CS 6222 Grad Crypto September 9th, 2025

Lecture 5: Stretching Pseudorandomness

Lecturer: Jack Doerner *Scribe: William Bradford*

1 Topics Covered

- Review of Definitions and Lemmas
- One-Bit Stretch Implies Polynomial Stretch

2 Review of Definitions and Lemmas

Definition 1 (Pseudorandom Generator)**.** *Let U_n be a the distribution over $\{0,1\}^n$. A function $G : \{0,1\}^n \to \{0,1\}^{\ell(n)}$ is a PRG if all three of the following hold:*

1. *G is deterministic and polynomial time (implicitly, $\ell(n)$ must be a polynomial)*
2. *$\ell(u) > n$*
3. *$\{G(U_n)\}_{n\in\mathbb{N}} \approx_c \{U_{\ell(n)}\}_{n\in\mathbb{N}}$*

Note 1 (Notational Shorthand)**.** *In the above definition $G(U_n)$ denotes the distribution produced by applying the function G to samples drawn from U_n. We will use this shorthand notation from now on.*

In a previous lecture, we gave a formal definition for the concept of computational indistinguishability. The complementary condition is computational *distinguishability.* It will we useful to write it out explicitly:

Definition 2 (Non-Negligible Function)**.** *We say that a function $\delta : \mathbb{N} \to \mathbb{R}^+ \cup \{0\}$ is* non-negligible *if $\exists c$ such that for infinitely many $n \in \mathbb{N}$, $\delta(n) \geq \frac{1}{n^c}$.*

Definition 3 (Computational Distinguishability)**.** *We say that some NUPPT algorithm D* distinguishes *the ensemble $\mathcal{X} = \{X_n\}_{n\in\mathbb{N}}$ from $\mathcal{Y} = \{Y_n\}_{n\in\mathbb{N}}$ if there exists some non-negligible function δ such that for all $n \in \mathbb{N}$,*

$$|\Pr[D(1^n, X_n) = 1] - \Pr[D(1^n, Y_n) = 1]| \geq \delta(n)$$

Note 2. *We sometimes say that an algorithm D* distinguishes *a specific pair of distributions X and Y. This means something slightly different than the above definition: it is not an asymptotic statement and only makes sense with respect to some* specific constant *δ such that*

$$|\Pr[D(X) = 1] - \Pr[D(Y) = 1]| \geq \delta$$

Lemma 1 (The Hybrid Lemma)**.** *Let* $\{X_i\}_{i\in[m]}$ *be a sequence of distributions. If there exists some PPT algorithm* D *and* $\delta \in \mathbb{R}$ *such that*

$$|\Pr\left[D\left(1^n, X_1\right)=1\right]-\Pr\left[D\left(1^n, X_m\right)=1\right]|\geq \delta$$

then $\exists i \in [m-1]$ *such that*

$$|\Pr\left[D\left(1^n, X_i\right)=1\right]-\Pr\left[D\left(1^n, X_{i+1}\right)=1\right]|\geq \frac{\delta}{m-1}$$

3 One-Bit Stretch Implies Polynomial Stretch

The two tools available to us in this proof are the hybrid lemma, and the PRG security of $G:\{0,1\}^n \to \{0,1\}^{n+1}$. Therefore, we want to find some $G':\{0,1\}^n \to \{0,1\}^{\ell(n)}$ such that we can define a sequence of hybrid distributions with the following properties:

- $H_n^0 = G'\left(U_n\right)$
- $H_n^m = U_{\ell(n)}$ (for some m to be defined)
- For all $i \in [m]$, some instance of $G\left(U_n\right)$ in H_n^{i-1} is replaced by U_{n+1} in H_n^i.

We'll call the distributions H_n^i and H_n^{i+1} *neighbors*. Intuitively, if we define ensembles of such distributions (i.e. $\mathcal{H}^i = \{H_n^i\}_{n\in\mathbb{N}}$) then neighbor ensembles $\mathcal{H}^i$ and $\mathcal{H}^{i+1}$ should be computationally indistinguishable by the security of G and the closure of computational indistinguishability under NUPPT post-processing.

