Privacy amplification

Randomness extractors

Outline

Basic concepts

Statistical distance
Min-entropy
Randomness extractors

Leftover Hash lemma

An efficient extractor based on universal hash functions

Average-case extractors

Randomness extraction in presence of side information

Quantum-proof extractors

Extraction in presence of quantum side information

Outline

Basic concepts

Statistical distance
Min-entropy
Randomness extractors

Leftover Hash lemma

An efficient extractor based on universal hash functions

Average-case extractors

Randomness extraction in presence of side information

Quantum-proof extractors

Extraction in presence of quantum side information

Why is “perfect” randomness needed?

Randomness: resource (example: key distribution)

Reality: random sources not perfect

Question for today’s presentation: can we turn bits generated by imperfect source into (almost) uniform bits?

Always?

If not, when and why not?

Examples

IID-Bit source: $X_1X_2\cdots X_n \in \{0,1\}$ **identical and independent, but biased:**

For each i, $\Pr[X_i = 1] = \delta$ **(unknown)**

Idea: Consider X **in pairs,**

Due to Von Neumann

$$X_iX_{i+1} = \begin{cases} 01 \implies & \text{output } 0 \\ 10 \implies & \text{output } 1 \\ 00/11 \implies & \text{discard} \end{cases}$$

Examples

IID-Bit source: $X_1X_2\cdots X_n \in \{0,1\}$ **identical and independent, but biased:**

For each i, $\Pr[X_i = 1] = \delta$ **(unknown)**

Idea: **Consider** X **in pairs,**

$$X_iX_{i+1} = \begin{cases} 01 \implies & \text{output } 0 \\ 10 \implies & \text{output } 1 \\ 00/11 \implies & \text{discard} \end{cases}$$

Independent bit source: $X = X_1X_2\cdots X_n \in \{0,1\}$ **independent but with**

Different biased $\Pr[X_i = 1] = \delta_i$ **for different** δ_i**, where** $0 < \delta \le \delta_i \le 1 - \delta$ **(constant** δ**)**

Idea: Output parity of each t **bits**

$$\left|\Pr[\oplus_{i=1}^{t} X_i = 1] - \frac{1}{2}\right| \le 2^{-\Omega(t)}$$

Examples

IID-Bit source: $X_1X_2\cdots X_n \in \{0,1\}$ **identical and independent, but biased:**

For each i**,** $\Pr[X_i = 1] = \delta$ **(unknown)**

Idea: **Consider** X **in pairs,**

$$X_iX_{i+1} = \begin{cases} 01 \implies & \text{output } 0 \\ 10 \implies & \text{output } 1 \\ 00/11 \implies & \text{discard} \end{cases}$$

Independent bit source: $X = X_1X_2\cdots X_n \in \{0,1\}$ **independent but with**

Different biased $\Pr[X_i = 1] = \delta_i$ **for different** δ_i**, where** $0 < \delta \le \delta_i \le 1 - \delta$ **(constant** δ**)**

Idea: Output parity of each t **bits**

$$\left| \Pr[\oplus_{i=1}^{t} X_i = 1] - \frac{1}{2} \right| \le 2^{-\Omega(t)}$$

$$\{(1-p)+p\}^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} p^{2k}(1-p)^{n-2k} + \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k+1} p^{2k+1}(1-p)^{n-(2k+1)}$$

$$= P\{X = \text{even}\} + P\{X = \text{odd}\}$$

$$\{(1-p)-p\}^n = \sum_{k=0}^{n} \binom{n}{k} (-p)^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} p^{2k}(1-p)^{n-2k} - \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k+1} p^{2k+1}(1-p)^{n-(2k+1)}$$

$$= P\{X = \text{even}\} - P\{X = \text{odd}\}$$

$$\{(1-p)+p\}^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} p^{2k} (1-p)^{n-2k} + \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k+1} p^{2k+1} (1-p)^{n-(2k+1)}$$

$$= \boxed{P\{X = \text{even}\} + P\{X = \text{odd}\}}$$

$$\{(1-p)-p\}^n = \sum_{k=0}^{n} \binom{n}{k})(-p)^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} p^{2k} (1-p)^{n-2k} - \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k+1} p^{2k+1} (1-p)^{n-(2k+1)}$$

$$= \boxed{P\{X = \text{even}\} - P\{X = \text{odd}\}}$$

$$P\{X = \text{even}\} = \frac{1}{2}\left(1 + (1-2p)^n\right)$$

$$\{(1-p)+p\}^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} p^{2k} (1-p)^{n-2k} + \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k+1} p^{2k+1} (1-p)^{n-(2k+1)}$$

$$= \boxed{P\{X = \text{even}\} + P\{X = \text{odd}\}}$$

$$\{(1-p)-p\}^n = \sum_{k=0}^{n} \binom{n}{k})(-p)^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} p^{2k} (1-p)^{n-2k} - \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k+1} p^{2k+1} (1-p)^{n-(2k+1)}$$

$$= \boxed{P\{X = \text{even}\} - P\{X = \text{odd}\}}$$

$$P\{X = \text{even}\} = \frac{1}{2}\left(1 + (1-2p)^n\right) \approx \frac{1}{2} + \frac{1}{2} e^{-2np}$$

$$\{(1-p)+p\}^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} p^{2k}(1-p)^{n-2k} + \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k+1} p^{2k+1}(1-p)^{n-(2k+1)}$$

$$= \boxed{P\{X = \text{even}\} + P\{X = \text{odd}\}}$$

$$\{(1-p)-p\}^n = \sum_{k=0}^{n} \binom{n}{k})(-p)^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} p^{2k}(1-p)^{n-2k} - \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k+1} p^{2k+1}(1-p)^{n-(2k+1)}$$

$$= \boxed{P\{X = \text{even}\} - P\{X = \text{odd}\}}$$

$$P\{X = \text{even}\} = \frac{1}{2}\left(1 + (1-2p)^n\right) \approx \frac{1}{2}\left(1 + (-1+2p')^n\right) = \frac{1}{2} + (-1)^n \frac{1}{2} e^{-2np'}$$

Randomness extraction

Source: Random variable X over $\{0,1\}^n$ in certain class $\mathscr{C}$

$\mathrm{IndBits}_{n,\delta} : X_1X_2\cdots X_n \in \{0,1\}$ **independent bits,** $\Pr[X_i = 1] = \delta_i$ **where** $0 < \delta \leq \delta_i \leq 1 - \delta$

IndBits$_{n,\delta}$: additionally assume all δ_i are equal.

