Minentropy and its Variations for Cryptography

Leonid Reyzin

BOSTON UNIVERSITY

May 23, 2011

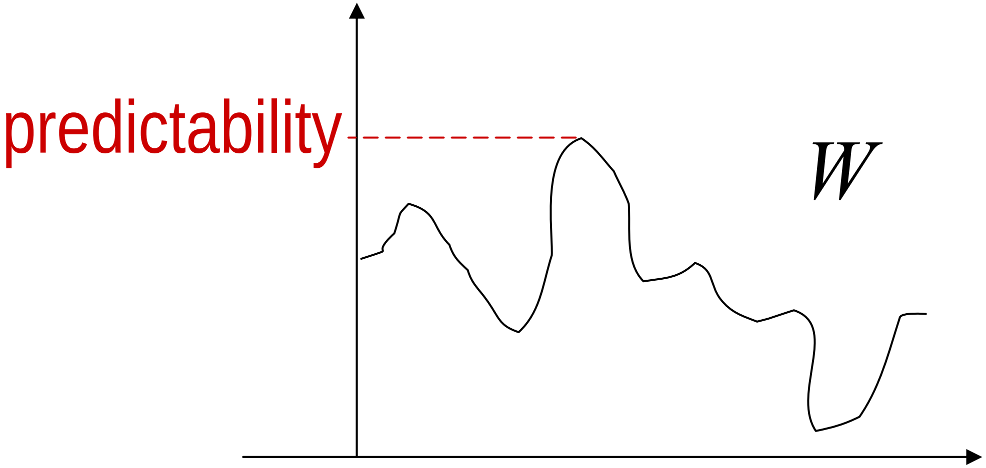
5th International Conference on Information Theoretic Security

guessability and entropy

- Many ways to measure entropy
- If I want to guess your password,

  which entropy do I care about?
- This talk:

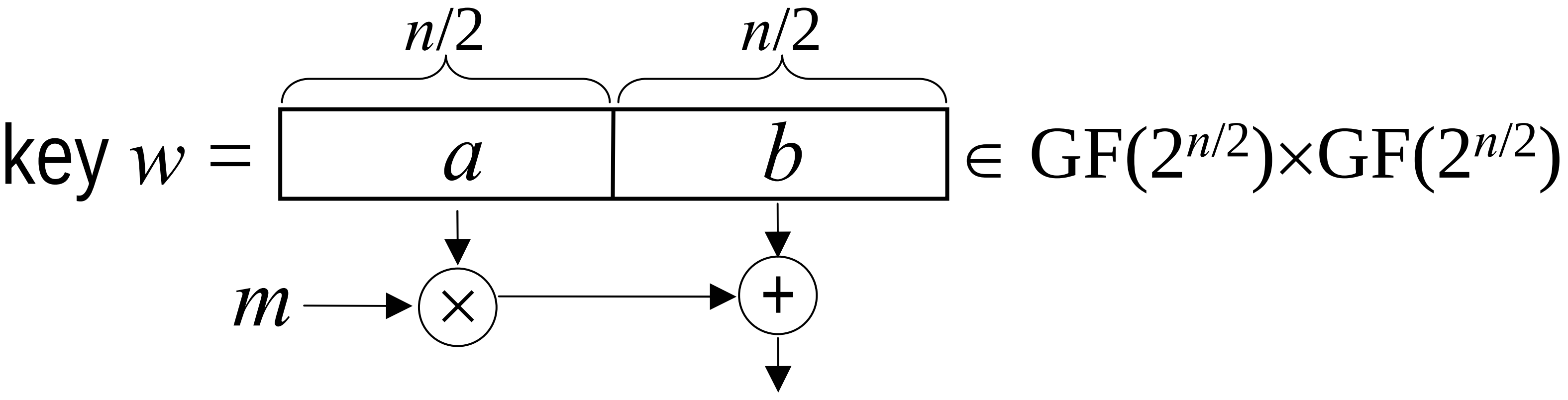
  minentropy = – log (Pr [adversary predicts sample])



$$H_\infty(W) = -\log \max_w \Pr[w]$$

what is minentropy good for?

- Passwords
- Message authentication

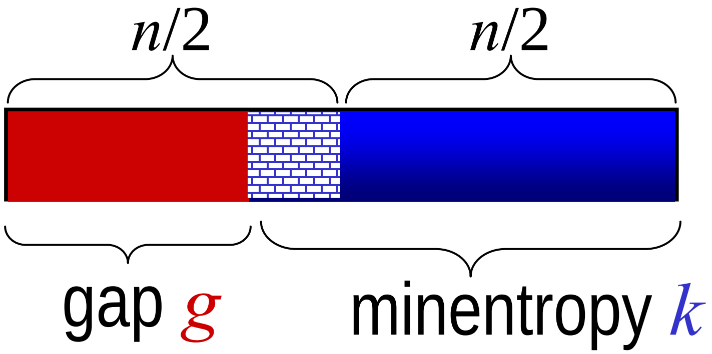


$$\mathrm{MAC}_{a,b}(m) = \sigma = am + b$$

[Wegman-Carter ‘81]

what is minentropy good for?

- Passwords
- Message authentication



key w =

$n/2$ $n/2$

gap g minentropy k

$$\mathrm{MAC}_{a,b}(m) = \sigma = am + b$$

Let $|a,b| = n,\ H_\infty(a,b) = k$

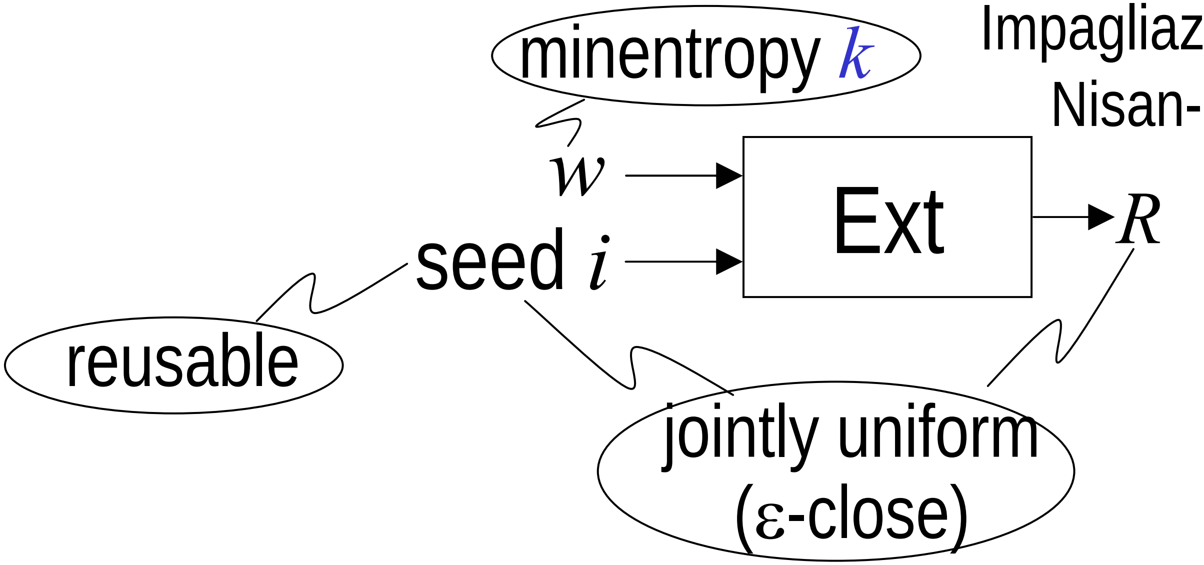
Let “entropy gap” $n - k = g$

Security: $k - n/2 = n/2 - g$ [Maurer-Wolf ’03]

