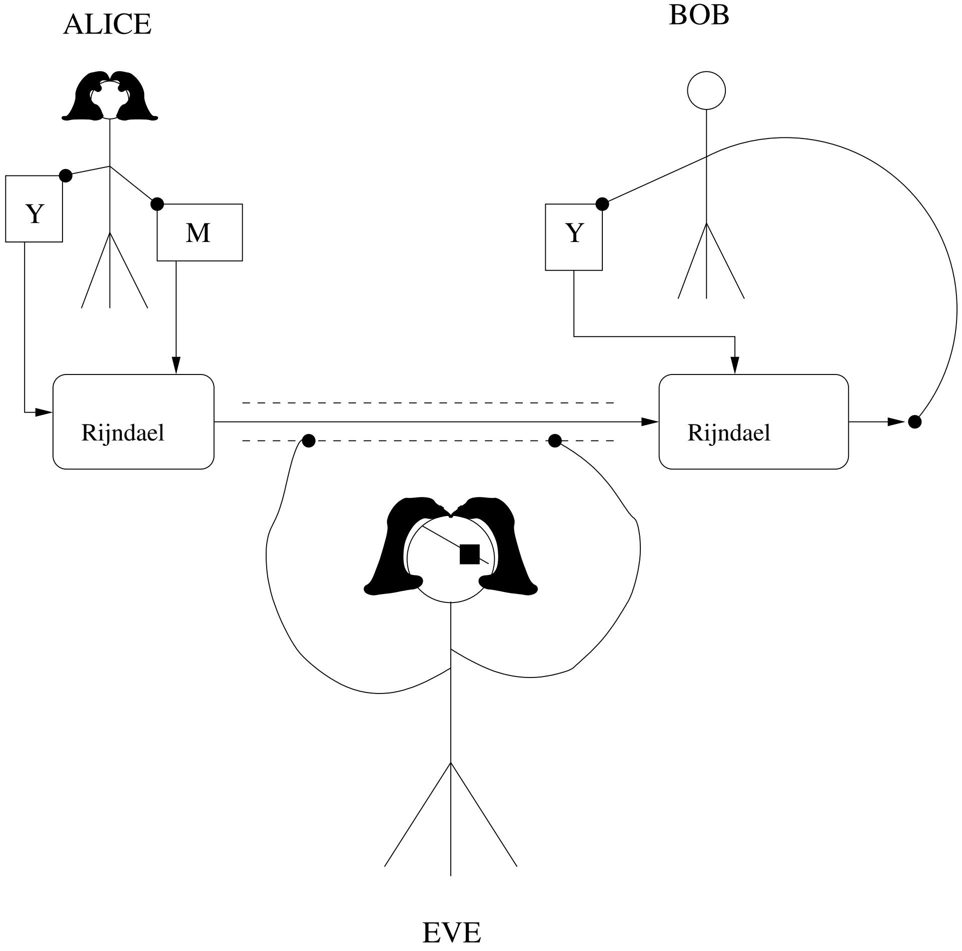
Cryptographic Applications of Extractors

