Statistically Hiding Commitment from One-Way Functions via Inaccessible Entropy

Class 2

Handout Mode

Benny Applebaum & Iftach Haitner, Tel Aviv University

Tel Aviv University.

June 14, 2017
Talk plan

- Inaccessible entropy generator from any one-way function
- Hashing protocols
Section 1

Inaccessible Entropy Generator from OWF
The generator

**Definition 1**

Given a function \( f : \{0, 1\}^n \mapsto \{0, 1\}^n \), let \( G \) be the \((n + 1)\)-block generator

\[
G(x) = f(x)_1, \ldots, f(x)_n, x
\]

**Lemma 2**

Assume that \( f \) is a OWF then \( G \) has accessible max-entropy (at most) \( n - \log n \).

- Recall \( f \) is OWF if
  \[
  \Pr_{x \leftarrow \{0,1\}^n} \left[ \text{Inv}(f(x)) \in f^{-1}(f(x)) \right] = \neg(n) \text{ for any PPT Inv.}
  \]
- The real entropy of \( G \) is \( H(f(U_n)_1, \ldots, f(U_n)_n, U_n) = H(U_n) = n \)
- Hence, inaccessible entropy is \( \log n \)
- Number of blocks can be reduced to \( n / \log n \)
- Proof idea
Proving Lemma 2

Assume \( \exists \) an efficient \( G \)-consistent online \((n+1)\)-block generator \( \tilde{G} \) such that

\[
\Pr_{t=(r_1,y_1,...) \leftarrow \tilde{T}} \left[ \text{AccH}_{\tilde{G}}(t) = \sum_{i=1}^{n+1} H_{\tilde{Y}_i \mid \tilde{R}_{<i}}(y_i \mid r_{<i}) \geq n - c \log n \right] \geq \varepsilon = \frac{1}{\text{poly}(n)}
\]

for some \( c > 0 \), and \( \tilde{T} = (\tilde{R}_1, \tilde{Y}_1, \ldots, \tilde{R}_m, \tilde{Y}_m) = T_{\tilde{G}}(1^n) \).

We show how to use \( \tilde{G} \) to invert \( f \).

- For simplicity assume \( c = 1 \)
- Assume for simplicity that the random strings used by \( \tilde{G}(1^n) \) are all of length \( n \).
- We use the notation \( \Pr_{X \mid Y} [x \mid y] = \Pr [X = x \mid Y = y] \)
The distribution $\tilde{T} = (\tilde{R}_1, \tilde{Y}_1, \ldots, \tilde{R}_m, \tilde{Y}_m) = T_G(1^n)$

- Fix $t = (r_1, y_1, \ldots, r_{n+1}, y_{n+1}) \in \text{Supp}(\tilde{T})$

- Let $P(t) := \prod_{i=1}^{n+1} \Pr_{\tilde{R}_i|\tilde{R}_{<i}, \tilde{Y}_i} [r_i|r_{<i}, y_i]$

\[
\Pr_{\tilde{T}}[t] = \Pr_{\tilde{Y}_1} [y_1] \cdot \Pr_{\tilde{R}_1|\tilde{Y}_1} [r_1|y_1] \cdot \Pr_{\tilde{Y}_2|\tilde{R}_1} [y_2|r_1] \cdot \Pr_{\tilde{R}_2|\tilde{R}_1, \tilde{Y}_2} [r_2|r_1, y_2] \cdots \\
= P(t) \cdot \Pr_{\tilde{Y}_1} [y_1] \cdot \Pr_{\tilde{Y}_2|\tilde{R}_1} [y_2|r_1] \cdot \Pr_{\tilde{Y}_3|\tilde{R}_{<3}} [y_3|r_{<3}] \cdots \\
= P(t) \cdot 2^{-\sum_{i=1}^{n+1} H_{\tilde{Y}_i|\tilde{R}_{<i}}(y_i|r_{<i})} \\
= P(t) \cdot 2^{-\text{Acc}H_G(t)}
The inverter

Algorithm 3 (\(f\)-inverter Inv)

Input: \(z \in \{0, 1\}^n\).

