Electronic Colloquium on Computational Complexity, Comment 2 on Report No. 174 (2016)

One-way Communication and Linear Sketching for Uniform Distribution

Swagato Sanyal

December 1, 2016

Abstract

This note is prepared based on the article titled "Linear Sketching over $\mathbb{F}_2$"(ECCC TR16-174) by Sampath Kannan, Elchanan Mossel and Grigory Yaroslavtsev. We quantitatively improve the parameters of Theorem 1.4 of the above work, as well as an earlier bound of ours. In particular, our result implies that for every $\epsilon \in (0, \frac{1}{2})$ and every constant $\delta > 0$, the one-way communication complexity of any Boolean function $f^+(x, y) := f(x \oplus y)$ corresponding to the uniform distribution over the input domain $\{+1, -1\}^n \times \{+1, -1\}^n$ and error ϵ, is asymptotically lower bounded by the linear sketch complexity of $f(x)$ corresponding to the uniform distribution over the input domain $\{+1, -1\}^n$ and error $(2 + \delta)\epsilon$. Our proof is information theoretic; our improvement is obtained by studying the mutual information between Alice's message and the evaluation of certain parities in the Fourier support of f on her input.

We recall the definition of approximate Fourier dimension by Kannan et al. (TR16-174).

Definition 1 (η-approximate Fourier dimension, Kannan et al. 2016)**.** *The η-approximate Fourier dimension of a Boolean function $f(x) = \sum_S \widehat{f}(S)\chi_S(x)$ is defined to be the smallest dimension of any linear subspace $\mathcal{A} \in \mathbb{F}_2^n$ such that $\sum_{S \in \mathcal{A}} \widehat{f}(S)^2 \geq \eta$.*

We will need the following basic fact about the Shannon entropy of ± 1-valued random variables, that can be easily proved by considering the Taylor expansion of the binary entropy function $H(p)$ about $p = \frac{1}{2}$.

Fact 2. *There is a universal constant $k \in (0, 1)$ such that for any random variable X supported on $\{+1, -1\}$, $H(X) \leq 1 - k(\mathbb{E}X)^2$.*

ISSN 1433-8092

1 Notation

For the rest of the note, fix an arbitrary $f : \{+1,-1\}^n \to \{+1,-1\}$. Let $f(x) = \sum_S \widehat{f}(S)\chi_S(x)$. Let $f^+(x,y) = f(x \oplus y)$. Fix an error parameter $\epsilon \in (0, \frac{1}{2})$. Let Π be a one-way protocol for $f^+(\cdot,\cdot)$ with error probability at most ϵ with respect to uniform distribution on inputs, and of cost $c_\Pi = D_\epsilon^{\to,U}(f^+)$. Let M be the random message sent by Alice to Bob in Π. Let U_n denote the uniform distribution over $\{+1,-1\}^n$. For $\eta \in (0,1)$, we denote the η-approximate Fourier dimension of f by $d_\eta(f)$.

Let $X_1, \ldots, X_m$ be jointly distributed random variables. We refer to their joint distribution as $(X_1, \ldots, X_m)$. We will abuse notation and use the same symbol (for example X_1) to refer both to a random variable and its distribution. For a distribution $\mathcal{D}$ and an event $\mathcal{E}$, we shall denote the distribution $\mathcal{D}$ conditioned on $\mathcal{E}$ by $\mathcal{D}\,|_{\mathcal{E}}$.

Let Z be a real-valued random variable. We define

$$V(Z \mid X_1, \ldots, X_m) \stackrel{\triangle}{=} \mathbb{E}_{(x_1,\ldots,x_m)\sim(X_1,\ldots,X_m)} \left(\mathbb{E}[Z \mid (X_1,\ldots,X_m) = (x_1,\ldots,x_m)]\right)^2. \tag{1}$$

For our applications, Z will be a ± 1-valued random variable. In that case, $1 - V(Z \mid X_1, \ldots, X_m)$ can be seen to be the expected conditional variance of Z conditioned on $X_1, \ldots, X_m$. We next show that $V(\cdot \mid \cdot)$ does not decrease under conditioning.

