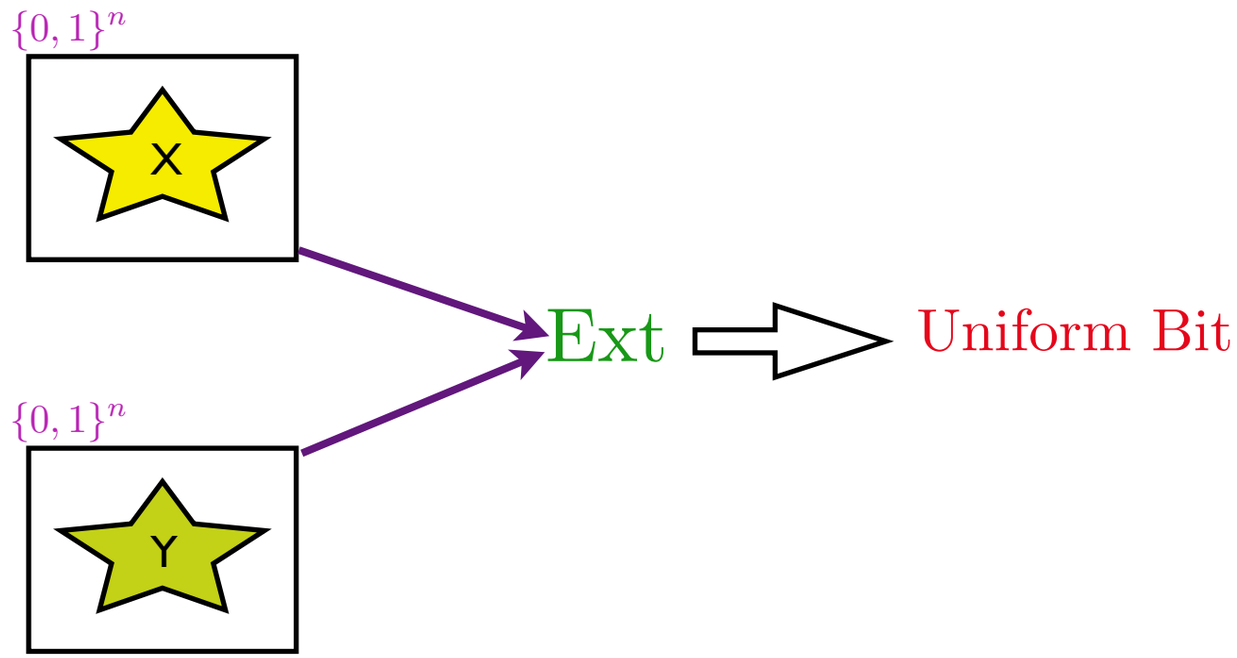
Part 3: Explicit 2-Source Extractors

Eshan Chattopadhyay
IAS & Cornell

STOC 2018: Extractor Workshop

2-Source Extractor

X, Y are independent (n,k)-sources



More formally,

$$\left| \Pr[\mathrm{Ext}(X,Y) = 1] - \frac{1}{2} \right| \leq \epsilon$$

Ramsey Graphs

K-Ramsey graph: No independent set or clique of size K.

