How to extract useful randomness from unreliable sources

Divesh Aggarwal   Maciej Obremski   **João Ribeiro**   Luisa Siniscalchi   Ivan Visconti

CQT & National University of Singapore

Imperial College London

University of Salerno
→ Aarhus University

University of Salerno

Eurocrypt 2020

Randomness and cryptography

Perfect randomness

Cryptography

In practice, randomness sources X are **not** perfect!

Weaker assumption: min-entropy lower bound

$$\boxed{\max_x \Pr[X = x] \leq 2^{-k}}$$

$$\mathbf{H}_\infty(X) \geq k$$

k bits of min-entropy

Randomness extraction

$$\boxed{\max_x \Pr[X = x] \leq 2^{-k}}$$

$$\mathbf{H}_\infty(X) \geq k$$

k bits of min-entropy

IMPOSSIBLE!

$$X \longrightarrow \boxed{\text{Ext}} \longrightarrow Y \approx_\epsilon \text{Unif}$$

(arbitrary weak k-source)

(statistically close to uniform)

Multi-source extraction: combine several **independent** weak sources (e.g., sampled from different devices/locations)

Multi-source randomness extraction

$$X_1 \qquad X_2 \qquad X_3 \qquad \cdots \qquad X_{\ell-1} \qquad X_\ell$$

$$\mathbf{H}_\infty(X_i) \geq k \quad \forall i$$ + independence

Need to trust several devices at different locations!

(especially when dealing with *public* randomness!)

What happens if some sources are corrupted?

SHELA sources:
Multi-source randomness extraction *without trust*

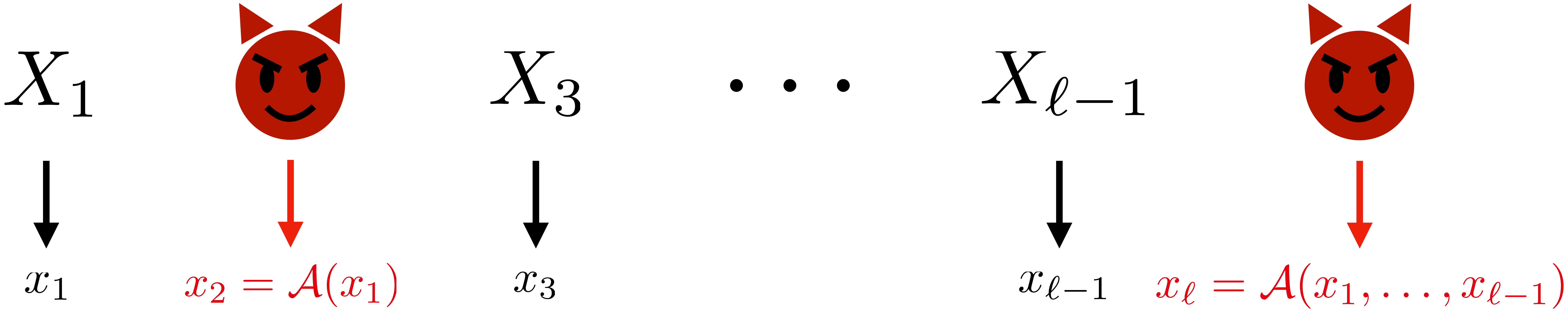
(t, ℓ, k)-**SHELA source:** Somewhere-Honest Entropic Look Ahead

$$X_1 \quad X_2 \quad X_3 \quad \cdots \quad X_{\ell-1} \quad X_\ell$$

1. Adversary chooses $\ell - t$ blocks to corrupt

SHELA sources:
Multi-source randomness extraction *without trust*

(t, ℓ, k)-**SHELA source:** Somewhere-Honest Entropic Look Ahead



1. Adversary chooses $\ell - t$ blocks to corrupt
2. Adversary fixes corrupted block based on previous samples

Adversary knows **positions** and **distributions** of **honest** blocks
Honest X_i's are independent of each other and satisfy $\mathbf{H}_\infty(X_i) \geq k$

Some other adversarial source models

Old:

Santha-Vazirani sources

Bit-fixing sources

[Dodis 2001]: *Bias-control limited sources*

Recent:

[Austrin, Chung, Mahmoody, Pass, Seth 2014]: *p-tampering attacks*

[Bentov, Gabizon, Zuckerman 2016]: *p-resettable sources*

[Chattopadhyay, Goodman, Goyal, Li 2019]: *Multi sources w/ local dependence*

[Dodis, Vaikuntanathan, Wichs 2019]: *Extractor-dependent sources*

[Ball, Goldreich, Malkin 2019]: *Somewhat-dependent sources*

Can we extract perfect randomness from SHELA sources?