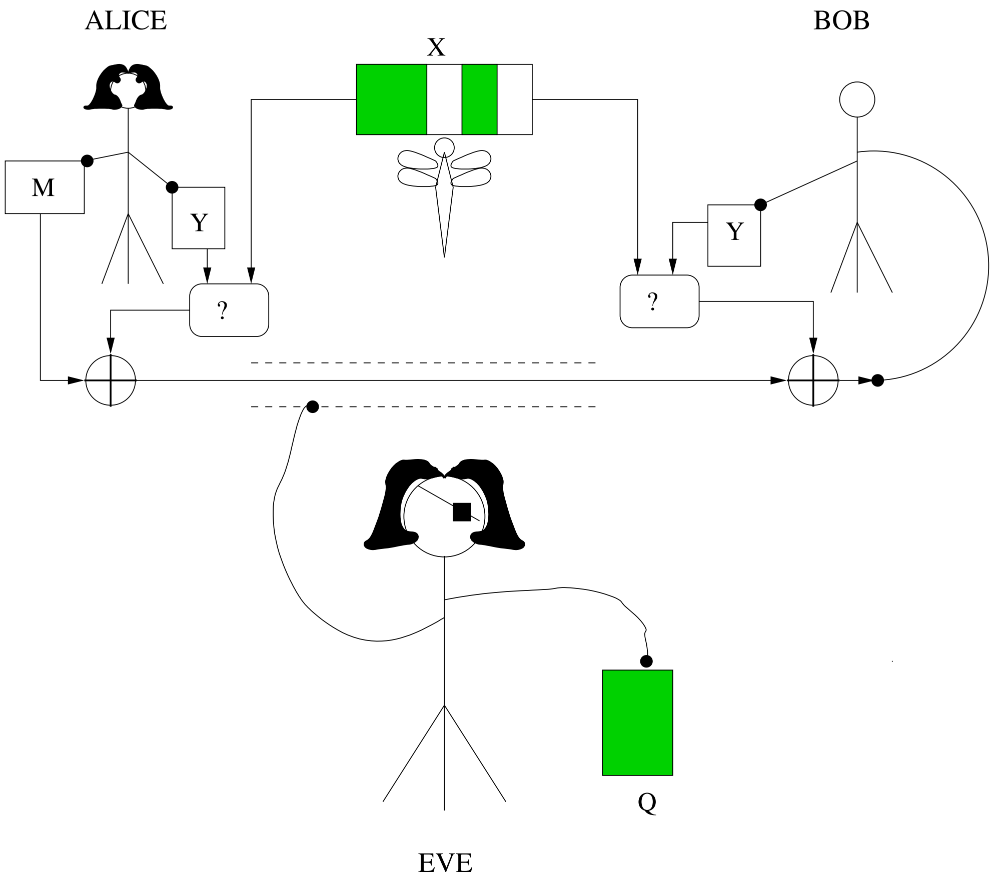
Atri Rudra

Private-Key Cryptography



- Based on Hardness Assumptions.

Bounded Storage Model



- Information Theoretic security.
- Ever-lasting Security.

Lets get Semi-Formal

- Security Parameter k
  - ◊ Eve has *advantage* $\leq 2^{-k}$.
- Eve's memory $\leq \nu|X| = \nu n$, $\nu < 1$.
- Need Everlasting Security.
- Minimize $|Y|$.

What’s in store for us ?

- Chi-Jen Lu, “Encryption against Storage-Bounded Adversaries from On-Line Strong Extractors". In *CRYPTO 2002*.

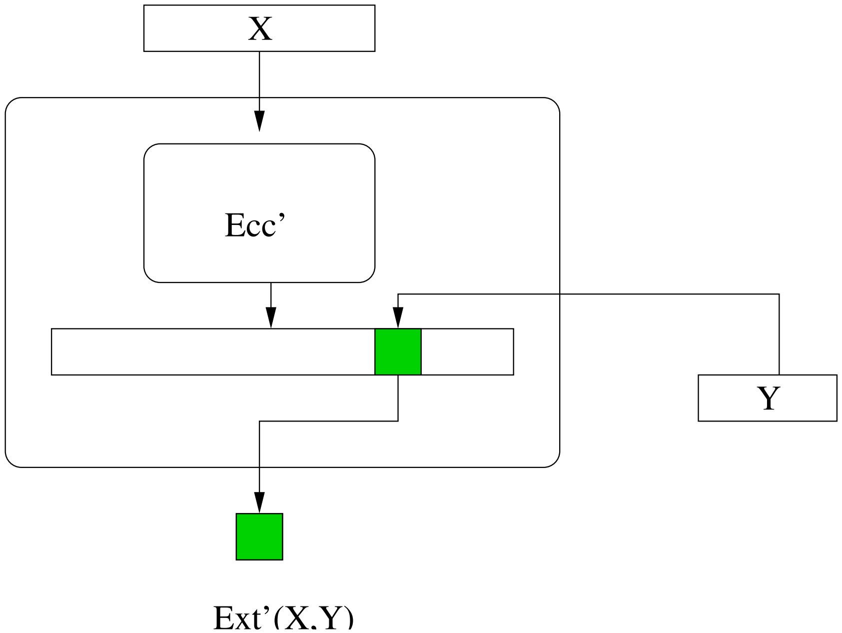
- Salil P. Vadhan, “On Constructing Locally Computable Extractors and Cryptosystems in the Bounded Storage Model".