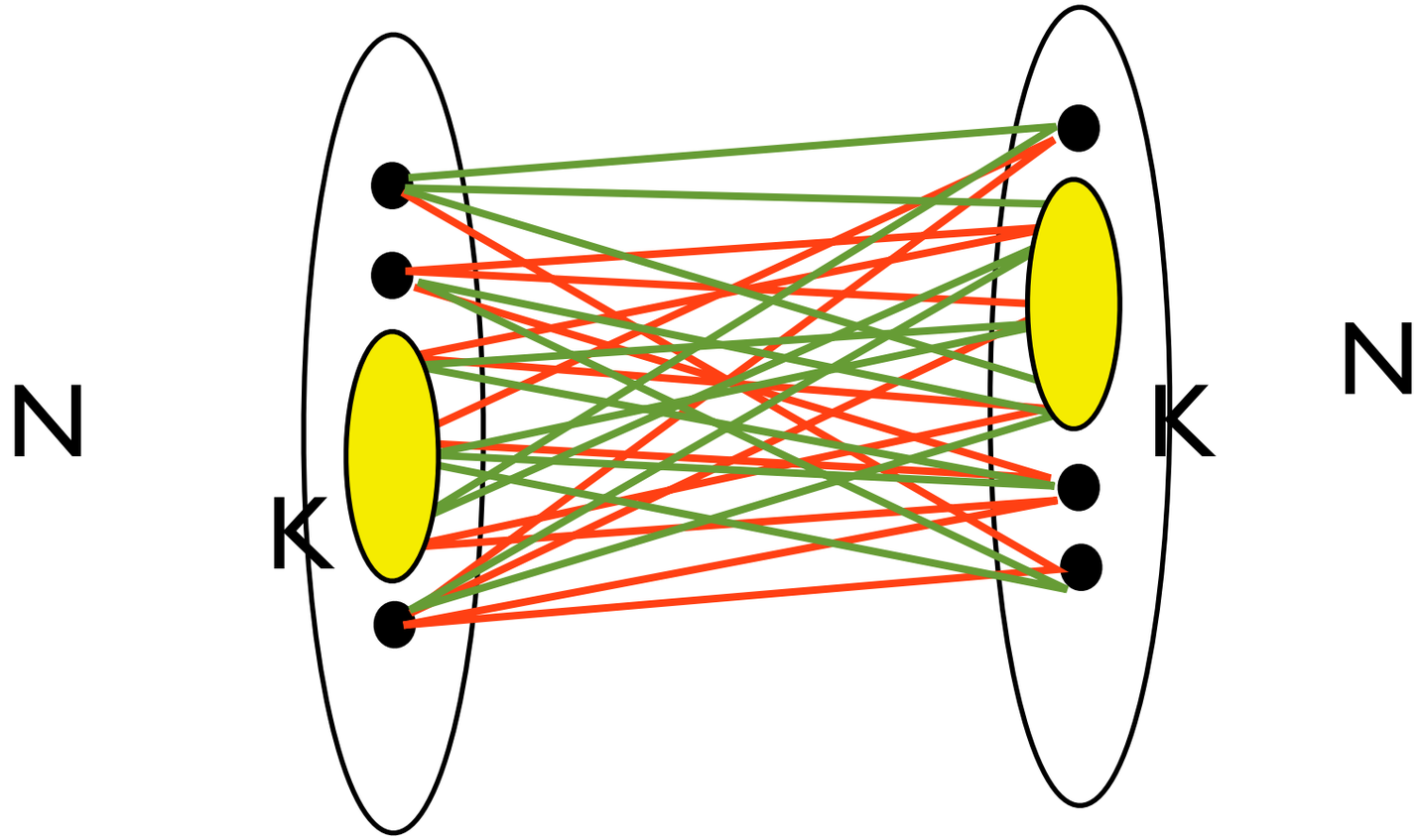
Bipartite K-Ramsey graph: Bipartite graph with no complete or empty K×K sub-graph.

Erdos (1947): Existence of K-Ramsey graphs on N vertices for K> (2+o(1)) log N.

Explicit Constructions?

Ramsey Graphs

$N=2^n$, $K=2^k$



Explicit 2-Source Extractors

Reference	k_1
Chor-Goldreich 88	>0.5n
Bourgain 05	≥0.499n
Recent advances	log n o(log log n)

Explicit Ramsey Graphs ($N=2^n$, $K=2^k$)

Reference	K	Bipartite
Erdös 47 (existential)	$\geq 2 \log N$	Yes
Hadamard Matrix	$\sqrt{N}$	Yes
Frankl-Wilson81	$2^{\sqrt{\log N}}$	No
Best 2-Source Extractors	$(\log N)^{o(\log \log \log N)}$	Yes

Tools and Techniques

Resilient Function

$f: \{0,1\}^n \rightarrow \{0,1\}$

(q,ε)-resilient function: For any subset of q co-ordinates, probability f is fixed on uniform sampling of the remaining co-ordinates is $\geq 1-\varepsilon$.

Example: MAJORITY is $(n^{0.49},\varepsilon)$-resilient.

PARITY is NOT (q, ε)-resilient for any $q>0$, $\varepsilon < 1$.

Resilient Functions

- [C-Zuckerman] Explicit $(n^{0.99},\varepsilon)$-resilient functions.
- [Meka] Explicit $(n/\log^2 n,\varepsilon)$-resilient functions.

Strong Seeded Extractors

$\mathrm{Ext}:\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$

X: (n,k)-source Y: U_d

$\mathrm{Ext}(X,Y), Y \approx_\varepsilon U_m, Y$

Explicit Construction: $d=O(\log(n/\varepsilon))$, $m=\Omega(k)$.

t-Non-Malleable Extractors

$nmExt:\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$,
non-malleable

$\exists\ S \subseteq \{0,1\}^d, \quad |S| > (1-\varepsilon)2^d,$
for any distinct seeds $s_1,\ldots,s_t$ in S

