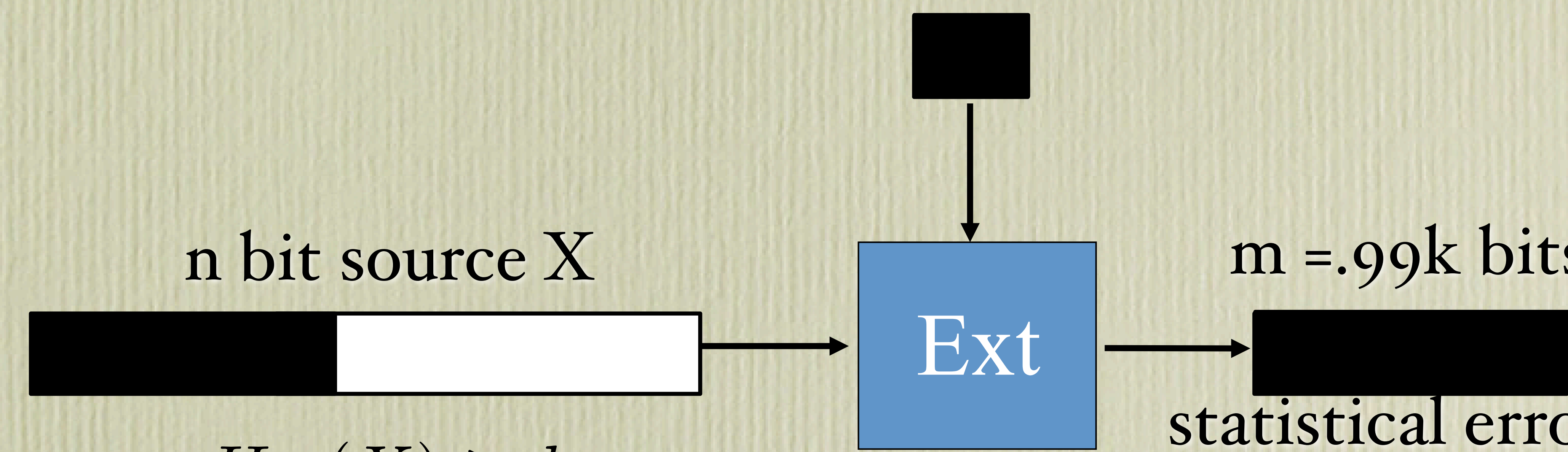
Non-Malleable Extractors

Xin Li
Johns Hopkins University

Seeded Extractor

[Nisan-Zuckerman '93,…, Guruswami-Umans-Vadhan '07, DW'08, DKSS'09]

d=O(log (n/ε)) uniform bit seed Y



$H_\infty(X) \geq k$

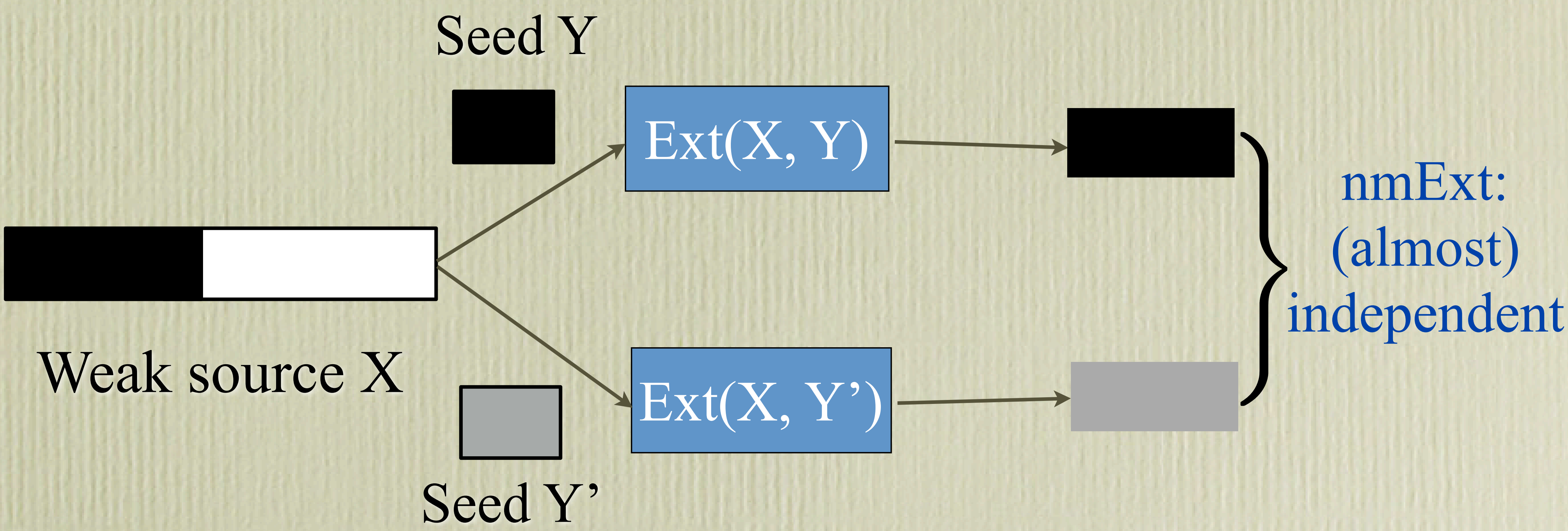
Strong extractor: (Ext(X,Y),Y) (m+d) bits

