PCP - Huji Course, Ex 1.

David Ponarovsky

May 28, 2023

1 Ex 1.

Let A be a random matrix in $M(\mathbb{F}_2^{k\times n})$. For any non-zero $x \in \mathbb{F}$, we have that Ax is distributed uniformly.

Claim 1.

Proof. By the fact that $x \neq 0$, there exists at least one coordinate $i \in [k]$ such that $x_i \neq 0$. Thus, we have

$$\begin{aligned}(Ax)_j &= \sum_k A_{jk}x_k = \sum_{i\neq k} A_{jk}x_k + A_{ji}x_i \\ &= \sum_{i\neq k} A_{jk}x_k + A_{ji}\end{aligned}$$

Notice that due to the fact that $\mathbb{F}_2$ is a field, there is exactly one assignment that satisfies the equation conditioned on all the values A_{jk} where $j \neq k$.

$$\begin{aligned}\mathbf{Pr}\left[(Ax)_j = 1\right] &= \sum_{A_{jk};k\neq i} \mathbf{Pr}\left[(Ax)_j = 1 \mid A_{jk}; k\neq i\right]\mathbf{Pr}\left[A_{jk}; k\neq i\right] \\ &= \frac{1}{2}\end{aligned}$$

Therefore, any coordinate of Ax is distributed uniformly $\Rightarrow$ Ax is distributed uniformly. $\square$

By the uniformity of Ax, we obtain that the expected Hamming weight of Ax is:

$$\mathbf{E}\left[|Ax|\right] = \mathbf{E}\left[\sum_i^n (Ax)_i\right] = \frac{1}{2}n$$

As the coordinates of A_x are independent (each row of A is sampled separately), we can use the Hoff's bound to conclude that:

$$\mathbf{Pr}\left[\left||Ax| - \mathbf{E}\left[|Ax|\right]\right| \geq \left(\frac{1}{2} - \delta\right)n\right] \leq e^{-n\left(\frac{1}{2}-\delta\right)^2}$$

Now, we will use the union bound to show that any $x \in \mathbb{F}_2^k$, Ax is of weight at least δ.

$$\mathbf{Pr}\left[|Ax| \geq \delta : \forall x \in \mathbb{F}_2^k\right] \geq 1 - |\mathbb{F}_2^k| \cdot e^{-n\left(\frac{1}{2}-\delta\right)^2}$$

Denote $k = \rho n$ and notice that the above holds when $\rho \geq \left(\frac{1}{2} - \delta\right)^2$.

2 Ex 2.

Claim 2. *Let $v_1, v_2, \ldots, v_m$ be unit vectors in an inner-product space such that $\langle v_i, v_j \rangle \leq -2\varepsilon$ for all $i \neq j$, then $m \leq \frac{1}{2\varepsilon} + 1$.*

Proof. Let us bound the norm of the summation $|\sum_i v_i|$ from both sides. As the norm is non-negative by definition, we will bound it from the left by 0. On the other hand, we have that:

$$0 \leq |\sum_{v_i} v_i| = m + 2\sum_{i,j} \langle v_i, v_j \rangle \leq m - 2 \cdot \frac{m\,(m-1)}{2} \cdot 2\varepsilon$$

Thus, we obtain $m\,(2(m-1)\varepsilon - 1) \leq 0$, namely, $m \leq \frac{1}{2\varepsilon} + 1$ □

Now, define the following product for $u, v \in \mathbb{F}_2^n$, $\langle v, u \rangle = \sum_i (-1)^{v_i}(-1)^{\bar{u}_i}$ and observe that:

1. $\langle v, v \rangle = \sum_i 1 = n \geq 0$.
2. $\langle v, u \rangle = \langle u, v \rangle$.
3. $\langle ax + by, z \rangle = (-1)^a \langle x, z \rangle + (-1)^b \langle y, z \rangle$.

Now, if the v's correspond to a code with distance at least d, then, for any codewords v and u that disagree on at least d coordinates, we have that $\langle v, u \rangle \leq$ `agree`$-$ `disagree = n - 2 disagree` $= n - 2d$. Now consider the normal codewords $\tilde{v_1}..\tilde{v_n}$ and assume that

$$\langle \tilde{v_i}, \tilde{v_j} \rangle = (1 - 2\delta) = \frac{1}{n}\,(n - 2d(v_i, v_j)) \leq \varepsilon$$

Therefore, if $d \geq \frac{1}{2} + \varepsilon$, we obtain the condition of the above claim.