$$nmExt(X,s_1),\ldots,nmExt(X,s_t) \approx_\varepsilon U_t$$

Explicit: $k = t \log(n/\varepsilon) \log\log(n/\varepsilon)$,

$d = O(t^2 \log(n/\varepsilon) (\log\log(n/\varepsilon))$

A Preliminary Attempt

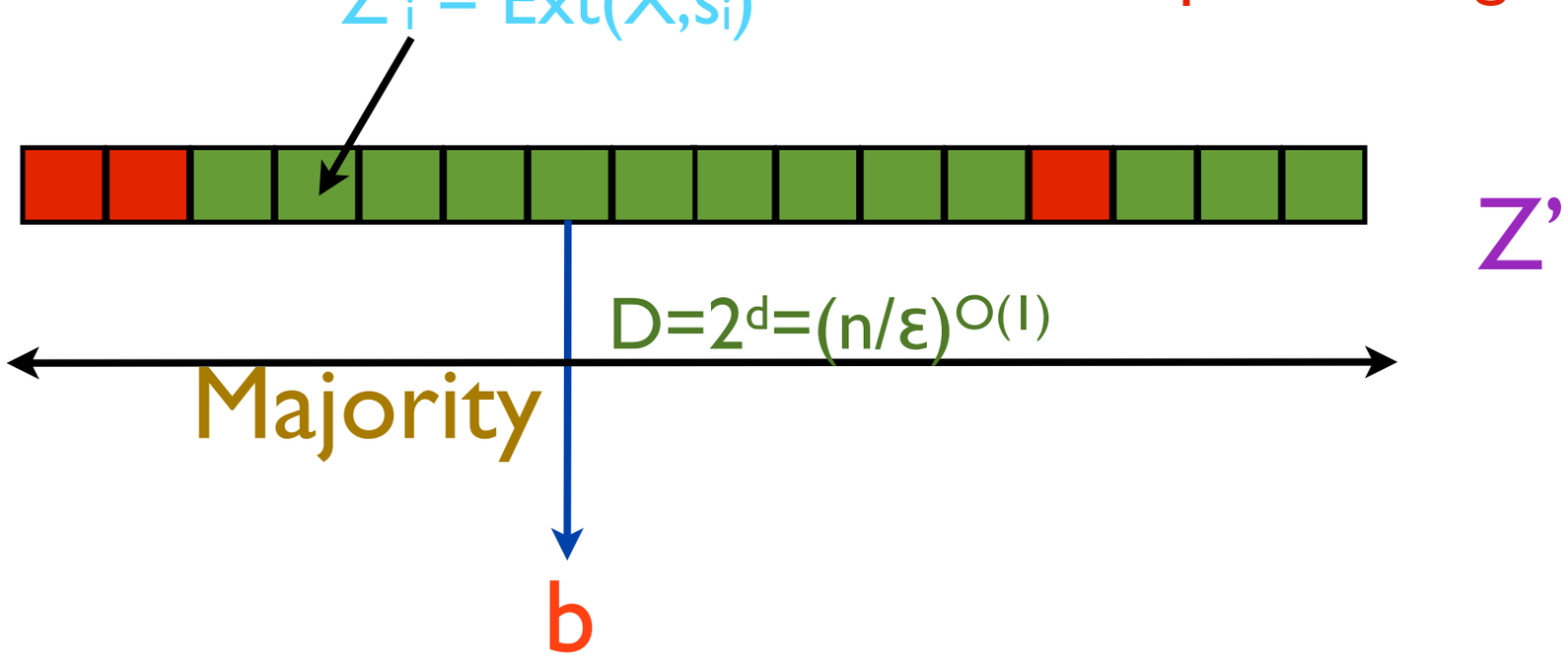
Ext:$\{0,1\}^n \times \{0,1\}^d \to \{0,1\}$: Strong-seeded extractor
Min-entropy k, error ε.

X: (n,k)-source

(1-ε) fraction of the bits in Z' are uniform

$Z'_i = Ext(X,s_i)$

Bad bit: Depends on good bits.