Claim 3. *Let Z, X_1, X_2 be jointly distributed random variables in a probability space, where Z is real-valued. Then,*

$$V(Z \mid X_1, X_2) \geq V(Z \mid X_1).$$

Proof.

$$\begin{aligned}
&V(Z \mid X_1, X_2) \\
&= \mathbb{E}_{(x_1,x_2)\sim(X_1,X_2)} \left(\mathbb{E}[Z \mid (X_1,X_2) = (x_1,x_2)]\right)^2 \\
&= \mathbb{E}_{x_1\sim X_1} \mathbb{E}_{x_2\sim X_2|_{X_1=x_1}} \left(\mathbb{E}[Z \mid (X_1,X_2) = (x_1,x_2)]\right)^2 \\
&\geq \mathbb{E}_{x_1\sim X_1} \left(\mathbb{E}_{x_2\sim X_2|_{X_1=x_1}} \mathbb{E}[Z \mid (X_1,X_2) = (x_1,x_2)]\right)^2 \\
&\qquad\qquad \text{(By Jensen's inequality applied to the convex function } x^2\text{.)} \\
&= \mathbb{E}_{x_1\sim X_1} \left(\mathbb{E}[Z \mid X_1 = x_1]\right)^2 \\
&= V(Z \mid X_1).
\end{aligned}$$

□

The following claim relates $V(\cdot \mid \cdot)$ to the conditional entropy of ± 1-valued random variables.

Claim 4. *Let $Z, X_1, \ldots, X_m$ be jointly distributed random variables in a probability space, where Z is ± 1-valued. Let k be the constant from Fact 2. Then,*

$$H(Z \mid X_1, \ldots, X_m) \leq 1 - k \cdot V(Z \mid X_1, \ldots, X_m).$$

Proof.

$$\begin{aligned}
&H(Z \mid X_1, \ldots, X_m) \\
&= \mathbb{E}_{(x_1,\ldots,x_m)\sim(X_1,\ldots,X_m)} H(Z \mid X_1 = x_1, \ldots, X_m = x_m) \\
&\leq 1 - k \cdot \mathbb{E}_{(x_1,\ldots,x_m)\sim(X_1,\ldots,X_m)} \left(\mathbb{E}[Z \mid X_1 = x_1, \ldots, X_m = x_m]\right)^2 \quad \text{(Fact 2)} \\
&= 1 - k \cdot V(Z \mid X_1, \ldots, X_m).
\end{aligned}$$

□

2 Relating c_Π to Linear Sketch Complexity

In this section we prove our main result. Recall that M will be used interchangeably to refer to the random message sent by Alice to Bob, and the distribution of Alice's messages.

Theorem 5.

$$4\epsilon \geq \sum_S \widehat{f}(S)^2 \cdot (1 - V(\chi_S(x)|M)).$$

where x is uniformly chosen from $\{+1, -1\}^n$.

Proof. Let $\mathcal{D}_m$ denote the distribution of Alice's input x conditioned on the event that $M = m$ (i.e. $\mathcal{D}_m = U_n \mid_{M=m}$). For any fixed input y of Bob, define $\epsilon_m^{(y)} := \mathbb{P}_{x\sim\mathcal{D}_m}[\Pi(x, y) \neq f^+(x, y)]$. Thus,

$$\epsilon = \mathbb{E}_{m\sim M} \mathbb{E}_{y\sim U_n} \epsilon_m^{(y)}. \tag{2}$$

Observe that

$$\epsilon_m^{(y)} \geq \min_{b\in\{0,1\}} \mathbb{P}_{x\sim\mathcal{D}_m}[f^+(x, y) = b] \geq \frac{\mathrm{Var}_{x\sim\mathcal{D}_m} f^+(x, y)}{4}. \tag{3}$$