Seed=Catalyst

Non-Malleable Extractor

[Dodis-Wichs 2009]

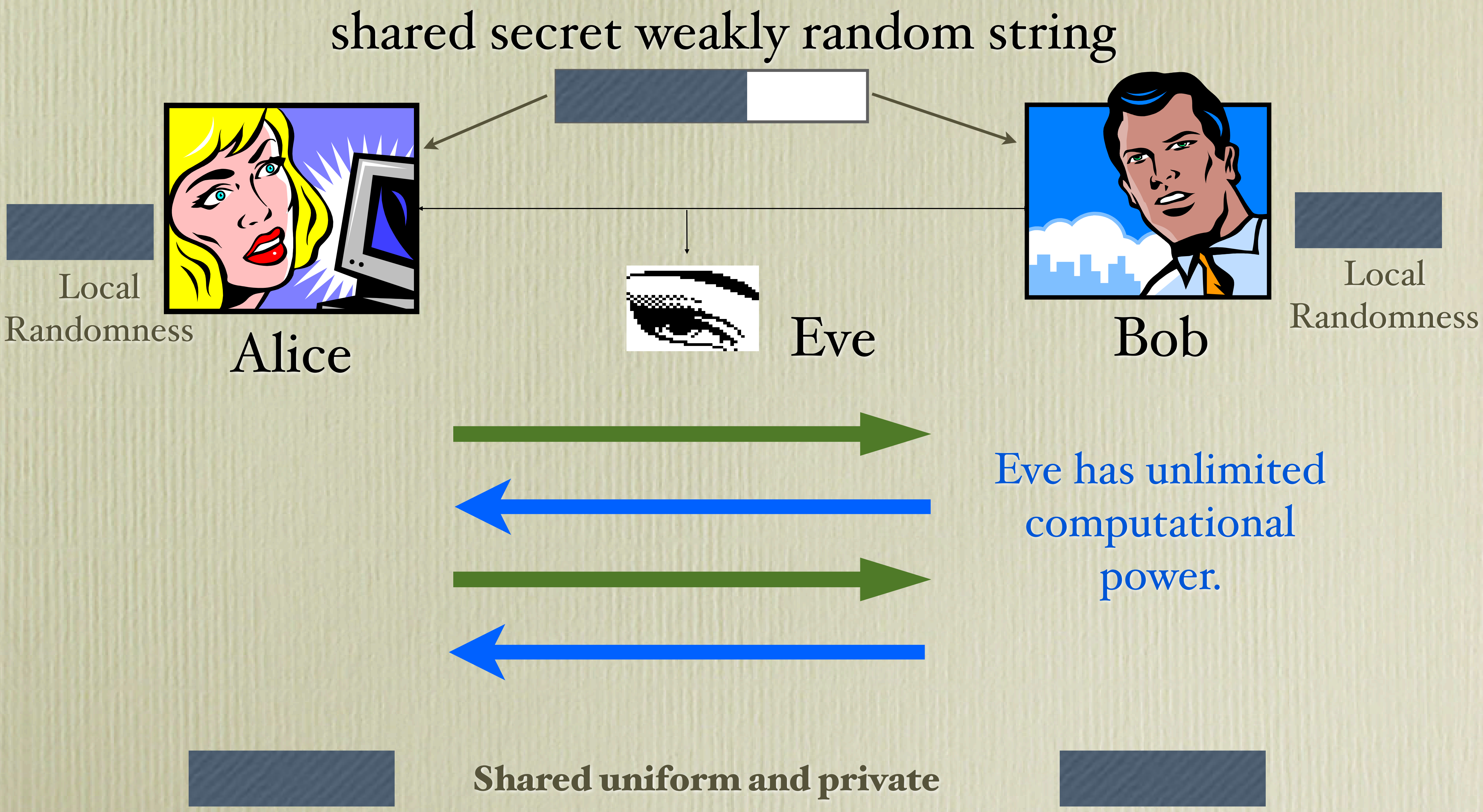
An adversary changes the seed Y to Y'≠Y.



How correlated are the two outputs?

