Minentropy and its Variations for Cryptography

Leonid Reyzin

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:
  \[ \text{min entropy} = - \log (\Pr \text{[adversary predicts sample]}) \]

\[ H_\infty(W) = - \log \max_w \Pr[w] \]
what is minentropy good for?

- Passwords
- Message authentication

\[
\text{key } w = \begin{array}{c}
\text{n/2} \\
a \\
\text{n/2} \\
b
\end{array} \in \text{GF}(2^{n/2}) \times \text{GF}(2^{n/2})
\]

\[
m \rightarrow \times \rightarrow + \rightarrow MAC_{a,b}(m) = \sigma = am + b
\]

[Wegman-Carter ‘81]
what is minentropy good for?

- Passwords
- Message authentication

\[
\text{key } w = \begin{array}{c}
\text{gap } g \\
\text{minentropy } k
\end{array}
\]

\[ \text{MAC}_{a,b}(m) = \sigma = am + b \]

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 \( \text{MAC}_{a,b}(m) = \sigma = am + b \)
• Secret key extraction (⇒ encryption, etc.)

\[
\text{minentropy } k
\]

\[
\text{Ext}
\]

\[
\text{reusable}
\]

\[
\text{jointly uniform (\(\varepsilon\)-close)}
\]

is it good for privacy amplification?

Goal: from a partial secret $\omega$ agree on a uniform secret $R$ [Bennett-Brassard-Robert '85]

Simple solution: use an extractor

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

 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($W$)
  \[
  \Pr[Y = n/2] > 1/(2^{\sqrt{n}}) \implies H_\infty(W \mid Y = n/2) \geq n - \frac{1}{2} \log n - 1
  \]
  \[
  \Pr[Y = n] = 2^{-n} \implies H_\infty(W \mid Y = 0) = 0
  \]

- 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 $Y$?
  \[
  H_\infty(W \mid Y) = - \log \mathbb{E} \max_w \Pr[w \mid Y=y]
  \]

“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 ($\Rightarrow$ encryption, etc.)
  – All extractors work [Vadhan ‘11]

\[
H_\infty(W \mid Y) = k
\]

\[w \rightarrow Ext \rightarrow R\]

seed $i$

jointly uniform given $Y$
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 ($\Rightarrow$ encryption, etc.)**
  - All extractors work [Vadhan ‘11]
  - Therefore, privacy amplification!

[Diagram showing Alice, Bob, and Eve with partially secret $W$ and $i$.]
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 $C$:** $\{0,1\}^m \rightarrow \{0,1\}^n$

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

\[
\begin{array}{cccccc}
\bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet \\
\end{array}
\]
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]:
  \[ s = w \oplus \text{ECC}(r) \]
- Recover: $\text{dec}(w' \oplus s) \oplus s$

\[ s = S(w) \]
what about information reconciliation?

**Lemma:** \( H_\infty(W \mid Y, S) \geq H_\infty(W, S \mid Y) - \text{bit-length } (S) \)

\[
H_\infty(W \mid Y) + |r| \quad \text{||} \quad n
\]

\[
H_\infty(W \mid Y) + m - n
\]

- Entropy loss for a code from \( m \) bits to \( n \) bits: \( n - m \)
active adversary

- 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]

Alice

$w$

$w$\xrightarrow{} Ext \xrightarrow{} t

challenge $x$

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

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

Bob

$w$

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

$w$\xrightarrow{} Ext \xrightarrow{} t'

$x'$

$w$\xrightarrow{} Ext \xrightarrow{} t

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

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

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

![Diagram]

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

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

-- Make $|t| = \text{constant}$.
-- 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**

![Diagram of Alice and Bob](image)

- **Challenge**: `x` to `Ext_x(w) → t`
- **Response**: `t = Ext_x(w)` if `b=1`; otherwise, just send `0`

**Problem**: For λ security, `|t| ≈ \lambda`, so each round loses \lambda entropy.

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

- Make `|t| = \text{constant}`.
- Now Eve can change/insert/delete at most constant fraction of bits.

**How to prove?**

Can we use $H_\infty(\text{Ext}(W;X) \mid X,Y) \geq \min (|t|, \log \frac{1}{\varepsilon}) - 1$? It talks about unpredictability of a single value; but doesn’t say anything about independence of two
improving entropy loss

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

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

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

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

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

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

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

Two more parameters relating to what $\approx$ means

-- maximum size $s$ of distinguishing circuit $D$

-- maximum advantage $\delta$ with which $D$ will distinguish
what is HILL entropy good for?