Now,

$$\begin{aligned}
&\mathrm{Var}_{x\sim\mathcal{D}_m} f^+(x,y) = 1 - \left(\mathbb{E}_{x\sim\mathcal{D}_m} f^+(x,y)\right)^2 \\
&= 1 - \left(\sum_S \widehat{f}(S)\chi_S(y)\mathbb{E}_{x\sim\mathcal{D}_m}\chi_S(x)\right)^2 \\
&= 1 - \Bigg(\sum_S \widehat{f}(S)^2 \left(\mathbb{E}_{x\sim\mathcal{D}_m}\chi_S(x)\right)^2 \\
&\qquad + \sum_{\{S_1,S_2\}:S_1\neq S_2} 2\widehat{f}(S_1)\widehat{f}(S_2)\chi_{S_1\triangle S_2}(y)(\mathbb{E}_{x\sim\mathcal{D}_m}\chi_{S_1}(x))(\mathbb{E}_{x\sim\mathcal{D}_m}\chi_{S_2}(x))\Bigg).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\mathbb{E}_{y\sim U_n}\mathrm{Var}_{x\sim\mathcal{D}_m} f^+(x,y) \\
&= 1 - \sum_S \widehat{f}(S)^2 \left(\mathbb{E}_{x\sim\mathcal{D}_m}\chi_S(x)\right)^2. \qquad (4)
\end{aligned}$$

Taking expectation over messages it follows from (2), (3) and (4) that,

$$\begin{aligned}
4\epsilon &\geq 1 - \sum_S \widehat{f}(S)^2 \cdot \mathbb{E}_{m\sim M}\left(\mathbb{E}_{x\sim\mathcal{D}_m}\chi_S(x)\right)^2 \\
&= \sum_S \widehat{f}(S)^2 \cdot \left(1 - \mathbb{E}_{m\sim M}\left(\mathbb{E}_{x\sim\mathcal{D}_m}\chi_S(x)\right)^2\right). \qquad (5)
\end{aligned}$$

The theorem follows. □

Fix any $\Delta \in (0,1)$ and construct a set $\mathcal{T}$ of parities by the following iterative algorithm. Recall that x is distributed uniformly over $\{+1,-1\}^n$.

Step 1: $\mathcal{T} := \emptyset, i = 0$;

Step 2: Let $\mathcal{T} = \{T_1,\ldots,T_i\}$. If for every $T \notin \mathsf{span}\ \mathcal{T}$, $V(\chi_T(x) \mid M, \chi_{T_1}(x),\ldots,\chi_{T_i}(x)) < 1-\Delta$, return $\mathcal{T}$;

Step 3: Let $T \notin \mathsf{span}\ \mathcal{T}$ be such that $V(\chi_T(x) \mid M, \chi_{T_1}(x),\ldots,\chi_{T_i}(x)) \geq 1-\Delta$. Then, $\mathcal{T} \leftarrow \mathcal{T}\cup\{T\}$; Go to step 2;

Let the set produced by this algorithm be $\mathcal{T} = \{T_1,\ldots,T_\ell\}$, and let T_i be the parity included in $\mathcal{T}$ in the i-th iteration. From the construction it is clear that:

$$\text{For each } S \notin \mathsf{span}\ \mathcal{T}, V(\chi_S(x) \mid M, \chi_{T_1}(x),\ldots,\chi_{T_\ell}(x)) < 1-\Delta. \qquad (6)$$

Claim 6.

$$\sum_{S \notin \mathsf{span}\ \mathcal{T}} \widehat{f}(S)^2 < \frac{4\epsilon}{\Delta}$$