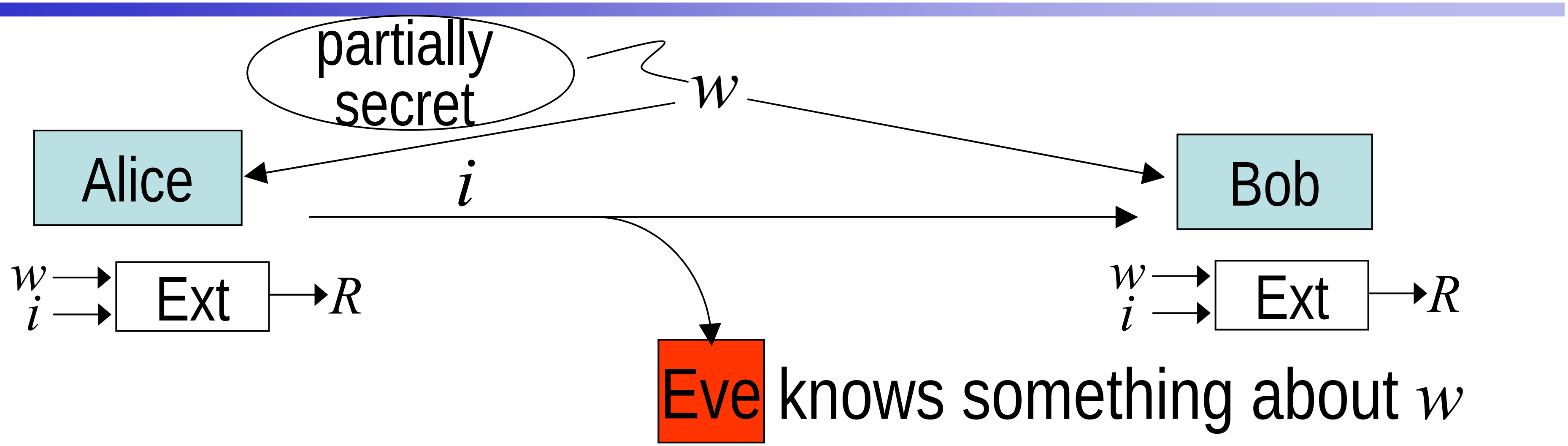
what is minentropy good for?

- Passwords
- Message authentication $\mathrm{MAC}_{a,b}(m) = \sigma = am + b$
- Secret key extraction (⇒ encryption, etc.)

[Bennett-Brassard-Robert ’85,
Impagliazzo-Levin-Luby ’89,
Nisan-Zuckerman ’93]



is it good for privacy amplification?



knows something about w

Goal: from a partial secret w

agree on a uniform secret R [Bennett-Brassard-Robert ’85]

Simple solution: use an extractor

But wait! What is the right value for $H_\infty(w)$?

Depends on Eve’s knowledge Y

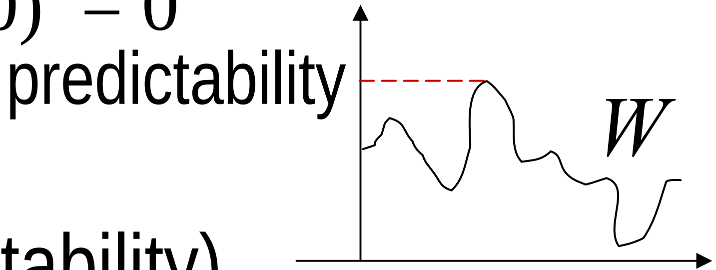
So how do we know what Ext to apply?

defining conditional entropy $H_\infty(W \mid Y)$

- E.g., W is uniform, Y = Hamming Weight(W)

$$\Pr[Y = n/2] > 1/(2\sqrt{n}) \;\Rightarrow\; H_\infty(W \mid Y = n/2) \geq n - ½ \log n - 1$$

$$\Pr[Y = n] = 2^{-n} \;\Rightarrow\; H_\infty(W \mid Y = 0) = 0$$



- But what about $H_\infty(W \mid Y)$?

- Recall: minentropy = – log (predictability)

$$H_\infty(W) = -\log \max_{w} \Pr[w]$$

- What's the probability of predicting W given Y?

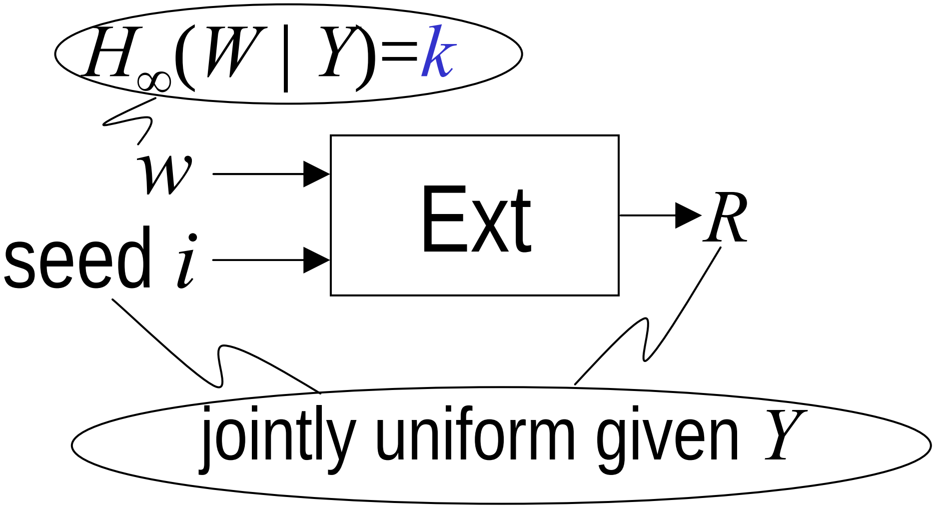
$$H_\infty(W \mid Y) = -\log \mathop{\mathrm{E}}_{y} \max_{w} \Pr[w|Y{=}y]$$

[Dodis-Ostrovsky
-R-Smith '04]

"average minentropy" but not average of minentropy:
if min-entropy is 0 half the time, and 1000 half the time,
you get $\log\, (2^0+2^{-1000})/2 \approx -\log 1/2 = 1$.