(Deterministic) extractor: a function $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m \ \mathrm{s.t.} \forall$

$$X \xrightarrow{\mathrm{Ext}} \mathrm{Ext}(X)$$

$X \in \mathscr{C}$ **"ϵ-close" to uniform.**

Randomness extraction

Source: **Random variable X over $\{0,1\}^n$ in certain class $\mathscr{C}$**

$\mathrm{IndBits}_{n,\delta} : X_1X_2\cdots X_n \in \{0,1\}$ **independent bits,** $\Pr[X_i = 1] = \delta_i$ **where** $0 < \delta \le \delta_i \le 1 - \delta$

IndBits$_{n,\delta}$: **ume all δ_i are equal.**

Deterministic extractor not possible

(Determinis **a function** $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m \text{ s.t. } \forall$

$$X \xrightarrow{\mathrm{Ext}} \mathrm{Ext}(X)$$

$X \in \mathscr{C}$ **"ϵ-close" to uniform.**

Randomness extraction

Source: Random variable X over $\{0,1\}^n$ in certain class $\mathscr{C}$

$\mathrm{IndBits}_{n,\delta}: X_1X_2\cdots X_n \in \{0,1\}$ **independent bits,** $\Pr[X_i = 1] = \delta_i$ **where** $0 < \delta \leq \delta_i \leq 1 - \delta$

IndBits$_{n,\delta}$: additionally assume all δ_i ar

Criterion for extraction

(Deterministic) extractor: a function E $\{0,1\}^m$ s.t. $\forall$

$$X \xrightarrow{\mathrm{Ext}} \mathrm{Ext}(X)$$

$X \in \mathscr{C}$

"ϵ-close" to uniform.

Randomness extraction

Source: Random variable X over $\{0,1\}^n$ in certain class $\mathscr{C}$

$\mathrm{IndBits}_{n,\delta} : X_1 X_2 \cdots X_n \in \{0,1\}$ independent bits, $\Pr[X_i = 1] = \delta_i$ where $0 < \delta \le \delta_i \le 1 - \delta$

IndBits$_{n,\delta}$: additionally assume all δ_i are equal.

Length of random string and degree of randomness

(Deterministic) extractor: a function $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m \text{ s.t.}$

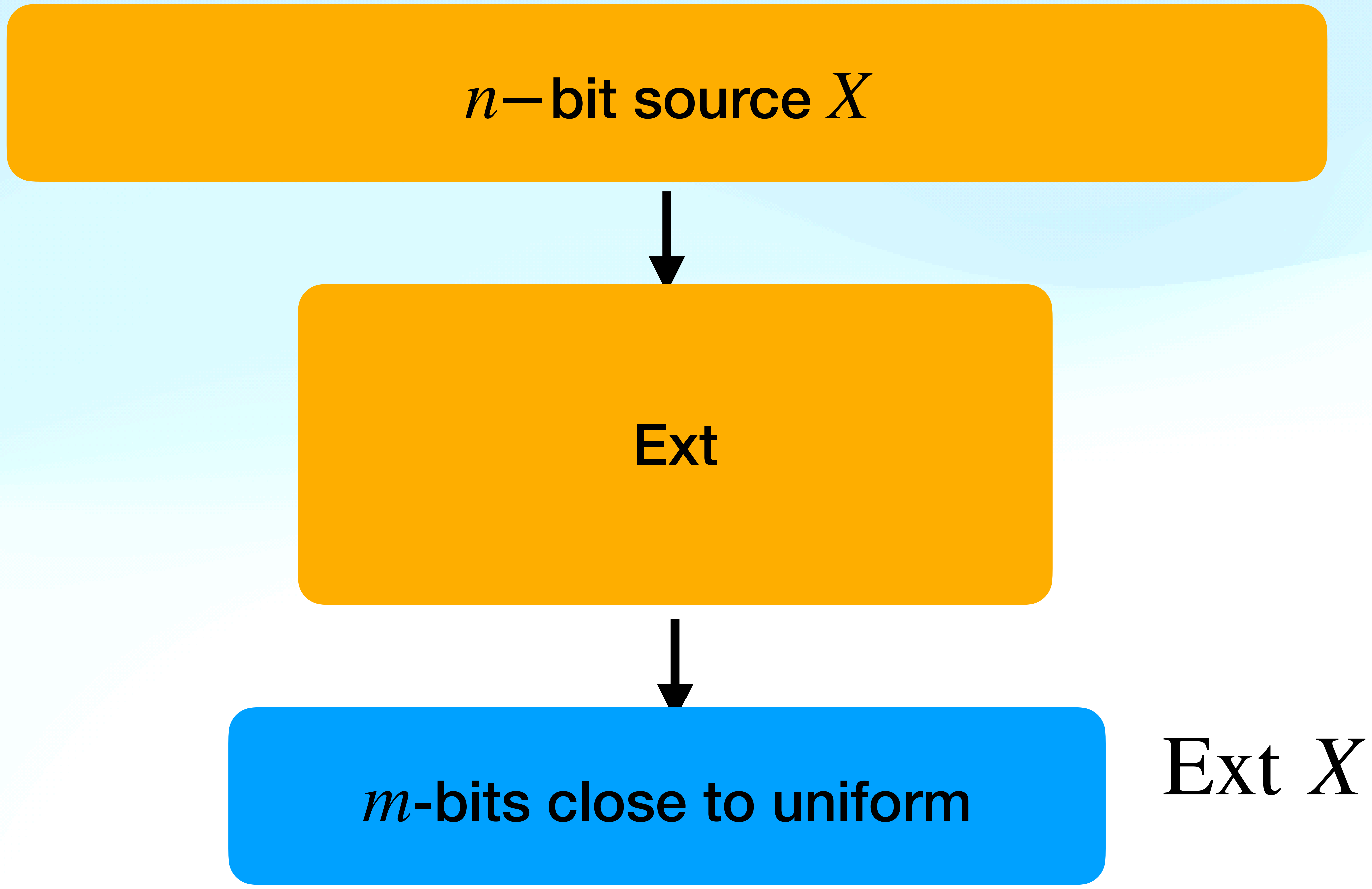
$$X \xrightarrow{\mathrm{Ext}} \mathrm{Ext}(X)$$

$X \in \mathscr{C}$

"ϵ-close" to uniform.