1. For \(i = 1\) to \(n\), do the following for \(n^2/\varepsilon\) times:
   1.1 Sample \(r_i \leftarrow \{0, 1\}^n\) and let \(y_i = \tilde{G}(r_1, \ldots, r_i)_i\).
   1.2 If \(y_i = z_i\), move to next value of \(i\).
   1.3 Abort, if the maximal number of attempts is reached.

2. Sample \(r_{n+1} \leftarrow \{0, 1\}^n\), and output \(y_{n+1} = \tilde{G}(r_1, \ldots, r_{n+1})_{n+1}\).

- We start by assuming that Inv is unbounded (replace \(n^2/\varepsilon\) with \(\infty\)).
- Let \(\hat{T} = (\hat{R}_1, \hat{Y}_1, \ldots, \hat{R}_{n+1}, \hat{Y}_{n+1})\) be the (final) values of \((r_1, y_1, \ldots, r_{n+1}, y_{n+1})\) in a random execution of Inv\((f(U_n))\).
- Since \(\tilde{G}\) is \(G\)-consistent, \(\hat{T} \in \text{Supp}(\tilde{T}) \implies \text{Inv}(\hat{Y}_{\leq n}) \in f^{-1}(\hat{Y}_{\leq n})\).
The distribution $\hat{T} = (\hat{R}_1, \hat{Y}_1, \ldots, \hat{R}_{n+1}, \hat{Y}_{n+1})$ induced by $\text{Inv}(f(U_n))$

- Fix $t = (r_1, y_1, \ldots, r_{n+1}, y_{n+1}) \in \text{Supp}(\hat{T})$
- Recall $P(t) = \prod_{i=1}^{n+1} \Pr_{\hat{R}_i|\hat{R}_{<i}, \hat{Y}_i} [r_i|r_{<i}, y_i]$

\[
\Pr[t] = \Pr_{\hat{Y}_1}[y_1] \cdot \Pr_{\hat{R}_1|\hat{Y}_1}[r_1|y_1] \cdot \Pr_{\hat{Y}_2|\hat{R}_1}[y_2|r_1] \cdot \Pr_{\hat{R}_2|\hat{R}_1,\hat{Y}_2}[r_2|r_1, y_2] \cdots
\]

\[
= \Pr_{f(U_n)}[y_{\leq n}] \cdot \Pr_{\hat{Y}_{n+1}|\hat{R}_{\leq n}}[y_{n+1}|r_{\leq n}] \cdot \Pr_{\hat{R}_1|\hat{Y}_1}[r_1|y_1] \cdot \Pr_{\hat{R}_2|\hat{R}_1,\hat{Y}_1}[r_2|r_1, y_1] \cdots
\]

- $\hat{R}_i|_{(\hat{R}_{<i}, \hat{Y}_i)=(r_{<i}, y_i)} \equiv \hat{R}_i|_{(\hat{R}_{<i}, \hat{Y}_i)=(r_{<i}, y_i)}$

$\Rightarrow \Pr_{\hat{T}}[t] = \Pr_{f(U_n)}[y_{\leq n}] \cdot \Pr_{\hat{Y}_{n+1}|\hat{R}_{\leq n}}[y_{n+1}|r_{\leq n}] \cdot P(t)$

Hence,

\[
\Pr[t] = \frac{\Pr_{f(U_n)}[y_{\leq n}] \cdot \Pr_{\hat{Y}_{n+1}|\hat{R}_{\leq n}}[y_{n+1}|r_{\leq n}]}{2^{-\text{AccH}_G(t)}} \cdot \Pr_{\hat{T}}[t]
\]  \hspace{1cm} (1)
Recall $\Pr_{\tilde{T}}[t] = \frac{\Pr_{f(U_n)}[y_{\leq n}]}{2^{-\text{AccH}_{\tilde{G}}(t)}} \cdot \Pr_{\tilde{Y}_{n+1}\mid \tilde{R}_{\leq n}}[y_{n+1}\mid r_{\leq n}] \cdot \Pr_{\tilde{T}}[t]$

