CS 830: Randomness and Complexity 01/27/05

Lecture 4: Weak to Strong one way: construction 3

instructor: Jin-Yi Cai Scribe: Scott Diehl, Jeff Kinne

Given a $w(n)$-weak one-way function $f$, we would like to construct a strong one-way function $h$. We have already seen two ways in which $h$ can be constructed, namely as a list of $f$ evaluated at several points, and by iterating a process of applying $f$ and then shifting by a completely independent bit vector. We have seen that both constructions convert a weak one-way function into a strong one-way function. However, the first construction increased the security parameter in the construciton, and the second construction used very many random bits. The construction that we analyze today produces a function with the same security parameter as $f$ and does not use too many public random bits.

## 1 Description of Construction

Today we consider the recursively defined construction described in the last lecture. Let $\ell = O(\log \frac{1}{w(n)})$, and the public input to $g$ be given by the shift vectors $\pi = \langle \pi_1, \pi_2, \ldots, \pi_l \rangle$ where $\pi \in \{0,1\}^n$. To define $g$, we first define the functions $s_i$ for $1 \leq i \leq \ell$, where

$$
\begin{aligned}
s_1(x, \pi_{\{1\}}) &= f(f(x) \oplus \pi_1) \\
s_2(x, \pi_{\{1,2\}}) &= s_1(s_1(x, \pi_{\{1\}}) \oplus \pi_1, \pi_{\{1\}}) \\
&\cdots \\
s_i(x, \pi_{\{1,\ldots,i\}}) &= s_{i-1}(s_{i-1}(x, \pi_{\{1,\ldots,i-1\}}) \oplus \pi_i, \ldots \pi_{\{1,\ldots,i-1\}}).
\end{aligned}
$$

Then we let $s(x, \pi) = s_\ell(x, \pi)$ and can finally define $g$ to be

$$g(x, \pi) = (s(x, \pi), \pi).$$

The function $s_i$ can be thought of as applying the process $s_{i-1}$, using the first $i-1$ shifts, then injecting some "fresh randomness" by shifting by $\pi_i$. We now essentially have a random new starting point, on which $s_{i-1}$ again uses the first $i-1$ shifts. The intuition behind this approach is that we use the same random process twice to apply $f$ at various points based on the first $i-1$ shifts. However, we'd like the points evaluated in each run of this process to be independent while reusing the same random shifts for each run. In order to achieve this, we apply a new random shift vector to the output of the first run, causing the starting point of the second run to be uniformly distributed and independent of the starting point of the first run. We can see that this technique will apply $f$ at exponentially many points (in the number of shift vectors) which are essentially independent of each other, offering a savings in the number of public bits over the iterative shift and apply method of last time (where the number of evaluation points is linear in the number of shifts).

Also recall that we can unwind the recursive definition of $s$ and explicitly compute this process iteratively as:

$$\begin{aligned} y_0 &= x \\ y_1 &= f(y_1) \oplus \pi_1 \\ y_2 &= f(y_2) \oplus \pi_2 \\ y_3 &= f(y_3) \oplus \pi_1 \\ &\cdots \\ g(x,\pi) &= f(y_{2^\ell - 1}) \end{aligned}$$

where $y_i = f(y_{i-1}) \oplus \pi_{m_i}$, and $m_i$ is the minimum bit position in the binary representation of $i$ that is a 1. So, the sequence of $\pi$ that are used is $1, 2, 1, 3, 1, 2, 1....$

As with the previous constructions, we will construct a bipartite graph for this construction and demonstrate that it is forward (and thus reverse) expanding. We will use this result to show that if $f$ is $w(n)$ weak one-way, then $g$ is $\frac{1}{8}$ weak one-way, and will then need to employ one further step to arrive at a strong one-way function that doesn't use too many public random bits.

We define the bipartite graph $(L, R, E)$ with $L$ the set of inputs $x$ to $f$ and $R$ the set of inputs $z, \pi$ to $g$. We then connect $x \in L$ to $\langle z, \pi \rangle \in R$ if $x$ appears as one of the $y_i$'s in the above computation of $s(z, \pi)$.

Note that such a setup has the two properties that we have desired for our previous constructions. Namely, that we can compute the final value of $g$ of a random neighbor of $f(x)$, and that the ability to invert such a value gives us the ability to invert $f(x)$. The former can be done by randomly choosing an $i, \pi$ and acting as though $y_i = f(x)$ in the $i^{th}$ step of computing $s$, computing forward to obtain the value. The latter can be done by using the inverted value $z, \pi$ as a starting point and computing forward to $y_{i-1}$ and shifting back by the appropriate $\pi$. Therefore, this graph is suited to use in the proof paradigms introduced in the previous lectures.

## 2 Expansion Result

Now our goal is to prove the technical step needed for our argument, namely

LEMMA 2.1. *$H = (L, R, E)$ has $(\frac{w(n)}{2}, \frac{3}{4})$-forward expansion.*

*Proof.* Suppose we have an arbitrary set $S \subseteq L$ of measure $\epsilon = w(n)/2$. To prove the result, we must show that the measure of the neighbor set of $S$ is at least $1/4$.

