A LOWER BOUND ON THE AVERAGE ENTROPY OF A FUNCTION DETERMINED UP TO A DIAGONAL LINEAR MAP ON $\mathbb{F}^n_q$

YARON SHANY AND RAM ZAMIR

Abstract. In this note, it is shown that if $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is any function and $A = (A_1, \ldots, A_n)$ is uniformly distributed over $\mathbb{F}_q^n$, then the average over $(k_1, \ldots, k_n) \in \mathbb{F}_q^n$ of the Rényi (and hence, of the Shannon) entropy of $f(A) + (k_1A_1, \ldots, k_nA_n)$ is at least about $\log_2(q^n) - n$ bits. In fact, it is shown that the average collision probability of $f(A) + (k_1A_1, \ldots, k_nA_n)$ is at most about $2^n/q^n$.

1. Introduction

Suppose that $f : \mathbb{F}_q \to \mathbb{F}_q$ is an arbitrary function (where $q$ is a prime power and $\mathbb{F}_q$ is the finite field of $q$ elements). Let $A$ be a random variable uniformly distributed over $\mathbb{F}_q$. Clearly, $f(A)$ may be far from uniform, while $kA$ is uniform for all $k \in \mathbb{F}_q^*$. Is $f(A) + kA$ nearly uniform for most values of $k \in \mathbb{F}_q^*$? More generally, given a positive integer $n$, for an arbitrary $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ and for $A$ uniformly distributed over $\mathbb{F}_q^n$, is $f(A) + (k_1A_1, \ldots, k_nA_n)$ nearly uniform for most values of $k \in \mathbb{F}_q^n$?

Recall that the Shannon entropy $H(B)$ of a random variable $B$ taking values in a finite set $S$ is defined by $H(B) := -\sum_{s \in S} \Pr(B=s) \log(\Pr(B=s))$, while the collision probability of $B$, $\text{cp}(B)$, is defined by $\text{cp}(B) := \sum_{s \in S} \Pr(B=s)^2 = \Pr(B = B')$, where $B'$ is an independent copy of $B$. The Rényi entropy of $B$, $H_2(B)$, is defined by $H_2(B) := -\log(\text{cp}(B))$. A straightforward application of Jensen’s inequality shows that $H_2(B) \leq H(B)$.

Since both the Rényi entropy and the Shannon entropy measure randomness (where for both entropies the maximum possible value of $\log(|S|)$ is equivalent to having uniform distribution, and the minimum possible value of 0 is equivalent to being deterministic), a possible formal phrasing of the above question on $f(A) + (k_1A_1, \ldots, k_nA_n)$ is: How much smaller than $\log(q^n)$ might the average over $k$ of the Rényi (or Shannon) entropy be?

1Throughout, we write $x_i$ for the $i$th coordinate of a vector $x$. Also, for a function $f$ with codomain $\mathbb{F}_q^n$, we will write $f_i$ for the $i$th component of $f$ (post-composition of $f$ with the $i$th projection).

2From this point on, all logarithms are to the base of 2.
The collision probability itself is yet an additional measure of randomness, where the minimum collision probability of $1/|S|$ is equivalent to having uniform distribution and the maximum possible collision probability of 1 is equivalent to being deterministic. So, another possible formal phrasing of the question on $f(A) + (k_1A_1, \ldots, k_nA_n)$ is: How much larger than $1/q^n$ might the average over $k$ of the collision probability be?

The main motivation for this question is a certain side-information problem in information theory [8]. Several neighboring questions were considered in the literature. For example, the case $n = 1$ of Theorem 1 ahead extends Lemma 21 of [6], stating that for any $f : \mathbb{F}_q \rightarrow \mathbb{F}_q$ there exists $k \in \mathbb{F}_q$ for which $|\{f(x) + kx \mid x \in \mathbb{F}_q\}| > q/2$. The same case of Theorem 1 ahead also extends the main theorem of [1], which states that the average over $k \in \mathbb{F}_q$ of $|\{f(x) + kx \mid x \in \mathbb{F}_q\}|$ (for $f$ a polynomial of degree $< \text{char}(\mathbb{F}_q)$) is at least $q/(2 - 1/q)$. In addition, a somewhat similar question, concerning the min-entropy of $a_1 \cdot f(A) + a_2 \cdot A$ for random $a_1$ and $a_2$ in $\mathbb{F}_q$ and for large $q$ was implicitly considered in the merger literature [3], see, e.g., Sec. 3.1 of [3], and Theorem 18 of [4].

The main contributions of the current note are the following two theorems.