Note $\Pr_{f(U_n)}[y_{\leq n}] = \left| f^{-1}(y_{\leq n}) \right| / 2^n$

Assuming $\text{AccH}_{\tilde{G}}(t) \geq n - \log n$:

$$\Pr_{\tilde{T}}[t] \geq \frac{2^n \cdot \Pr_{f(U_n)}[y_{\leq n}]}{n} \cdot \Pr_{\tilde{Y}_{n+1}\mid \tilde{R}_{\leq n}}[y_{n+1}\mid r_{\leq n}] \cdot \Pr_{\tilde{T}}[t]$$

$$\Pr_{\tilde{T}}[t] \geq \Pr_{\tilde{T}}[t] \cdot \frac{f^{-1}(y_{\leq n})}{n} \cdot \Pr_{\tilde{Y}_{n+1}\mid \tilde{R}_{\leq n}}[y_{n+1}\mid r_{\leq n}].$$

If $H_{\tilde{Y}_{n+1}\mid \tilde{R}_{\leq n}}(y_{n+1}\mid r_{\leq n}) \leq \log \left| f^{-1}(y_{\leq n}) \right| + k$ for some $k \in O(\log n)$:

$$\Pr_{\tilde{T}}[t] \geq \Pr_{\tilde{T}}[t] \cdot \frac{f^{-1}(y_{\leq n})}{n} \cdot \frac{2^{-k}}{|f^{-1}(y_{\leq n})|} = \frac{\Pr_{\tilde{T}}[t]}{2^k \cdot n} \geq \frac{\Pr_{\tilde{T}}[t]}{\text{poly}(n)}$$

We show that such transcripts happen whp by $\tilde{T}$
The good set $S$

Let $S \subseteq \text{Supp}(\tilde{T})$ denote set of the transcripts $t = (r_1, y_1, \ldots, r_{n+1}, y_{n+1})$ with

1. $\text{AccH}_{\tilde{G}}(t) \geq n - \log n$
2. $H_{\tilde{Y}_{n+1} | \tilde{R} \leq n}(y_{n+1} | r \leq n) \leq \log(4/\epsilon) + \log |f^{-1}(y_{\leq n})|$
3. $H_{\tilde{Y}_i | \tilde{R} < i}(y_i | r < i) \leq \log(4n/\epsilon)$ for all $i \in [n]$

- $\Pr_{\tilde{T}} \left[ H_{\tilde{Y}_{n+1} | \tilde{R} \leq n}(y_{n+1} | r \leq n) > \log(4/\epsilon) + \log |f^{-1}(y_{\leq n})| \right] \leq \epsilon/4$
  (see Lemma 1 from last class)

- $\Pr_{\tilde{T}} \left[ \exists i \in [n]: H_{\tilde{Y}_i | \tilde{R} < i}(y_i | r < i) > \log(4n/\epsilon) \right] \leq n \cdot \frac{\epsilon}{4n} = \epsilon/4$

- Hence,

$$\Pr_{\tilde{T}} [S] \geq \Pr_{\tilde{T}} [\text{AccH}_{\tilde{G}}(T) \geq n - \log n] - 2 \cdot \epsilon/4 \geq \epsilon/2 \quad (4)$$

- By Eq. (3): $\Pr_{\tilde{T}} [S] \geq \frac{\epsilon/4}{n} \cdot \Pr_{\tilde{T}} [S] \geq \epsilon^2 / 8n$

- Hence, unbounded Inv inverts $f$ with probability at least $\epsilon^2 / 8n$

- What about bounded Inv?
(bounded) $\text{Inv}$’s success probability

- Let $\hat{T}'$ denote the final value of $\tilde{G}$’s coins and output blocks, induced by the bounded version of $\text{Inv}$ (set to $\bot$ if $\text{Inv}$ aborts).
- The third property of $S$ yields that for every $t \in S$
  
  $$\Pr_{\hat{T}'}[t] \geq \Pr_{\hat{T}}[t] \cdot \left(1 - n \cdot \left(1 - \frac{\varepsilon}{4n}\right)^{n^2/\varepsilon}\right) \geq \Pr_{\hat{T}}[t] \cdot \left(1 - O(n \cdot 2^{-n})\right) \geq \Pr_{\hat{T}}[t] / 2$$

