Remarks on Bias Correctors

Maciej Skorski

DELL

June 24, 2019

Overview

- Bias correctors decrease input bias, at the price of compression; fundamental for TRNGs!
- 1-bit output example: celebrated XOR [Dav02]
- Multiple-bit output: non-linear maps [Dic07] or resilient linear codes [Lac08]
- How about optimality of assumptions and bounds?

What is exactly bias?

Fix a distribution $X \in \{0,1\}^n$ and a candidate $f : \{0,1\}^n \to \{0,1\}^m$ for a bias corrector.
We study one of the following metrics for $Z = f(X)$

- $\textsc{max-bias}(Z) = \max_y |\Pr[Z = y] - 2^{-m}|$
- $\textsc{total-bias}(Z) = \frac{1}{2}\sum_y |\Pr[Z = y] - 2^{-m}|$

For 1-bit Z simplifies to $\textsc{bias}(Z) = |\Pr[Z = 1] - \Pr[Z = 0]| = |\mathbb{E}(-1)^Z|$.

Bias by Fourier Analysis

- Consider linear spaces $\mathbf{F}_2^n \equiv \{0,1\}^n$ and $\mathbf{F}_2^m \equiv \{0,1\}^m$
- Compute bias for *all linear combinations of the outupt f*

$$f.u(x) = \sum f(x)_i \cdot u_i = \bigoplus_{i:\ u_i \neq 0} f(x)_i$$

$$\Delta(u) = \mathbb{E}_{x \sim X}(-1)^{f.u(x)}|$$

- One-dimensional biases $\Delta(u)$ are connected to the original bias [Lac08, Gol95]
- Bias for a single-valued $g : \mathbf{F}_2^n \to \mathbf{F}_2$ also computed by the Fourier expansion

$$(-1)^g = \sum_I \hat{g}_I \prod_{i \in I} (-1)^{x_i}$$

- Works very well under the *independent bits model*

Bias by Fourier Analysis

The folowing result shows how to connect single- and multidimensional biases.

Theorem (Multidimensional Output Bias / XOR Lemma)

Let Δ be as before, then

- $\textsc{max-bias}(f(X)) \leqslant \max_u \|\Delta(u)\|_\infty$ *[Lac08, Gol95]*
- $\textsc{total-bias}(f(X)) \leqslant \frac{1}{2} \cdot 2^{m/2} \max_u \|\Delta(u)\|_\infty$ *[Gol95]*

Compute Bias with Fourier Analysis - Examples

Example (single-bit output)

If X has independent bits each with bias ϵ, then $f(x) = \bigoplus_i x_i$ has bias of $\frac{1}{2} \cdot (2\epsilon)^n$.

Example (multi-bit output from linear codes)

If X has independent bits each with bias ϵ, and $f : \{0,1\}^n \to \{0,1\}^m$ is a linear code with distance d then $|\Delta(u)| \leqslant \frac{1}{2} \cdot (2\epsilon)^d$ for each u.

Example (reslient codes)

A linear (n, m, d) code is $t = d - 1$ resilient because with $n - t$ unbiased bits the output is unbiased.

Better Bias Analysis

Using total bias and sharper bounds on fourier transforms, one gets better bounds than Lacharme.

Consider inputs with bias $\epsilon = \frac{1}{4}$ and a (n, m, t)-resilient linear code. Then

- $\text{MAX-BIAS} = 2^{-t}$, equivalently min-entropy is $m - \log(1 + 2^{m-t})$
- $\text{TOTAL-BIAS} = 2^{\frac{m}{2}-t}$, closeness to the uniform distribution (smooth min-entropy)

Comparison

- In both cases resilience t large enough compared to m.
- But the second one preferable for indistinguishability applications (e.g. ciphers) from the theoretical perspective.

Removing Independence

Under the Markov model one can work with the *conditional input bias* defined as

$$\mathrm{BIAS}(X) = \max_{x_{<n}} \left| \mathbb{E}(-1)^{X_n} | X_{<n} = x_{<n} \right|$$

and then previous results hold true with $\epsilon = \mathrm{BIAS}(X)$.

Beating XOR Extractors

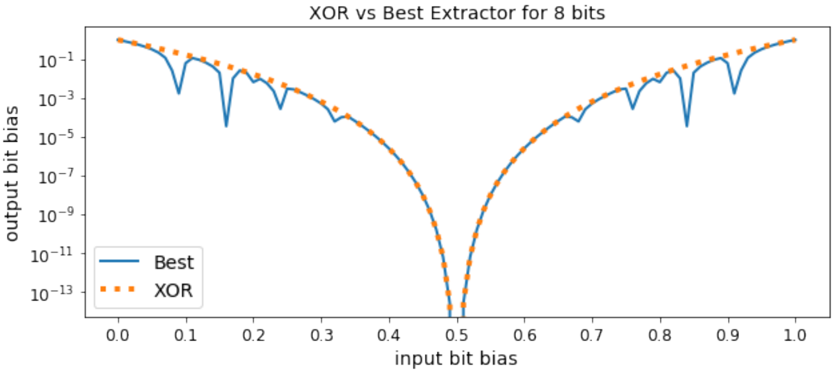
For a given class of distributions (e.g. bias prior region in the IID model), how to build an optimal (min-max) corrector?

Some observations for 1-bit correctors [1]

- XOR is optimal for small bias; for some bias values one can do better!
- Dimensionality reduction: under IID bits and with sufficiently many bias possibilities, the corrector depends on the hamming weight. Search space shrinks from 2^{2^n} to 2^n.
- Dimensionality reduction: under IID bits and symmetric bias prior, the corrector is symmetric w.r.t. the hamming weight. Search space shrinks to $2^{n/2}$
- ...

[1] Unpublished work

Beating XOR Extractors



Summary

- Other stochastic models for discrete sources?
- Trade resilience for entopy (to get condensers)?
- Solve min-max for multidimensional outputs?
- ...

References I

- Robert B Davies, *Exclusive or (xor) and hardware random number generators*, HYPERLINK" http://www. robertnz. net/pdf/xor2. pdf" http://www. robertnz. net/pdf/xor2. pdf (2002).
- Markus Dichtl, *Bad and good ways of post-processing biased physical random numbers*, International Workshop on Fast Software Encryption, Springer, 2007, pp. 137–152.
- Oded Goldreich, *Three xor-lemmas-an exposition*, Electronic Colloquium on Computational Complexity (ECCC, Citeseer, 1995.
- Patrick Lacharme, *Post-processing functions for a biased physical random number generator*, International Workshop on Fast Software Encryption, Springer, 2008, pp. 334–342.