Privacy Amplification

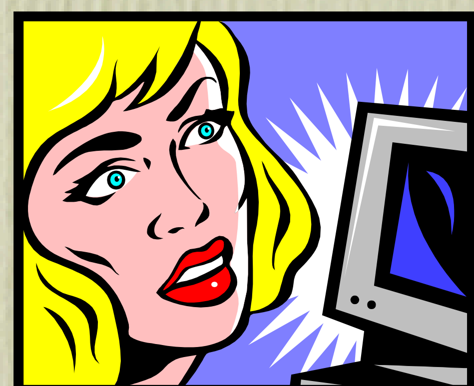
[Bennett, Brassard, Robert 1985]



Privacy Amplification with Passive Adversary

[Bennett, Brassard, Robert 1985]

shared X



Alice

Eve

Bob

Pick random Y → Y

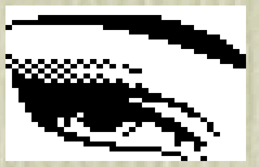
Ext(X,Y) Ext(X,Y)

Shared uniform and private

Seeded Extractor Fails for Active Adversary

shared X

Alice



Eve

Bob

Pick random Y → Y'

Ext(X,Y)

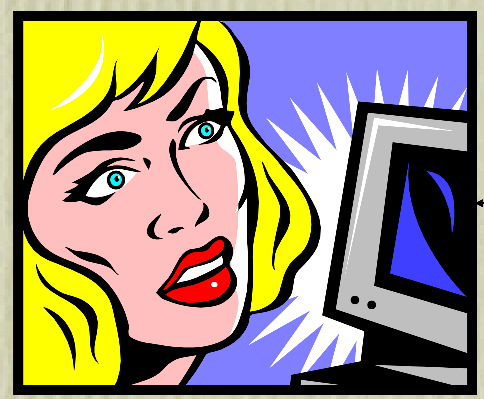
May be ≠

Ext(X,Y')

Active adversary: can arbitrarily insert, delete, reorder messages

Privacy Amplification with nmExt
[Dodis-Wichs'09]

(n,k)-weak source



Alice

Eve

Bob

Y

$R = \mathsf{nmExt}(X, Y)$ → Y' $R' = \mathsf{nmExt}(X, Y')$

$W', T' = MAC_{R'}(W')$

W, T ←

Reject if
$MAC_R(W) \neq T$

$Z = \mathsf{Ext}(X, W)$

$Z' = \mathsf{Ext}(X, W')$

Non-Malleable Extractor
[Dodis-Wichs 2009]

- No one-round protocol if $k<n/2$, and optimal 2-round protocols follow from non-malleable extractors.
- If Eve is passive, then the protocol succeeds.
- If Eve is active, then the protocol detects the tampering and aborts w.h.p.
- Another important application: independent source (e.g., two-source) extractors.

Error correcting codes

message Y → Y'

codeword M=Enc(Y) → M’

Y=Dec(M’)

Error correcting codes

- However, the type of error one can correct is limited—symbol erasure/modification.
- How to handle more general error?
- Error detection — however, cannot even detect a function that changes all codewords into a fixed string.

Non-Malleable (NM) Codes [Dziembowski, Pietrzak and Wichs 2010]

- Fix a family of tampering functions F on $\{0,1\}^n$.
- Non-malleable code C on $\{0,1\}^n$ against F consists of:
  - Randomized encoder: Enc: $\{0,1\}^m \to \{0,1\}^n$
  - Deterministic decoder: Dec: $\{0,1\}^n \to \{0,1\}^m$
    - 1) For all s, Dec(Enc(s)) = s.
    - 2) For any f in F, either Dec(f(Enc(s))) = s, or is a probability distribution independent of s.