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

- Many uses: indistinguishability is a powerful notion.
- In the proofs, substitute \( Z \) for \( W \);
  a bounded adversary won’t notice

\[ W \rightarrow \text{Ext} \rightarrow R \]

looks \((\varepsilon+\delta)\)-close to uniform to circuits of size \( s \)
**what about conditional?**

**Very common:**

<table>
<thead>
<tr>
<th>entropic secret: $g^{ab}$</th>
<th>observer knows $g^a, g^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>entropic secret: $SK$</td>
<td>observer knows leakage</td>
</tr>
<tr>
<td>entropic secret: $\text{Sign}_{SK}(m)$</td>
<td>observer knows $PK$</td>
</tr>
<tr>
<td>entropic secret: $\text{PRG}(x)$</td>
<td>observer knows $\text{Enc}(x)$</td>
</tr>
</tbody>
</table>
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, \( Y = \text{Hamming Weight}(W) \)

\[ \Pr[Y = n/2] > 1/(2\sqrt{n}) \quad \Rightarrow \quad H_\infty(W \mid Y = n/2) \geq n - \frac{1}{2} \log n - 1 \]
conditioning \(H_{\text{ILL}}\) 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:

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

[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 \(H_{\text{HILL}}\)!

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

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

(moreover, \(D\) is limited to deterministic distinguishers)

It can be converted to \(H_{\text{HILL}}\) with a loss in circuit size \(s\)

[Barak, Shaltiel, Wigderson 03]
conditioning HILL entropy on a fixed event

Long story, but simple message:

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

It can be converted to $H^{\text{HILL}}$ 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: $\text{Sign}_{SK}(m)$ | observer knows $PK$
entropic secret: $\text{PRG}(x)$ | observer knows $\text{Enc}(x)$

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 $> \text{HILL}$

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:

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

Follows from

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

Can work with $\text{metric*}$ and then covert to HILL when needed (loss in $s$)
conditioning the conditional

\[ H_{\delta b, s}^{\text{metric}^*} (W \mid Y) \geq H_{\delta, s}^{\text{metric}^*} (W) - b \]

The theorem can be applied multiple times, of course:

\[ H_{\delta b_1 + b_2, s}^{\text{metric}^*} (W \mid Y_1, Y_2) \geq H_{\delta, s}^{\text{metric}^*} (W) - b_1 - b_2 \]

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

But we can’t prove:

\[ H_{\delta b_2, s}^{\text{metric}^*} (W \mid Y_1, Y_2) \geq H_{\delta, s}^{\text{metric}^*} (W \mid Y_1) - b_2 \]

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

Note: Gentry-Wichs ’11 implies:

\[ H_{2\delta, s/\text{poly}(\delta, 2^{b_2})}^{\text{HILL-relaxed}} (W \mid Y_1, Y_2) \geq H_{\delta, s}^{\text{HILL-relaxed}} (W \mid Y_1) - b_2 \]

**Defn:** \( H_{\delta, s}^{\text{HILL-relaxed}} (W \mid Y) \geq k \) if \( \exists (Z, T) \) such that \( H_{\infty}(Z \mid T) = k \) and \( (W, Y) \approx (Z, T) \)
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]

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

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

Corollary: Reconstructive extractors work for \( H_{\text{Unp}}^{\text{Unp}} \)

Lemma: \( H_{s}^{\text{Unp}}(W | Y_1, Y_2) \geq H_{s}^{\text{Unp}}(W, | Y_1) - b_2 \)
what is it good for?

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

Examples: 
- Diffie-Hellman: \[ g^{ab} | g^a, g^b \]
- One-Way Functions: \[ x | f(x) \]
- Signatures: \[ \text{Sign}_{SK}(m) | PK \]

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
  - \[ H^\text{Unp}(X|Z) + \text{reconstructive extractors } \Rightarrow \text{simple generalization language} \]
- Leakage-resilient crypto (assuming strong hardness)
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 $H^{\text{Unp}}$) only make sense conditionally
• Separations and conversions between definitions exist
• Still, can simply proofs!