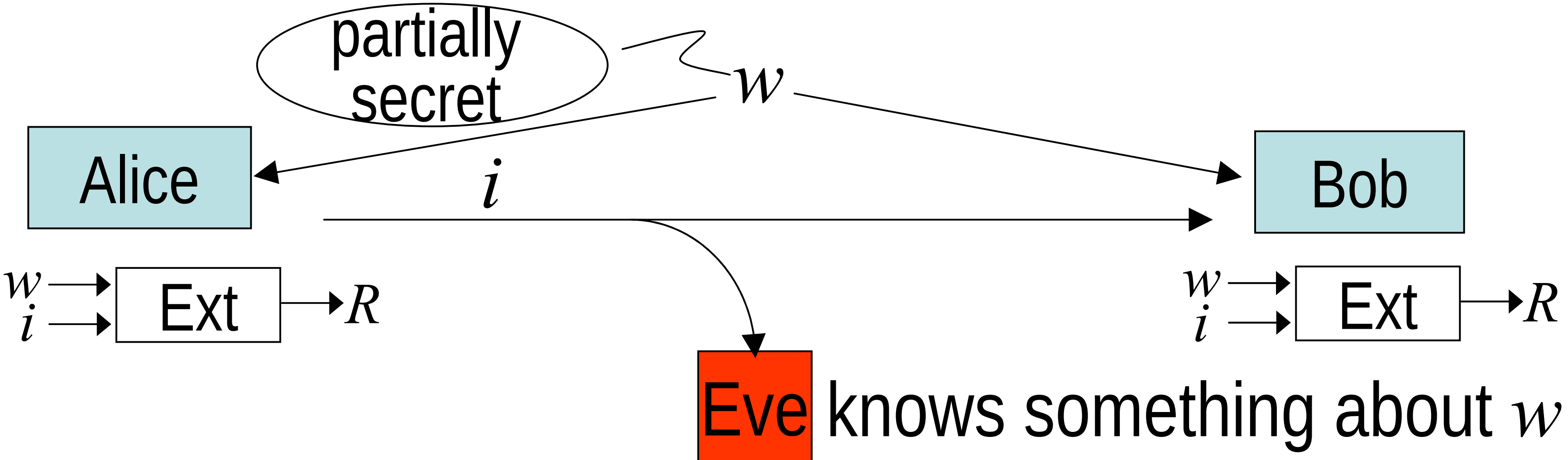
what is $H_\infty(W \mid Y)$ good for?

- Passwords
  - Prob. of guessing by adversary who knows Y: $2^{-H_\infty(W \mid Y)}$
- Message authentication
  - If key is W and adversary knows Y: security $H_\infty(W \mid Y) - n/2$
- Secret key extraction (⇒ encryption, etc.)
  - All extractors work [Vadhan '11]



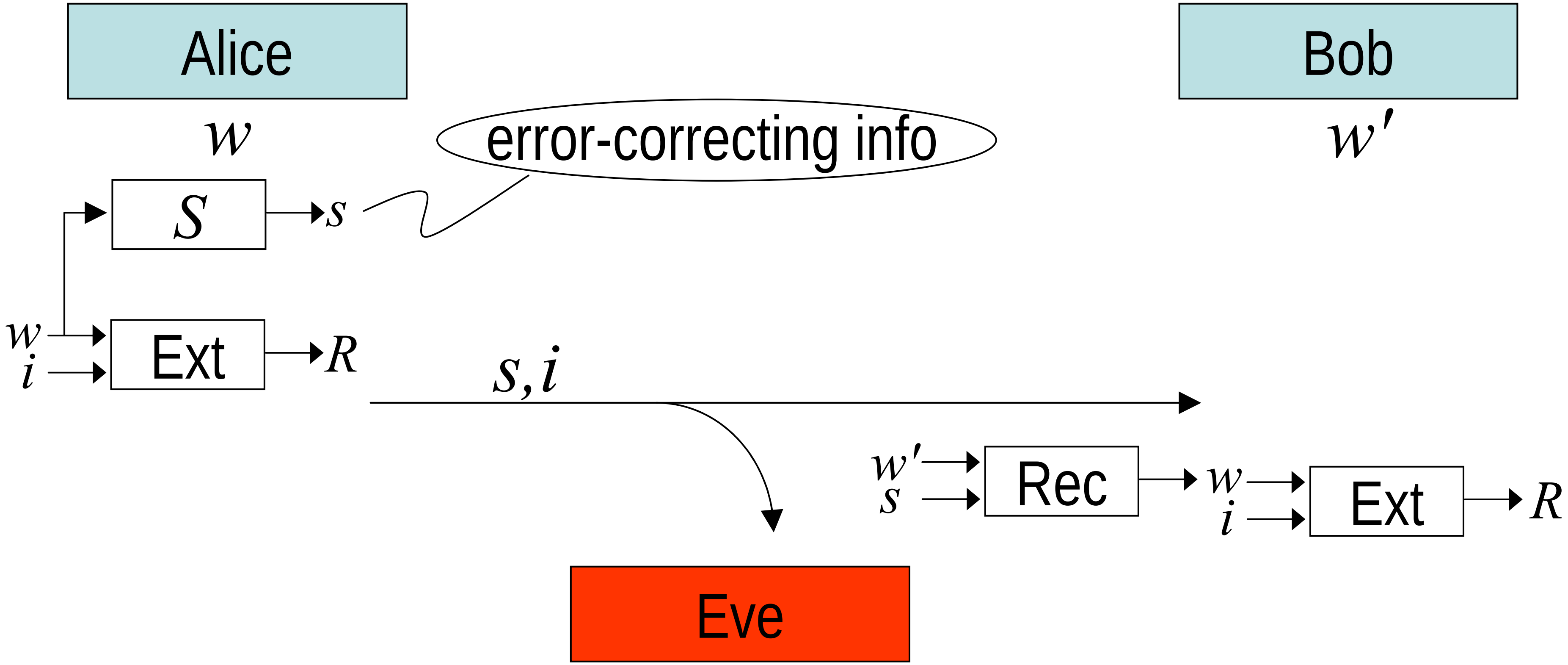
what is $H_\infty(W \mid Y)$ good for?

- Passwords
  - Prob. of guessing by adversary who knows Y: $2^{-H_\infty(W \mid Y)}$
- Message authentication
  - If key is W and adversary knows Y: security $H_\infty(W \mid Y) - n/2$
- Secret key extraction (⇒ encryption, etc.)
  - All extractors work [Vadhan '11]
  - Therefore, privacy amplification!



Eve knows something about w

what about information reconciliation?