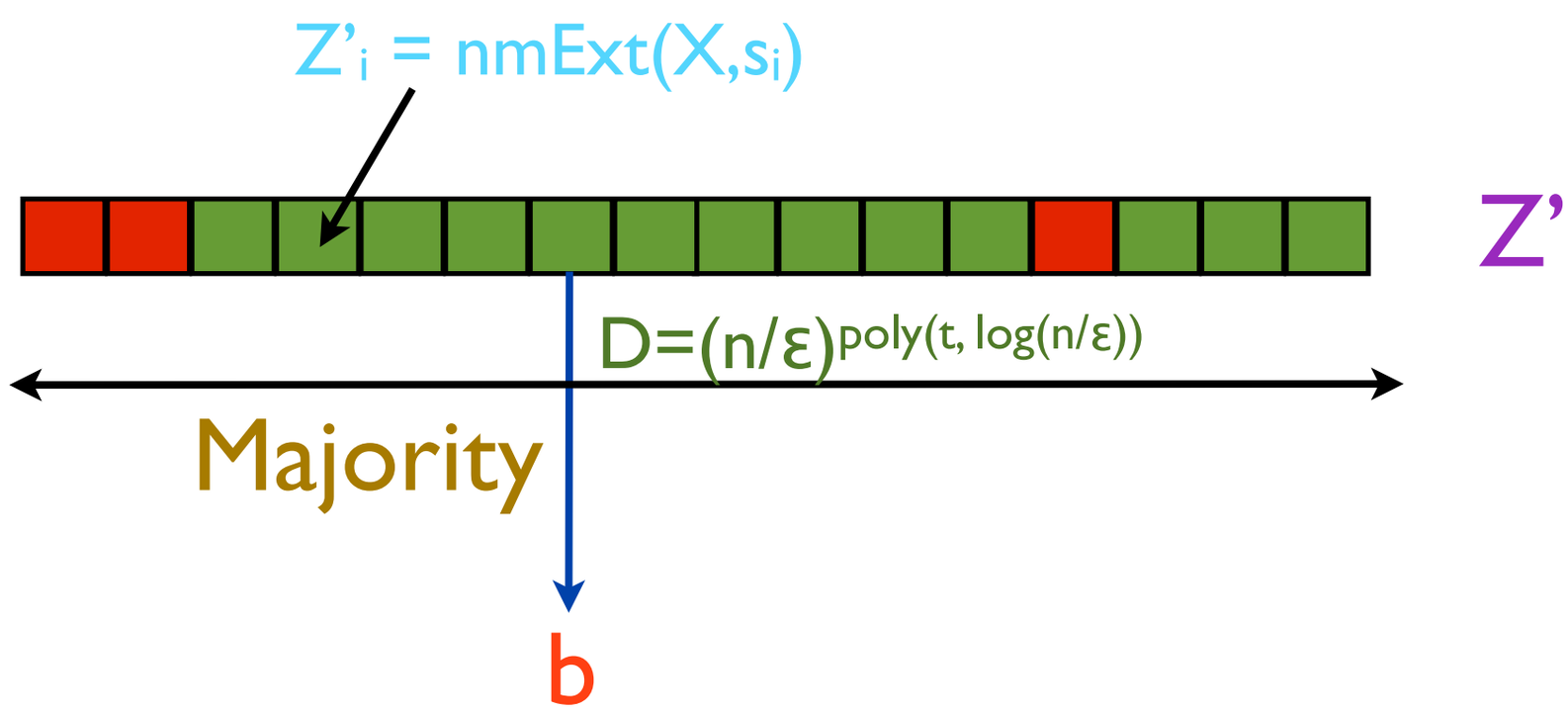
Does not work: The uniform bits are arbitrarily correlated