No

Regime of interest: $t = \gamma \cdot \ell$ (constant fraction of corruptions), ℓ larger than some constant

impossibility for p-resettable sources
[Bentov, Gabizon, Zuckerman 2016]

Follows from impossibility for special subset of Santha-Vazirani sources

$\Rightarrow$

impossibility for SHELA sources

must have error $\epsilon = \Omega(1-\gamma)$

Holds even if honest blocks are uniform!

Can we extract “useful” randomness from SHELA sources?

The next best thing: somewhere-random sources

(t', ℓ')-**SR source:** Somewhere-Random

$$\boxed{Y_1} \quad \boxed{Y_2} \quad \boxed{Y_3} \quad \cdots \quad \boxed{Y_{\ell'}}$$

Guarantee: There exist $i_1 < i_2 < \cdots < i_{t'}$ such that $(Y_{i_1}, Y_{i_2}, \ldots, Y_{i_{t'}}) =$ unif

Interested in **convex combinations** of SR sources $\longrightarrow$ **convSR** sources

convSR sources are very useful!

SHELA $\longrightarrow$ great convSR sources

SR sources and one-sided error

$\mathcal{A}$ randomized algorithm with one-side error

$x \in \mathcal{L}$ Always outputs YES

$x \notin \mathcal{L}$ Outputs NO with probability 2/3, YES otherwise

Only guaranteed under uniform randomness!

SR source

$\boxed{Y_1} \longrightarrow \mathcal{A}(x, Y_1)$

$\boxed{Y_2} \longrightarrow \mathcal{A}(x, Y_2)$

⋮

$\boxed{Y_{\ell'}} \longrightarrow \mathcal{A}(x, Y_{\ell'})$

YES if all output YES

NO otherwise

Also one-sided error!

Runtime: $O(\underline{\ell'} \cdot t_{\mathcal{A}})$

wish to:

i) Minimize ℓ'

ii) Maximize length of Y_i's

Crypto applications of SR sources

Overall: non-interactive primitives with a “somewhere-random CRS”

We construct (from generic complexity assumptions):

- Non-interactive witness indistinguishable proof systems
- Non-interactive commitments

Elsewhere:

- Publicly-verifiable proof systems [Scafuro, Siniscalchi, Visconti 2019]

“Somewhere-extraction” from SHELA sources

Goal: Design $\mathsf{Ext} : \{0,1\}^{\ell \cdot n} \to \{0,1\}^{\ell' \cdot m}$ such that for every (t, ℓ, k)-SHELA source X,

$$\mathsf{Ext}(X) \approx_\epsilon \mathsf{convSR}$$

Want: #output blocks ℓ' and error ϵ **small**, output block length m **large**

Naive approach: apply 2-source extractor $2\mathsf{Ext}$ to every pair of blocks of X

Why? If X_i and X_j are honest, then $2\mathsf{Ext}(X_i, X_j) \approx \mathsf{unif}$

Cons:

i) $\ell' = \Omega(\ell^2)$

ii) Non-negligible error when $k < 0.44n$

Can we do better?

Better somewhere-extraction from SHELA sources

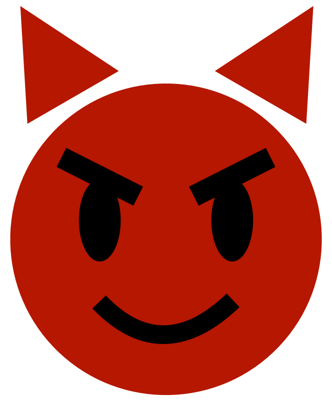
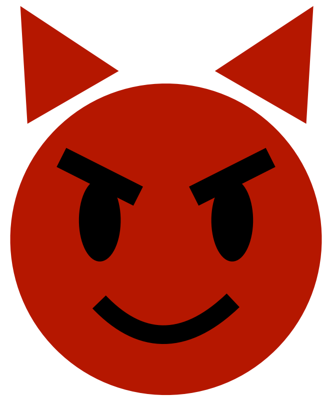
$X_i \in \{0,1\}^n$
$k \geq 0.51n$