**Theorem 1.** Let $n \geq 1$ be an integer, let $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^n$ be an arbitrary function, and for $k \in \mathbb{F}_q^n$, let $g_k : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^n$ be defined by $g_k(x) := f(x) + (k_1x_1, k_2x_2, \ldots, k_nx_n)$.

Suppose that a random variable $A$ is uniformly distributed over $\mathbb{F}_q^n$. Then

$$
\frac{1}{q^n} \sum_{k \in \mathbb{F}_q^n} H_2(g_k(A)) \geq \log(q^n) - n \log \left(2 - \frac{1}{q}\right).
$$

The point of the theorem is that the average over $k$ of $H_2(g_k(A))$ is at most about $n$ bits below the entropy of a uniform distribution over $\mathbb{F}_q^n$, regardless of $q$ and $f$. Of course, since the Shannon entropy is not smaller than the Rényi entropy, we may replace $H_2$ by $H$ in Theorem 1. In fact, a stronger result is proven:

3It should be noted that in this case ($n = 1$), the result follows immediately from the Leftover Hash Lemma as described, e.g., in Lemma 7.1 of [7], or in Theorem 8 of [2].

4The distribution of $a_1$ and $a_2$ depends on whether the merger in question is the linear merger or the curve merger, see, e.g., the introduction of [5]. For example, for the curve merger of [5], it was shown in [4] that for any $\varepsilon, \delta > 0$, the weighted sum is $\varepsilon$-close (in statistical distance) to having min-entropy $(1 - \delta) \cdot n \cdot \log(q)$, as long as $q \geq (4/\varepsilon)^{1/\delta}$. 


Theorem 2. Using the terminology of Theorem 1, we have
\[
\frac{1}{q^n} \sum_{k \in \mathbb{F}_q^n} cp(g_k(A)) \leq \frac{1}{q^n} \left(2 - \frac{1}{q}\right)^n,
\]
with equality if for all \(i\), \(f_i(x)\) depends only on \(x_i\).

Note that by Jensen,
\[
\frac{1}{q^n} \sum_{k \in \mathbb{F}_q^n} H_2(g_k(A)) = - \sum_{k \in \mathbb{F}_q^n} \frac{1}{q^n} \log(cp(g_k(A))) \geq - \log \left(\frac{1}{q^n} \sum_{k \in \mathbb{F}_q^n} cp(g_k(A))\right),
\]
and hence Theorem 2 implies Theorem 1.

As stated in the theorem itself, the bound of Theorem 2 is tight. The bound of Theorem 1 is also tight, as seen by the following proposition.

Proposition 3. For the function \(f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^n\) defined by \(f(x) := (x_1^2, \ldots, x_n^2)\), we have (using the terminology of Theorem 7)
\[
\frac{1}{q^n} \sum_{k \in \mathbb{F}_q^n} H(g_k(A)) = \log(q^n) - n \left(1 - \frac{1}{q}\right),
\]
and
\[
\frac{1}{q^n} \sum_{k \in \mathbb{F}_q^n} H_2(g_k(A)) = \begin{cases} 
\log(q^n) - n \left(1 - \frac{1}{q}\right) & \text{if } q \text{ is even,} \\
\log(q^n) - n \log \left(2 - \frac{1}{q}\right) & \text{otherwise.}
\end{cases}
\]

2. Proof of Theorem 2

The proof begins as the proof of the Leftover Hash Lemma as appearing in [2]. Letting \(K\) and \(A'\) be random variables uniformly distributed over \(\mathbb{F}_q^n\) such that \(A, K\) and \(A'\) are jointly independent, the left-hand side of (1) can be written as
\[
\frac{1}{q^n} \sum_{k \in \mathbb{F}_q^n} cp(g_k(A)) = \sum_{k \in \mathbb{F}_q^n} \Pr(K = k) \cdot \Pr(g_k(A) = g_k(A'))
\]
\[
= \sum_{k \in \mathbb{F}_q^n} \Pr(K = k) \cdot \Pr(g_K(A) = g_K(A') | K = k)
\]
\[
= \Pr(g_K(A) = g_K(A')).
\]

It follows that Theorem 2 is an immediate consequence of the following Lemma.

Lemma 4. Using the above notation,
\[
\Pr(g_K(A) = g_K(A')) \leq \frac{1}{q^n} \left(2 - \frac{1}{q}\right)^n
\]
with equality if for all i, $f_i(x)$ depends only on $x_i$.

