Foundations and Frontiers of Probabilistic Proofs (July 2023)
Worksheet 10: Polynomial-Size PCP
Date: July 21, 2023

An ϵ-*biased generator* over a finite field $\mathbb{F}$ is a function $G\colon \mathbb{F}^k \to \mathbb{F}^m$ such that for every non-zero vector $v \in \mathbb{F}^m$ it holds that $\Pr_{x\in\mathbb{F}^k}[\langle v, G(x)\rangle = 0] \le \epsilon$, where $\langle v, G(x)\rangle$ is the inner product over $\mathbb{F}^m$.

Problem 1. (Biased generators from linear codes) Let $C\colon \mathbb{F}^m \to \mathbb{F}^n$ be a linear error correcting code with relative distance $1-\epsilon$. Recall that a linear error correcting code can be represented by a generator matrix $A \in \mathbb{F}^{n\times m}$, which is such that $C(v) = A\cdot v$ for all $v \in \mathbb{F}^m$. Construct an ϵ-biased generator from the matrix A. You may assume $|\mathbb{F}|^k$ is a multiple of n for some k.

In lecture we reduced the satisfiability of a system of quadratic polynomials $(p_1, p_2, \ldots, p_m)$ to the satisfiability of a single quadratic polynomial q. One of the suggested options was to sample a random $r \in \mathbb{F}^m$ and set $q := \sum_{i\in[m]} r_i p_i$. This method has a small soundness error, $O(\frac{1}{|\mathbb{F}|})$, but uses too much randomness, $\Omega(m \log|\mathbb{F}|)$. Instead, we identified $[m]$ with $H_e^{s_e}$ where $s_e := \frac{\log m}{\log|H_e|}$ where H_e is a subset of $\mathbb{F}$ of size $O(\log m)$, and we set $q := \sum_{0\le i_1,\ldots,i_{s_e}<|H_e|} r_1^{i_1}\cdots r_{s_e}^{i_{s_e}}\cdot p_{i_1,\ldots,i_{s_e}}$ with each $r_j \in \mathbb{F}$ uniformly. This achieves a soundness error of $O(\frac{s_e|H_e|}{|\mathbb{F}|})$ and uses $O(s_e \log|\mathbb{F}|)$ random bits. The field can then be chosen to be large enough so the soundness error is at most a constant and small enough so that the amount of randomness is logarithmic in m.

Problem 2. (Randomized reduction from biased generators)

1. Prove that choosing $r \in \mathbb{F}^m$ according to an ϵ-biased generator $G\colon \mathbb{F}^k \to \mathbb{F}^m$ we can reduce the randomness requirements from $O(m\log|\mathbb{F}|)$ to $O(k\log|\mathbb{F}|)$, while incurring a soundness error of ϵ.

2. Show that the strategy outlined above is a particular case of the previous item, i.e., that taking $k = s_e$, the mapping $G\colon \mathbb{F}^k \to \mathbb{F}^m$ where the $(i_1,\ldots,i_{s_e})^{\text{th}}$ coordinate of $G(r_1,\ldots,r_{s_e})$ is $r_1^{i_1}\cdots r_{s_e}^{i_{s_e}}$ is an $O(\frac{s_e|H_e|}{|\mathbb{F}|})$-biased generator.

Problem 3. (Settings for logarithmic randomness) In lecture we chose $|H_e| = O(\log m)$. Suppose we had instead chosen $|H_e| = 2$ (e.g., $H_e = \{0,1\}$).

1. What is the number of variables s_e in this case?

2. How large should $\mathbb{F}$ be to achieve soundness error $\frac{1}{2}$?

3. Taking the field size from the previous item, how much randomness do we need to sample r? Explain why this is "too much" randomness for the construction.

Similar considerations also apply for the size of H_v.