Deterministic extractors

- **(Determinstic) extractor:** a function $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m$ s.t. $\forall$ source $X \in \mathscr{C}$, $\mathrm{Ext}(X)$ is "$\epsilon-$close" to uniform



- **Single function works for all sources in $\mathscr{C}$**
- **Only one sample X is available**
- **Need to define "ϵ-close" to uniform**

Statistical distance

Definition: Let X, Y be random variables over the range U, statistical distance between X, Y is defined as

$$\boxed{\Delta(X, Y) \equiv \frac{1}{2} \sum_{u \in U} \left| \Pr(X = u) - Pr(Y = u) \right|}$$

View X, Y as vectors over $\mathbf{R}^{|U|}$, it is simply the L_1 distance.

Definition: X is $\epsilon-$ close to Y if

$$\boxed{\Delta(X, Y) \leq \epsilon\,.}$$

Important properties

Operational meaning: **max advantage to distinguish** X, Y

$$\Delta(X, Y) \equiv \max_{T \in U}(\Pr[X \in T] - \Pr[Y \in T])$$

- **If** X **is** $\epsilon-$**close** $Y,$ **then for any event** T,

$$\Pr(X \in T) \leq \Pr(Y \in T) + \epsilon$$

Important properties

Post-processing inequality: for any function f,

$$\Delta(f(X), f(Y)) \leq \Delta(X, Y)$$

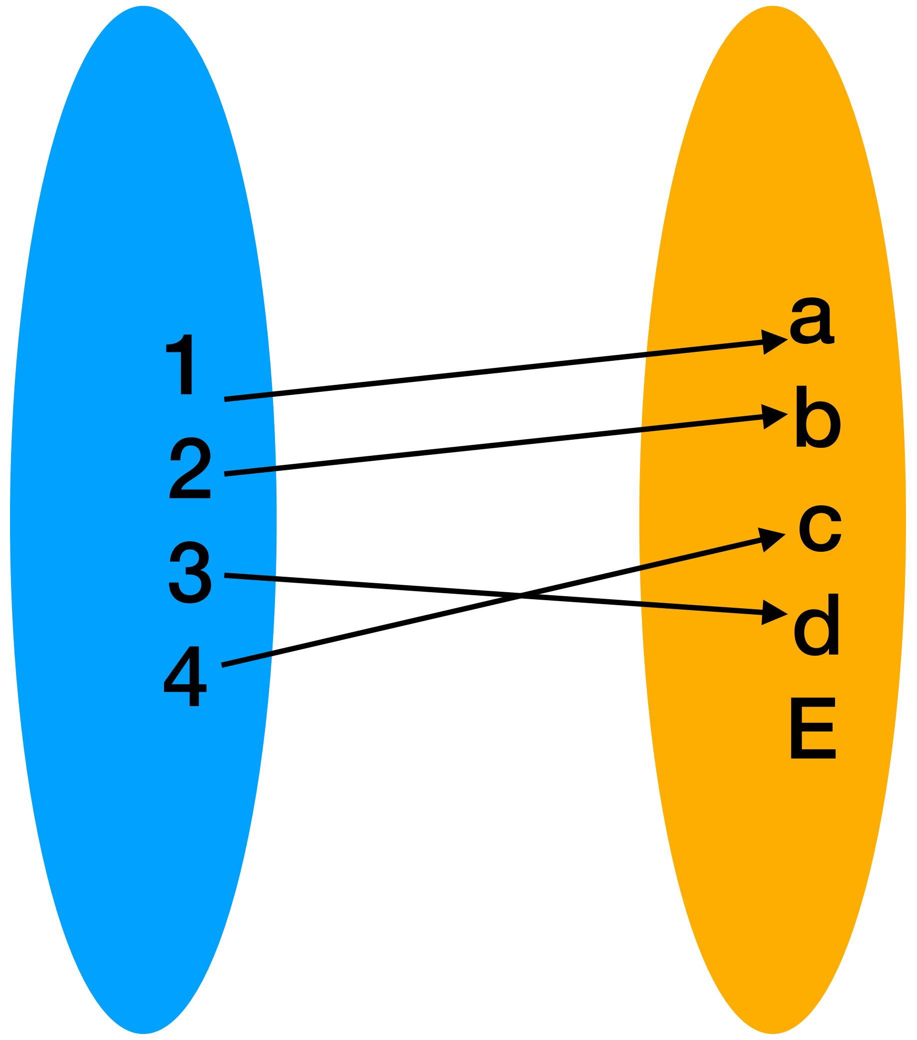
- post-processing only decreases statistical distance.
- Equality holds when f is injective.

Important properties

Post-processing inequality: for any function f,

$$\Delta(f(X), f(Y)) \leq \Delta(X, Y)$$

- post-processing only decreases statistical distance.
- Equality holds when f is injective.



Extractor for $\text{IndBits}_{n,\delta}$: One example

Theorem: $\forall$ constant δ, $\forall n, m \in \mathbf{N}, \exists \ \mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m$ for $\text{IndBits}_{n,\delta}$ source with error $\epsilon = m \cdot 2^{-\Omega(n/m)}$

$\mathrm{Ext}(X)$ breaks X into m blocks of length $\lfloor n/m \rfloor$ and outputs the parity of each block.

$n-$bit source X

$\downarrow$

Ext

$\downarrow$

m-bits close to uniform $\quad \mathrm{Ext}\ X$

Extractor for $\text{IndBits}_{n,\delta}$: One example

Theorem: $\forall$ constant δ, $\forall n, m \in \mathbf{N}, \exists\ \text{Ext} : \{0,1\}^n \to \{0,1\}^m$ for $\text{IndBits}_{n,\delta}$ source with error $\epsilon = m \cdot 2^{-\Omega(n/m)}$

$\text{Ext}(X)$ breaks X into m blocks of length $\lfloor n/m \rfloor$ and outputs the parity of each block.

$n-$bit source X

$\downarrow$

Ext

$\downarrow$

Deterministic!!!

m-bits close to uniform

$\text{Ext } X$

Extractor for $\text{IndBits}_{n,\delta}$: One example

Theorem: $\forall$ constant δ, $\forall n, m \in \mathbf{N}, \exists\ \mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m$ for $\text{IndBits}_{n,\delta}$ source with error $\epsilon = m \cdot 2^{-\Omega(n/m)}$

$\mathrm{Ext}(X)$ breaks X into m blocks of length $\lfloor n/m \rfloor$ and outputs the parity of each block.

$n-$bit source X

↓

Ext

↓

m-bits close to uniform

$\mathrm{Ext}\ X$

Response to Ravi sir's question on inhomogeneous randomness

Extractor for general source?

Can we extract truly uniform bits from any source?

No, if the source is not random, e.g., $X = 0^n$ with probability 1.

Hope: Ext works whenever X has sufficient "entropy".