Lu's Basic Idea

- $X|Q$ is still "random".
  - ◊ Eve stores νn bits, Q ($|X| = n$).
- Extractors!
- Strong Extractors
  - ◊ $\langle Ext(X, Y), Y \rangle \sim_\epsilon U^m \times U^s$.
- Ever-lasting Security.
- Key re-use.

Lu's Basic Idea (Contd.)

- Online strong extractors.
- List-decodable Codes.
- Uses Trevisan's technique.



- ◊ $Ecc' : (n) \to (\overline{n})$.
- ◊ $Ext' : (n) \times [\overline{n}] \to (1)$.

Definitions

- $\{-1,1\}$ for binary values.
- U^n is uniform over $\{-1,1\}^n = (n)$.
- $(\frac{1}{2}-\epsilon, L)$- list code, $C : (n) \to (\overline{n})$
  - $\forall z \in (\overline{n}), |\{x | d_H(x,z) \le (\frac{1}{2}-\epsilon)\overline{n}\}| \le L$
- $\|A - B\| = \frac{1}{2}\sum_x |Pr[A = x] - Pr[B = x]|$.
- $H_\infty(X) = \min_x \log \frac{1}{Pr[X=x]}$.
- (r,ϵ)-strong extractor, $Ext : (n) \times (s) \to (m)$
  - $\forall X$ s.t. $H_\infty(X) \ge r, Y = U^s$:
    $\langle Ext(X,Y), Y\rangle \sim_\epsilon U^m \times U^s$.

Let's Prove Stuff

Lemma 1: A ind. of B, B', function g
$\|\langle g(A,B), B\rangle - \langle g(A,B'), B'\rangle\| = \|B - B'\|$.

Proof : $\|\langle g(A,B), B\rangle - \langle g(A,B'), B'\rangle\|$

$= \frac{1}{2}\sum_b \sum_c |Pr[g(A,B) = c \wedge B = b] - Pr[g(A,B') = c \wedge B' = b]|$

$= \frac{1}{2}\sum_b \sum_c |Pr[g(A,b) = c]Pr[B = b] - Pr[g(A,b) = c]Pr[B' = b]|$

$= \frac{1}{2}\sum_b \sum_c Pr[g(A,b) = c]|Pr[B = b] - Pr[B' = b]|$

$= \frac{1}{2}\sum_b |Pr[B = b] - Pr[B' = b]|$

$= \|B - B'\|$.

More proofs on the slides

Lemma 2: $\langle A, B\rangle$ is ϵ-random
$Pr_{b\in B}[\langle A|B=b\rangle$ not $\sqrt{\epsilon}-$random$] \leq 2\sqrt{\epsilon}$.

Proof : $E[\|\langle A|B=b\rangle - U\|]$

$= \sum_b Pr[B=b]\frac{1}{2}\sum_a |Pr[A=a|B=b] - \frac{1}{|A|}|$

$= \frac{1}{2}\sum_{a,b}|Pr[(A,B)=(a,b)) - Pr[B=b]\frac{1}{|A|}|$

$\leq \frac{1}{2}\sum_{a,b}|Pr[(A,B)=(a,b)) - \frac{1}{|A||B|}| + \frac{1}{2}\sum_{a,b}|\frac{1}{|A||B|} - Pr[B=b]\frac{1}{|A|}|$

$= \|\langle A, B\rangle - U\| + \|B - U\|$

$\leq 2\epsilon$.

More Definitions

- A and B agree on Y.
- A and B : Compute Z_Y.
- B $\rightarrow$ A : $M \oplus Z_Y$.
- Everlasting security of degree k
  - ◊ $Pr_{y \in Y, q \in Q}[\langle Z_Y | Y = y, Q = q\rangle$ not 2^{-k}-random $] \leq 2^{-\Omega(k)}$.

