Electronic Colloquium on Computational Complexity, Report No. 205 (2015)

Quadratic maps are hard to sample

Emanuele Viola*

August 20, 2015

Abstract

This note proves the existence of a quadratic GF(2) map $p : \{0,1\}^n \to \{0,1\}$ such that no constant-depth circuit of size poly(n) can sample the distribution $(u, p(u))$ for uniform u.

We continue the study of sampling lower bounds [Vio12a, LV12, DW12, Vio14b, BIL12, Vio12b, BCS14, vio14a]. The paper [Vio14b] exhibits an explicit function $f : \{0,1\}^n \to \{0,1\}$ such that the distribution $(u, f(u))$, for uniform $u \in \{0,1\}^n$, is not the output distribution of any poly(n)-size, constant-depth (AC^0) circuit evaluated on a uniform input. We say that the circuit *cannot sample* $(u, f(u))$. Although explicit, this function is somewhat complicated and so it leaves a gap in our understanding of what can be sampled in AC^0.

This note makes a step towards closing this gap by proving the existence of a quadratic GF(2) map $p : \{0,1\}^n \to \{0,1\}$ such that $(u, p(u))$ cannot be sampled in AC^0. The degree bound on p is tight because for every degree-1 map p there exists an AC^0 circuit that samples $(x, p(x))$ [Bab87]. Moreover, the quadratic map Inner Product, $x_1x_2 + x_2x_3 + \ldots, x_{n-1}x_n$ modulo 2, can be sampled in AC^0 [IN96]. Also related is the result [LV12] which gives a non-boolean, linear transformation $f : \{0,1\}^n \to \{0,1\}^m$ such that $f(u)$ cannot be sampled in AC^0. The quadratic map in this work is the composition of Inner Product with a random linear transformation.

Theorem 1. *For any d and all sufficiently large n:*
There exists a quadratic map $p : \{0,1\}^n \to \{0,1\}$ such that for any AC^0 circuit $C : \{0,1\}^{n^d} \to \{0,1\}^{n+1}$ of depth d and size n^d, the output distribution of $C(v)$ for uniform $v \in \{0,1\}^{n^d}$ is different from the distribution $(u, p(u))$ for uniform $u \in \{0,1\}^n$.

The proof goes by using the fact that any AC^0 source (a.k.a. distribution) is a convex combination of bit-block sources [Vio14b], and then noticing that one can extract from the latter by a quadratic map. Here we do not gain from bounding the size of the blocks.

Bit-block sources are a special case of affine sources, and for the proof we will have to work with the latter. As affine source S over $\{0,1\}^n$ is a random variable that is uniform over

*Supported by NSF grant CCF-1319206. Email: viola@ccs.neu.edu

ISSN 1433-8092

an affine subspace of the vector space given by $\{0,1\}^n$ with component-wise addition modulo 2. The min-entropy of S is the dimension of the space. We say that $f : \{0,1\}^n \to \{0,1\}$ is an ϵ-extractor for a class of sources if for every source S in the class we have $|E[(-1)^{f(S)}| \leq \epsilon$.

Definition 2 (Bit-block source)**.** *A random variable $Y = (Y_1, Y_2, \ldots, Y_n)$ over $\{0,1\}^n$ is a* bit-block source *if every Y_i belongs to $\{0, 1, X_1, X_2, \ldots, X_n, 1 - X_1, 1 - X_2, \ldots, 1 - X_n\}$. A sample from the source is obtained by selecting the X_i uniformly and independently in $\{0,1\}$.*

Note that the entropy of a bit-block source equals the number of different variables X_i (negated or not) that appear in the output.

The following result reduces our task to that of extracting from bit-block sources.

Lemma 3 ([Vio14b])**.** *Suppose that $f : \{0,1\}^n \to \{0,1\}$ is a $o(1)$-extractor for bit-block sources of min-entropy $n^{0.99}$. Then AC^0 circuits of size n^d and depth d cannot sample $(u, f(u))$, for any constant d.*

Proof. Note that $\Pr[f(u) = 1] \geq 1/2 - o(1)$. Suppose that C' is a circuit that samples $(u, f(u))$. Construct the circuit C that first runs C' to obtain a sample $(x, b) \in \{0,1\}^n \times \{0,1\}$. Then, if $b = 1$ it outputs x, otherwise it output a uniform string in $\{0,1\}^n$. Note that C samples a source which has entropy $k \geq n - O(1)$ and on which f is biased. Specifically, $\Pr[f(S) = 1] \geq 1/2 + \Omega(1)$. By Theorem 1.8 in [Vio14b], the output distribution of C is $o(1)$-close to a convex combination of bit-block sources with min-entropy $n^{0.9}$. Hence, f should be nearly unbiased on the output of C, which is a contradiction. □

Thus to prove Theorem 1 it only remains to construct a quadratic map which extracts from bit-block sources. This is given by the next theorem.

Theorem 4. *There exists a quadratic map $f : \{0,1\}^n \to \{0,1\}$ that is a $o(1)$-extractor for bit-block sources with min-entropy $k = n^{1/2+\Omega(1)}$.*

To prove Theorem 4 we first notice that there are few bit-block sources, then we show the existence of a linear map which condenses any such source to an affine source whose entropy is more than half the length of the string, at which point we use the folklore fact that Inner Product extracts from such sources.

The proof of the following claim is immediate from the definition.

Claim 5. *The number of bit-block sources over n bits is at most $(2 + 2n)^n$.*

The next claim gives the condenser.