$$X_1 \qquad X_2 \qquad X_3 \qquad X_4 \qquad X_5$$

Better somewhere-extraction from SHELA sources

$X_i \in \{0,1\}^n$

$k \geq 0.51n$

X_1  X_3  X_5

unbalanced 2-source extractors

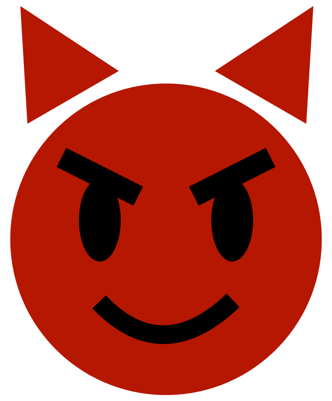
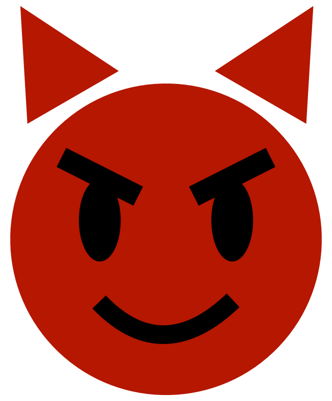
(left source: low entropy, right source: high entropy)

$\mathsf{Ext}_i : \{0,1\}^{i \cdot n} \times \{0,1\}^n \to \{0,1\}^m$

Better somewhere-extraction from SHELA sources

$X_i \in \{0,1\}^n$

$k \geq 0.51n$

X_1  X_3  X_5

left source **right source**

unbalanced 2-source extractors

(left source: low entropy, right source: high entropy)

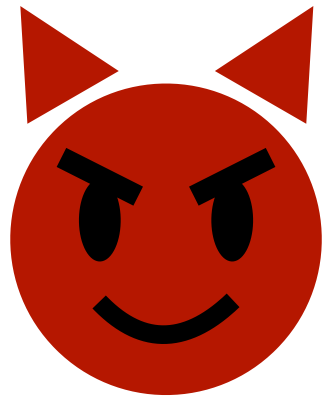
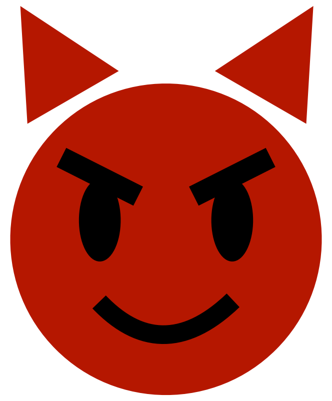
$\text{Ext}_i : \{0,1\}^{i \cdot n} \times \{0,1\}^n \to \{0,1\}^m$

$Y_1 = \text{Ext}_1(X_1,$  $)$

Better somewhere-extraction from SHELA sources

$X_i \in \{0,1\}^n$

$k \geq 0.51n$

X_1  X_3  X_5

left source **right source**

unbalanced 2-source extractors

(left source: low entropy, right source: high entropy)