A Second Attempt

Idea: Make the uniform bits almost t-wise independent

Use a t-non-malleable extractor

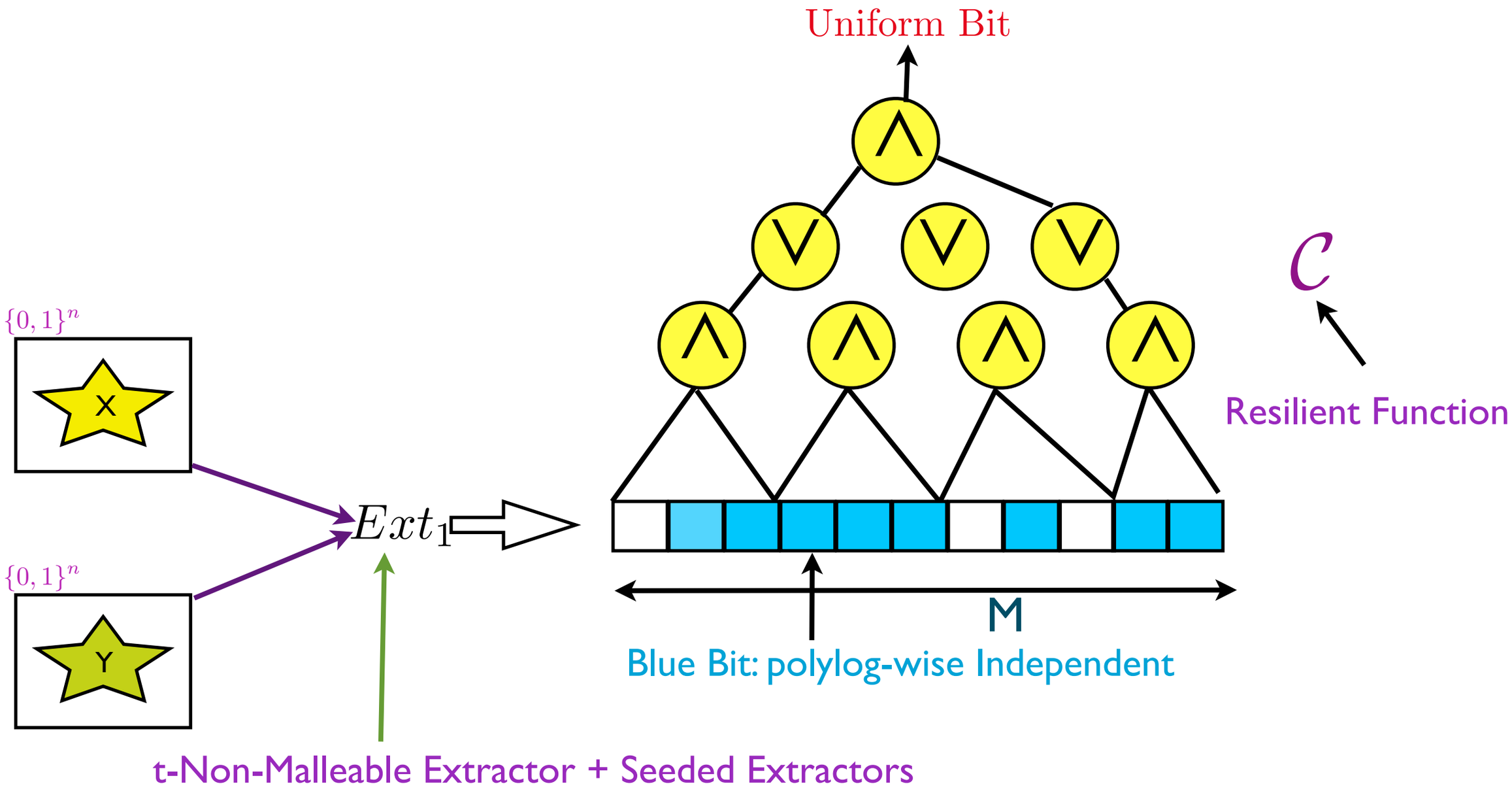
X: (n,k)-source



Does not work: $>D^{0.5}$ bad bits.

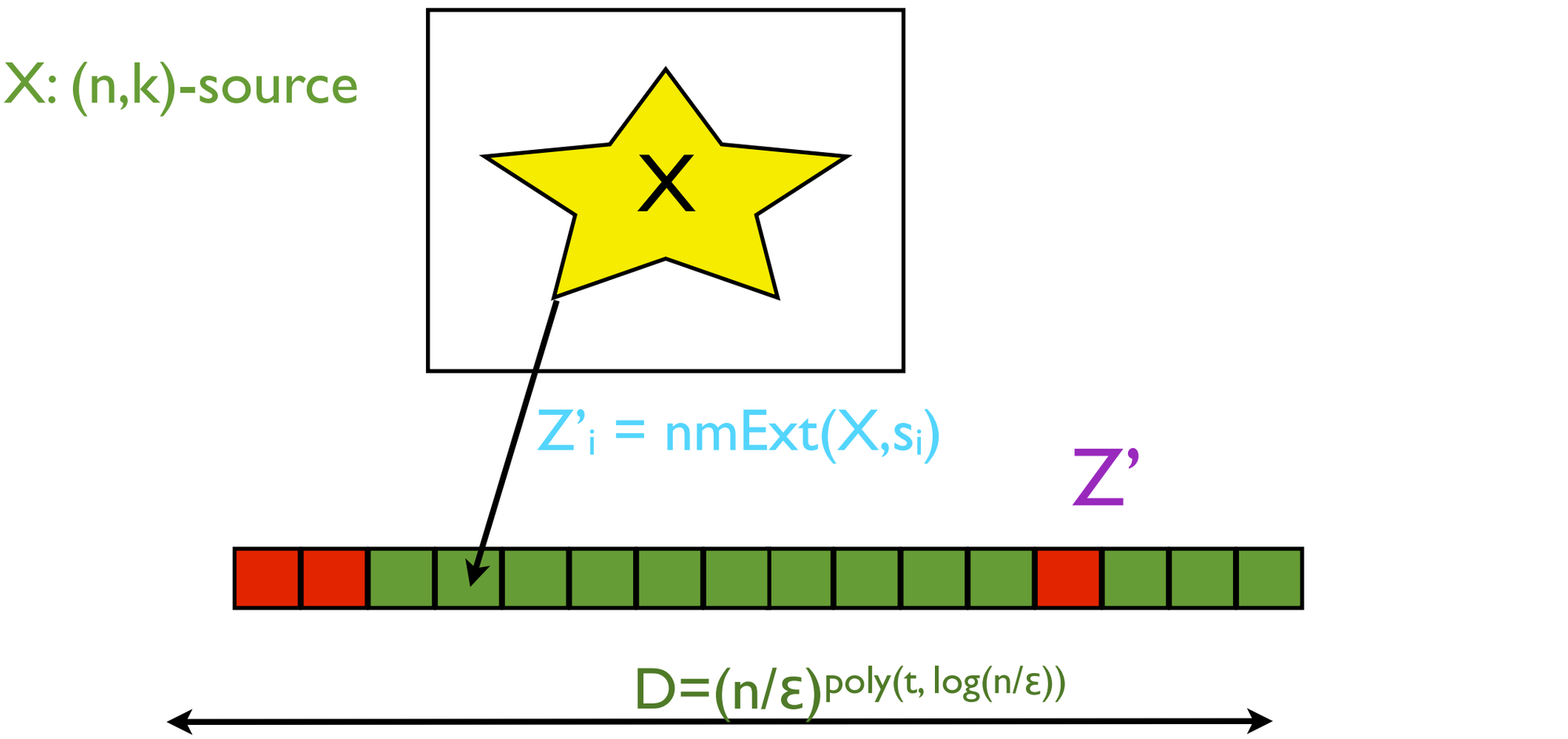
(not surprisingly! since we have 1 source)