Everlasting Sec. frm Str. Extractors

- $Q = f(X)$ $(|Q| = \nu|X| = \nu n)$.
- $\epsilon = 2^{-2k}$ $(k \ll n)$.
- $\mu \leq 1 - \nu$ s.t. $2^{-(1-\mu-\nu)} \leq \frac{\epsilon}{2}$.
- $r = \mu n$.
- $(r, \frac{\epsilon}{2})$-strong $Ext : (n) \times (s) \rightarrow (m)$.
- $Z = Ext(X, Y)$, Y chosen from U^s.

Lemma 3: $\|\langle f(X), Y, Ext(X, Y)\rangle - \langle f(X), U^s, U^m\rangle\| \leq \epsilon$.

Reusing the same Private Key

- $X_1, \ldots, X_K$ s.t. $H_\infty(X_i | X_{i-1}, \ldots, X_1) \geq n$.
- $\forall i \in [K]$, $Z_i = Ext(X_i, Y), C_i = M_i \oplus Z_i$.
- Eve has $Z_{[i-1]} = (Z_1, \ldots, Z_{i-1})$ and $Q_i = f_i(X_i, Q_{i-1}, Z_{[i-1]})$.
- Goal: Eve gets C_K and is given Y, she should not "guess" M_K.

Reusing the Key (Contd.)

- $\epsilon = 2^{-3k}$, $r = \mu n$.
- Use strong $(r, \frac{\epsilon}{2})$ extractor.

Lemma 4: $\forall i \in \mathbb{N}$,
$\| \langle Z_{[i-1]}, Q_i, Y, Z_i \rangle - \langle Z_{[i-1]}, Q_i, U^s, U^m \rangle \| \leq i\epsilon$.

Theorem 1: Any $(\mu n, 2^{-3k})$ extractor gives encryption scheme with security degree k. Key can be re-used 2^k times.

On-line List-Decodable Code

Lemma: G is a (d, λ)-expander graph.

$$B \subseteq V(G), |B| \leq \alpha V(G).$$

$Pr[t\text{-step RW on } G \text{ misses } B] \leq \beta_a = (1 - \alpha + (\frac{\lambda}{d})^2)^{\frac{t}{2}}$

- $Ecc_t : (n) \to (\overline{n}), \overline{n} = nd^t 2^t.$

Input: $x \in (n)$.
Output Index: $(v, w, b) \in [n] \times [d]^t \times \{0, 1\}^t$.
Algorithm:
Do a t-step walk on G to get $W_G(v, w) = (v_1, \ldots, v_t)$.
Output the bit indexed by (v, w, b) as