Proof. For $x, x' \in \mathbb{F}_q^n$, let $d_H(x, x')$ be the Hamming distance between $x$ and $x'$ (number of coordinates $i$ for which $x_i \neq x'_i$) and let $x \odot x' := (x_1x'_1, \ldots, x_nx'_n)$. We have

$$\Pr(g_K(A) = g_K(A')) = \Pr(A = A') + \sum_{d=1}^{n} \Pr(d_H(A, A') = d) \cdot \Pr(g_K(A) = g_K(A') | d_H(A, A') = d).$$

(2)

Now, $\Pr(g_K(A) = g_K(A') | d_H(A, A') = d)$ (probability over $K, A$ and $A'$) is the average over pairs of vectors $a, a' \in \mathbb{F}_q^n$ of Hamming distance $d$ of expressions like

$$\Pr(f(a) + K \odot a = f(a') + K \odot a')$$

(probability over $K$). The last expression is either 0 (if $f_i(a) \neq f_i(a')$ for some $i$ for which $a_i = a'_i$) or $q^{n-d}/q^n$ otherwise ($d$ entries of $K$ are determined by the equation, and the other $n - d$ entries are free). So, in either case, the expression in (3) is $\leq q^{-d}$ (with equality if for all $i$, $f_i$ depends only on the $i$th argument), and hence so is the average of these expressions. Substituting in (2), we get

$$\Pr(g_K(A) = g_K(A')) \leq cp(A) + \sum_{d=1}^{n} \frac{q^n(n)(q-1)^d}{q^{2n}} q^{-d}$$

$$= \frac{1}{q^n} + \frac{1}{q^n} \sum_{d=1}^{n} \binom{n}{d} \left(1 - \frac{1}{q}\right)^d$$

$$= \frac{1}{q^n} \left[ \left(2 - \frac{1}{q}\right)^n - 1 \right]$$

$$= \frac{1}{q^n} \left(2 - \frac{1}{q}\right)^n,$$

with equality if for all $i$, $f_i(x)$ depends only on $x_i$.

□

3. Proof of Proposition 3

The assertion regarding the average Shannon entropy will follow immediately from the chain rule for conditional Shannon entropy if we

---

5Note that this cannot happen if for all $i$, $f_i(x)$ depends only on $x_i$. This will show that for such functions we have equality in the proposition.
prove that for \( n = 1 \) and for the function \( f: \mathbb{F}_q \to \mathbb{F}_q \) defined by \( f(x) = x^2 \), we have

\[
\frac{1}{q} \sum_{k \in \mathbb{F}_q} H(g_k(A)) = \log(q) - \left(1 - \frac{1}{q}\right)
\]

for \( A \) uniformly distributed on \( \mathbb{F}_q^* \).

Suppose first that \( q \) is even. Then \( g_0 = (x \mapsto x^2) \) is a permutation on \( \mathbb{F}_q \) (in fact, an automorphism), and so \( H(g_0(A)) = \log(q) \). For \( k, y \in \mathbb{F}_q^* \), let \( X_{k,y} := g_k^{-1}(y) \). We claim that for all \( k \in \mathbb{F}_q^* \) and for all \( y \in \mathbb{F}_q \) with \( X_{k,y} \neq \emptyset \), there are exactly 2 elements in \( X_{k,y} \). On one hand, there are at most two solutions to a quadratic equation, and on the other hand, for \( x \in X_{k,y} \), \( x + k \) is different from \( x \) and satisfies \( g_k(x + k) = g_k(x) \), which means that \( x + k \in X_{k,y} \). Hence in the case of characteristic 2, the average entropy is \((1/q) \cdot \log(q) + (1 - 1/q) \cdot \log(q/2)\), as desired.

For odd \( q \), we claim that for all \( k \in \mathbb{F}_q^* \), there is a single \( y \) with \( |X_{k,y}| = 1 \), and \((q - 1)/2\) values of \( y \) with \( |X_{k,y}| = 2 \): Fix \( k \), take \( y \) with \( X_{k,y} \neq \emptyset \), and let \( x \in X_{k,y} \). Clearly, \( g_k(-k - x) = g_k(x) \), and if \( x \neq -k/2 \), then \(-k - x \neq x \), which implies that \( |X_{k,y}| = 2 \). For \( y \) with \(-k/2 \in X_{k,y} \), \( |X_{k,y}| \) must therefore be odd, and hence necessarily equal \(1\). Hence in the case of odd characteristic, the average entropy is \((1/q) \cdot \log(q) + (1 - 1/q) \cdot \log(q/2)\), as in (4).