$\mathsf{Ext}_i : \{0,1\}^{i \cdot n} \times \{0,1\}^n \to \{0,1\}^m$

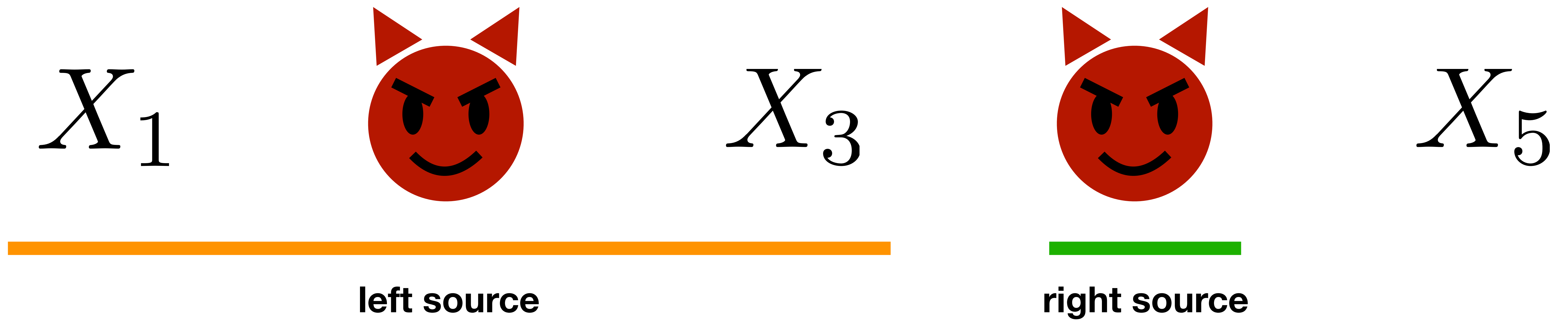
$Y_1 = \mathsf{Ext}_1(X_1,$  $)$

$Y_2 = \mathsf{Ext}_2(X_1$  $, X_3)$

Better somewhere-extraction from SHELA sources

$X_i \in \{0,1\}^n$

$k \geq 0.51n$



unbalanced 2-source extractors

(left source: low entropy, right source: high entropy)

$\mathsf{Ext}_i : \{0,1\}^{i \cdot n} \times \{0,1\}^n \to \{0,1\}^m$

$Y_1 = \mathsf{Ext}_1(X_1,$ 😈 $)$

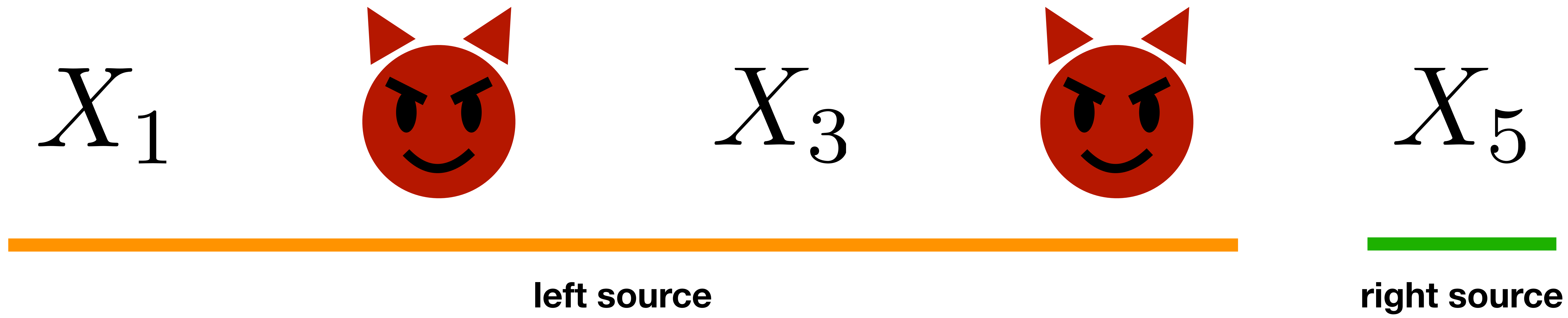
$Y_2 = \mathsf{Ext}_2(X_1$ 😈 $, X_3)$

$Y_3 = \mathsf{Ext}_3(X_1$ 😈 $X_3,$ 😈 $)$

Better somewhere-extraction from SHELA sources

$X_i \in \{0,1\}^n$

$k \geq 0.51n$



unbalanced 2-source extractors
(left source: low entropy, right source: high entropy)

$\mathsf{Ext}_i : \{0,1\}^{i \cdot n} \times \{0,1\}^n \to \{0,1\}^m$

$Y_1 = \mathsf{Ext}_1(X_1, \text{😈})$

$Y_2 = \mathsf{Ext}_2(X_1 \text{😈}, X_3)$

$Y_3 = \mathsf{Ext}_3(X_1 \text{😈} X_3, \text{😈})$

$Y_4 = \mathsf{Ext}_4(X_1 \text{😈} X_3, \text{😈}, X_5)$

Better somewhere-extraction from SHELA sources

$X_i \in \{0,1\}^n$
$k \geq 0.51n$

X_1 😈 X_3 😈 X_5

left source **right source**

unbalanced 2-source extractors
(left source: low entropy, right source: high entropy)