Note 3 (Notation for Concatenation)**.** $a\|b$ *is the concatenation of* a *and* b*. So for example, if* $a=$ *"pseudo",* $b=$ *"random",* $a\|b=$ *"pseudorandom". Similarly, we can use this notation to indicate destructuring. If* $c = 1011$ *and* $a\|b := c$ *such that* $b \in \{0,1\}$*, then* $a = 101$ *and* $b = 1$*.*

Construction 1. *Given* $G:\{0,1\}^n \to \{0,1\}^{n+1}$ *and polynomial* ℓ*, define* $G':\{0,1\}^n \to \{0,1\}^{\ell(n)}$ *such that*

$$G': s \mapsto b_1\left\|b_2\right\|\ldots\left\|b_{\ell(n)}: x_0 := s, \forall i \in [\ell(u)], x_i\right\| b_i := G\left(x_{i-1}\right)$$

In other words, there are $\ell(n)$ steps, and at the i^{th} step we use G to stretch an n-bit value x_{i-1} into an $(n+1)$-bit value $x_i\|b_i$. The single bit b_i is contributed to the output of G', and the n-bit value x_i is fed back into G to repeat the process.

Theorem 1. *If* $G:\{0,1\}^n \to \{0,1\}^{n+1}$ *is a PRG and* ℓ *is a polynomial such that* $\ell(n) > n$*, then* $G':\{0,1\}^n \to \{0,1\}^{\ell(n)}$ *as specified in Construction 1 is also a PRG.*

Proof Overview: Notice first of all that the number of recursive calls to G depends upon n. It follows that the number of necessary hybrid distributions (the to-be-defined m above) depends upon n. The following two ensembles are therefore *not* separated by a constant number of hybrids:

$$\mathcal{H}^0 = \left\{H_n^0\right\}_{n\in\mathbb{N}} = \left\{G'\left(U_n\right)\right\}_{n\in\mathbb{N}}$$
$$\mathcal{H}^\infty = \left\{H_n^{\ell(n)}\right\}_{n\in\mathbb{N}} = \left\{U_{\ell(n)}\right\}_{n\in\mathbb{N}}$$

For any fixed n, the number of neighbor distributions over which we must apply the hybrid lemma is polynomial in n, but as $n \to \infty$, there are $\ell(n) \to \infty$ neighbor distributions between H_n^0 and $H_n^{\ell(n)}$. We must therefore take care in setting up our proof to ensure we only apply the hybrid lemma to polynomially-long sequences. The main steps are as follows:

1. $\forall i \in \mathbb{N}$ define the "hybrid experiment" $\mathcal{H}^i = \{H_n^i\}_{n\in\mathbb{N}}$ in a way that is consistent with the criteria described above.

2. Use hybrid lemma to prove that if there exists any NUPPT algorithm D that distinguishes $\mathcal{H}^0$ from $\mathcal{H}^\infty$ with non-negligible advantage, then there exists some non-negligible function δ such that for infinitely many $n \in \mathbb{N}$ there exists some $i_n \in [\ell(n)]$ and some PPT algorithm D_n such that D_n distinguishes H_n^{i-1} from H_n^i with advantage no less than $\delta(n)$.

3. Prove that if G is a PRG, then for all $i \in \mathbb{N}^+$, $\mathcal{H}^{i-1} \approx_c \mathcal{H}^i$. In particular, we will prove that a *lossless* reduction exists.

4. Combine Steps 2 and 3 to complete the proof by contraposition: if G' is not a PRG, then there exists a NUPPT algorithm that distinguishes $\{G(U_n)\}_{n\in\mathbb{N}}$ from $\{U_{\ell(n)}\}_{n\in\mathbb{N}}$ with non-negligible advantage, which implies that G is not a PRG.

Proof of Theorem 1. We begin by defining our hybrid distributions, using a family of helper functions $G^i : \{0,1\}^n \to \{0,1\}^{n+1}$ for $i \in \mathbb{N}$. For every $n \in \mathbb{N}$ we have:

$$\begin{aligned}
&G^0 : x \mapsto \varnothing \\
&G^i : x \mapsto b \| G^{i-1}(x) : x\|b := G(x) && \text{for } i \in \mathbb{N}^+ \\
&H_n^i = U_i \| G^{\ell(n)-i}(U_n) && \text{for } i \in [0, \ell(n)] \\
&H_n^i = H_n^{i-1} && \text{for } i \in \mathbb{N} \text{ s.t. } i > \ell(n)
\end{aligned}$$

Intuitively, each H_n^i is the concatenation of a truly random i-bit number and a PRG output of length $\ell(n)-i$, where the input of the PRG is drawn from U_n. In other words, each successive H_n^i peels away an additional layer of recursion from G', and replaces the output bit produced by that layer with a uniformly-random bit. Once the output is completely replaced by uniform bits (at step $i = \ell(n)$), further distributions H_n^i for $i > \ell(n)$ are identical (i.e. they all consist exclusively of uniform bits). These hybrids are illustrated in Figure 1.