A (very) High Level Idea of our Construction



Executing Step 1

X: (n,k)-source



The good bits of Z are almost t-wise independent

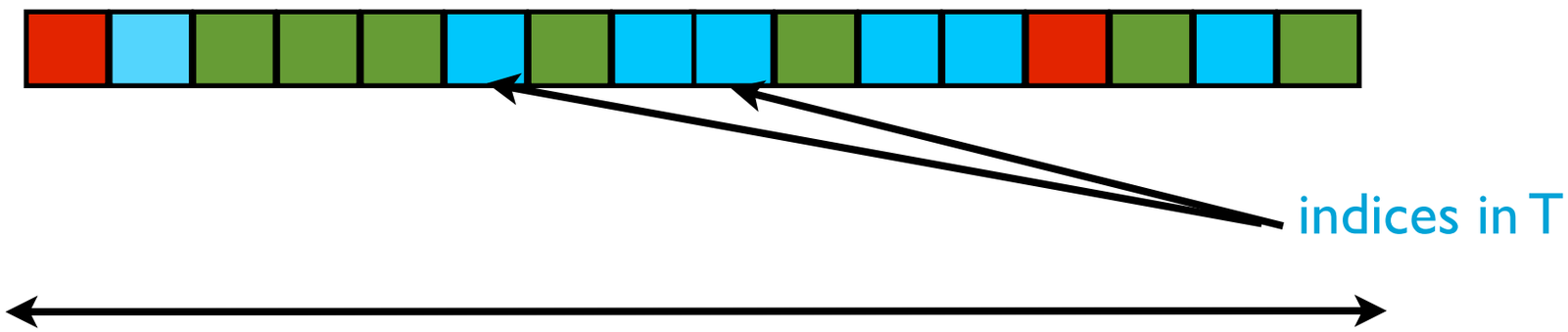
# of bad indices $\leq \varepsilon D$

Idea: Sample a pseudorandom subset T of [D] using Y.

Executing Step 1

$Z'_i = nmExt(X,s_i)$

Z'



$D=2^{poly(t, \log(n/\varepsilon))}$

$Z=Z'_T$

Pseudorandom Subset:

$T= \{Ext(Y,r_1),...,Ext(Y,r_M)\}$, $M = 2^{O(\log(n/\varepsilon'))}$

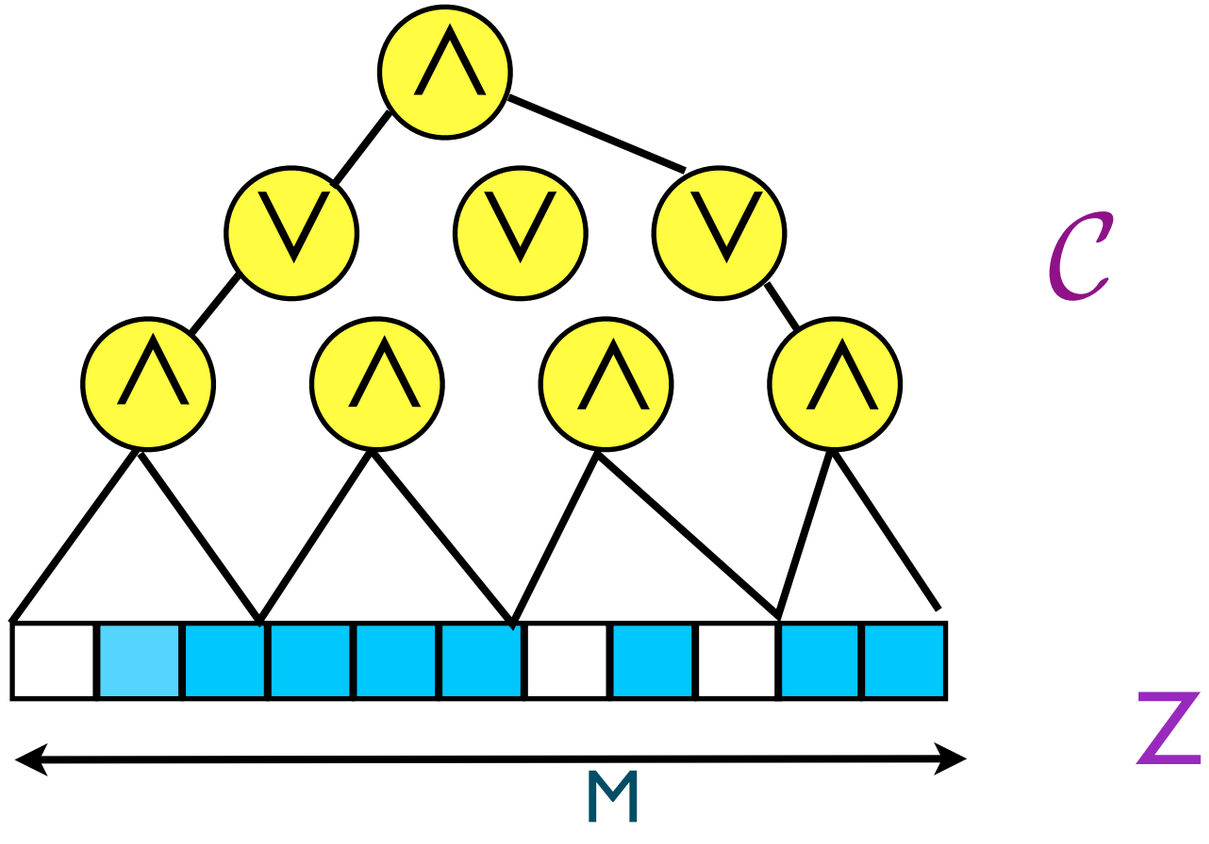
No. of bad indices in T: $(\varepsilon+\varepsilon')M<M^{1-\delta}$