- We conclude that
  
  $$\Pr_{z \leftarrow f(U_n)}[\text{Inv}(z) \in f^{-1}(z)] = \Pr_{z \leftarrow f(U_n)}[\text{Inv}(z) \text{ does not abort}]$$

  $$\geq \Pr_{\hat{T}'}[S]$$

  $$\geq \frac{1}{2} \cdot \Pr_{\hat{T}}[S]$$

  $$\geq \frac{\varepsilon^2}{16n}.$$
Section 2

Hashing Protocols
An information theoretic hashing protocol

- Let $\mathcal{H}^1$ be $\ell$-wise independent family mapping $\ell$-bit strings to $k$-bit strings
- Let $\mathcal{H}^2$ be 2-universal family mapping $\ell$-length strings to $n$-bit strings

### Protocol 4 ($((S_{IH}(x \in \{0,1\}^\ell), R_{IH})^{\mathcal{H}_1, \mathcal{H}_2})$)

1. $R_{IH}$ sends $h^1 \leftarrow \mathcal{H}^1$ to $S_{IH}$
2. $S_{IH}$ sends $y^1 = h^1(x)$ to $R_{IH}$
3. $R_{IH}$ sends $h^2 \leftarrow \mathcal{H}^2$ to $S_{IH}$
4. $S_{IH}$ sends $y^2 = h^2(x)$ to $R_{IH}$

### Claim 5 (Binding of $(S_{IH}, R_{IH})^{\mathcal{H}_1, \mathcal{H}_2}$)

Let $\tilde{S}$ be an algorithm and let $Y^1, Y^2, H^1, H^2$ be value of $y^1, y^2, h^1, h^2$ in a random execution of $(\tilde{S}, R_{IH})$. Then for any set $2^k$-size $T \subseteq \{0,1\}^\ell$:

$$\Pr\left[\exists x \neq x' \in T : H^1(x) = H^1(x') = Y^1 \land H^2(x) = H^2(x') = Y^2\right] \in 2^{-\Omega(n)}.$$

Proof: ? (warmup $k < \ell/2 - n$) HW1 Can we do it in a single round? HW2 Hiding?
**Target collision-resistant functions**

**Definition 6 (target collision-resistant functions (TCR))**

An efficient function family \( \mathcal{F} = \{ \mathcal{F}_n : \{0, 1\}^{\ell(n)} \mapsto \{0, 1\}^n \} \) is **target collision resistant**, if

\[
\Pr_{(x,s) \leftarrow A_1(1^n) ; f \leftarrow \mathcal{F}_n ; x' \leftarrow A_2(s,f)} [x \neq x' \land f(x) = f(x')] = \text{neg}(n)
\]

for any pair of PPT \( A_1, A_2 \).

Relaxed variant of collision resistant.

**Theorem 7**

**OWFs imply TCRs for any** \( \ell \in \text{poly} \).
A computational hashing protocol

Let $\mathcal{H}^1 = \{\mathcal{H}^1_n\}$, $\mathcal{H}^2 = \{\mathcal{H}^2_n\}$ and $\mathcal{F} = \{\mathcal{F}_n\}$ be function families, mapping strings of length $\ell(n)$ to strings of length $k(n)$, $n$ and $n$, respectively. Such that $\mathcal{H}^1_n$ is $\ell(n)$-wise ind, $\mathcal{H}^2_n$ is pair-wise ind, and $\mathcal{F}$ be a TCR.

