Intro remarks: Joseph M. Renes : A Theory of Information (?)

1. Physicists are sorta lazy: witness the principle of least action!

2. but really what this means is that we want simple, coherent, powerful theories of physical phenomena. Such as
   a. Maxwell's equations (second Wrangler; also at Trinity)
   b. Newton's Laws (since we're in Cambridge)
   c. if we have to mention it: general relativity

One-shot framework invites us to ~~think~~ consider a "theory of information" not just "information theory", by which I mean whatever results you can manage to find regarding information transmission & processing, but rather the logic behind the whole thing.

One-shot language allows us to see and to express this structure.

Because:

1. it allows us to build up more complicated protocols from simple primitives in a transparent way
2. ~~it makes sense of the entropy measures motivates / explains~~
   it allows us to see the relationships between protocols by looking at their entropic characterization

Outline

- Primitives
- Entropies
- Protocols
  - channel coding
  - channel simulation
  - quantum: PA = IR

Primitives

I'll argue that the two primitives are:

- Privacy Amplification

- Information Reconciliation (Data Compression / Slepian-Wolf)

For experts: these are "state" (↗ or prob. distribution) versions of resolvability & channel coding

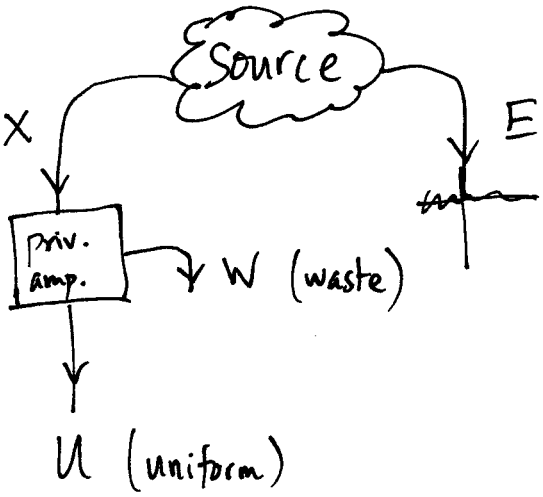
[They will prove more convenient to work with.] (covering) (packing)

PA By "state" (↗ or "source") problem, I mean we're trying to transform a given (probability distribution) random variable / quantum state into a desired form: (It's always a bipartite RV)

PA: remove correlations

IR: ~~create~~ perfect correlations

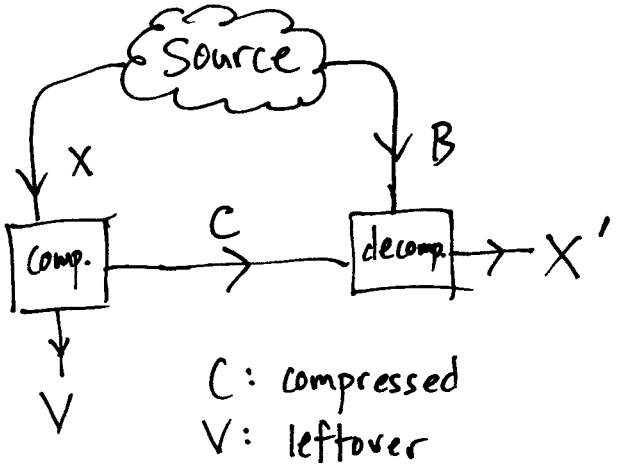
PA



U is uniform, uncorrelated with E

$$\frac{1}{2}\| \rho_{UE} - \pi_U \otimes \rho_E \|_1 \leq \varepsilon$$

IR



C: compressed
V: leftover

X' is essentially identical to X

~~$\frac{1}{2}$~~

~~$X' = D(C(X))$~~

$X' = \mathrm{Dec}\left(\mathrm{Cmp}(X), B\right)$

$\Pr[X \neq X'] \leq \varepsilon$

$\frac{1}{2}\| \rho_{XX'} - \rho_{X\otimes} \|_1 \leq \varepsilon$

Entropies: $\rho_{XB} = \sum_x p_x [x] \otimes \varphi_x$ $\quad D_{\max}(\rho\|\sigma) = \min\{\lambda \in \mathbb{R} : \rho \leq 2^\lambda \sigma\}$

1. $H_{\min}(X|B)_\rho := -\min_{\sigma_B} D_{\max}(\rho_{XB} \| \mathbb{1}_X \otimes \rho_B)$ $\quad$ (∞ norm $\sigma^{-1/2}\rho\sigma^{-1/2}$)