rate of the code: m/n

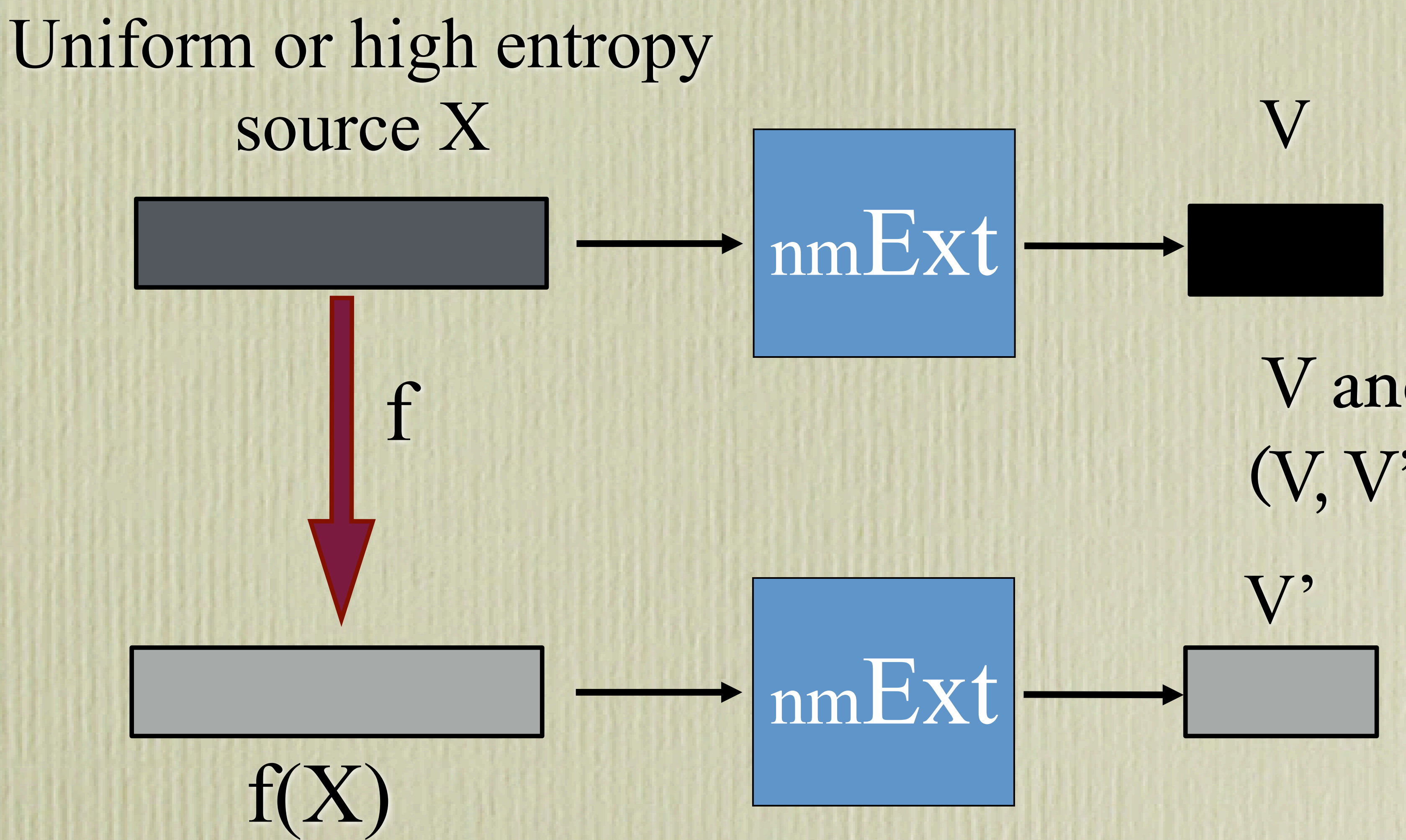
Existential Result
[Cheraghchi-Guruswami’14a]

- If the size of the class of tampering functions is limited: $|\mathcal{F}| \leq 2^{2^{\alpha n}}$

- There exists non-malleable codes against F with rate close to 1-α with exponentially small error.

- Explicit constructions known for: split-state tampering, NC0, AC0, affine functions…

Connections to nm Extractors [Cheraghchi-Guruswami'14b]

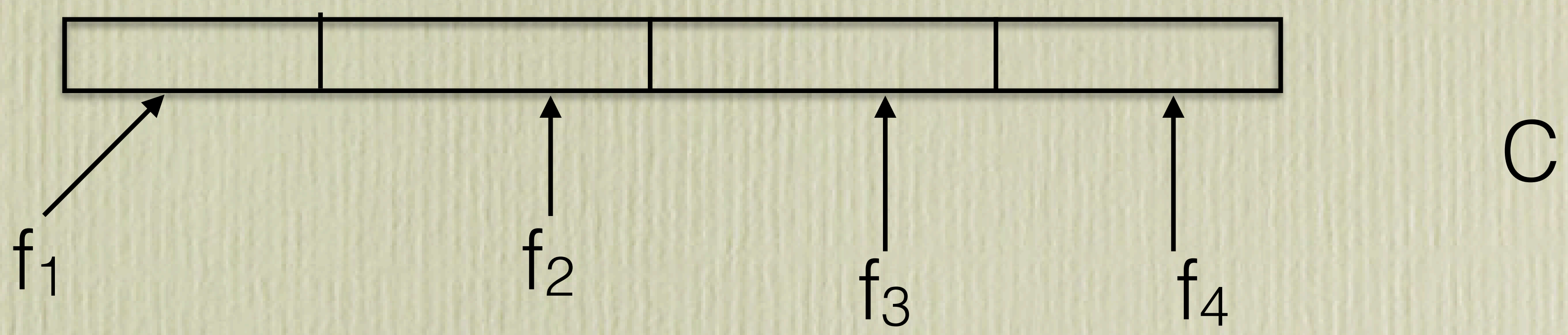


V and V' each has m bits.
(V, V') is ε-close to (U, V').