3 Ex 3.

Consider the following process for decoding a,

for $t \in [\tau]$ **do**
  for $i \in [n]$ **do**
    $x \sim_u \mathbb{F}_2^n$
    $a_i^{(t)} \leftarrow w\,(x) + w\,(\sigma_i(x))$
  end
end
for $i \in [n]$ **do**
  $\hat{a}_i \leftarrow [\frac{1}{\tau} \sum_t^\tau a_i^{(t)}]$
end
return $\hat{a}_0, \hat{a}_1, \hat{a}_2..\hat{a}_n$

Claim 3. *For $\tau = \Omega\left(\frac{1}{\varepsilon^4} \log(n)\right)$ The above decoding success to decode $w\,(x)$ with probability $\geq 1 - \frac{1}{n}$.*

Proof. In this question we will say that w agree on $x, \sigma_i(x)$ if both $x, \sigma_i(x)$ were either filliped or unflipped. Clearly if $w(x)$ agree with $w(\sigma_i(x))$ than

$$\begin{aligned} w\,(x) + w\,(\sigma_i(x)) &= H_a(x) + H_a(\sigma_i(x)) \\ &= \sum_{i \neq j} a_j(x_j + x_j) + a_i(x_j + 1 + x_j) = a_j \quad \text{(neither of them were flipped.)} \\ &= 1 + H_a(x) + 1 + H_a(\sigma_i(x)) = a_i \quad \text{(both flipped.)} \end{aligned}$$

Thus we can bound the probability that $w\,(x) + w\,(\sigma_i(x)) \neq a_{_i}$ by the probability that w disagree on x and that append at probability:

$$\xi := (1 - f(x)) f(\sigma_i(x)) + (1 - f(\sigma_i(x))) f(x)$$

Now as we want bound ξ we could think about the maximization problem under the restrictions that $f(x), f(\sigma_i(x) \leq \frac{1}{2} - \varepsilon$. We know that the maximum lay on the boundary so we can assign $\frac{1}{2} - \varepsilon$ for each of the probabilities to obtain an upper bound. That will yield $\xi \leq 2 \cdot \left(\frac{1}{2} - \varepsilon\right)\left(\frac{1}{2} + \varepsilon\right)$, namely $\xi \leq \frac{1}{2} - 2\varepsilon^2$. Now the probability that a coordinate i will rounded to the opposite side, that it $\hat{a}_i \neq a_i$ mean that arithmetic mean over τ experiments were $2\varepsilon^2$ far from the expectation. Which by Hoff' bound is bounded by: $e^{\tau 4 \varepsilon^4}$. So using the union bound we obtain:

$$\mathbf{Pr}\left[\text{ decoding success }\right] \geq 1 - n \cdot e^{\tau 4 \varepsilon^4}$$

Therefore it's enough to take $\tau = O(\frac{1}{\varepsilon^4}\log(n))$ to obtain a decoder which run at time $O(\frac{1}{\varepsilon^4} n \log(n))$ and success with heigh probability. □

4 Ex 3.

Consider the following process for decoding a:

for $t \in [\tau]$ **do**
  for $i \in [n]$ **do**
    $x \sim_u \mathbb{F}_2^n$
    $a_i^{(t)} \leftarrow w(x) + w(\sigma_i(x))$
  end
end
for $i \in [n]$ **do**
  $\hat{a}_i \leftarrow [\frac{1}{\tau}\sum_t^{\tau} a_i^{(t)}]$
end
return $\hat{a}_0, \hat{a}_1, \hat{a}_2..\hat{a}_n$

Claim 4. *For $\tau = \Omega\left(\frac{1}{\varepsilon^4}\log(n)\right)$ the above decoding succeeds in decoding $w(x)$ with probability $\geq 1 - \frac{1}{n}$.*

Proof. In this question we will say that w agrees on $x, \sigma_i(x)$ if both $x, \sigma_i(x)$ were either flipped or unflipped. Clearly, if $w(x)$ agrees with $w(\sigma_i(x))$ then

$$\begin{aligned} w(x) + w(\sigma_i(x)) &= H_a(x) + H_a(\sigma_i(x)) \\ &= \sum_{i \neq j} a_j(x_j + x_j) + a_i(x_j + 1 + x_j) = a_j \quad (\text{ neither of them were flipped. }) \\ &= 1 + H_a(x) + 1 + H_a(\sigma_i(x)) = a_i \quad (\text{ both flipped.}) \end{aligned}$$

Thus, we can bound the probability that $w(x) + w(\sigma_i(x)) \neq a_i$ by the probability that w disagrees on x and that is at probability:

$$\xi := (1 - f(x)) f(\sigma_i(x)) + (1 - f(\sigma_i(x))) f(x)$$

Now, to bound ξ, we can think about the maximization problem under the restrictions that $f(x), f(\sigma_i(x) \leq \frac{1}{2} - \varepsilon$. We know that the maximum lies on the boundary, so we can assign $\frac{1}{2} - \varepsilon$ for each of the probabilities to obtain an upper bound. That will yield $\xi \leq 2 \cdot \left(\frac{1}{2} - \varepsilon\right)\left(\frac{1}{2} + \varepsilon\right)$, namely $\xi \leq \frac{1}{2} - 2\varepsilon^2$. Now, the probability that a coordinate i will be rounded to the opposite side, i.e. $\hat{a}_i \neq a_i$, means that the arithmetic mean over τ experiments is $2\varepsilon^2$ far from the expectation. According to Hoff's bound, this is bounded by $e^{\tau 4 \varepsilon^4}$. Thus, using the union bound, we obtain:

$$\mathbf{Pr}\left[\text{ decoding success }\right] \geq 1 - n \cdot e^{\tau 4 \varepsilon^4}$$