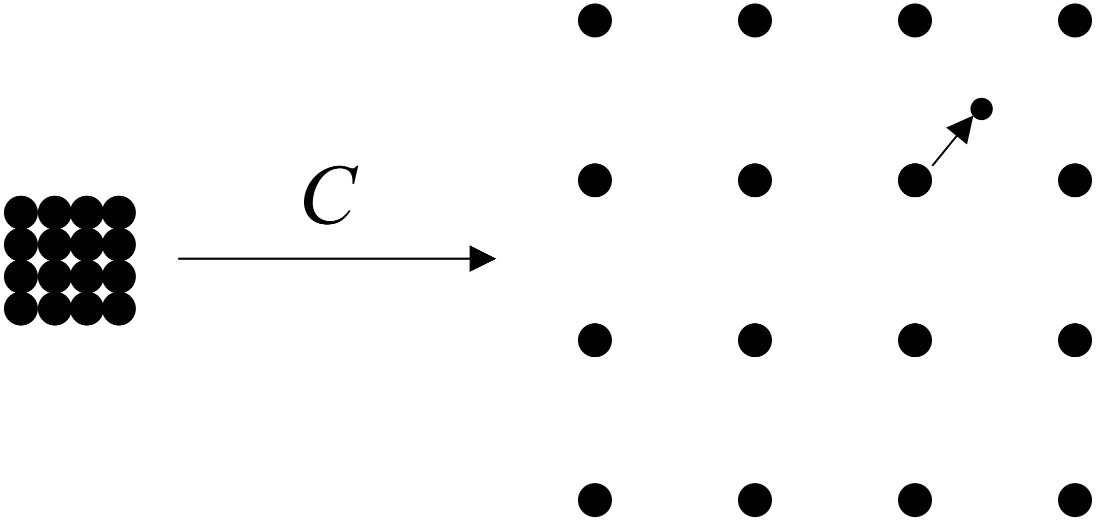


- How long an R can you extract?
- Depends on $H_\infty(W \mid Y, S)$!
- Lemma: $H_\infty(W \mid Y, S) \geq H_\infty(W, S \mid Y) - \text{bit-length}\,(S)$

how to build S?

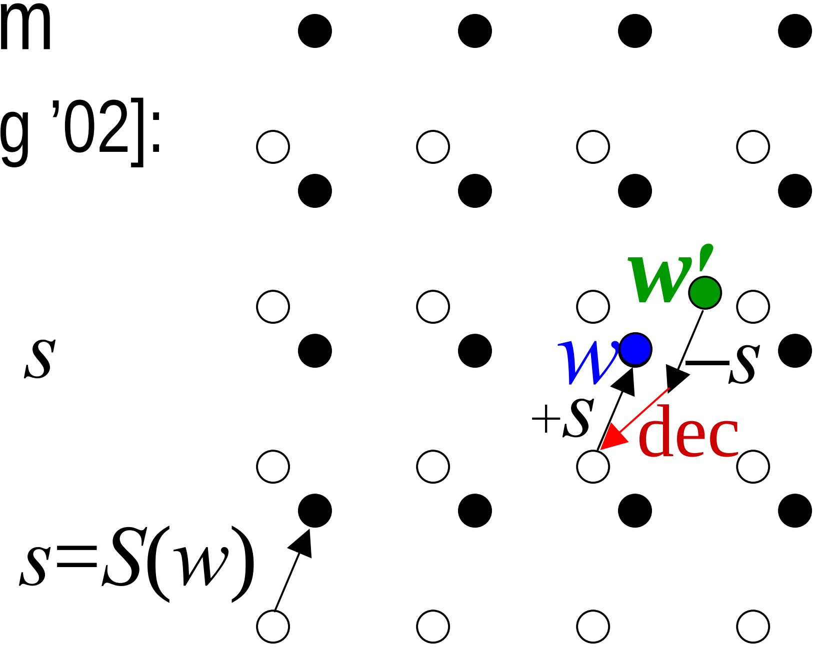
Code C: $\{0,1\}^m \to \{0,1\}^n$

- encodes m-bit messages into n-bit codewords
- any two codewords differ in at least d locations
  - fewer than $d/2$ errors ⇒ unique correct decoding

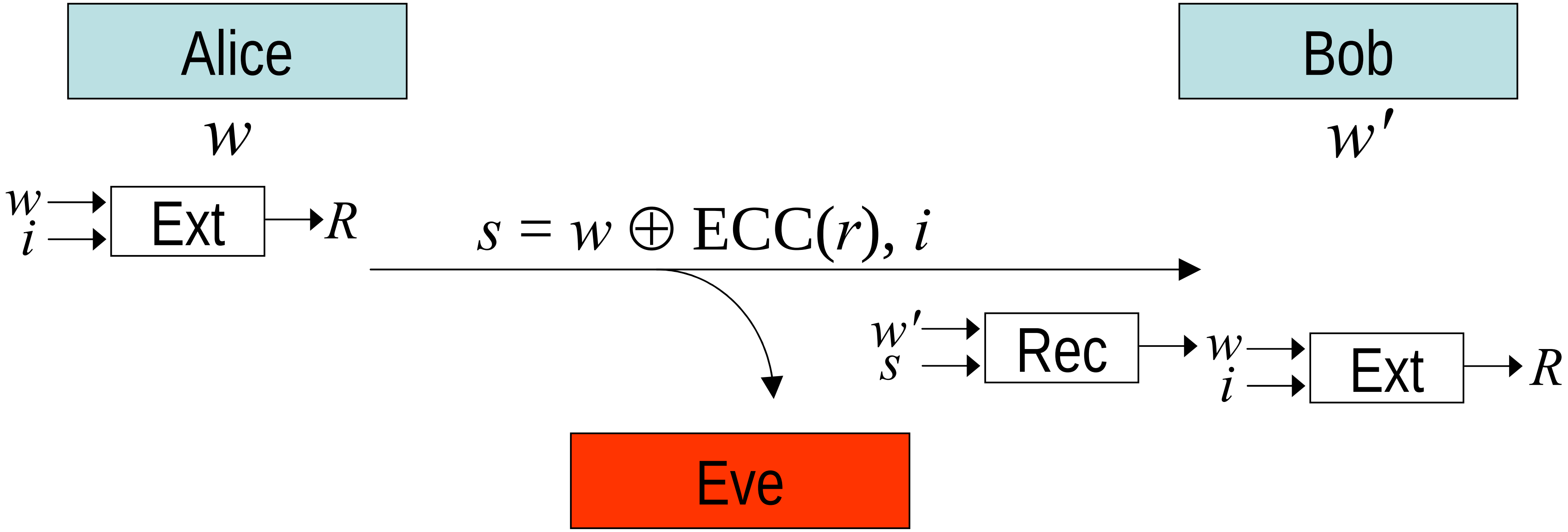


how to build *S*?