$n-$bit source X

↓

Ext

↓

m-bits close to uniform $\quad \mathrm{Ext}\ X$

Is Shannon entropy a good candidate?

1^{st} attempt : Shannon entropy

$$H_{\text{sh}}(X) \equiv \sum_x P(X=x)\log\left[\frac{1}{P(X=x)}\right]$$

Not good,

Example: X defined as follows:

With probability $\frac{1}{2}$, $X = 0^n$

With probability $\frac{1}{2}$, sample $X =$ uniform on $\{0,1\}^n$

$$\boldsymbol{H_{\text{sh}}(X) \geq \frac{n}{2}}$$

But $\mathbf{Pr}[X = 0^n] > \frac{1}{2}$; can't extract from X.

1^{st} attempt : Shannon entropy

$$H_{\text{sh}}(X) \equiv \sum_x P(X = x)\log\left[\frac{1}{P(X = x)}\right]$$

Not good,

Example: X **defined as follows:**

With probability $\frac{1}{2}$, $X = 0^n$

With probability $\frac{1}{2}$, **sample** $X =$ **uniform on** $\{0,1\}^n$

$$\boldsymbol{H_{\text{sh}}(X) \geq \frac{n}{2}}$$

But $\mathbf{Pr}[X = 0^n] > \frac{1}{2}$; **can't extract from** X.

On an average, more than $50\,\%$ **of the time, it will yield string of zeros.**

2^{nd} attempt: Min-Entropy

Def. Min entropy $H_{\min}(X) \equiv \min_x \left(\log \frac{1}{P(X = x)} \right)$

$H_{\min}(X) \geq k$ If for every x, $\Pr[X = x] \leq 2^{-k}$

Worst-case notion; possible for extraction.

Def.: X is $k-$ source if $H_{\min}(X) \geq k$.

Extractor for the class of $k-$sources?

Flat $k-$sources: have uniform distribution on a set $S \subset \{0,1\}^n$ with $|S| = 2^k$.

Every $k-$source is a convex combination of flat $k-$sources (provided that $2^k \in \mathbf{N}$), i.e., $X = \sum_i p_i X_i$ with $0 \le p_i \le 1$, $\sum_i p_i = 1$ and all the X_i are flat $k-$ sources.

Every $k-$source is a convex combination of flat $k-$sources (provided that $2^k \in \mathbf{N}$), i.e., $X = \sum_i p_i X_i$ with $0 \leq p_i \leq 1$, $\sum_i p_i = 1$ and all the X_i are flat $k-$ sources.

Simply use convex sum argument.

$X : k-$**source on** $[N]$

$X : k-$**source on** $[N]$

View X **as partitioning a circle of unit circumference into** N **intervals.**

$X : k-$**source on** $[N]$

View X **as partitioning a circle of unit circumference into** N **intervals.**

Length of the t**th interval=** $\Pr(X = t)$

$X : k-$**source on** $[N]$

View X **as partitioning a circle of unit circumference into** N **intervals.**

Length of the t**th interval=** $\Pr(X = t)$

If we associate the points on the circle with $[0,1)$**, then the** t**th interval is** $[\Pr(X < t), \Pr(X < t))$

$X : k-$source on $[N]$

View X as partitioning a circle of unit circumference into N intervals.

Length of the tth interval= $\Pr(X = t)$

If we associate the points on the circle with $[0,1)$, then the tth interval is $[\Pr(X < t), \Pr(X < t))$

Consider a set S of K equally spaced points on the circle.

$X : k-$source on $[N]$

View X as partitioning a circle of unit circumference into N intervals.

Length of the tth interval= $\Pr(X = t)$

If we associate the points on the circle with $[0,1)$, then the tth interval is $[\Pr(X < t), \Pr(X < t))$

Consider a set S of K equally spaced points on the circle.

Since each interval is half-open and has length at most $\frac{1}{K}$, each interval contains at most one point from S,

$X : k-$source on $[N]$

View X as partitioning a circle of unit circumference into N intervals.

Length of the tth interval= $\Pr(X = t)$

If we associate the points on the circle with $[0,1)$, then the tth interval is $[\Pr(X < t), \Pr(X < t))$

Consider a set S of K equally spaced points on the circle.

Since each interval is half-open and has length at most $\frac{1}{K}$, each interval contains at most one point from S,

So the uniform distribution on the set $T(S) = \{t : S \cap I_t \neq \phi\}$ is a flat $k-$source.

$X : k-$source on $[N]$

View X as partitioning a circle of unit circumference into N intervals.

Length of the tth interval= $\Pr(X = t)$

If we associate the points on the circle with $[0,1)$, then the tth interval is $[\Pr(X < t), \Pr(X < t))$

Consider a set S of K equally spaced points on the circle.

Since each interval is half-open and has length at most $\frac{1}{K}$, each interval contains at most one point from S,

So the uniform distribution on the set $T(S) = \{t : S \cap I_t \neq \phi\}$ is a flat $k-$source.

If we perform a uniformly random rotation of S on the circle to obtain a rotated set R and then choose

$X : k-$**source on** $[N]$

View X **as partitioning a circle of unit circumference into** N **intervals.**

Length of the t**th interval=** $\Pr(X = t)$

If we associate the points on the circle with $[0,1)$**, then the** t**th interval is** $[\Pr(X < t), \Pr(X < t))$

Consider a set S **of** K **equally spaced points on the circle.**

Since each interval is half-open and has length at most $\frac{1}{K}$**, each interval contains at most one point from** S**,**

So the uniform distribution on the set $T(S) = \{t : S \cap I_t \neq \phi\}$ **is a flat** $k-$**source.**

If we perform a uniformly random rotation of S **on the circle to obtain a rotated set** R **and then choose**

A uniformly random element of $T(R),$ **the probability that we output any value** $t \in [N]$ **Is equal to the length of** I_t.

$X : k-$source on $[N]$

View X as partitioning a circle of unit circumference into N intervals.

Length of the tth interval= $\Pr(X = t)$

If we associate the points on the circle with $[0,1)$, then the tth interval is $[\Pr(X < t), \Pr(X < t))$

Consider a set S of K equally spaced points on the circle.

Since each interval is half-open and has length at most $\frac{1}{K}$, each interval contains at most one point from S,

So the uniform distribution on the set $T(S) = \{t : S \cap I_t \neq \phi\}$ is a flat $k-$source.

If we perform a uniformly random rotation of S on the circle to obtain a rotated set R and then choose

A uniformly random element of $T(R),$ the probability that we output any value $t \in [N]$ Is equal to the length of I_t.