Executing Step 2

No. of bad indices < $M^{1-\delta}$

The good bits of Z are (t=polylog(M))-wise independent.

$Pr_{x\sim\mathcal{D}}[\mathcal{C}(x) \text{ is not fixed}] =?$ $\mathcal{D}$

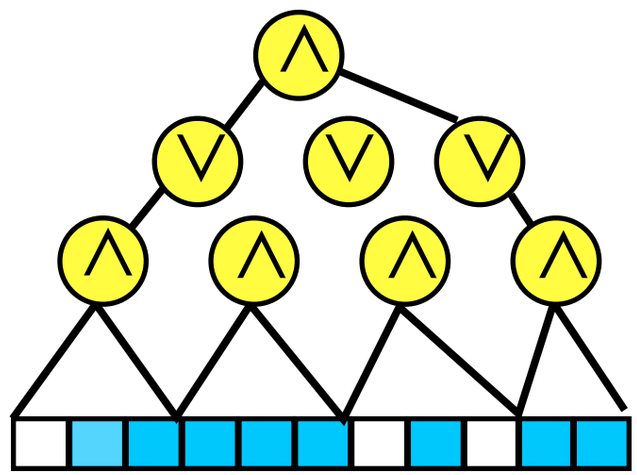


$\mathcal{C}$

Z

Executing Step 2

Step 2a: Good bits can be assumed to uniform, independent



$$\mathcal{C}'(x) := (\mathcal{C}(x \circ \vec{0}) \neq \mathcal{C}(x \circ \vec{1}))$$

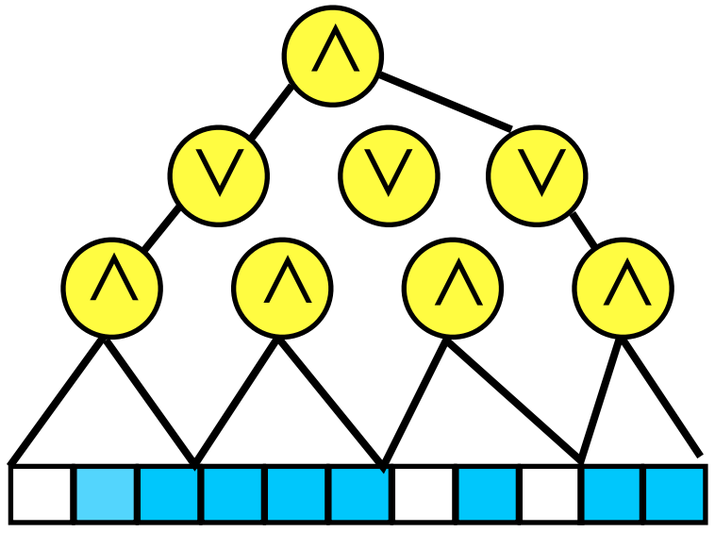
If $\mathcal{C}$ is monotone,

$$Pr_{x \sim \mathcal{D}}[\mathcal{C}(x) \text{is not fixed}] = Pr_{x \sim \mathcal{D}}[\mathcal{C}'(x) = 1]$$