- Idea: what if w is a codeword in an ECC?
- Decoding finds w from w'
- If w not a codeword, simply shift the ECC
- $S(w)$ is the shift to random codeword [Juels-Watenberg '02]: $s = w \oplus \mathrm{ECC}(r)$
- Recover: $\mathrm{dec}(w' \oplus s) \oplus s$

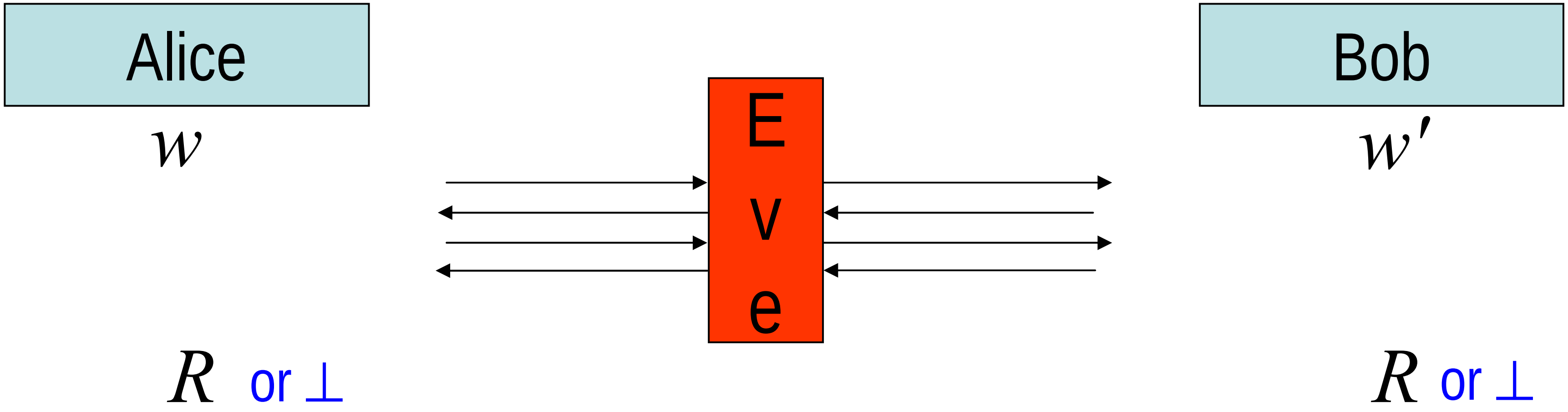


what about information reconciliation?



- <u>Lemma</u>: $H_\infty(W \mid Y, S) \geq \underset{\substack{\parallel \\ H_\infty(W \mid Y) + \underset{\substack{\parallel \\ m}}{|r|}}}{H_\infty(W, S \mid Y)} - \underset{\substack{\parallel \\ n}}{\text{bit-length}\ (S)}$
- $H_\infty(W \mid Y, S) \geq H_\infty(W \mid Y) + m - n$
- Entropy loss for a code from m bits to n bits: $n - m$

active adversary



Alice

w

Eve

Bob

w'

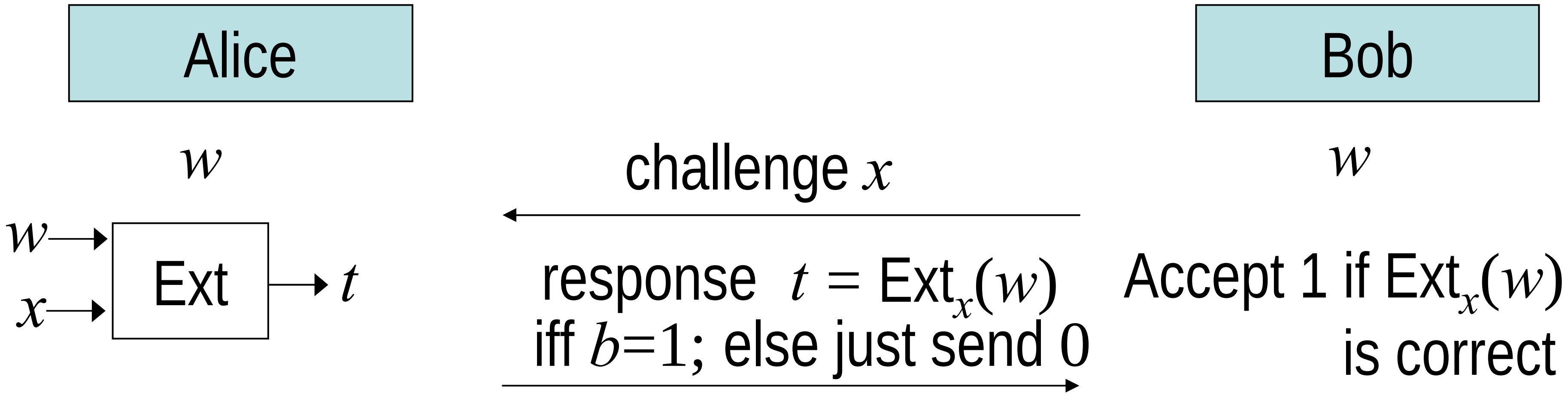
R or $\perp$

R or $\perp$

- Starting in Maurer and Maurer-Wolf 1997
- Interesting even if $w = w'$
- Basic problem: authenticate extractor seed i
- Problem: if $H_\infty(W|Y)<n/2$, w can't be used as a MAC key
- Idea [Renner-Wolf 2003]: use interaction,

  one bit in two rounds

authenticating a bit b [Renner-Wolf 03]