Protocol 8 ($(S_H(x \in \{0, 1\}^{\ell(n)}, R_H)(1^n)$)

1. The two parties interact in $(S_{IH}(x), R_{IH})^{\mathcal{H}^1_n, \mathcal{H}^2_n}$, with $S_H$ and $R_H$ taking the role of $S_{IH}$ and $R_{IH}$ respectively.

2. $R_H$ sends $f \leftarrow \mathcal{F}_n$ to $S_H$.

3. $S_H$ sends $y = f(x)$ back to $R_H$.

Claim 9 (Binding of $(S_H, R_H)$)

Let $\tilde{S}$ be a PPT, let $Y^1, Y^2, Y^3, H^1, H^2, F$ be value of $y^1, y^2, y, h^1, h^2, f$ in a random execution of $(\tilde{S}, R_H)(1^n)$. Let $X$ and $X'$ be the two strings $\tilde{S}$ outputs at the end of the protocol. Then for any set $2^{k(n)}$-size $T \subseteq \{0, 1\}^{\ell(n)}$:

$$\Pr[X \neq X' \land \{X, X'\} \cap T \neq \emptyset \land H^1(X) = H^1(X') = Y^1 \land H^2(X) = H^2(X') = Y^2 \land F(X) = F(X') = Y_3] \leq \text{neg}(n).$$
Proving the computational binding of \((S_H, R_H)\)

Assume \(\exists \text{ PPT } \tilde{S}\) that breaks the binding of \((S_H, R_H)\) with probability \(\varepsilon(n) > 1 / \text{poly}(n)\). We use it to find collision in \(\mathcal{F}\).

**Algorithm 10 (collision finder \(A_1(1^n)\))**

1. Emulate \((\tilde{S}, R_H)(1^n)\) till the *end of the embedded execution of \((S_{IH}, R_{IH})\).* Let \(\text{state}\) be the *state* (parties’ views) of emulated protocol.
2. Continue the execution \((\tilde{S}, R_H)\) till it ends.
   Let \(\{x_0, x_1\}\) be the two values \(\tilde{S}\) outputs at the end of the emulation.
3. Output \((x \leftarrow \{x_0, x_1\}, \text{state})\).

**Algorithm 11 (collision finder \(A_2(x \in \{0, 1\}^\ell, \text{state} \in \{0, 1\}^*, f \in \mathcal{F}_n)\))**

1. Emulate a random execution of \((\tilde{S}, R_H)(1^n)\) conditioned on \(\text{state}\) and \(f\).
   Let \(x_0\) and \(x_1\) be the output of \(\tilde{S}\) in the end of the emulation.
2. Output \(x' \in \{x_0, x_1\} \setminus \{x\}\).
Let $Z$ be the value of the element of $\mathcal{T}$ that is consistent with the embedded execution of $(S_{IH}, R_{IH})$ in $(\tilde{S}, R_H)(1^n)$. Setting it to $\bot$ if number of consistent elements is not one.

Claim 5 yields that the probability $\tilde{S}$ breaks the binding and $Z \neq \bot$, is at least $\varepsilon - 2^{-\Omega(n)} > \varepsilon/2$.

Let $q(s)$ be the probability that $\tilde{S}$ breaks the binding in $(\tilde{S}, R_H)(1^n)$ and $Z \neq \bot$, conditioned that its state after the execution of $(S_{IH}, R_{IH})$ is $s$.

A($s, \cdot, F$) finds a collision with probability $\geq q(s)^2/2$:

- $\Pr [x \in \mathcal{T} \text{ and } x \text{ is the only consistent element in } \mathcal{T}] \geq q(s)/2$
- $\Pr [x_0 \neq x_1 \land \{x_0, x_1\} \cap \mathcal{T} \neq \emptyset \text{ and both are consistent}] \geq q(s)$

Let $\text{State}$ be the value of state in a random execution of $A_1(1^n)$.

$E[q(\text{State})] \geq \varepsilon/2$

$A(1^n)$ finds a collision with probability at least

$$E\left[\frac{q(\text{State})^2}{2}\right] \geq (\text{by Jensen}) E[q(\text{State})]^2/2 \geq 1/8p(n)^2.$$