Thus, we have decomposed X as a convex sum of flat $k-$sources (Specifically, $X = \sum_T p_T U_T,$ where the sum is over subsets $T \subset [N]$ of size $K,$ and $p_T = \Pr_R[T(R) = T]$.

Impossibility of deterministic extraction

Theorem: **For any** $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}$, **there exists an** $(n-1)$**- source** X **such that** $\mathbf{Ext}(X) = \mathbf{constant}$

What is variable? Source
What is fixed? Extractor

Consequence: No deterministic
extractor possible.

For any function $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}$**, there exists an** $(n-1)$**- source** X **such that** $\mathrm{Ext}(X)$ **is constant.**

Proof: Let $b \in \{0,1\}$ **be such that** $|S_b| > \dfrac{2^n}{2}$ **with** $S_b = \{x \,|\, \mathrm{Ext}(x) = b\}$.

Choose a subset $S' \subset S_b$ **such that** $|S'| = 2^{n-1}$**.**

Define X **by the following distribution:**

$$p_x = \begin{cases} \dfrac{1}{2^{n-1}} & \text{if } x \in S' \\ 0 & \text{otherwise} \end{cases}$$

$H_{\min}(X) = n - 1$, **but** $\mathrm{Ext}(\mathrm{X}) = \mathrm{b}$ **is a constant!**

For any function $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}$, there exists an $(n-1)$- source X such that $\mathrm{Ext}(X)$ is constant.

Pro $\{x \mid \mathrm{Ext}(x) = b\}$.

2^{n-1}.

In words: If one knows the randomness extractor function's domain and range, what One does in response is to look at only pre-image of one value. That's it!!!

$H_{\min}($ s a constant!

For every $n, k, m \in \mathbf{N}$, **every** $\epsilon > 0$ **and every flat** $k-$**source** X**, if we choose a random function** $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m$ **With** $m = k - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$**, then** $\mathrm{EXT}(X)$ **will be** $\epsilon-$**close to** U_m **with probability** $1 - 2\Omega(K\epsilon^2)$**, where** $K = 2^k$**.**

For every $n, k, m \in \mathbf{N}$, **every** $\epsilon > 0$ **and every flat** $k-$**source** X, **if we choose a random function** $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m$ **With** $m = k - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$, **then** $\mathrm{EXT}(X)$ **will be** $\epsilon-$**close to** U_m **with probability** $1 - 2\Omega(K\epsilon^2)$, **where** $K = 2^k$.

Marks output alphabet length

Closeness to uniform distribution

For every $n, k, m \in \mathbf{N}$, every $\epsilon > 0$ and every flat $k-$source X, if we choose a random function $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m$ With $m = k - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$, then $\mathrm{EXT}(X)$ will be $\epsilon-$close to U_m with probability $1 - 2\Omega(K\epsilon^2)$, where $K = 2^k$.

For all $T \subset [M]$, $\left| \Pr\left[Ext(X) \in T\right] - \Pr\left[U_m \in T\right] \right| \leq \epsilon$.

For every $n, k, m \in \mathbf{N}$, every $\epsilon > 0$ and every flat $k-$source X, if we choose a random function $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m$ With $m = k - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$, then $\mathrm{EXT}(X)$ will be $\epsilon-$close to U_m with probability $1 - 2\Omega(K\epsilon^2)$, where $K = 2^k$.

For all $T \subset [M]$, $\left| \Pr\left[Ext(X) \in T\right] - \Pr\left[U_m \in T\right] \right| \leq \epsilon$.

OR

$\frac{1}{K}\left| \{x \in \mathrm{Supp}(X) : EXT(X) \in T\} \right|$ differs from the density $\mu(T)$ by almost ϵ .

For every $n, k, m \in \mathbf{N}$, every $\epsilon > 0$ and every flat $k-$source X, if we choose a random function $\mathrm{Ext}: \{0,1\}^n \to \{0,1\}^m$ With $m = k - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$, then $\mathrm{EXT}(X)$ will be $\epsilon-$close to U_m with probability $1 - 2\Omega(K\epsilon^2)$, where $K = 2^k$.

For all $T \subset [M]$, $\left|\Pr\left[Ext(X) \in T\right] - \Pr\left[U_m \in T\right]\right| \leq \epsilon$.

OR

$\frac{1}{K}\left|\{x \in \mathrm{Supp}(X) : EXT(X) \in T\}\right|$ **differs from the density** $\mu(T)$ **by almost** ϵ.

For each point $x \in Supp(X)$, **the probability that** $Ext(X) \in T$ **is** $\mu(T)$**, and these events are independent.**

For every $n, k, m \in \mathbf{N}$, every $\epsilon > 0$ and every flat $k-$source X, if we choose a random function $\mathrm{Ext} : \{0,1\}^n \to \{0,1\}^m$ With $m = k - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$, then $\mathrm{EXT}(X)$ will be $\epsilon-$close to U_m with probability $1 - 2\Omega(K\epsilon^2)$, where $K = 2^k$.

For all $T \subset [M]$, $\left| \Pr\left[Ext(X) \in T\right] - \Pr\left[U_m \in T\right] \right| \leq \epsilon$.

OR

$\frac{1}{K}\left| \{x \in \mathrm{Supp}(X) : EXT(X) \in T\} \right|$ differs from the density $\mu(T)$ by atmost ϵ .

For each point $x \in Supp(X)$, the probability that $Ext(X) \in T$ is $\mu(T)$, and these events are independent.

By the Chernoff Bound for each fixed T, the condition holds with a probability of at least $1 - 2^{-\Omega(K\epsilon^2)}$.

For every $n, k, m \in \mathbf{N}$, every $\epsilon > 0$ and every flat $k-$source X, if we choose a random function $\mathrm{Ext}:\{0,1\}^n \to \{0,1\}^m$ With $m = k - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$, then $\mathrm{EXT}(X)$ will be $\epsilon-$close to U_m with probability $1 - 2\Omega(K\epsilon^2)$, where $K = 2^k$.

For all $T \subset [M]$, $\left|\Pr\left[Ext(X) \in T\right] - \Pr\left[U_m \in T\right]\right| \leq \epsilon$.

OR

$\frac{1}{K}\left|\{x \in \mathrm{Supp}(X) : EXT(X) \in T\}\right|$ differs from the density $\mu(T)$ by almost ϵ.

For each point $x \in Supp(X)$, the probability that $Ext(X) \in T$ is $\mu(T)$, and these events are independent.

