Carnegie Mellon University Instructor: Venkatesan Guruswami
15-859Y: Coding Theory Term: Fall 2018

PROBLEM SET 3

Due date: Tuesday, November 6, 2018

INSTRUCTIONS

- You are allowed to collaborate with up to two students taking the class in solving problem sets. But here are some rules concerning such collaboration:

  1. You should think about each problem by yourself for at least 30 minutes before commencing any collaboration.
  2. Collaboration is defined as discussion of the lecture material and solution approaches to the problems. Please note that *you are not allowed to share any written material and you must write up solutions on your own*. You must clearly acknowledge your collaborator(s) in the write-up of your solutions.
  3. Of course, if you prefer, you can also (and are encouarged to) work alone.

- Solutions typeset in LaTeX are encouraged, but not required. If you are submitting handwritten solutions, please write clearly and legibly (you might want to first write the solution sketch in rough, before transferring it to the version you turn in).
- You should not search for solutions on the web. More generally, you should try and solve the problems without consulting any reference material other than the course notes and what we cover in class. If for some reason you feel the need to consult some source, *please acknowledge the source* and try to articulate the difficulty you couldn't overcome before consulting the source and how it helped you overcome that difficulty. Alternatively, before turning to any such material, we encourage you to ask the instructor for hints or clarifications.
- Please start work on the problem set early. There are **five** problems, for a total of 120 points.

1. (25 points) Recall the definition of an ϵ-biased space: it is a (multi)-set $S \subseteq \mathbb{F}_2^m$ such that for every $a \in \mathbb{F}_2^m$, $a \neq 0$,
$$\left| \Pr_{x \in S}[a \cdot x = 1] - \Pr_{x \in S}[a \cdot x = 0] \right| \leq \epsilon \ .$$
Below we think of $\epsilon > 0$ as a fixed but arbitrarily small constant, and m as growing.

   (a) Argue that for a suitable constant $C < \infty$, a random set of size $\lceil Cm/\epsilon^2 \rceil$ is an ϵ-biased set with high probability (tending to 1 as $m \to \infty$).

   (b) Using Reed-Solomon codes concatenated with Hadamard codes, show how one can explicitly construct (i.e., list all elements of S in $\mathrm{poly}(|S|)$ time) an ϵ-biased set $S \subseteq \mathbb{F}_2^m$ of size at most $O(m^2/\epsilon^2)$.

   (c) Consider the Hermitian curve over $\mathbb{F}_q$ for $q = r^2$ defined as $Y^r + Y = X^{r+1}$ which we discussed in class. Suppose we use as messages polynomials of total degree at most $s < r$, and evaluate it at all r^3 points on the curve. What are the parameters of the resulting code (it suffices to give a good lower bound on the distance)?

(d) By concatenating the codes from the previous part with Hadamard codes and a suitable choice of parameters, show how one can construct ϵ-biased sets in $\mathbb{F}_2^m$ of size at most $O((m/\epsilon^2)^{5/4})$.

2. (25 points) Recall the two invertible linear transforms P_n, R_n on $\mathbb{F}_2^n$, n a power of two, that we defined recursively:

   - $P_0(x) = x$, and $P_n(U, V) = (P_{n/2}(U + V), P_{n/2}(V))$ where U (resp. V) is the first $n/2$ (resp. last $n/2$) bits of the n-bit input to P_n.
   - $R_0(x) = x$, and $R_n(U, V)_{2i} = R_{n/2}(U)_i + R_{n/2}(V)_i$ and $R_n(U, V)_{2i+1} = R_{n/2}(V)_i$, for $i = 0, 1, \ldots, n/2 - 1$, where the X_i denotes the i'th bit of vector X, and we index the bits starting at 0.

   Recall that we used the transform P_n to conveniently present and analyze the successive cancellation decoder, and the transform R_n to conveniently analyze the polynomially strong polarizing property of the transform. In this exercise, we will relate these transforms to conclude that the polarizing property of R_n implies that of P_n, as alluded to in class.

   (a) Prove that P_n is its own inverse.

   (b) Let us use the notation P_n to also denote the $n \times n$ matrix so that the transform is given by $Z \mapsto P_n Z$ where $Z \in \mathbb{F}_2^n$ is a column vector. Prove that $P_n = P_2^{\otimes(\lg n)}$ where $P_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

   (c) Prove that $R_n = B_n P_n$ where $B_n : \mathbb{F}_2^n \to \mathbb{F}_2^n$ permutes the coordinates via "bit-reversal," that is, it maps location i with binary representation $b_m b_{m-1} \cdots b_2 b_1$ (where $m = \lg n$) to location j with binary representation $b_1 b_2 \cdots b_{m-1} b_m$.