Therefore, it is enough to take $\tau = O(\frac{1}{\varepsilon^4}\log(n))$ to obtain a decoder that runs in time $O(\frac{1}{\varepsilon^4} n \log(n))$ and succeeds with high probability. □

5 Ex 4.

5.1 (a)

We will prove that if for any x, f interpolate well on $x, x+1, .., x+d+1$ than any it interpolate well on every coordinates set at size $d+1$. Denote by $J \subset \mathbb{F}_q$ at size $d+1$. Let's continue by induction on $\max J$. The base case $\max J = d+1 \Rightarrow J = \{1, 2.., d+1\}$ follow straightforwardly from the assumption. Assume the correctness for any J such that $\max J \leq x_0$ and consider J' such that $\max J' = x_0+$. Now it given that $S = \{x_0 - d, x_0 - d + 1, ..x_0 + 1\}$ is well interpolating set, so there exists coefficients a_1, a_{d+1} such that $a_{d+1} f(x_0 + 1) = \sum_{x_i \in S/(x_0+1)} a_i f(x_i)$ On the overhand, for ant any $x_i \in S/(x_0 + 1)$ the union $K = x_i \cup J/(x_0 + 1)$ is subset of $\mathbb{F}_q$ at size $d+1$ such that $\max K \leq x_0$. Hence by induction assumption K is well interpolating set and we can exchange any $f(x_i)$ for $x_i \in S$ by a linear combination of $f(x_i)$ for $x_i \in J/(x_0 + 1)$. So in overall we obtain that J is depended set, namely f is well interpolate on J.

5.2 (b)

Define the function $g(x) = f(t^{-1}(x - s))$. Note that q is prime, thus $(\mathbb{F}_q/0, \cdot)$ and $(\mathbb{F}_q, +)$ are groups and the inverse elements $-s, t^{-1}$ are exist and uniqs. Suppose that y is a zero of $g \Rightarrow f(t(y+s)) = g(y) = 0$, Hence the number of zeros of f equals to the number of zeros of g, which means that their degree are equal $\Rightarrow g$ is also a polynomial at degree at most d, $\Rightarrow a_1, a_2..a_d$ are also the interpolation coefficients respecting to the interpolation set $\{tx_1 + s, tx_2 + s, .., tx_d + s\}$.

6 Ex (5).

6.1 (a)

As shown in the previous section, by the fact that q is prime, we have that $g_{u,v}$ acts on $\mathbb{F}_q^m$ by $g_{u,v}(x) = u + vx$, (for any $v \neq 0$). Thus, $f(g_{u,v}(x))$ is just a permutation over the values of f. As the number of zeros remains the same, we have that $f(g_{u,v}(x))$ is also a degree d polynomial. Therefore, the restricted polynomial $f|_L$ corresponds to the restriction L' of another polynomial obtained by taking $u' = 0$ and v to be supported only on a single coordinate. Hence, the restricted polynomial can have at most d zeros.

6.2 (b)

As we have that f is a polynomial of degree exactly d, there must be a monomial $x_1^{d_1} x_2^{d_2} .. x_k^{d_k}$ such that $\sum_i d_i = d$. Denote by g the sum of all those monomials, and by $v \in \mathbb{F}_q^n$ a coordinate on which $g(v) \neq 0$ (if there is no such v, then we could write f as a sum of monomials, each of degree at most $d-1$).

Now, as $g(v), t \neq 0$, we obtain that $g(vt)$ is equal exactly to $t^d \cdot c$ where c is the sum of coefficients of each monomial of g (also $c = g(1)$). As $f(x) - g(x)$ is a polynomial of degree $d-1$, it holds from the previous section that $(f-g)(vt)$ is also a polynomial of degree at most $d-1$. Thus, it cannot zero out $g(vt) \Rightarrow f(vt) = (f-g)(vt) + g(vt)$ is also a polynomial of degree d.

7 Ex (6).

7.1 (a) and (b).

Notice that the polynomial f could be written as $\sum_{i,j} a_i a_j x^i y^j$. Thus, the assignments of $(d+1)^2$ points define $(d+1)^2$ equations over $(d+1)^2$ variables. In addition, as the determinant of the matrix equals

$$\sum_{\sigma \in S_n} (-1)^{\sigma(\pi)} \prod_{\sigma(x^{\sigma(n)_1} y^{\sigma(n)_2})} = \prod_{i<j} (x_i - x_j) \prod_{i<j} (y_i - y_j) \tag{1}$$

So the detriment is not zero, thus we can solve that system by gauss elimination and obtain unique solution .

7.2 (c).