Claim 1. *If there exists some $n \in \mathbb{N}$, some algorithm D_n, and some function δ such that*

$$\left|\Pr\left[D_n\left(H_n^0\right) = 1\right] - \Pr\left[D_n\left(H_n^{\ell(n)}\right) = 1\right]\right| \geq \delta(n)$$

then there exists some $i_n \in [\ell(n)]$ such that

$$\left|\Pr\left[D_n\left(H_n^{i_n-1}\right) = 1\right] - \Pr\left[D_n\left(H_n^{i_n}\right) = 1\right]\right| \geq \frac{\delta(n)}{\ell(n)}$$

Claim 2. *If there exists some NUPPT algorithm D and some function δ such that D distinguishes $\mathcal{H}^0$ from $\mathcal{H}^\infty$ with advantage at least $\delta(n)$ for all $n \in \mathbb{N}$, then $D_n = D(1^n, \cdot)$ satisfies Claim 1 with respect to δ. Furthermore, there is a single fixed polynomial such that the runtime of every D_n is bounded by that polynomial on n.*

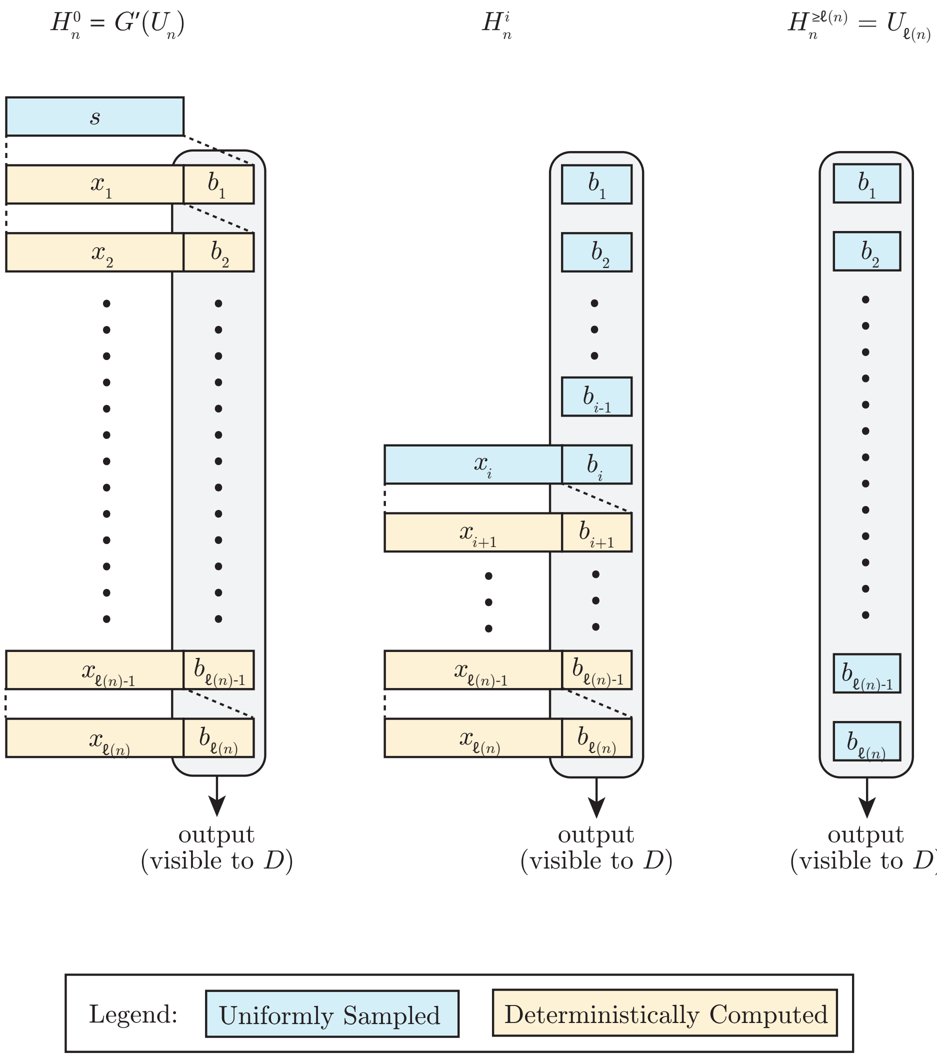

Figure 1: Illustration of the hybrid distributions used in the proof of Theorem 1