Claim 6. *Let S be an affine source on n bits with min-entropy k. Let M be a random $k \times n$ matrix. Then the probability that MS has min-entropy less than $0.9k$ is at most $2^{-\Omega(k^2)}$.*

Proof. Let $S = TX + b$, where T is an $n \times k$ full-rank matrix, X is uniform in $\{0,1\}^k$, and $b \in \{0,1\}^n$ is a shift. Consider a change-of-basis full-rank $n \times n$ matrix A such that AT is the $n \times k$ matrix which is identity in the first k rows and 0 everywhere else. Rewriting

M as $MA^{-1}A$, and noting that MA^{-1} is uniform for uniform M, we have that the min-entropy of MS equals the dimension of MAT, which in turn is the dimension of the span of k uniformly chosen vectors. The probability that this dimension is less than $0.9k$ is at most the probability that there exist $0.9k$ vectors such that every other vector lies in their span, which is at most

$$\binom{k}{0.9k}\left(\frac{2^{0.9k}}{2^k}\right)^{0.1k} \leq 2^{O(k)}2^{0.01k^2} = 2^{-\Omega(k^2)}.$$

□

The following result is folklore but we do not find a proof in the literature.

Lemma 7. *Let $f : \{0,1\}^n \to \{0,1\}$ be the Inner Product function modulo 2. Let X be uniform over an affine subspace of $\{0,1\}^n$ with dimension k. Then $|E[(-1)^{f(X)}| \leq 2^{n/2-k}$.*

Proof. Let $X = V + a$ where V is uniform over a vector space of dimension k and $a \in \{0,1\}^n$ is a shift. Let $g(x) := f(x + a)$. Now write

$$E[(-1)^{f(X)}] = E[(-1)^{g(V)}] = 2^{n-k}E[(-1)^{g(U)}1_V(U)],$$

where U is uniform in $\{0,1\}^n$, and 1_V is the 0/1 indicator function of V.

Now let $X^\perp$ be uniform over the orthogonal complement $V^\perp$ of V. Note that for every $a \in \{0,1\}^n$ we have

$$1_V(a) = E[(-1)^{\sum_i Y_i a_i}].$$

To verify this, notice that if $a \in V$ then the inner product $\sum_i y_i a_i$ equals 0 for any $y \in V^\perp$. While if a is not in V then the same inner product equals 1 for some $y \in V^\perp$. In this case consider sampling $X^\perp$ by selecting a uniform linear combination of a basis of $V^\perp$ that contains y. For any choice for the coefficients of the vectors different from y, the inner product $\sum_i Y_i a_i$ will be 0 for one choice of the coefficient for y, and 1 for the other.

Combining these two facts, and using the triangle inequality, we have

$$|E[(-1)^{f(X)}| = 2^{n-k}|E[(-1)^{g(U)}(-1)^{\sum_i Y_i U_i}]| \leq 2^{n-k}E_Y|E_U[(-1)^{g(U)}(-1)^{\sum_i Y_i U_i}]|.$$

The inner expectation is at most $2^{-n/2}$, as follows from noticing that, for every x, $g(x) = f(x) + \ell(x)$ modulo 2, where $\ell(x)$ is an affine function, and using the fact that f is *bent*, i.e., all of its Fourier coefficients haver absolute value $2^{-n/2}$. □

Proof of Theorem 4. By Claim 5 the number of bit-block sources is $O(n^n)$. The latter is $2^{o(k^2)}$ for $k \geq n^{1/2+\Omega(1)}$. By Claim 6 and a union bound, there exists a linear map M such that M condenses any bit-block source to a source on k bits with entropy $0.99k$. By Lemma 7, the evaluation of the Inner Product function on this source is nearly unbiased. Thus, the quadratic map obtained by composing the Inner Product function with M is the desired extractor. □

Acknowledgements. I am grateful to Eli Ben-Sasson for many discussions on this project.

References

[Bab87] László Babai. Random oracles separate PSPACE from the polynomial-time hierarchy. *Information Processing Letters*, 26(1):51–53, 1987.

[BCS14] Itai Benjamini, Gil Cohen, and Igor Shinkar. Bi-lipschitz bijection between the boolean cube and the hamming ball. In *IEEE Symp. on Foundations of Computer Science (FOCS)*, 2014.

[BIL12] Chris Beck, Russell Impagliazzo, and Shachar Lovett. Large deviation bounds for decision trees and sampling lower bounds for AC0-circuits. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:42, 2012.

[DW12] Anindya De and Thomas Watson. Extractors and lower bounds for locally samplable sources. *ACM Trans. Computation Theory*, 4(1):3, 2012.

[IN96] Russell Impagliazzo and Moni Naor. Efficient cryptographic schemes provably as secure as subset sum. *J. of Cryptology*, 9(4):199–216, 1996.

[LV12] Shachar Lovett and Emanuele Viola. Bounded-depth circuits cannot sample good codes. *Computational Complexity*, 21(2):245–266, 2012.

[Vio05] Emanuele Viola. On constructing parallel pseudorandom generators from one-way functions. In *20th IEEE Conf. on Computational Complexity (CCC)*, pages 183–197, 2005.

[Vio12a] Emanuele Viola. The complexity of distributions. *SIAM J. on Computing*, 41(1):191–218, 2012.

[Vio12b] Emanuele Viola. Extractors for turing-machine sources. In *Workshop on Randomization and Computation (RANDOM)*, 2012.

[vio14a] 2014. http://emanueleviola.wordpress.com/2014/11/09/is-nature-a-low-complexity-sampler.

[Vio14b] Emanuele Viola. Extractors for circuit sources. *SIAM J. on Computing*, 43(2):355–972, 2014.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de