The key concept in proving this result is that of calculating the size of the set of $x, \pi$ such that during the computation of $s(x, \pi)$, at least one $y_i \in S$. Such $x, \pi$ comprise the neighbor set of $S \subseteq L$, whose size we wish to bound.

Towards this end, we define

$$F_i = \{x \in \{0,1\}^n | \text{ at least one } y_j(x, \pi_{\{1,\ldots,i\}}) \in S \text{ for } 1 \le j \le 2^i\}.$$

By how $s$ is recursively defined, we can see that $F_i$ contains all the points of $F_{i-1}$ plus those in "$F_{i-1}$ shifted by $\pi_i$". More precisely, we add in all the points which land in $S$ during the computation of $s_{i-1}$ starting from the points of $F_{i-1}$ shifted by $\pi_i$. We write

$$F_i = F_{i-1} \cup s_{i-1}^{-1}(F_{i-1} \oplus \pi_i).$$

Fixing $\pi_{\{1,\ldots,i-1\}}$ with $\pi_i$ uniformly chosen, what is the resulting size of $F_i$? Note that fixing $\pi_{\{1,\ldots,i-1\}}$ makes $s_{i-1}^{-1}$ a one-to-one function so that $|s_{i-1}^{-1}(F_{i-1} \oplus \pi_i)| = |F_{i-1}|$. Since we can now write

$$\begin{aligned}
|F_i| &= |F_{i-1} \cup s_{i-1}^{-1}(F_{i-1} \oplus \pi_i)| \\
&= |F_{i-1}| + |s_{i-1}^{-1}(F_{i-1} \oplus \pi_i)| - |F_{i-1} \cap s_{i-1}^{-1}(F_{i-1} \oplus \pi_i)| \\
&= 2|F_{i-1}| - |F_{i-1} \cap s_{i-1}^{-1}(F_{i-1} \oplus \pi_i)|,
\end{aligned}$$

we must now focus on finding the intersection of $F_{i-1}$ with $s_{i-1}^{-1}(F_{i-1} \oplus \pi_i)$. From how the latter set is generated, we would think that this intersection is similar in size to the intersection of a set with a random shift of itself.

For now, we consider the expectation, which by linearity is

$$E_{\pi_i}(|F_i|) = 2|F_{i-1}| - E_{\pi_i}[|F_{i-1} \cap s_{i-1}^{-1}(F_{i-1} \oplus \pi_i)|]$$

Consider a point in $F_{i-1}$ as it is randomly shifted by $\pi_i$ and processed through $s_{i-1}$. Since $s_{i-1}$ is one-to-one, the resulting point is uniformly distributed. Therefore, each of the $|F_{i-1}|$ points has a $|F_{i-1}/2^n|$ chance of being in the intersection, so the expected size of the intersection is $|F_{i-1}|^2/2^n$. From this we can conclude that the expected measure of $F_i$ is

$$E_{\pi_i}[\mu(F_i)] = 2\mu(F_{i-1}) - \mu(F_{i-1})^2 = 1 - (1 - \mu(F_{i-1}))^2.$$

This is a convex function of $\mu$, so we can lower bound the expectation by its linear approximation. Specifically, if $\mu(F_{i-1}) \leq 1/2$ then $\mu(F_i) \geq \frac{3}{2}\mu(F_{i-1})$. Thus the expected size measure of the neighbor set grows by a constant factor (greater than one) with each stage. Note that if the measure of $F_i$ is ever greater than $1/2$, then we are already done because we are only trying to show that at the end it is at least of measure $1/4$. With this, we know that our construction has essentially succeeded: a constant gain in meausure is expected for each step, and after we work out the details we will see that after the final step the measure will be a constant fraction.

Specifically, let

$$\begin{aligned}
\epsilon_0 &= \mu(S) = \epsilon \\
\ldots \epsilon_i &= \mu(F_i).
\end{aligned}$$

By the above calculations, we have that

$$E_{\pi_i}[\epsilon_i/\epsilon_{i-1}] \geq 3/2.$$

We now bound the probability that at step $i$ there is some constant factor increase of the neighbor set slightly less than the expectation of $3/2$. Let $Z = \epsilon_i/\epsilon_{i-1}$. Clearly $Z$ is nonnegative, and since $|F_i|$ is no more than double $|F_{i-1}|$ but not less than $|F_{i-1}|$, we have that $1 \leq Z \leq 2$. Letting $p = \Pr[Z \leq 5/4]$, we can write

$$3/2 \leq E[Z] \leq 2 \leq 2 + p/4 = 5/4p + 2(1-p),$$

from which we can deduce that $p \leq 2/3$. Thus at each step, with probability at least $1/3$, the neighbor set increases by at least a factor of $5/4$.