This gives a non-malleable code against f with rate m/n and error $2^m\varepsilon$.

Encoding: uniformly sample the pre-image of V.
Decoding: compute the output of the extractor.

The split state model



- Non explicit: non-malleable codes exist in the 2-split state with constant rate and exponentially small error.
- 2-split state model corresponds to a non-malleable two-source extractor.

Constructions of Seeded nm Extractors

- Non explicit: $k=O(m+\log d+\log(1/\varepsilon))$, $d=O(\log (n/\varepsilon))$.

- Lower bound on k: $k=\Omega(\log \log n)$ [GS'17].

- Best constructions: either k or d can be optimal, the other has a $\log^{1+o(1)}(1/\varepsilon)$ dependence on ε, or both have $\log (1/\varepsilon)\log \log(1/\varepsilon)$ dependence on ε [L'17, L'18].

Constructions of nm codes in the split state model

- 2-split state model: [DKO’13, ADL’14, ADKO’15, CGL’16, L’17] give codes with rate 1/log n and exponentially small error.
- 3&4-split state model: [KOS’17, GMW’18] constant rate with negligible error.
- 10-split state model: [CZ’14] gives codes with constant rate and exponentially small error.
- 2-split state model: [L’18] gives codes with constant rate and arbitrarily small constant error.

Constructions of nm Extractors

- Early constructions use character sums [DLWZ11], small biased sample space [CRS12], and inner product [L'12].

- Only work for entropy rate at least 1/2 (or slightly below).

A Simple Construction of nmExt for k>n/2 [L'12]

- $Ext(x,y) = \langle x,y \rangle$ over F_2.
- Two-source extractor for (n, k_1) and (n, k_2) sources with $k_1+k_2 > n$.
- Let X be an $(n, k>n/2)$ source.
- Let Y be a uniform random seed with n/2 bits.
- View Y as an element in F_{2^n} and let $Enc(Y)=(Y, Y^3)$.
- $nmExt(x, y)=\langle x, Enc(y) \rangle$ over F_2.

Analysis

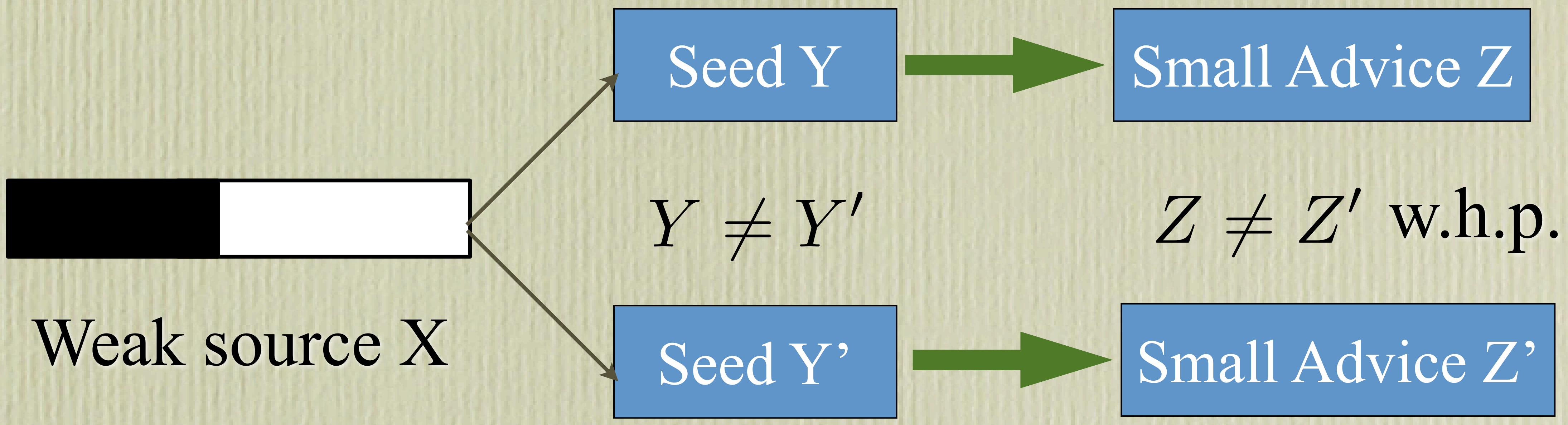
- Enc(Y)=(Y, Y^3) is injective =>Enc(Y) has entropy n/2=> nmExt(X, Y) is close to uniform.