Proof. From Theorem 5 and Claim 3, we have that

$$\begin{aligned}
4\epsilon &\geq \sum_{S} \widehat{f}(S)^2 \cdot (1 - V(\chi_S(x) \mid M, \chi_{T_1}(x), \ldots, \chi_{T_\ell}(x))) \\
&= \sum_{S \in \mathsf{span}\ \mathcal{T}} \widehat{f}(S)^2 \cdot (1 - V(\chi_S(x) \mid M, \chi_{T_1}(x), \ldots, \chi_{T_\ell}(x))) \\
&\qquad + \sum_{S \notin \mathsf{span}\ \mathcal{T}} \widehat{f}(S)^2 \cdot (1 - V(\chi_S(x) \mid M, \chi_{T_1}(x), \ldots, \chi_{T_\ell}(x))) \\
&> \Delta \cdot \sum_{S \notin \mathsf{span}\ \mathcal{T}} \widehat{f}(S)^2. \qquad \text{(From (6))}
\end{aligned}$$

The claim follows. □

Claim 7. *Let k be the constant from Fact 2. Then,*

$$\ell \leq \frac{c_\Pi}{k(1-\Delta)}.$$

Proof.

$$\begin{aligned}
c_\Pi &\geq I(M; \chi_{T_1}(x), \ldots, \chi_{T_l}(x)) \\
&= \sum_{i=1}^{\ell} I(M; \chi_{T_i}(x) \mid \chi_{T_1}(x), \ldots, \chi_{T_{i-1}}(x)) \\
&\qquad \text{(by the chain rule of mutual information)} \\
&= \sum_{i=1}^{\ell} \big(H(\chi_{T_i}(x) \mid \chi_{T_1}(x), \ldots, \chi_{T_{i-1}}(x)) - H(\chi_{T_i}(x) \mid M, \chi_{T_1}(x), \ldots, \chi_{T_{i-1}}(x))\big) \qquad (7)
\end{aligned}$$

By the construction of $\mathcal{T}$, for each i, $T_i \notin \mathsf{span}\ \{T_1, \ldots, T_{i-1}\}$. Hence, $H(\chi_{T_i}(x) \mid \chi_{T_1}(x), \ldots, \chi_{T_{i-1}}(x)) = 1$. Continuing from (7) we have,

$$\begin{aligned}
c_\Pi &\geq \ell - \sum_{i=1}^{\ell} H(\chi_{T_i}(x) \mid M, \chi_{T_1}(x), \ldots, \chi_{T_{i-1}}(x)) \\
&\geq \ell - \sum_{i=1}^{\ell} \big(1 - k \cdot V(\chi_{T_i}(x) \mid M, \chi_{T_1}(x), \ldots, \chi_{T_{i-1}}(x))\big) \quad \text{(Claim 4)} \\
&= \ell k(1-\Delta). \quad \text{(by the construction of } \mathcal{T})
\end{aligned}$$

The claim follows. □

Theorem 8.

$$D^{lin,U}_{2\epsilon/\Delta}(f) \leq \frac{D^{\rightarrow,U}_{\epsilon}(f^+)}{k(1-\Delta)}.$$

Proof.

$$\begin{aligned} D^{lin,U}_{2\epsilon/\Delta}(f) &\leq d_{1-4\epsilon/\Delta} \qquad \text{(Theorem 3.4 (Part (1)), Kannan et al.)} \\ &\leq \ell \qquad\qquad \text{(Claim 6)} \\ &\leq \frac{c_{\Pi}}{k(1-\Delta)} \qquad \text{(Claim 7)} \\ &= \frac{D^{\rightarrow,U}_{\epsilon}(f^+)}{k(1-\Delta)}. \end{aligned} \tag{8}$$

□

By setting $\Delta = \frac{2}{2+\delta}$ and invoking Theorem 8, we have the following corollary.

Corollary 9. *For every $\delta > 0$,*

$$D^{lin,U}_{(2+\delta)\epsilon}(f) = O_{\delta}\left(D^{\rightarrow,U}_{\epsilon}(f^+)\right).$$

Acknowledgement: I thank Nikhil Mande for useful comments, proofreading the manuscript, and assisting with the presentation. This work was done while I was an intern at the Centre for Quantum Technologies, National University Singapore. I thank all members of CQT for their support.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de