**Petar Maymounkov**
MIT 6.859 – Randomness and Computation, with Ronitt Rubinfeld
Collaborators: Benjamin Rossman

# Problem Set 5 – Solutions

## Problem 1

We have $\text{Ext} : \{0,1\}^n \to \{0,1\}$ and $0 \le \delta \le 1/2$. Observe that there is $D \subseteq \{0,1\}^n$ with $|D| = 2^n/2$ where Ext is constant. Fix one such $D$. Define the SV-source $X$ as:

$$\mathbf{Pr}[X = x] = \begin{cases} 2\delta/2^n, & \text{if } x \notin D \\ 2(1-\delta)/2^n, & \text{otherwise} \end{cases}$$

Note that if $\text{Ext}(D) = 0$ then $\mathbf{Pr}[\text{Ext}(X) = 1] \le \delta$, otherwise $\mathbf{Pr}[\text{Ext}(X) = 1] \ge 1 - \delta$. We have to show that $X$ is an SV-source with parameter $\delta$.

Given a fixed $i \in [n]$ and $x_1, \ldots, x_{i-1} \in \{0,1\}$, let:

$$U_b = \left\{ y \in \{0,1\}^n \;\middle|\; y_i = b \wedge \bigwedge_{k \in [i-1]} y_k = x_k \right\}$$

Also set $U_* = U_0 \cup U_1$. For shorthand, let $p_\beta = \mathbf{Pr}[X \in U_b]$, where $\beta \in \{0, 1, *\}$. Then:

$$\mathbf{Pr}\left[ X_i = 1 \;\middle|\; \bigwedge_{k \in [i-1]} X_k = x_k \right] = \frac{p_1}{p_*} = \frac{p_1}{p_0 + p_1} = \frac{1}{1 + p_0/p_1}$$

Now, observe that $(2\delta/2^n)|U_j| \le p_j \le (2(1-\delta)/2^n)|U_j|$ for $j \in \{0,1\}$. Since $|U_0| = |U_1|$, we get:

$$\frac{\delta}{1-\delta} \le \frac{p_0}{p_1} \le \frac{1-\delta}{\delta}$$

Which, in turn, implies that:

$$\delta \le \frac{1}{1 + p_0/p_1} \le 1 - \delta$$

## Problem 2

**Part (a)** Let $\text{Ext} : \{0,1\}^n \to \{0,1\}^m$ be a fixed extractor, and let $X$ be a fixed flat $k$-source on $K \subseteq \{0,1\}^n$ where $K \cong \{0,1\}^k$. If $d_{TV}(\text{Ext}(X), U_m) > \epsilon$, then there must exist $A \subseteq \{0,1\}^m$ with $|A| = 2^m/2$ such that $\mathbf{Pr}_x[\text{Ext}(x) \in A] > 1/2 + \epsilon/2$.

The latter is equivalent to $|\text{Ext}^{-1}(A)| > (1/2 + \epsilon/2)2^k$. For a fixed $A$ (of size $2^m/2$), let $Y_A = |\text{Ext}^{-1}(A)|$ (a random variable). And let $Z_i$, for $i \in K$, be the indicator that $\text{Ext}(i) \in A$. Then $Y_A = \sum_{i \in K} Z_i$, and therefore $\mathbf{E}[Y_A] = \sum_{i \in K} |A|/2^m = 2^k/2$.

Applying a Chernoff bound for $Y_A$ yields that:

$$\mathbf{Pr}_{\text{Ext}}\left[Y_A > (1/2 + \epsilon/2)2^k\right] < \exp\left(-\epsilon^2 2^{k-1}/3\right)$$

Next, applying a union bound over all $A \subseteq \{0,1\}^m$ with $|A| = 2^m/2$ (at most $2^{2^m}$ in count) and using that $m = k - 2\log 1/\epsilon - D$ (where $D = O(1)$) produces:

$$\mathbf{Pr}_{\text{Ext}}\left[\bigvee_{A \subseteq \{0,1\}^m} Y_A > (1/2 + \epsilon/2)2^k\right] < \exp\left(2^k \epsilon^2\big(-1/6 + (\ln 2)/2^D\big)\right)$$

Picking $D$ to be sufficiently large completes the proof.

**Part (b)** Build $\text{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ randomly. Observe that when Ext is fed a fixed flat $k$-source $X_k$ and a random number $U_d$, it can be treated as a function $\{0,1\}^{n+d} \to \{0,1\}^m$ which is fed a specific $(k+d)$-source $Y_{k+d} = Y_{k+d}(X_k)$. Therefore part (a) applies and the probabibility that $\text{Ext}(Y_{k+d})$ fails to be at most $\epsilon$-far from $U_m$ is $2^{\Omega(-2^{k+d}\epsilon^2)}$. Taking a union bound over all $X_k$ (equivalently over all $Y_{k+d}$) which are $\binom{2^n}{2^k} \leq 2^{k+k(n-k)\ln 2}$ in count, yields that the probability that Ext fails to be a $(k,\epsilon)$-extractor is $2^{\Omega(-2^k(n-k)C' + (n-k)kC'')}$. Choosing the constants $C'$ and $C''$ appropriately ensures that this probability is strictly less than 1, and therefore by the probabilistic method, such an extractor must exist.