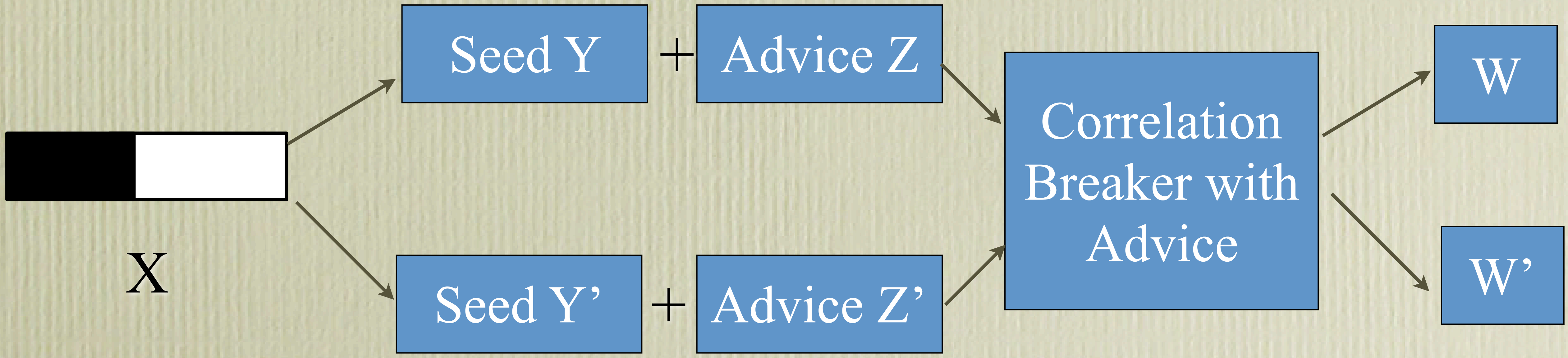
- Enc(Y)=(Y, Y^3) is 4-wise linearly independent over F_2 =>Enc(Y)+Enc(f(Y)) has entropy at least n/2-1.

- nmExt(X, Y) $\oplus$ nmExt(X, f(Y)) is close to uniform.

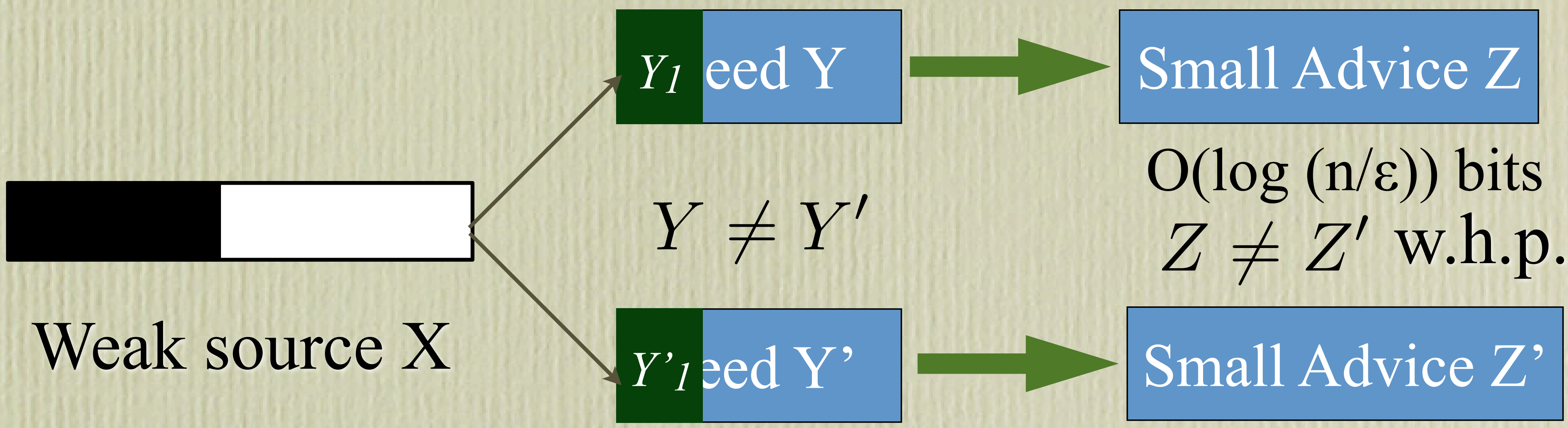
- Recently shown to be the first quantum-proof nm extractor [ACLV'17].

More Recent Constructions
[CGL'16, Cohen'17, L'17, L'18]





Advice Generation [CGL'16]