$$Ecc_t(x)(v, w, b) = \prod_{i \in [t]} x(v_i)^{b(i)}.$$

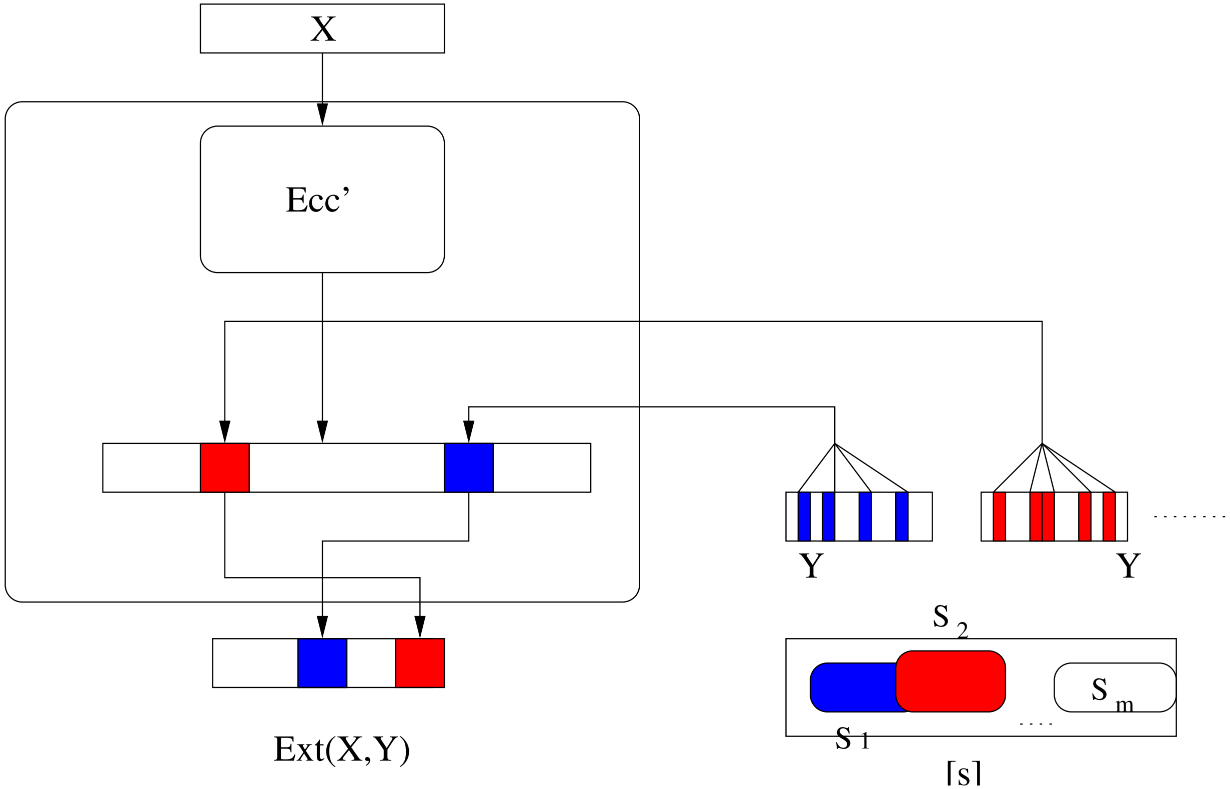
Stuff about Ecc_t

Lemma 7: $x, y \in (n)$ and $d_H(x,y) \geq \alpha n \Rightarrow d_H(Ecc_t(x), Ecc_t(y)) \geq (\frac{1}{2} - \beta_a)\overline{n}$.

Lemma 8: $Ecc \subseteq (\overline{n})$ s.t. $\forall c \in Ecc$, $|\{x | x \in Ecc, d_H(c,x) \leq (\frac{1}{2} - \beta)\overline{n}\}| \leq M \Rightarrow$
$\forall z \in (\overline{n})$, $\delta > \sqrt{\frac{\beta}{2}}$;
$|\{y | y \in Ecc, d_H(y,z) \leq (\frac{1}{2} - \delta)\overline{n}\}| \leq \frac{M}{4\delta^2 - 2\beta}$.

Corollary 1: Ecc_t is a $(\frac{1}{2} - \delta, \frac{2^{h(\alpha)n}}{2^{\delta^2}})$-list code, $\alpha = 1 + (\frac{\lambda}{d})^2 - \delta^{\frac{4}{t}}$.

Stuff from Trevisan's Extractor



- $Ecc' : (n) \rightarrow (\overline{n})$.
- $Ext : (n) \times (s) \rightarrow (m)$.

Stuff from Trevisan's Extractor

- $S_1, \ldots, S_m \subseteq [s]$ is a weak (l, ρ) design
  - $\forall i$, $|S_i| = l$
  - $\forall i$, $\sum_{j<i} 2^{|S_i \cap S_j|} \leq \rho(m-1)$.

Lemma : $S_1, \ldots S_m$ is a weak $(\log \overline{n}, \rho)$-design
$Ecc' : (n) \rightarrow (\overline{n})$, a $(\frac{1}{2} - \frac{\epsilon}{2m}, L)$-list code $\Rightarrow$
$Ext : (n) \times (s) \rightarrow (m)$,
$Ext(x, y) = (Ecc'(x)(y(S_1)), \ldots, Ecc'(x)(y(S_m)))$
is a strong (r, ϵ)-extractor $\forall r \geq \rho(m-1) + \frac{\log 2L}{\epsilon}$.

