# Small Bias Sets

## Not based on any particular source

Gil Cohen

December 21, 2020

# Overview

1. Pseudorandom generators
2. Small bias sets
3. Explicit construction of size $O(n^2/\varepsilon^2)$
4. Explicit construction of size $O(n/\varepsilon^8)$

# Pseudorandom generators

Let $C$ be a class of functions of the form $f : \{0,1\}^* \to \{0,1\}$. We partition $C = \cup_n C_n$ according to the input length.

## Definition

A family of functions $\mathbf{G}_n : \{0,1\}^s \to \{0,1\}^n$ is an $\varepsilon$ pseudorandom generator for $C$ if for every $n$ and $f \in C_n$,

$$\left|\mathbb{E}[f(\mathbf{G}_n(U_s))] - \mathbb{E}[f(U_n)]\right| \leq \varepsilon.$$

The parameter $s = s(n, \varepsilon)$ is called the seed length.

# Pseudorandom generators

A fundamental question in complexity theory is to construct PRG for natural classes with short seed.

For example, PRG with logarithmic seed length for the class of polynomial-time computable functions implies $\mathbf{BPP} = \mathbf{P}$.

## Proposition

*Every class* $C = \{C_n\}_{n \in \mathbb{N}}$ *has a PRG with seed length* $s = \log_2 \log |C_n| + 2 \log \frac{1}{\varepsilon} + O(1)$.

# Small bias sets

Recall that an $\mathbb{F}_2$ linear function is a function of the form $f(x_1, \ldots, x_n) = \sum_{i \in S} x_i$ for some $S \subseteq [n]$, where the sum is taken over $\mathbb{F}_2$.

## Proposition

*For every $n, \varepsilon$ there exists an $\varepsilon$-PRG for the class of $\mathbb{F}_2$ linear functions with seed length $s = \log_2 n + 2 \log_2 \frac{1}{\varepsilon} + O(1)$.*

# Overview

# Small bias sets

## Definition

Let $\mathbf{G}$ be an $\varepsilon$-PRG for the class of $\mathbb{F}_2$ linear functions. The distribution $\mathbf{G}(U_s)$ is called an $\varepsilon$-biased set. Hence, there exist $\varepsilon$-biased sets of size $O(n/\varepsilon^2)$.

It is a basic open problem in pseudorandomness to find explicit constructions of small-bias sets with such parameters. Known constructions have size:

# Small bias sets

- Naor-Naor (1990) $n/\varepsilon^{O(1)}$
- Alon-Goldreich-Hastad-Peralta (1990) $n^2/\varepsilon^2$
- Easy from AGHP $n \log^2 n/\varepsilon^3$
- Alon-Bruck-Naor-Naor-Roth (1992) / AGHP + Algebraic-Geometric codes gives $n/\varepsilon^3$
- Ben-Aroya Ta-Shma (2009) $(n/\varepsilon^2)^{5/4}$
- Ta-Shma (2017) $n/\varepsilon^{2+o(1)}$

# Small bias sets and Cayley graphs

**Proposition**

$S \subseteq \mathbb{F}_2^n$ *is* $\varepsilon$*-biased* $\implies$ $\mathrm{Cay}(\mathbb{F}_2^n, S)$ *is a* $(1-\varepsilon)$*-spectral expander.*

# Overview

# Explicit constructions

For $x, y \in \mathbb{F}_2^m \simeq \mathbb{F}_{2^m}$ define $s_{x,y} \in \mathbb{F}_2^n$ as follows: for every $0 \leq i < n$, $(s_{x,y})_i = x^i y$.

## Proposition

*The set $S = \{s_{x,y} \mid x, y \in \mathbb{F}_2^m\}$ is $n/2^m$-biased.*

## Corollary

*There exists an explicit $\varepsilon$-biased set in $\{0,1\}^n$ of size $O(n^2/\varepsilon^2)$.*

# Extra space for the proof

# Binary codes

## Definition

A function $C : \{0,1\}^n \to \{0,1\}^m$ is a code with distance $\delta$ if for every distinct $x, y \in \{0,1\}^n$ it holds that $|C(x) - C(y)| \geq \delta m$.

## Definition

A code $C$ is linear if it is an $\mathbb{F}_2$-linear function.

# Binary codes and small bias sets

## Proposition

*A binary code* $C : \{0,1\}^n \to \{0,1\}^m$ *with distance* $\delta$ *induces an* $\varepsilon = \frac{1-\delta}{1+\delta}$ *biased set in* $\{0,1\}^n$ *of size* $m$.

# Explicit construction of size $O(n/\varepsilon^8)$

Due to lack of time, we will assume we have explicit binary codes with constant distance $\delta$ and $m = O(n)$.

In the problem set you proved the following.

**Theorem**

*Let $G = (V, E)$ be a $1 - \omega$ spectral expander on $n$ vertices. Let $f : V \to \{0, 1\}$ be a labelling with bias $\mu$. Then,*

$$|\mathbb{E}_{v_1,\ldots,v_t}[(-1)^{f(v_1)+\cdots+f(v_t)}]| \leq (O(\mu^2 + \omega))^{t/4},$$

*where $v_1, \ldots, v_t$ is a random walk on $G$.*

(In fact a stronger statement holds in which the exponent is $t/2$.)

# Explicit construction of size $O(n/\varepsilon^8)$

We now describe an explicit construction of an $\varepsilon$-biased set of size $O(n/\varepsilon^8)$.

# Extra space for the proof