(Braverman 07) If $\mathcal{C}$ is in AC^0,

$$|Pr_{x \sim \mathcal{D}}[C'(x) = 1] - Pr_{x \sim U}[C'(x) = 1]| \leq \epsilon$$

Executing Step 2



C

z

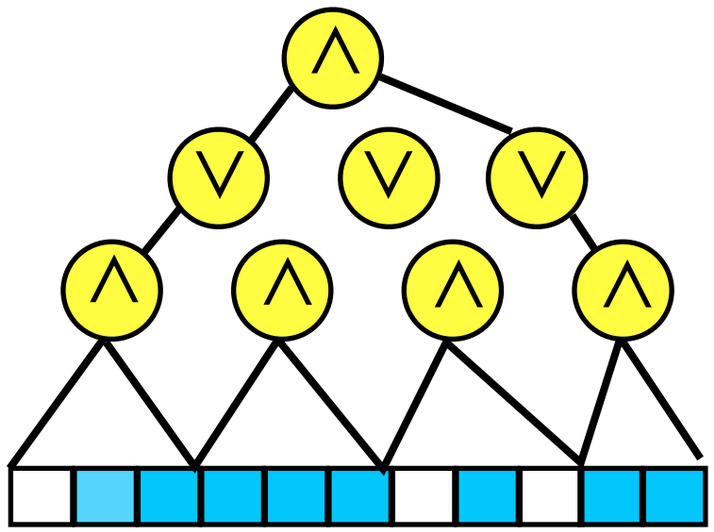
$$\mathcal{C}'(x) := (\mathcal{C}(x \circ \vec{0}) \neq \mathcal{C}(x \circ \vec{1}))$$

Thus, if $\mathcal{C}$ is in AC^0 and monotone,

$$|Pr_{x\sim\mathcal{D}}[C(x) \text{ is not fixed}] - Pr_{x\sim U}[C(x) \text{ is not fixed}]| \leq \epsilon$$

Thus, we can assume good bits are independent, uniform.

Executing Step 2



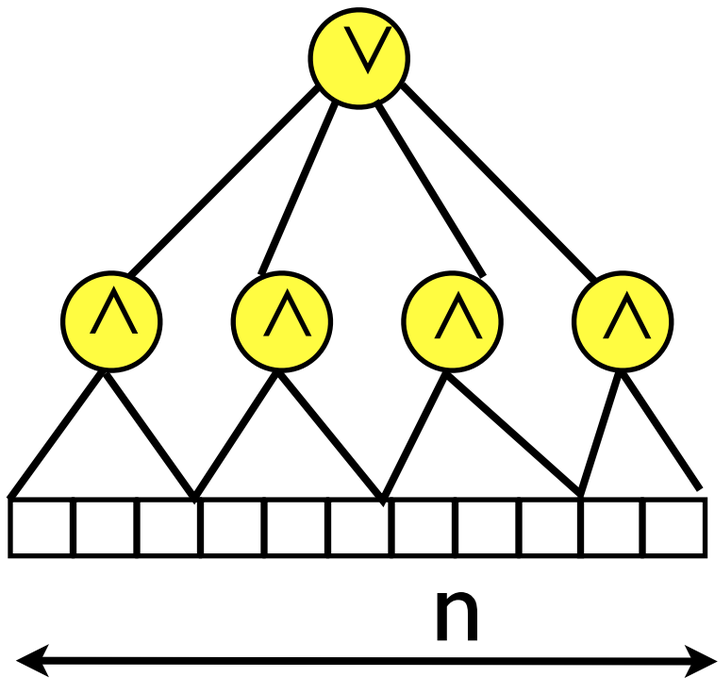
C

Remaining Task:

Explicit construction of a monotone C in AC^0 on n bits s.t:
(1) C is $(n^{1-\delta},\varepsilon)$-resilient and
(2) almost unbiased.

Ajtai-Linial Function

Tribes function:



AL Function: $T_1 \wedge T_2 ..... \wedge Tn$

T_i: randomly-negated Tribes on randomly chosen partitions.

Resilient to coalitions of size $O(n/\log^2 n)$.

Problems: (1) Probabilistic: Naive derandomization takes time $n^{O(n^2)}$.

(2) Not Monotone.

Derandomizing Ajtai-Linial

Key Ingredient: An explicit construction of a collection of partitions of [n] s.t:

(1) Any small subset of [n] has small intersection with most partitions.

Used to bound influence

(2) The partitions are pairwise pseudorandom:

- the intersection of any two blocks is bounded.

Used to bound bias

Derandomized via extractors!

Open Questions

- Negligible error: All constructions known achieve error $1/n^{\Omega(1)}$; not enough for cryptographic applications.
- Extractors for other sources: Low degree varieties, Circuit sources,…
  - Applications to circuit lower bounds
- Optimal Ramsey graphs
- Other applications of Non-malleable extractors
- Simpler constructions

Thank You!