Combining Stuff

- Use $(\frac{1}{2} - \delta, L)$-list code Ecc_t
  - $\delta = \frac{1}{2m2^{-3k}}$,
  - $t = O(k)$, $\frac{\lambda}{d} = O(1)$, s.t. $h(\alpha) < \mu$ for $\alpha = 1 + (\frac{\lambda}{d})^2 - \delta^{\frac{4}{t}}$
  - $L = \frac{2^{h(\alpha)n}}{2\delta^2} = 2^{h(\alpha)+6k+1}m^2$.

Theorem 2: $\forall$ const. $\gamma \in (0,1)$, $\exists$ strong $(\mu n, 2^{-3k})$ extractor, $s = O(\frac{(k+\log n)^2}{\log n})$ and $m = n^\gamma$.

Seven Year Journey of Extractors

- Introduced in [NZ96].
  - PRGs for space-bounded computations.
- "Natural" fit for the Bounded Storage Model.
- "fit" shown in [Lu02].
  - Locally Computable Extractors.
  - Plugged Locally computable List Decodable Code in Trevisan's Extractor.

This Paper

- Lemma of [NZ96] $\Rightarrow$ Locally Computable Extractors
  - ◊ High $H_\infty(X) \Rightarrow$ High H_∞(Random Sample(X)).
  - ◊ Refines the Lemma.
- Strong Extractor + (randomness-efficient) Sampler $\Rightarrow$ Locally Computable Strong Extractors.
- Plug in Known Constructions.

Reverse Block Source

- $X_1, X_2, \ldots$ [over (n)].
- Entropy Rate α.
- $\forall t \in \mathbb{N}, \forall x_{t+1}, x_{t+2}, \ldots$
  $(X_t | X_{t+1} = x_{t+1}, X_{t+2} = x_{t+2}, \ldots)$ is an αn source.

Pseudorandom Extraction Scheme

- $Ext : (n) \times (s) \rightarrow (m)$.
- A, B : $Y \in (s)$.
- $Z_t = Ext(X_t, Y)$.
- At $\geq (t+1)$; A, B : use Z_t.
  - Used as a one-time pad.
  - $(\geq t+1)$, X_t is not “available” to Eve.
- Storage rate: β.

Defintion of Security

- $X_1, X_2, \ldots$
- $Y \in_R (s)$.
- $S_t \in (\beta n)$. $S_0 = 0^{\beta n}$.
- Real Experiment (R.E.) $\forall t$, $Z_t = Ext(X_t, Y)$
  - ◊ $\forall t \in [T]$, $S_t = A(Z_{[t-1]}, S_{t-1}, X_t)$.
  - ◊ Output $A(Z_{[T]}, S_T, X_{[T+1,\infty)}, Y) \in (1)$.
- Ideal Experiment (I.E.) $\forall t$, $Z_t \in_R (m)$
  - ◊ $\forall t \in [T]$, $S_t = A(Z_{[t-1]}, S_{t-1}, X_t)$.
  - ◊ Output $A(Z_{[T]}, S_T, X_{[T+1,\infty)}, Y) \in (1)$.

Sec. Pseudorandom Extr. Scheme

- $Ext : (n) \times (s) \to (m)$, ϵ-secure, storage rate β, entropy rate α
  - $\forall$ reverse blk. src. (X_t), entropy rate α.
  - $\forall$ Adversary A, storage bound βn.
  - $\forall T > 0$.
  - A distinguishes R.E. and I.E. with advantage $\leq T\epsilon$.
- ϵ- Secure Pseudorandomness Extraction Scheme (SPES).

Table

Reference	Key Length	# bits read	Restrictions
[ADR02]	$O(m \cdot \log n \cdot \log(\frac{1}{\epsilon})$	$O(m \cdot \log(\frac{1}{\epsilon})$	
[Lu02]	$O(\log^2(\frac{n}{\epsilon})/\log n)$	$O(m \cdot \log(\frac{1}{\epsilon}))$	$m \leq n^{1-\Omega(1)}$
here	$O(\log n + \log(\frac{1}{\epsilon}))$	$O(m + \log(\frac{1}{\epsilon}))$	$\epsilon > e^{-n/2^{O(\log^* n)}}$

Strong Extractors

- $Ext : (n) \times (s) \to (m)$, strong (r, ϵ) extractor
  - ◊ $\forall r$-source X, $\langle U^s, Ext(X, U^s) \rangle \sim_\epsilon U^s \times U^m$.