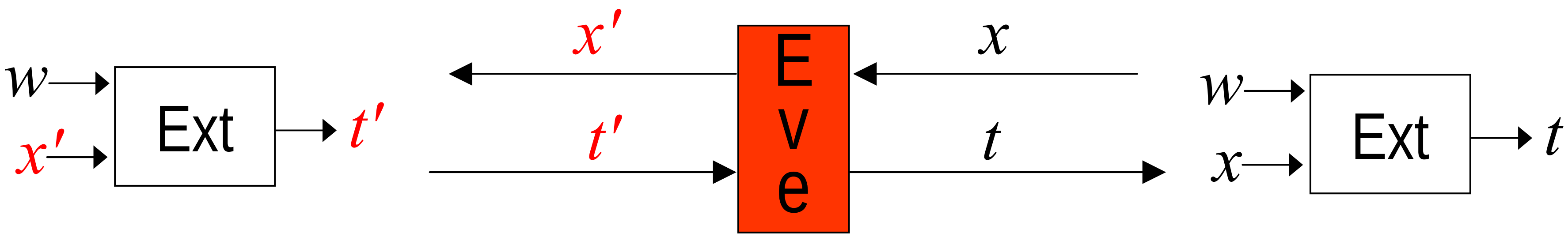


challenge x

response $t = \mathrm{Ext}_x(w)$ iff $b=1$; else just send 0

Accept 1 if $\mathrm{Ext}_x(w)$ is correct

Note: Eve can make Bob's view $\neq$ Alice's view

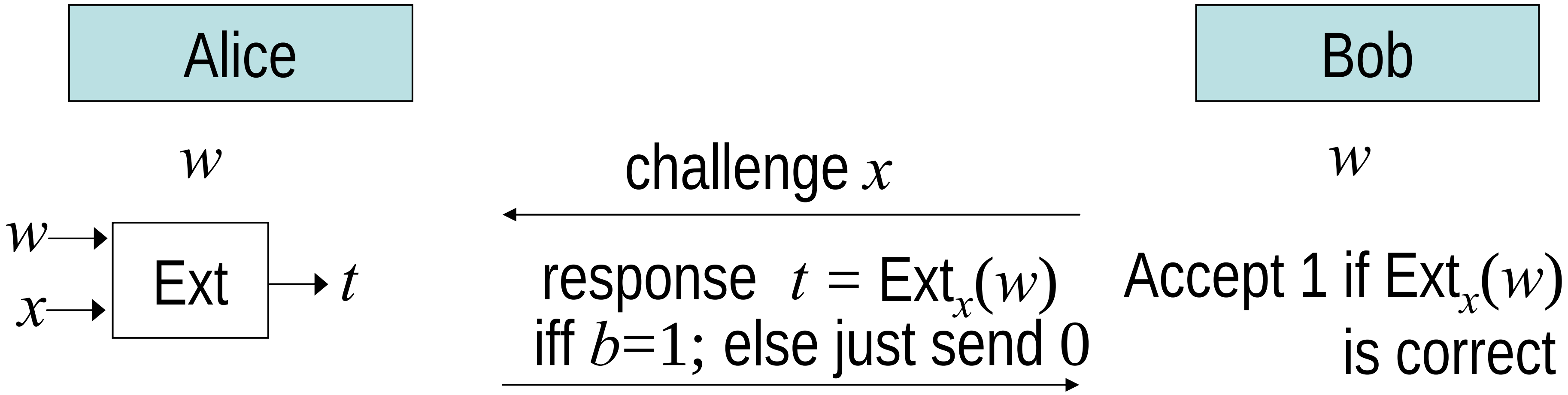


Claim: Eve can't change 0 to 1! (To prevent change of 1 to 0, make #0s = #1s)

Lemma [Kanukurthi-R. '09] $H_\infty(\mathrm{Ext}(W;X) \mid X,Y) \geq \min\,(|t|, \log \frac{1}{\varepsilon}) - 1$

As long as $H_\infty(W \mid Y)$ is high enough for Ext to ensure quality ε;
but we can measure it: each bit authenticated reduces it by $|t|$

improving entropy loss



Alice

Bob

w

challenge x

w

w → Ext → t

x →

response $t = \text{Ext}_x(w)$
iff $b=1$; else just send 0

Accept 1 if $\text{Ext}_x(w)$
is correct

Problem: For λ security, $|t| \approx \lambda$, so each round loses λ entropy

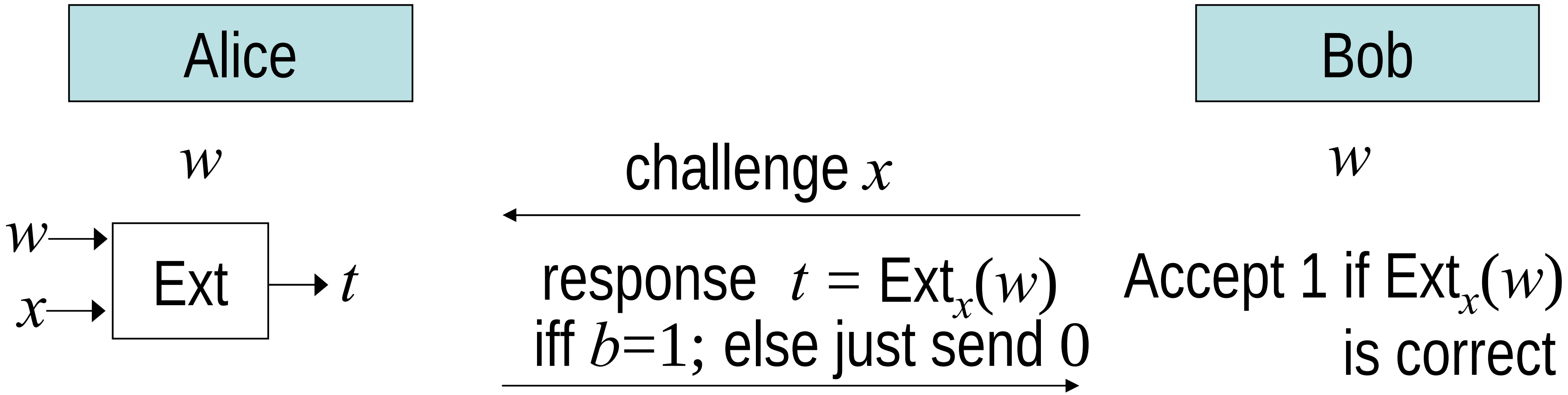
Getting optimal entropy loss [Chandran-Kanukurthi-Ostrovsky-R ’10]:

-- Make $|t|$ = constant.

-- Now Eve can change/insert/delete at most constant fraction of bits

-- Encode whatever you are sending in an edit distance code [Schulman-Zuckerman99] of const. rate, correcting constant fraction

improving entropy loss



<u>Problem:</u> For λ security, $|t| \approx \lambda$, so <u>each round</u> loses λ entropy

<u>Getting optimal entropy loss</u> [Chandran-Kanukurthi-Ostrovsky-R '10]:

-- Make $|t|$ = constant.

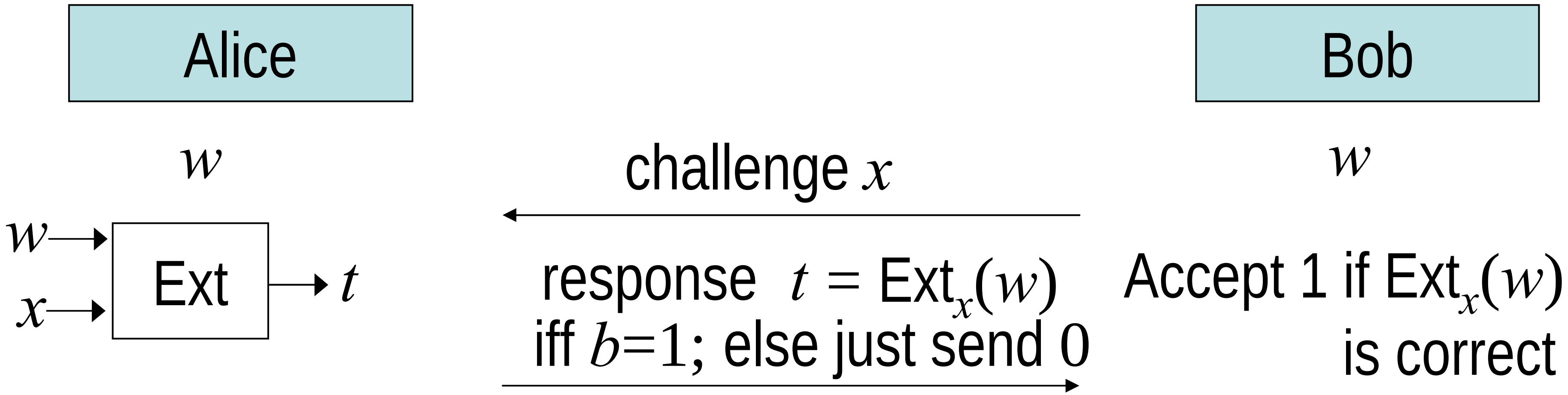
-- Now Eve can change/insert/delete at most constant fraction of bits

How to prove?