By the Chernoff Bound for each fixed T, the condition holds with a probability of at least $1 - 2^{-\Omega(K\epsilon^2)}$.

For every $n, k, m \in \mathbf{N}$, every $\epsilon > 0$ and every flat $k-$source X, if we choose a random function $\mathrm{Ext}:\{0,1\}^n \to \{0,1\}^m$ With $m = k - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$, then $\mathrm{EXT}(X)$ will be $\epsilon-$close to U_m with probability $1 - 2\Omega(K\epsilon^2)$, where $K = 2^k$.

For all $T \subset [M]$, $\left|\Pr\left[Ext(X) \in T\right] - \Pr\left[U_m \in T\right]\right| \leq \epsilon$.

OR

$\frac{1}{K}\left|\{x \in \mathrm{Supp}(X) : EXT(X) \in T\}\right|$ differs from the density $\mu(T)$ by almost ϵ.

For each point $x \in Supp(X)$, the probability that $Ext(X) \in T$ is $\mu(T)$, and these events are independent.

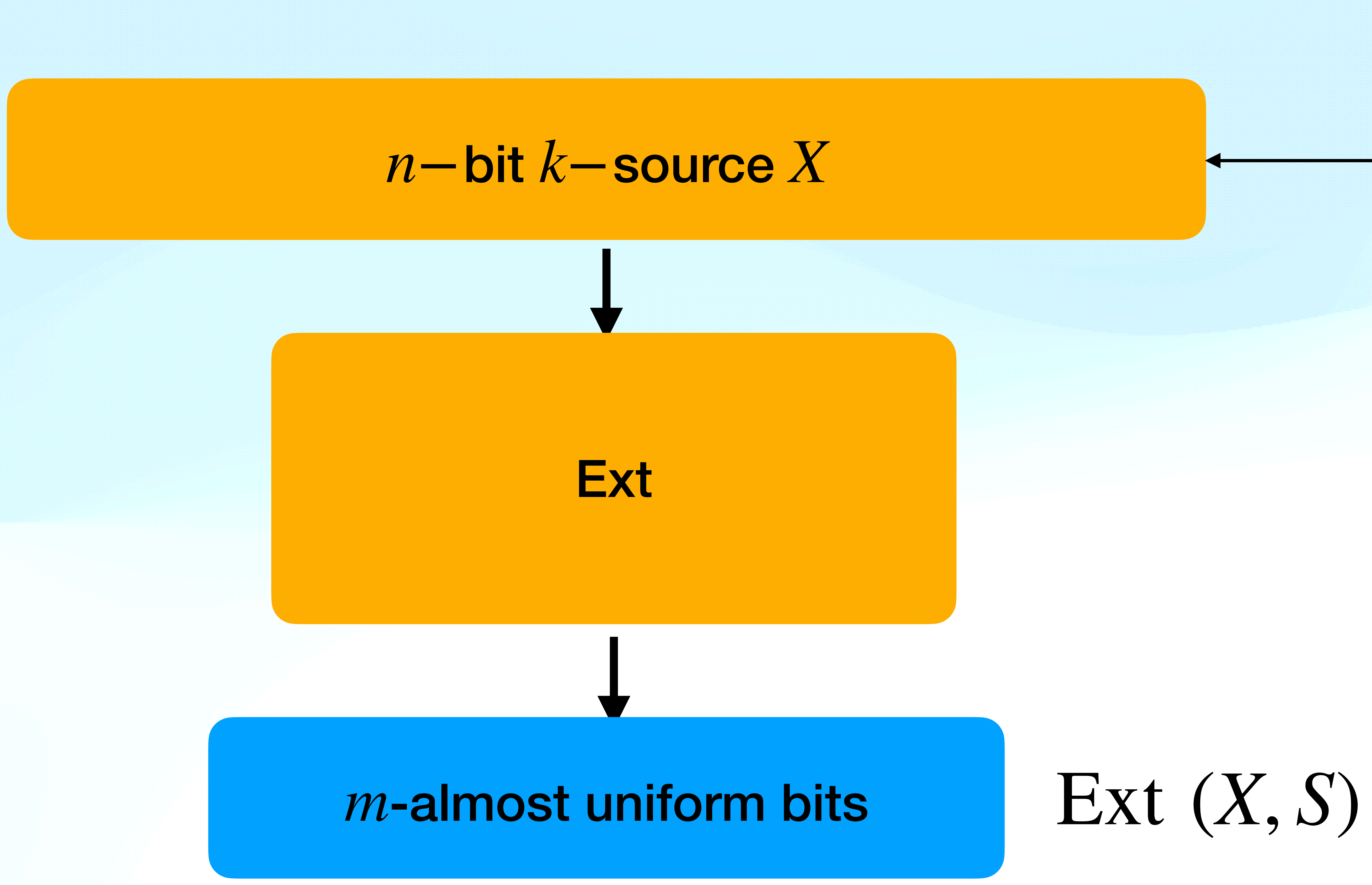
By the Chernoff Bound for each fixed T, the condition holds with a probability of at least $1 - 2^{-\Omega(K\epsilon^2)}$.

Then, the probability that condition is violated for at least one T is at most $2^M 2^{-\Omega(K\epsilon^2)}$, which is less than 1 for