Theorem 5 : (Implicit in [Lu02])

- ◊ $Ext : (n) \times (s) \to (m)$, strong $(\delta n - \log(\frac{1}{\epsilon}), \epsilon)$-extractor.
- ◊ $\forall \beta > 0$, Ext is 2ϵ- SPES
  - ▷ Storage rate β.
  - ▷ Entropy rate $\beta + \delta$.

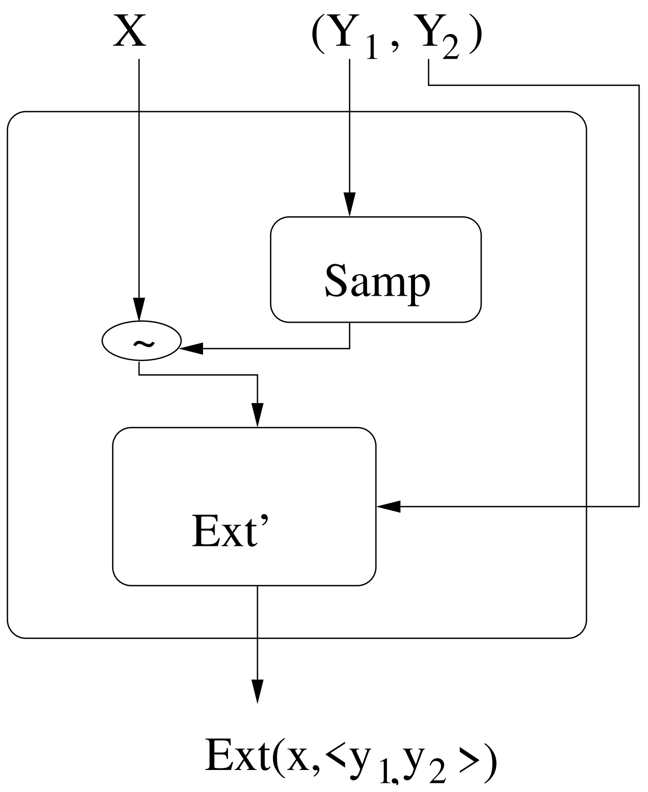
Locally Computable Extractors

- $Ext : (n) \times (s) \rightarrow (m)$, t-local extractor
  - $\forall r \in (s)$, $Ext(x, r)$ depends on t-bits of x.
- We need strong locally computable extractors!
- **Proposition 7 :** Ext being a t-local strong $(\delta n, \epsilon)$ extractor $\Rightarrow t \geq (1 - \epsilon)(\frac{1}{\delta})m - 1$.

Averaging Sampler

- $Samp : (r) \to [n]^t$ is (μ, θ, γ)-averaging sampler
  - ◊ $\forall f : [n] \to [0, 1]$, s.t. $\frac{1}{n} \sum_i f(i) \geq \mu$.
  - ◊ $Pr_{(i_1,\ldots,i_t) \in_R Samp(U^r)} [\frac{1}{t} \sum_{j=1}^{t} f(i_j) < \mu - \theta] \leq \gamma$
- Distinct Samples
  - ◊ $\forall x \in (r)$, samples of $Samp(x)$ are distinct.
  - ◊ Averaging Sampler with Distinct Samples (ASDS).
- "Optimal" parameter values
  - ◊ $t = O(\log(\frac{1}{\gamma}))$.
  - ◊ $r = O(\log n + \log(\frac{1}{\gamma}))$.

The Crux of the paper



- $Samp : (r) \to [n]^t$ ASDS.
- $Ext : (t) \times (s) \to (m)$ strong extractor.
- $Ext' : (n) \times (r+s) \to (m)$ t-local strong extractor.

Refinement of [NZ96] Lemma

Lemma 9 : $Samp : (r) \to [n]^t$, (μ, θ, γ) ASDS

- $\mu = \frac{\delta - 2\tau}{\log(\frac{1}{\tau})}$.
- $\theta = \frac{\tau}{\log(\frac{1}{\tau})}$.
- $\forall$ δn-source X on (n)
  $\langle U^r, X_{Samp(U^r)} \rangle \sim_{\gamma + 2^{-\Omega(\tau n)}} \langle A, B \rangle$.
  - ◊ $A = U^r$.
  - ◊ $\forall a \in (r)$, $B | A = a$ is a $(\delta - 3\tau)t$-source.