Can we use $H_\infty(\text{Ext}(W;X) \mid X,Y) \geq \min\left(|t|, \log \frac{1}{\varepsilon}\right) - 1$?

<u>It talks about unpredictability of a single value;</u>
<u>but doesn't say anything about independence of two</u>

improving entropy loss



Can we use $H_\infty(\text{Ext}(W;X) \mid X,Y) \ge \min(|t|, \log\frac{1}{\varepsilon}) - 1$?

It talks about unpredictability of a single value; but doesn't say anything about independence of two

avg entropy still useful

Step 1: $H_\infty(W \mid$ all variables Eve sees) is sufficient

Step 2: $H_\infty(W \mid$ a specific transcript) is sufficient with high prob

Step 3: $H_\infty(W \mid$ at every Ext step) is sufficient with high prob

Lemma: If $H_\infty(W)$ is sufficient, then $H_\infty(\text{Ext}(W;x) \mid x) \ge |t| - 1$ with high prob.

just minentropy

If (conditional) min-entropy

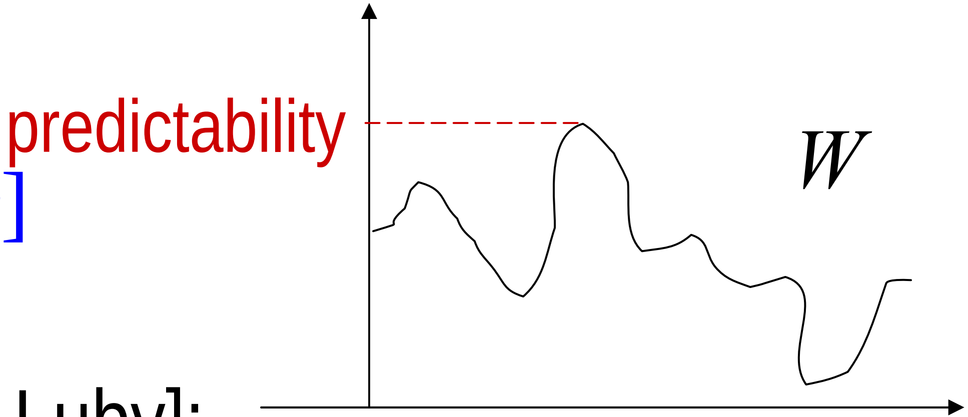
is so useful in information-theoretic crypto,

what about computational analogues?

computational entropy (HILL)

Min-Entropy

$$H_\infty(W) = -\log \max_{w \in W} \Pr[w]$$



[Håstad,Impagliazzo,Levin,Luby]:

$H^{\mathrm{HILL}}_{\delta,s}(W) \geq k$ if $\exists Z$ such that $H_\infty(Z) = k$ and $W \approx Z$

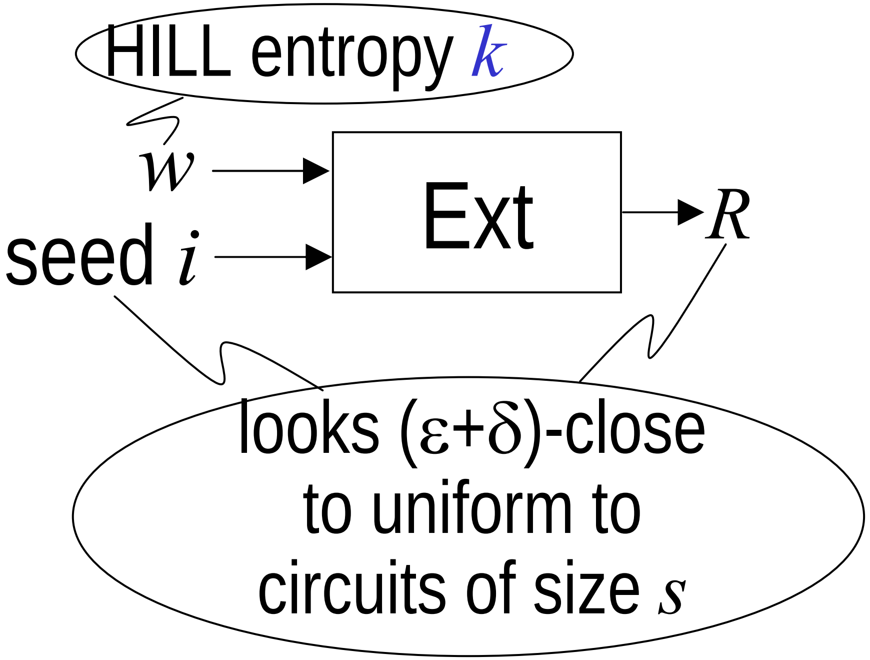
Two more parameters relating to what $\approx$ means

-- maximum size s of distinguishing circuit D

-- maximum advantage δ with which D will distinguish

what is HILL entropy good for?

$H^{\text{HILL}}_{\delta,s}(W) \geq k$ if $\exists Z$ such that $H_\infty(Z) = k$ and $W \approx Z$



HILL entropy k

w

seed i

Ext

R

looks (ε+δ)-close to uniform to circuits of size s

- Many uses: indistinguishability is a powerful notion.
- In the proofs, substitute Z for W;

  a bounded adversary won't notice

what about conditional?

Very common:

entropic secret: g^{ab} | observer knows g^a, g^b

entropic secret: SK | observer knows leakage

entropic secret: $\mathrm{Sign}_{SK}(m)$ | observer knows PK

entropic secret: $\mathrm{PRG}(x)$ | observer knows $\mathrm{Enc}(x)$

conditioning HILL entropy on a fixed event

Recall: how does conditioning reduce minentropy?

By the probability of the condition!

$$H_\infty(W \mid Y = y) \geq H_\infty(W) - \log 1/\Pr[y]$$

E.g., W is uniform, Y = Hamming Weight(W)

$$\Pr[Y = n/2] > 1/(2\sqrt{n}) \;\Rightarrow\; H_\infty(W \mid Y = n/2) \geq n - ½ \log n - 1$$

conditioning HILL entropy on a fixed event

Recall: how does conditioning reduce minentropy?

By the probability of the condition!

$$H_\infty(W \mid Y = y) \geq H_\infty(W) - \log 1/\Pr[y]$$

Theorem: same holds for computational entropy:

$$\boxed{H^{\text{metric*}}_{\delta/\Pr[y],s}(W \mid Y = y) \geq H^{\text{metric*}}_{\delta,s}(W) - \log 1/\Pr[y]}$$

[Fuller-R '11] (variant of Dense Model Theorem of
[Green-Tao '04, Tao-Ziegler '06,
Reingold-Trevisan-Tulsiani-Vadhan '08, Dziembowski-Pietrzak '08]

Warning: this is not H^{HILL}!