Note that the first two claims, above, follow directly from applying the hybrid lemma to the hybrid distributions and ensembles we have defined above. Next we consider a reduction R_n^i that uses any distinguisher for the neighbor distributions defined above to break the security of G. Claim 3 establishes that the reduction is lossless.

Construction 2 ($R_n^i : \{0,1\}^{n+1} \to \{0,1\}^{\ell(n)}$)**.** *On input x, R_n^i does the following:*

1. *Let $x' \| b := x$*
2. *Sample $y \leftarrow U_{i-1}$*
3. *Output $y \| b \| G^{\ell(n)-i}(x')$*

Claim 3. *For $i \in [\ell(n)]$,*

$$R_n^i\left(G\left(U_n\right)\right) = H_n^{i-1}$$
$$R_n^i(U_{n+1})) = H_n^i$$

Combining Claim 3 with the fact that $H_n^i = H_n^{\ell(n)}$ when $i \geq \ell(n)$, we can see that the PRG security of G implies that $\mathcal{H}^{i-1} \approx_c \mathcal{H}^i$ for $i \in \mathbb{N}^+$.[1] Combining Claims 1 and 3 yields:

Claim 4. *If there exists some $n \in \mathbb{N}$, some algorithm D_n, and some function δ such that*

$$\left|\Pr\left[D_n\left(H_n^0\right) = 1\right] - \Pr\left[D_n\left(H_n^{\ell(n)}\right) = 1\right]\right| \geq \delta(n)$$

then there exists some $i_n \in [\ell(n)]$ such that

$$\left|\Pr\left[D_n\left(R_n^{i_n}\left(G\left(U_n\right)\right)\right) = 1\right] - \Pr\left[D_n\left(R_n^{i_n}(U_{n+1})\right)\right) = 1\right]\right| \geq \frac{\delta(n)}{\ell(n)}$$

Now we can combine Claims 4 and 1 with the fact that G is polynomial time to find:

Claim 5. *If there exists some NUPPT algorithm D and some non-negligible function δ such that for all $n \in \mathbb{N}$,*

$$\left|\Pr\left[D\left(1^n, H_n^0\right) = 1\right] - \Pr\left[D\left(1^n, H_n^{\ell(n)}\right) = 1\right]\right| \geq \delta(n)$$

Then there exists some NUPPT algorithm D'[2] *such that for all $n \in \mathbb{N}$,*

$$\left|\Pr\left[D'\left(1^n, G(U_n)\right) = 1\right] - \Pr\left[D'\left(1^n, U_{n+1}\right) = 1\right]\right| \geq \frac{\delta(n)}{\ell(n)}$$

Finally, we observe that since ℓ is a polynomial, $\delta(n)/\ell(n)$ is negligible if and only if $\delta(n)$ is negligible. From this fact and the contraposition of Claim 5 it follows that

$$\{G(U_n)\}_{n\in\mathbb{N}} \approx_c \{U_{n+1}\}_{n\in\mathbb{N}} \Rightarrow \mathcal{H}^0 \approx_c \mathcal{H}^\infty \Rightarrow \left\{G'(U_n)\right\}_{n\in\mathbb{N}} \approx_c \left\{U_{\ell(n)}\right\}_{n\in\mathbb{N}}$$

and thus if G is a PRG, then G' is one as well. □

[1] This fact is not important for the rest of the proof, but we mention it in order to make it clear that Claim 3 corresponds to Step 3 of the proof overview.

[2] We can construct D' by taking the values of i_n in Claim 4 to be advice. That is, $D' = \{D'_n\}_{n\in\mathbb{N}}$ such that $D'_n = D\left(1^n, R_n^{i_n}(\cdot)\right)$