It remains to calculate the average Rényi entropy for \( f = (x \mapsto (x_1^2, \ldots, x_n^2)) \). It follows from the above discussion on the Shannon entropy that if \( q \) is even, then for all \( k \) and all \( i \), the collision probability of the \( i \)-th entry of \( g_k(A) \) equals \( 2/q \) if \( k_i \neq 0 \) (uniform distribution on \( q/2 \) elements), and 1/q if \( k_i = 0 \). As the collision probability of a vector of jointly independent random variables is the product of the individual collision probabilities, it follows that \( cp(g_k(A)) = 2^{w(k)/q^n} \), where \( w(k) \) is the Hamming weight of \( k \) (number of nonzero coordinates in \( k \)).

Since

\[
\sum_{k \in \mathbb{F}_q} w(k) = nq^n - nq^{n-1},
\]

we get

\[
\frac{1}{q^n} \sum_k H_2(g_k(A)) = \frac{1}{q^n} \sum_k (\log(q^n) - w(k))
\]

\[
= \log(q^n) - \frac{1}{q^n} (nq^n - nq^{n-1})
\]

\[
= \log(q^n) - n \left(1 - \frac{1}{q}\right),
\]

as desired.

---

\(^6\)Of course, the last \( y \) equals \(-k^2/4\), and the fact that \( |X_{k,y}| = 1 \) for this \( y \) may also be verified directly.

\(^7\)One way to verify the following identity is to note that the sum \( W_q(n) \) of the weights of all vectors in \( \mathbb{F}_q^n \) satisfies \( W_q(1) = q - 1 \) and \( W_q(n) = W_q(n - 1) + (q - 1) \cdot (W_q(n - 1) + q^{n-1}) \) for \( n \geq 2 \).
Finally, if \( q \) is odd, then it follows from the discussion in the beginning of the proof that for all \( k \), the collision probability of any entry of \( g_k(A) \) equals

\[
\frac{1}{q^2} + \frac{q - 1}{2} \left( \frac{2}{q} \right)^2 = \frac{2q - 1}{q^2}.
\]

Because the collision probability of \( g_k(A) \) is the product of the collision probabilities of the individual entries, it follows that for all \( k \),

\[
H_2(g_k(A)) = -\log \left( \frac{1}{q^{2n}} \cdot (2q - 1)^n \right) = -\log \left( \frac{1}{q^n} \cdot \left( 2 - \frac{1}{q} \right)^n \right),
\]

which completes the proof.

**Remark.** Note that in Proposition 3, the components \( f_i \) may be any quadratic functions \( x_i \mapsto a_i x_i^2 + b_i x_i + c_i \) with \( a_i \neq 0 \) for all \( i \) (eliminating \( a_i \) and \( c_i \) is done by an invertible function, and then the linear term is “absorbed” in the averaging over \( k_i \)).

**Acknowledgments**

We are grateful to Avner Dor for carefully reading several earlier drafts and for his helpful comments. We would also like to thank Simon Litsyn for pointing us to [1].

**References**

[1] L. Carlitz, *On the number of distinct values of a polynomial with coefficients in a finite field*, Proc. Japan Acad., 31, pp. 119–120, 1955.

[2] R. Cramer and S. Fehr, *The mathematical theory of information, and applications*, ver. 2.0. Course notes available online at [http://homepages.cwi.nl/~bouman/icc/InfTheory2.pdf](http://homepages.cwi.nl/~bouman/icc/InfTheory2.pdf)

[3] Z. Dvir, *From randomness extraction to rotating needles*, ECCC TR09-077, 2009.

[4] Z. Dvir, S. Kopparty, S. Saraf, and M. Sudan, *Extensions to the method of multiplicities, with applications to Kakeya sets and mergers*, arXiv:0901.2529v2.

[5] Z. Dvir and A. Wigderson, *Kakeya sets, new mergers and old extractors*, in Proc. FOCS 2008, pp. 625–633.

[6] S. Kopparty, V. F. Lev, S. Saraf, and M. Sudan, *Kakeya-type sets in finite vector spaces*, arXiv:1003.3736v1.

[7] D. R. Stinson, *Universal hash families and the leftover hash lemma, and applications to cryptography and computing*, J. Combin. Math. Combin. Comput., 42, pp. 3-31, 2002.

[8] R. Zamir, *Anti-structure problems*, in *Proc. Int. Zurich Seminar on Communications*, Feb. 29 – Mar. 2, 2012, pp. 91–94; available also as arXiv:1109.0414v1.

25B Sirkin St. Kfar Saba, Israel

E-mail address: yaron.shany@gmail.com

Department of EE-Systems, Tel Aviv University, Tel Aviv, Israel

E-mail address: zamir@eng.tau.ac.il