} + 2^{-n\varepsilon/2}.$$
   (Remember that we are assuming that $k \le n(1-p-\varepsilon)$).

(d) If you have not seen Chernoff bounds much before, you may use the fact that

$$\mathbb{P}\{|J| \geq n(p + \varepsilon/2)\} \leq 2^{-Cn\varepsilon^2}$$

for some constant C. (If you have seen Chernoff bounds before, convince yourself that this follows from a Chernoff bound). Either way, nothing to write down for this part, just be aware of this fact.

(e) Put it all together to prove the thing that this part asks you to prove, that for any $x \in \mathbb{F}_2^k$,

$$\mathbb{E}_G\left\{\mathbb{P}_{BEC_p}\left\{D(BEC_p(Gx)) \neq x \,|\, G\right\}\right\} \leq 2^{-\gamma n}$$

for some constant γ. (Note that γ is allowed to depend on ε and on the choice of the constant C). Hint: If it's been a while since probability, remember that for two random variables X, Y and some event $\mathcal{E}$ that depends on X and Y,

$$\mathbb{E}_Y[\mathbb{P}_X\{\mathcal{E} \,|\, Y\}] = \mathbb{E}_X[\mathbb{P}_Y\{\mathcal{E} \,|\, X\}].$$

3. In this part, we will show that:

Claim 1. *For all $\varepsilon > 0$, there is a code of rate at least $1 - p - \varepsilon$ so that the error probability on BEC_p is at most $2^{-\gamma n}$, for some γ (which depends on p and ε).*

(a) Explain why this is not obvious from part 2. That is, part 2 says that:

For all $\varepsilon > 0$, if G is a random generator matrix for a code of rate about $1 - p - \varepsilon$, then for all x,

$$\mathbb{E}_G[\mathbb{P}_{BEC_p}\{D(BEC_p(Gx)) \neq x \,|\, G\}] \leq 2^{-\gamma n}$$

for some constant γ.

My (incorrect) claim is that we are done! If the expected failure probability (that is, the expectation of $\mathbb{P}_{BEC_p}\{D(BEC_p(Gx)) \neq x\}$, which is what we are looking at) is less than $2^{-\gamma n}$, then there must exist a code that has failure probability less than $2^{-\gamma n}$. But that's exactly what Claim 1 says!

What's wrong with this reasoning?

(b) We will show Claim 1 by "throwing out" some bad codewords of our random code, and showing that this new (slightly smaller) code will satisfy the statement that we want. Towards that end, imagine choosing $x \in \mathbb{F}_2^k$ uniformly at random. Explain why

$$\mathbb{E}_G\mathbb{E}_x\mathbb{P}_{BEC_p}\{D(BEC_p(Gx) \neq x) \,|\, G, x\} \leq 2^{-\gamma n},$$

and conclude that there exists some generator matrix G^* so that

$$\mathbb{E}_x\mathbb{P}_{BEC_p}\{D(BEC_p(G^*x) \neq x) \,|\, x\} \leq 2^{-\gamma n}.$$

(c) Show that there is some set $\mathcal{X} \subseteq \mathbb{F}_2^k$ so that $|\mathcal{X}| \geq 2^{k-1}$ and that for all $x \in \mathcal{X}$,

$$\mathbb{P}_{BEC_p}\{D(BEC_p(G^*x)) \neq x\} \leq 2 \cdot 2^{-\gamma n}.$$

(Hint: Markov's inequality says that, for any non-negative random variable Z, $\mathbb{P}[Z \geq 2\mathbb{E}[Z]] \leq \frac{1}{2}$.)

(d) Explain why the above establishes Claim 1, and thus proves that the capacity of the BEC_p is at least $1 - p$.

4. **(Bonus, if you have extra time)** So far we've seen that the capacity of the BEC_p is at least $1 - p$. Now we'll show that it's *exactly* $1 - p$.

Claim 2. *For any code with rate at least $1-p+\varepsilon$, the error probability on BEC_p must be at least $1/2$.*

Prove Claim 2. Notice that we must prove the result for any code (not necessarily linear) and for any decoding map (not necessarily the one proposed above).

Hint: try the following steps. Let $\mathcal{C}$ be a binary code with encoding map $E : \{0,1\}^k \to \{0,1\}^n$, so that $k/n \geq 1-p+\varepsilon$. Let $D : \{0,1,\perp\} \to \{0,1\}^k$ be any decoding map for $\mathcal{C}$.

For ease of notation, let $\hat{x} = D(BEC_p(E(x)))$ be the guess at x that the decoder recovers.

(a) Let J again denote the set of indices erased by the BEC. Explain why

$$\mathbb{P}\left\{|J| < n(p-\varepsilon/2)\right\} \leq 2^{-Cn\varepsilon^2}$$

for some constant C.

(b) Choose an x at random. Explain why

$$\mathbb{E}_x \mathbb{P}_{BEC_p}\left\{\hat{x} \neq x\right\} \geq (1-2^{-Cn\varepsilon^2})\mathbb{E}_J\left[\mathbb{P}_x\left\{\hat{x} \neq x \mid |J| \leq n(p-\varepsilon/2)\right\}\right]$$

(c) Fix any set J of size at least $n(p-\varepsilon/2)$. Show that, if the BEC deletes the set J, then

$$\mathbb{P}_x\left\{\hat{x} = x\right\} \leq \frac{1}{2^k} \cdot 2^{n(1-p+\varepsilon/2)}.$$

(Notice that the probability here is over x. Since we are fixing J, the behavior of the BEC is fixed).

(Hint: Write out the definition of that probability and play around with the order of summations. Indicator random variables are your friends.)

(d) Prove Claim 2.