$$m = k - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$$

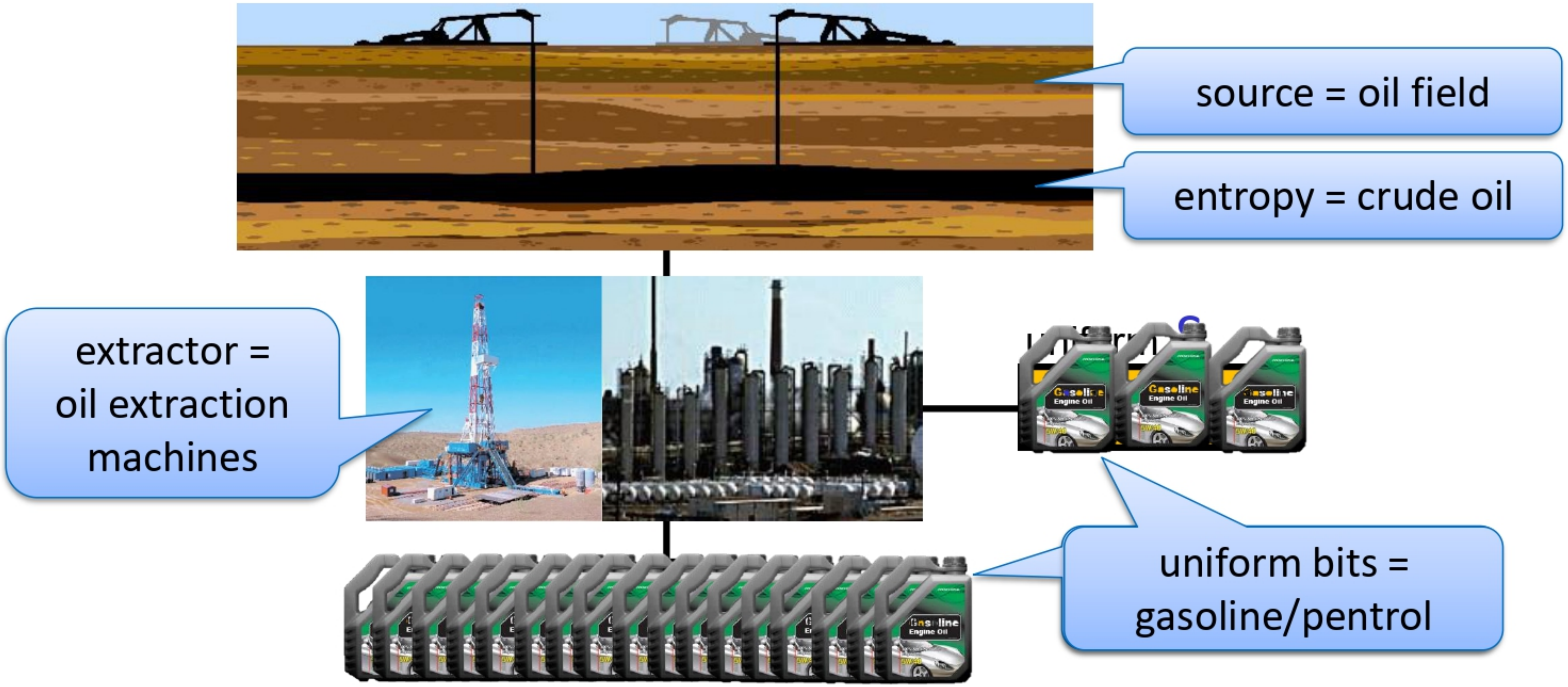
Seeded extractor

Add short uniform seed as catalyst for extraction.



Ext: $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ **is** $(k, \epsilon)-$ **seeded extractor**
if $\forall k-$ **source** $X, \mathrm{Ext}(X; S)$ **is** $\epsilon-$ **close uniform** U_m .

An Analogy: Oil Extraction



Ext: $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is (k,ε)-*seeded extractor if*

$\forall$ k-source X, $Ext(X; S)$ is ε-close uniform U_m

Aims

Minimize seed length d

Minimise initial gasoline investment.

Maximise output length m, ideally close to min-entropy k

Extract and distill all crude oil to gasoline.

Extraction even for small entropy rate $\frac{k}{n}$

i.e., even oil field has low crude oil content.

Explicit construction: efficient polynomial time extractor

Cost-efficiency of oil extraction machines

Seeded extractor

A function $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ **is a** $(k, \epsilon)-$**extractor if every** $k-$**source** X **on** $\{0,1\}^n, \mathrm{EXT}(X, U_d)$ **is** $\epsilon-$ **close to** U_m .

Seeded extractor

For every $n \in \mathbf{N},\ k \in [0,n]$ **and** $\epsilon > 0$, **there exists a** $(k,\epsilon)-$ **extractor** $\mathrm{EXT} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ **with**

$$m = k + d - 2\log\frac{1}{\epsilon} - O(1) \text{ and}$$

$$d = \log(n-k) + 2\log\frac{1}{\epsilon} + O(1)\,.$$

Seeded extractor

For every $n \in \mathbf{N},\ k \in [0,n]$ **and** $\epsilon > 0$, **there exists a** $(k,\epsilon)-$ **extractor** $\mathrm{EXT} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ **with**

$$m = k + d - 2\log\frac{1}{\epsilon} - O(1) \text{ and}$$

$$d = \log(n-k) + 2\log\frac{1}{\epsilon} + O(1)\,.$$

Proof: Sufficient to work out for flat sources.

Seeded extractor

For every $n \in \mathbf{N},\ k \in [0,n]$ and $\epsilon > 0$, there exists a $(k,\epsilon)-$ extractor $\mathrm{EXT} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with

$$m = k + d - 2\log\frac{1}{\epsilon} - O(1) \text{ and}$$

$$d = \log(n-k) + 2\log\frac{1}{\epsilon} + O(1)\,.$$

Proof: Sufficient to work out for flat sources.

Choose the extractor EXT at random.

Seeded extractor

For every $n \in \mathbf{N},\ k \in [0,n]$ and $\epsilon > 0$, there exists a $(k,\epsilon)-$ extractor $\mathrm{EXT} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with

$$m = k + d - 2\log\frac{1}{\epsilon} - O(1) \text{ and}$$

$$d = \log(n-k) + 2\log\frac{1}{\epsilon} + O(1)\,.$$

Proof: Sufficient to work out for flat sources.

Choose the extractor EXT at random.

Then the probability that the extractor fails is at most the number of flat k-sources times the probability EXT fails for a fixed flat k-source.

By the above proposition, the probability of failure for a fixed flat k-source is at most $2^{-\Omega(KD\epsilon^2)}$

Seeded extractor

For every $n \in \mathbf{N},\ k \in [0,n]$ and $\epsilon > 0$, there exists a $(k,\epsilon)-$ extractor $\mathrm{EXT} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with

$$m = k + d - 2\log\frac{1}{\epsilon} - O(1) \text{ and}$$

$$d = \log(n-k) + 2\log\frac{1}{\epsilon} + O(1)\,.$$

Proof: Sufficient to work out for flat sources.

Choose the extractor EXT at random.

Then the probability that the extractor fails is at most the number of flat k-sources times the probability EXT fails for a fixed flat k-source.

By the above proposition, the probability of failure for a fixed flat k-source is at most $2^{-\Omega(KD\epsilon^2)}$

Since (X, U_d) **is a flat** $(k+d)-$ **source and** $m = k + d - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$**. Thus the total probability is at most**

$$\binom{N}{K} 2^{-\Omega(KD\epsilon^2)} \leq \left(\frac{Ne}{K}\right)^K 2^{-\Omega(KD\epsilon^2)}$$

Seeded extractor

For every $n \in \mathbf{N}$, $k \in [0,n]$ and $\epsilon > 0$, there exists a $(k, \epsilon)-$ extractor $\mathrm{EXT} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with

$$m = k + d - 2\log\frac{1}{\epsilon} - O(1) \text{ and}$$

$$d = \log(n-k) + 2\log\frac{1}{\epsilon} + O(1)\,.$$

Proof: Sufficient to work out for flat sources.