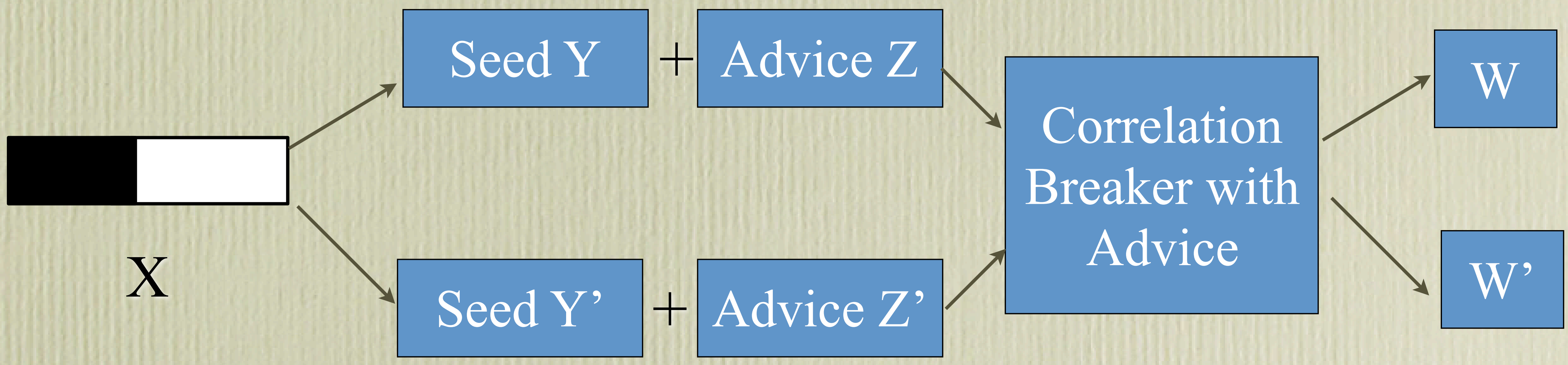
Take a small slice Y_1 of Y, and Y'_1 of Y'

Compute $V=Ext(X, Y_1)$ and $Z=(Sample(Enc(Y), V), Y_1)$

If $Y_1 \neq Y_1'$, done.

Otherwise V=V', Enc(Y) and Enc(Y')
has a large distance, so $Z \neq Z'$ w.h.p.