We now consider the process of the tossing of $i$ coins which lands on heads with probability 1/3. When $\epsilon_i$ is at most 1/2 (where our linear estimate of the expectation holds), the probability that we get at least $j$ heads in such an experiment is a lower bound on the probability that the measure of $F_i$ is at least $(5/4)^j \epsilon$ (i.e. that we increase by a factor of 5/4 at least $j$ times). We can apply the Chernoff bound to such a process, yielding

$$\Pr_{\pi_{\{1,\ldots,\ell\}}}[\epsilon_\ell \geq 1/2] \geq 1 - e^{-c\ell},$$

for some constant $c$, when we pick $i = \ell$ and $j = \ell/6 = \log_{5/4} 1/w(n)$.

We need to relate this information back to what we ultimately want to know about - $\mu(\Gamma(S))$. Notice that

$$E_\pi[\epsilon_\ell] = \int_\pi \Pr_x[x \in \ldots] = \Pr_{\pi,x}[\exists j \in \{1, \ldots, N\} st. y_j(x, \pi) \in S] = \mu(\Gamma(S)).$$

By Markov's inequality, we know that

$$\Pr[\epsilon_\ell \geq 1/2] \leq E[\epsilon_\ell] / \frac{1}{2}.$$

Having shown that the left-hand side is exponentially close to 1, we can now more or less say that

$$\mu(\Gamma(S)) \geq 1/4.$$

□

We can now use this lemma and use the forward-to-reverse expansion lemma to show that $H$ is reverse expanding. With appropriate choice of parameters, this yields that $g$ is a 1/8-weak one-way function when $f$ is a $w(n)$-weak one way function. Now that we have a constant fraction of the space as the "hard-core", we can apply the shift and repeat algorithm of the last lecture to transform $g$ into a strong one-way function $h$.

It remains to see how many public bits are used in this trasformation. For each iteration of $g$, we need $\ell = O(\log 1/w(n))$ shifts of length $n$. From last time, we saw that the iterative shift method we will use to construct $h$ from $g$ requires $N = 2n/w'(n)$ iterations, each requiring a $n'$ random bits, where $w'(n)$ is the measure of the hardness of the weak one-way function $g$ and $n'$ is the amount of randomness needed by $g$. Since $w'(n)$ is now a constant, we can see that $h$ will use $O(n^2 \log 1/w(n))$ public coins. Note that this is a considerable savings over the straightforward application of the iterative shift method to $f$ when $w(n) = O(1/\text{poly})$.

# 3 Expanders

Now we will introduce concepts which will help us achieve an even more efficient reduction. Informally, an expander graph is a graph so that for any subset $S$ of vertices, the neighbor set of $S$ is some constant amount greater than the size of $S$.

Definition 3.1 (Expander graph). *Let $G = (V, E)$ be a $d$-regular graph (i.e. each vertex has degree $d$). For $S \subseteq V$, let*

$$N(S) = \{y \in V | \exists x \in S, d(x, y) \leq 1\}$$

*where $d(x, y)$ is the distance between vertices. Then $G$ is an expander if*

$$\forall S \subseteq V \text{ and } |S| \leq \frac{1}{2}|V|, |N(S)| \geq (1+\theta)|S|$$

*for some constant $\theta$.*

Notice that since each vertex has the same degree, we can think of each edge coming from a vertex being labeled with an integer $j, 1 \leq j \leq d$.

There is another formulation of expander graphs in terms of bipartite graphs. It can be shown that the two formulations are equivalent.

Definition 3.2 (Expander graph, alternate). *Let $G = (L, R, E)$ be a $d$-regular bipartite graph with edges going from $L$ to $R$. Then $G$ is an expander if*

$$\forall S \subseteq L, |N(S)| \geq (1 + \theta(1 - \frac{|S|}{|L|}))|S|.$$

Let us consider an example of a graph that has been shown to be an expander under this definition. Let $L = R = (\mathbb{Z}/n)^2$. We can think of each vertex in $L$ and each vertex in $R$ as being a point on an $n \times n$ grid. A point on either side of the graph is identified by its grid coordinates. Consider the matrices

$$\begin{pmatrix} 1 & \pm 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \pm 1 & 1 \end{pmatrix}.$$

We define an edge from $(i, j) \in L$ to a point in $R$ by multiplying one of the matrices by $(i, j)$. The degree of each vertex in $L$ will thus be 5. It can be shown that such a graph is an expander. This construction was one of the first explicit expander graph constructions - rather than proving that there exists some expander graph or giving an algorithm to compute an expander graph, the graph is explicitly given.

We will use expander graphs to modify the weak to strong one-way construction that was given in this lecture. In today's construction, we were able to decrease the number of public random bits that are needed. We can use expanders to decrease this even more. We will replace the random bit strings $\pi$ with much shorter random bit strings $\rho \in [d]$. Recall that each $\pi$ was used to shift a $z \in \{0, 1\}^n$. We will use each $\rho$ to do something similar: pick a random neighbor of a vertex in an expander graph. Due to the properties of expander graphs, taking random neighbors is similar to taking a random walk in the graph, and will be sufficient to achieve a weak to strong one-way construction. The advantage of this approach is that there is significant savings in random bits used: each $\pi$ is a linear number of bits whereas each $\rho$ is only a constant number of bits. We will fill in the details of this construction next lecture.