   (d) Prove that P_n and B_n commute, i.e., $B_n P_n = P_n B_n$. (Note that this is equivalent to showing that R_n is its own inverse.)

   (e) Recap why the above implies that R_n and P_n have identical polarization properties.

3. (25 points) Let $Z \in \mathbb{F}_2^n$ be sampled as n i.i.d copies of Ber(p) and let $W = P_n(Z)$.

   (a) Show how, given $i \in \{1, 2, \ldots, n\}$, one can efficiently sample the random variable $W_{<i}$, i.e., W restricted to the first $i - 1$ coordinates.

   (b) Given an arbitrary string $w_{<i} \in \mathbb{F}_2^{i-1}$, show how one can compute the distribution of W_i conditioned on $W_{<i} = w_{<i}$. (Hint: Use the approach behind the successive cancellation decoder.)

   (c) Using the above parts, show that there is a randomized algorithm running in poly$(n, 1/\gamma)$ time which with probability at least $1 - 1/n^2$ outputs an estimate of the conditional entropy $H(W_i | W_{<i})$ within an additive error of γ.

   (d) Suppose that P_n is $(\epsilon, 1/n^2)$-polarizing, i.e., the set $\{i \mid H(W_i \mid W_{<i}) \geq 1/n^2\}$ has size at most $(h(p) + \epsilon)n$ (we proved this property in class provided n is at least a sufficiently big polynomial function of $1/\epsilon$).

   Combine the steps above to give a poly(n) time randomized algorithm that with probability at least $1 - 1/n$ outputs a set S of size at most $n(h(p) + \epsilon)$ such that for all $i \notin S$, $H(W_i \mid W_{<i}) \leq 3/n^2$.

4. (20 points) Let $p \in (0, 1/2)$. Suppose $H \in \mathbb{F}_2^{m \times n}$ is a linear compression matrix and Decompress : $\mathbb{F}_2^m \to \mathbb{F}_2^n$ is such that

$$\Pr_{Z \sim \text{Ber}(p)^{\otimes n}} \Big[\text{Decompress}(HZ) \neq Z \Big] \leq 2^{-t} \, .$$

Prove that the code with parity check matrix H has minimum distance at least $\frac{t}{\lg(1/p)}$.

(Thus linear compression schemes for i.i.d Bernoulli sources (equivalently linear codes for BSC) with very good error probability imply codes with good minimum distance.)

Hint: The optimal decompression algorithm is the maximum likelihood algorithm, that on input $w \in \mathbb{F}_2^m$, outputs $\arg\min_{x:Hx=w} \mathrm{wt}(x)$ (Why?)

5. (25 points) For a correlated random variables (U, A) with U taking values in $\mathbb{F}_2$ and A in any finite set and joint distribution p_{UA}, define the quantity

$$B(U \mid A) = 2 \sum_{a \in \mathrm{supp}(A)} \sqrt{p_{UA}(0,a) \cdot p_{UA}(1,a)}$$

where $\mathrm{supp}(A)$ is the support of A.

Prove that

(a) Prove that there is a function $D : \mathrm{supp}(A) \to \mathbb{F}_2$ such that

$$\Pr_{(u,a)\sim(U,A)} [D(a) \neq u] \leq B(U \mid A) .$$

That is, if $B(U \mid A)$ is small, then we can predict U well from A.

(b) The binary entropy function satisfies $4x(1-x) \leq h(x) \leq 2\sqrt{x(1-x)}$ for $x \in [0,1]$. Using this prove that

$$B(U \mid A)^2 \leq H(U \mid A) \leq B(U \mid A) .$$

Suppose (U_1, A_1) and (U_2, A_2) are two i.i.d copies of (U, A).

(c) Prove that

$$B(U_1 + U_2 \mid (A_1, A_2)) \leq 2 \cdot B(U \mid A) .$$

(d) Prove that

$$B(U_2 \mid (A_1, A_2, U_1 + U_2)) = B(U \mid A)^2 . \tag{1}$$

Comment: Using the above facts, one can use the quantity $B(\cdot \mid \cdot)$ in place of conditional entropy in the analysis of the second phase of the polarization (where we amplified moderate polarization to values very close to 0), to drive the polarized entropies within $2^{-n^{0.49}}$ (rather than $n^{-\omega(1)}$) of the boundary values $\{0, 1\}$. This is because (1) shows that the "good branch" *squares* $B(\cdot \mid \cdot)$ instead of dropping it by a large multiplicative factor.

(e) **Bonus**: Formalize and provide the details of the analysis alluded to in the above comment.