Main Theorem

Theorem 10 :

- $Samp : (r) \to [n]^t$, (μ, θ, γ) ASDS $\mu = \frac{\delta - 2\tau}{\log(\frac{1}{\tau})}$, $\theta = \frac{\tau}{\log(\frac{1}{\tau})}$.
- $Ext : (t) \times (s) \to (m)$ strong $((\delta - 3\tau)t, \epsilon)$ extractor.
- $Ext' : (n) \times (r + s) \to (m)$ s.t. $Ext'(x, (y_1, y_2)) = Ext(x_{Samp(y_1)}, y_2)$.

$\Rightarrow$ Ext' is a t-local strong $(\delta n, \epsilon + \gamma + 2^{-\Omega(\tau n)})$ extractor.

Explicit Constructions

- Local Extractor work on min-entropy $(\alpha - \beta)n - \log(\frac{1}{\epsilon})$.
- Small ϵ.
- # bits read is important.
- Not necessary to extract "all" entropy.

The "Chosen" Extractor

Lemma 18 : [Zuc97]

- $\forall$ const. $\delta > 0$, $\forall n \in \mathbb{N}$, $\epsilon > e^{-n/2^{O(\log^* n)}}$.
- explicit strong $(\delta n, \epsilon)$ $Ext : (n) \times (r) \to (m)$
  - ◊ $r = O(\log n + \log(\frac{1}{\epsilon})$.
  - ◊ $m = \frac{\delta n}{2}$.

The Sampler

- Expander sampler, $Samp : (r) \to [n]^t$.
  - Expander garph, G. $|V(G)| = n$.
  - Take a t-step RW on G.
- Non-distinct Samples
  - Discard repeated steps.
  - Works well.
  - Use amplification.

Lemmas on Sampler

Lemma 19 : $\forall\, 0 < \theta < \mu < 1, \gamma > 0, n \in \mathbb{N}$ explicit (μ, θ, γ) ASDS, $Samp : (r) \rightarrow [n]^t$.

- $t \in [O(\frac{1}{\theta^2} \cdot \log(\frac{1}{\gamma})), n]$.
- $r = \log n + O(t \cdot \log(\frac{1}{\theta}))$.

Lemma 20 : Explicit (μ, θ, γ) ASDS, $Samp : (r) \rightarrow [n]^t \Rightarrow$ $\forall m \in \mathbb{N}, \exists (\mu, \theta, \gamma)$ ASDS $Samp' : (r) \rightarrow [m \cdot n]^{m \cdot t}$.

Finally.....

Lemma 21 :$\forall n \in \mathbb{N}, 1 > \mu > \theta > 0, \gamma > 0, \exists$ explicit (μ, θ, γ) ASDS, $Samp : (r) \to [n]^t$

- $t \in [O(\frac{1}{\theta^2} \cdot \log(\frac{1}{\gamma})), n]$.
- $r = \log(\frac{n}{t}) + \log(\frac{1}{\gamma}) \cdot poly(\frac{1}{\theta})$.

Theorem 22 : $\forall$ const. $\delta > 0, n \in \mathbb{N}, \epsilon > e^{-n/2^{O(\log^* n)}}$, $m \leq \frac{\delta n}{4}$, explicit t-local strong $(\delta n, \epsilon)$ $Ext : (n) \times (s) \to (m)$.

- $s = O(\log n + \log(\frac{1}{\epsilon}))$.
- $t = O(m + \log(\frac{1}{\epsilon}))$.

For the Sake of Comparison

Theorem 17 : Non-constructive Local Strong Extractors

- $s = \log n + 3\log(\frac{1}{\epsilon}) + \log\log(\frac{1}{\delta}) + O(1)$.
- $t = O(\frac{m + \log(\frac{1}{\delta})\log(\frac{1}{\epsilon})}{\delta})$