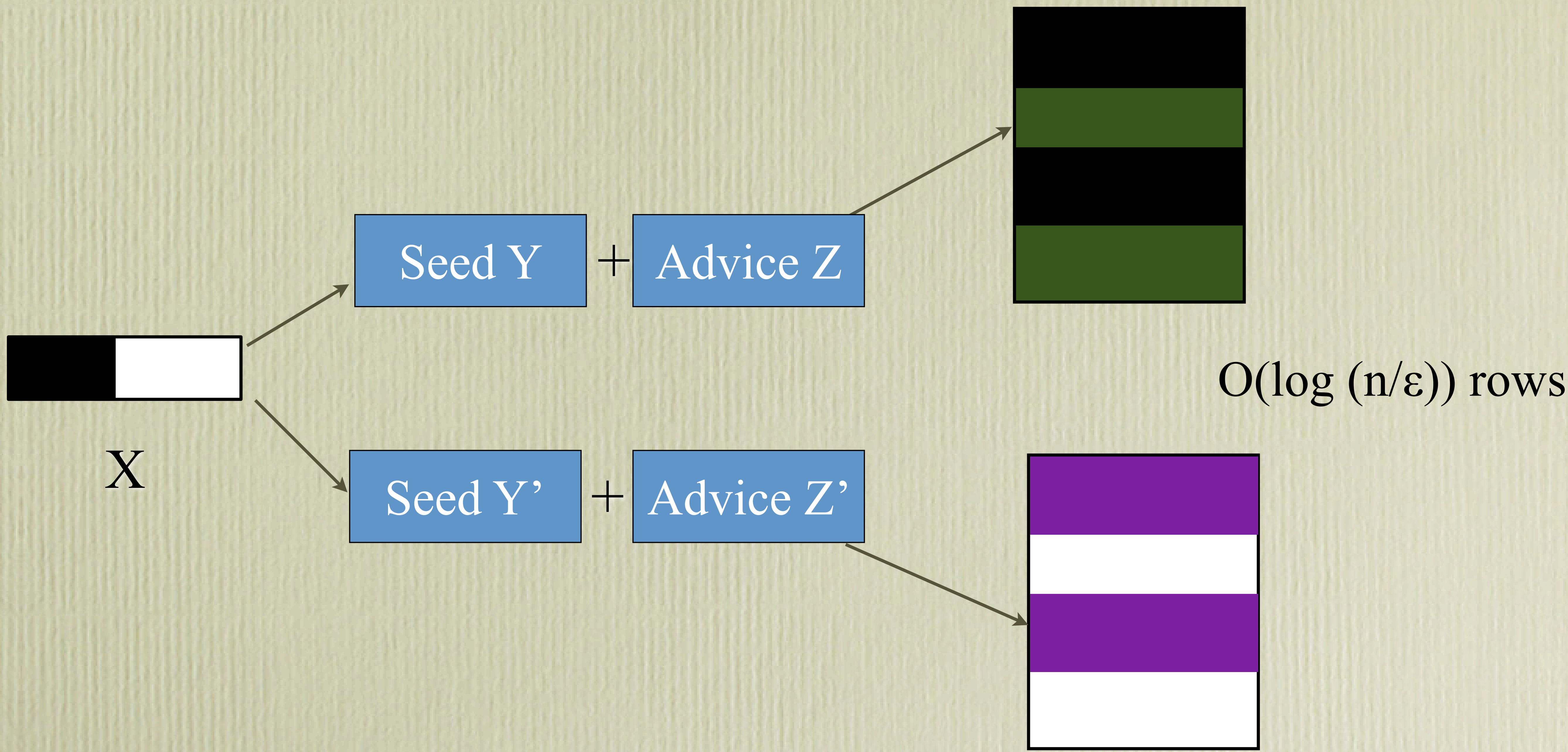
Correlation Breaker with Advice



Many Constructions of Correlation Breakers

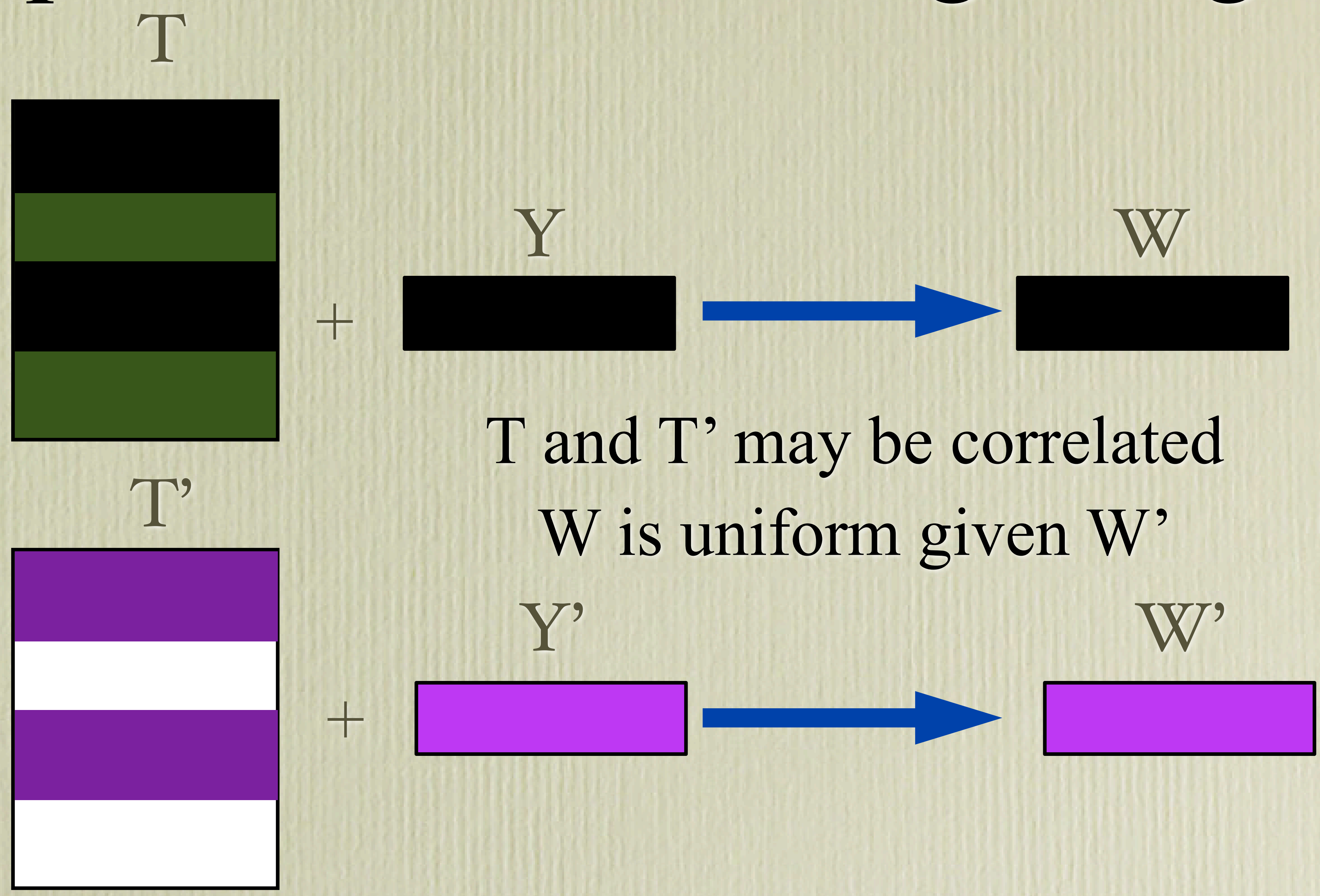
The most efficient one uses independence preserving mergers.

Correlation Breaker: First Step



Use each bit of Z (Z’) to do a flip-flop extraction

Independence Preserving Merger



Every row of T is uniform, and

∃ i s.t. T_i is uniform given T'_i (by flip-flop extraction)