$\mathsf{Ext}_i : \{0,1\}^{i \cdot n} \times \{0,1\}^n \to \{0,1\}^m$

$Y_1 = \mathsf{Ext}_1(X_1, \text{😈})$

$\color{blue}{Y_2} = \mathsf{Ext}_2(X_1 \text{😈}, X_3)$

$Y_3 = \mathsf{Ext}_3(X_1 \text{😈} X_3, \text{😈})$

$\color{blue}{Y_4} = \mathsf{Ext}_4(X_1 \text{😈} X_3, \text{😈}, X_5)$

Better somewhere-extraction from SHELA sources

$X_i \in \{0,1\}^n$

$k \geq 0.51n$

X_1 😈 X_3 😈 X_5

left source **right source**

unbalanced 2-source extractors

(left source: low entropy, right source: high entropy)

$\mathrm{Ext}_i : \{0,1\}^{i \cdot n} \times \{0,1\}^n \to \{0,1\}^m$

i) $Y_2 \approx U_m$ **ii)** whp over fixing of Y_2, $Y_4 \approx U_m$

$Y_1 = \mathrm{Ext}_1(X_1,$ 😈$)$

contains enough min-entropy

$Y_2 = \mathrm{Ext}_2(X_1$ 😈$, X_3)$ **independent high min-entropy**

$Y_3 = \mathrm{Ext}_3(X_1$ 😈 $X_3,$ 😈$)$

independent high min-entropy

$Y_4 = \mathrm{Ext}_4(X_1$ 😈 $X_3,$ 😈$, X_5)$

contains enough min-entropy given Y_1, Y_2, Y_3

Y is ϵ-close to (t', ℓ')-convSR source in $\{0,1\}^{\ell' \cdot m}$

$\ell' = \ell - 1; \quad t' = t - 1;$

$\epsilon = 2^{-\Omega(n)}; \quad m = \Omega(n);$

works with only 2 honest blocks!

Somewhere-extraction from **low-entropy** SHELA sources

Want: Somewhere-extractor for $(t, \ell, k = \delta n)$-SHELA, for δ **arbitrarily small constant**

Idea: Combine previous high-entropy construction with **somewhere-condensers**

[Raz 2005], [Barak, Kindler, Shaltiel, Sudakov, Wigderson 2005], [Zuckerman 2007], [Li 2011]

Essentially the same parameters:

$\ell' = O_\delta(\ell); \quad t' = t - 1;$

$\epsilon = 2^{-\Omega_\delta(n)}; \quad m = \Omega_\delta(n);$

works with only 2 honest blocks!

Somewhere-extraction from a weak source

Can we extract useful convSR sources **without** exploiting structure of SHELA sources?

Treat (t, ℓ, k)-SHELA source $X \in \{0,1\}^{\ell \cdot n}$ as weak $(n' = \ell \cdot n, k' = t \cdot k)$-source

Naive somewhere-extractor:

Ext : $\{0,1\}^{n'} \times [2^d] \to \{0,1\}^m$ any strong seeded extractor

$\boxed{\text{Ext}(X,1)}$ $\boxed{\text{Ext}(X,2)}$ $\cdots$ $\boxed{\text{Ext}(X,2^d)}$ $\approx_\epsilon$ **SR source**

Problem: Superpolynomial #blocks if error ϵ is negligible!

Can we do better?

Somewhere-extraction from a weak source

Can we extract useful convSR sources **without** exploiting structure of SHELA sources? **No!**

Treat (t, ℓ, k)-SHELA source $X \in \{0,1\}^{\ell \cdot n}$ as weak $(n' = \ell \cdot n, k' = t \cdot k)$-source

Somewhere-extractor for (n', k')-sources with error ϵ, output block length m

#output blocks $\geq \dfrac{n' - k'}{\epsilon + 2^{-m}}$

If m isn't small and ϵ is negligible, need superpolynomial #output blocks

Proof: Somewhere-extractor $\longrightarrow$ disperser, so can apply well-known lower bounds

[Radhakrishnan, Ta-Shma 2000]

Open Q: Prove analogous result when $m = 1$

Summing up

- SHELA sources model multiple randomness sources corrupted by strong adversary
- Can’t extract perfect randomness
- **Can** extract great SR sources from low-entropy SHELA sources **(only need 2 honest blocks!)**
- SR sources are very useful (algorithms + crypto)
- Can’t extract useful SR sources **without** exploiting structure of SHELA source

Thanks for watching!