Choose the extractor EXT at random.

Then the probability that the extractor fails is at most the number of flat k-sources times the probability EXT fails for a fixed flat k-source.

By the above proposition, the probability of failure for a fixed flat k-source is at most $2^{-\Omega(KD\epsilon^2)}$

Since (X, U_d) is a flat $(k+d)-$ source and $m = k + d - 2\log\left(\frac{1}{\epsilon}\right) - O(1)$. Thus the total probability is at most

$$\binom{N}{K} 2^{-\Omega(KD\epsilon^2)} \leq \left(\frac{Ne}{K}\right)^K 2^{-\Omega(KD\epsilon^2)}$$

What we can achieve?

Non-constructively, $\forall n, k, \epsilon, \exists (k, \epsilon)-$seeded extractor with seed length

Seed length $d = \log(n-k) + 2\log(1/\epsilon) + O(1)$

Output length $d = k + d - 2\log(1/\epsilon) - O(1)$

-Use logarithmic-length seed

-Extract almost all min-entropy out

-For any small entropy rate

-However, not an explicit construction

Extractor example: Universal Hash Functions

Let $\mathscr{H} = \{h : \{0,1\}^n \rightarrow \{0,1\}^m\}$ be a family of Hash functions.

Let H denote a random hash function from $\mathscr{H}$

Definition: $\mathscr{H}$ is universal if for every $x \neq x' \in \{0,1\}^n$,

$$\Pr\left[H(x) = H(x')\right] \leq \frac{1}{2^m}$$

I.e., probability of hash collision on x and x' is small for every $x \neq x'$

Universal Hash Functions

Let $\mathscr{H} = \{h : \{0,1\}^n \to \{0,1\}^m\}$ **be a family of Hash functions.**

Let H **denote a random hash function from** $\mathscr{H}$

Definition: $\mathscr{H}$ **is universal if for every** $x \neq x' \in \{0,1\}^n$,

$$\Pr\left[H(x) = H(x')\right] \leq \frac{1}{2^m}$$

I.e., probability of hash collision on x **and** x' **is small for every** $x \neq x'$

Example

$\mathscr{H} = \{h_s : s \in GF(2^n)\}$, **where** $h_s(x) =$ **first** m **bits of** $s \cdot x$

Note that $h_s(x) = h_s(x')$ **implies** $s \cdot (x - x') = 0^m z$ **for some** $z \in \{0,1\}^{n-m}$.

Each z **determines** $s = \dfrac{0^m z}{x - x'}$, **so at most** 2^{n-m} **out of** $2^n h_s$.

So, $\Pr[H(x) = H(x')] \leq \dfrac{2^{n-m}}{2^n} = \dfrac{1}{2^m}$.

Extractor construction

Let $\mathscr{H} = \{h : \{0,1\}^n \to \{0,1\}^m\}$ **be a family of hash functions.**

Let H **denote a random hash function from** $\mathscr{H}$

Def.: We say $\mathscr{H}$ **is universal if for every** $x \neq x' \in \{0,1\}^n$,

$$\Pr\left[H(x) = H(x')\right] \leq \frac{1}{2^m}$$

I.e., probability of hash collision on x **and** x' **is small for every** $x \neq x'$

Define $\mathrm{Ext} : \{0,1\}^{\mathrm{n}} \times \{0,1\}^{\mathrm{d}} \to \{0,1\}^{\mathrm{m}}$ **by** $\mathrm{Ext}(x, h) = h(x)$

i.e., use seed h **to select a hash function to hash**

Why does it work?

Define $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ **by** $\mathrm{Ext}(x, h) = h(x)$, **where** h **is from universal hash family**

$$\mathscr{H} = \{h : \{0,1\}^n \to \{0,1\}^m\}$$

$$\Pr\left[H(x) = H(x')\right] \leq \frac{1}{2^m} \textbf{ for every } x \neq x' \in \{0,1\}^n$$

Want to show $\left(\mathrm{Ext}(X; H), H\right) \approx_\epsilon (U_m, H)$ **or** $\left(H, H(X)\right) \approx_\epsilon (H, U_m)$

Analyse via "collision probability"

Step 1: Z **has small "collision probability"** $\implies$ Z **is close to uniform**

Step 2: Show $\left(H, H(X)\right)$ **has small "collision probability".**

Collision probability

Def.: Let Z be a random variable over $[M]$. Collision probability of Z

$CP(Z) \equiv \Pr(Z = Z')$**, where Z' is an independent copy of Z.**

e.g., for uniform distribution $U_{[M]}$, $CP\left(U_{[M]}\right) = \dfrac{1}{M}$

View Z as vector $v \in \mathbf{R}^M$, i.e., $v_i = \Pr[Z = i]$, then $CP(Z)$ is the square of L_2-norm of v.

$$CP(Z) = Pr[Z = Z'] = \sum_i Pr[Z = Z' = i] = \sum_i v_i^2 = ||V||_2^2$$

Intuition: uniform distribution minimise collision probability. If $CP(Z) \approx CP(U_{[M]})$, then Z is close to $U_{[M]}$

Small CP $\implies$ Close to uniform

Lemma: $CP(Z) \leq \frac{1+\epsilon}{M} \implies \Delta(Z, U_{[M]}) \leq \frac{\sqrt{\epsilon}}{2}$

Proof: Define $w \in \mathbf{R}^M$ **by** $w_i = \left(v_i - \frac{1}{M}\right)$.

Note $\Delta(Z, U_{[M]}) = \frac{1}{2} \cdot ||w||_1$

Let's compute $||w||_2^2 = \sum_i \left(v_i - \frac{1}{M}\right)^2$

$$= \sum_i v_i^2 - \sum_i \left(\frac{2v_i}{M}\right) + \sum_i \left(\frac{1}{M}\right)^2$$

$$= CP(Z) - \frac{1}{M}$$

Thus, $||w||_2^2 \leq \frac{\epsilon}{M}$, **or** $||w||_2 \leq \sqrt{\frac{\epsilon}{M}}$

By relation between L_1 **and** L_2 **norm** $||w||_1 \leq \sqrt{M} \cdot ||w||_2 \leq \sqrt{\epsilon}$

So, $\Delta(Z, U_{[M]}) \leq \frac{\sqrt{\epsilon}}{2}$