Weaker entropy notion: a different Z for each distinguisher ("metric*")

$H^{\text{metric*}}_{\delta,s}(W) \geq k$ if $\forall$ distinguisher D $\exists Z$ s.t. $H_\infty(Z) = k$ and $W \approx_D Z$

(moreover, D is limited to deterministic distinguishers)

It can be converted to H^{HILL} with a loss in circuit size s

[Barak, Shaltiel, Wigderson 03]

conditioning HILL entropy on a fixed event

Long story, but simple message:

$$\boxed{H^{\text{metric*}}_{\delta/\Pr[y],s}(W \mid Y = y) \geq H^{\text{metric*}}_{\delta,s}(W) - \log 1/\Pr[y]}$$

It can be converted to H^{HILL} with a loss in circuit size s

[Barak, Shaltiel, Wigderson 03]

what about conditioning on average?

entropic secret: g^{ab} | observer knows g^a, g^b

entropic secret: SK | observer knows leakage

entropic secret: $\mathrm{Sign}_{SK}(m)$ | observer knows PK

entropic secret: $\mathrm{PRG}(x)$ | observer knows $\mathrm{Enc}(x)$

Again, we may not want to reason about specific values of Y

[Hsiao-Lu-R '04]:

Def: $H^{\text{HILL}}_{\delta,s}(W \mid Y) \geq k$ if $\exists Z$ such that $H_\infty(Z \mid Y) = k$ and $(W, Y) \approx (Z, Y)$

Note: W changes, Y doesn't

What is it good for? Original purpose: negative result

Computational Compression (Yao) Entropy can be > HILL

Hasn't found many uses because it's hard to measure
(but it can be extracted from by reconstructive extractors!)

conditioning HILL entropy on average

Recall: suppose Y is over b-bit strings

$$H_\infty(W \mid Y) \geq H_\infty(W) - b$$

Average-Case Entropy Version of Dense Model Theorem:

$$\boxed{H^{\text{metric*}}_{\delta 2^b,s}(W \mid Y) \geq H^{\text{metric*}}_{\delta,s}(W) - b}$$

Follows from $H^{\text{metric*}}_{\delta/\Pr[y],s}(W \mid Y = y) \geq H^{\text{metric*}}_{\delta,s}(W) - \log 1/\Pr[y]$

Can work with metric* and then covert to HILL when needed (loss in s)

conditioning the conditional

$$\boxed{H^{\text{metric*}}_{\delta 2^b,s}(W \mid Y) \geq H^{\text{metric*}}_{\delta,s}(W) - b}$$

The theorem can be applied multiple times, of course:

$$H^{\text{metric*}}_{\delta 2^{b_1+b_2},s}(W \mid Y_1, Y_2) \geq H^{\text{metric*}}_{\delta,s}(W) - b_1 - b_2$$

(where support of Y_i has size 2^{b_i})

But we can't prove: $H^{\text{metric*}}_{\delta 2^{b_2},s}(W \mid Y_1, Y_2) \geq H^{\text{metric*}}_{\delta,s}(W \mid Y_1) - b_2$

(bad case: W = plaintext, $Y_1 = PK$;
because for any given y_1, W has no entropy!)

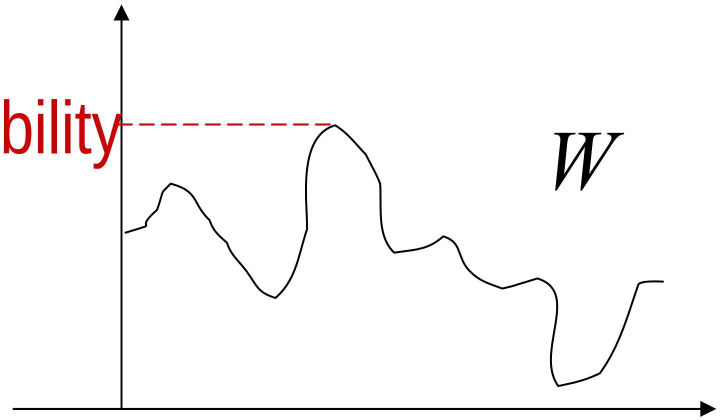
Note: Gentry-Wichs '11 implies:

$$H^{\text{HILL-relaxed}}_{2\delta,s/\text{poly}(\delta,\, 2^{b_2})}(W \mid Y_1, Y_2) \geq H^{\text{HILL-relaxed}}_{\delta,s}(W \mid Y_1) - b_2$$

<u>Defn</u>: $H^{\text{HILL-relaxed}}_{\delta,s}(W|Y) \geq k$ if $\exists (Z, T)$ such that $H_\infty(Z \mid T) = k$
and $(W, Y) \approx (Z, T)$

unpredictability entropy

Why should computational min-entropy be defined through indistinguishability? Why not model unpredictability directly?



$$H_\infty(W) = -\log \max_{w \in W} \Pr[w]$$

[Hsiao-Lu-R. '04]

$H_s^{\text{Unp}}(W|Z) \geq k$ if for all $\forall A$ of size s, $\Pr[A(z) = w] \leq 2^{-k}$

Lemma: $H^{\text{Yao}}(W|Z) \geq H^{\text{Unp}}(W|Z) \geq H^{\text{HILL}}(W|Z)$

Corollary: Reconstructive extractors work for H^{Unp}

Lemma: $H_s^{\text{Unp}}(W \mid Y_1, Y_2) \geq H_s^{\text{Unp}}(W, \mid Y_1) - b_2$

what is it good for?

$H_s^{\mathsf{Unp}}(W\,|\,Z) = k$ if for all $\forall A$ of size s, $\Pr[A(z) = w] \le 2^{-k}$

Examples:

Diffie-Hellman: $g^{ab} \mid g^a, g^b$

$H^{\mathsf{HILL}}=0$ $\left\{\begin{array}{l}\text{One-Way Functions: } x \mid f(x) \\ \text{Signatures: } Sign_{SK}(m) \mid PK\end{array}\right.$

Why bother?

- Hardcore bit results (e.g., [Goldreich&Levin,Ta-Shma&Zuckerman]) are typically stated only for OWF, but used everywhere
  - They are actually reconstructive extractors
  - $H^{\mathsf{Unp}}(X\,|\,Z)$ + reconstructive extractors $\Rightarrow$ simple generalization language
- Leakage-resilient crypto (assuming strong hardness)

the last slide

Minentropy is often the right measure

Conditional Entropy useful natural extension

Easy to use because of simple bit counting

Computational Case is trickier

- A few possible extensions
- Bit counting sometimes works
- Some definitions (such as H^{Unp}) only make sense conditionally
- Separations and conversions between definitions exist
- Still, can simply proofs!