$\Rightarrow$ guessing probability $p_{\mathrm{guess}}(X|B)_\rho = 2^{-H_{\min}(X|B)_\rho}$

2. $H_{\max}(X|B)_\rho := \max_\sigma \log \left\| \sqrt{\rho_{XB}} \sqrt{\mathbb{1}_X \otimes \sigma_B} \right\|_1^2$ $\quad$ (1-norm)

$\Rightarrow$ proximity to uniform, uncorrelated state

$H_{\max}(X|B)_\rho = \max_\sigma \log \frac{1}{|X|} F(\rho_{XB}, \pi_X \otimes \sigma_B)^2$ $\quad$ (eqn after 4.8 in MT omits the 2)

3. Smoothing: minimize $H_{\max}$, maximize $H_{\min}$ by looking in ε-neighborhood of actual state

in quantum case, best to use purified distance

Entropic Characterizations of Protocols

PA target output: $H_{max}(U|E) \approx 0 \Rightarrow H^{\varepsilon}_{max}(U|E) = 0$

"rate": $\log|U| \approx H^{\varepsilon}_{min}(X|E)$ ← should be $\log|U|$

IR target output $H^{\varepsilon}_{min}(X|X') = 0$

"rate" $\log|C| \approx H^{\varepsilon}_{max}(X|B)$

Proofs: & universal hashing & leftover hash lemma PA
pretty good measurement IR

A unifying entropy — hypothesis-testing entropy

1211.3141

also used by Wang/Renner, Toma/Haya, Buscemi/Datta, Brandão/Datta

$$2^{-D^{\varepsilon}_{H}(\rho\|\sigma)} = \min \frac{1}{\varepsilon} \operatorname{tr} Q\sigma = \max \mu - \frac{1}{\varepsilon} \operatorname{tr} R$$

$\operatorname{tr} Q\rho \geq \varepsilon$ $\quad$ $\mu\rho \leq \sigma + R$

$0 \leq Q \leq \mathbb{1}$ $\quad$ $R \geq 0$

← already looks a bit like Dmax

$$H^{\varepsilon}_{H}(X|B)_{\rho} = -\min_{\sigma} D^{\varepsilon}_{H}(\rho_{XB} \| \mathbb{1}_X \otimes \sigma_B)$$

Point being: $H^{\varepsilon}_{H}(X|B)_{\rho} \approx H^{\varepsilon}_{min}(X|B)_{\rho}$

$H^{1-\varepsilon}_{H}(X|B)_{\rho} \approx H^{\varepsilon}_{max}(X|B)_{\rho}$

Protocols

Renes/Renner
1012.4814 TIT 57, 7377 (2011)

Noisy channel coding (CQ) $X \to [N] \to B$

$M \to [\text{Enc}] \to [N] \to [\text{Dec}] \to M'$ M uniform

Can show $\Pr[M \neq M'] \leq \varepsilon$

expurgate to get worst-case, i.e. identity channel simul.

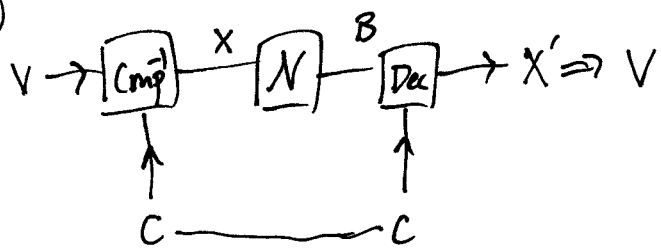
$$\log |M| \approx H_{\min}^{\varepsilon}(X) - H_{\max}^{\varepsilon}(X|B)$$

$\max_X$

Yet another coding thm proof!

(1) (examine case $X = U$) use IR to simulate state

(a)



(b) imagine Cmp is linear, then if X is uniform, so are C, V

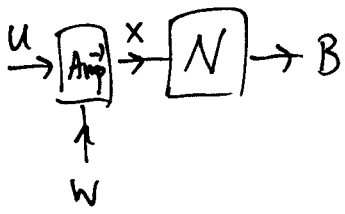
$$\Rightarrow |V| \cdot |C| = |X| \Rightarrow |V| = \frac{|X|}{|C|} \Rightarrow \log|V| = \log|X| - \log|C|$$

$$\Rightarrow \log|V| \approx \log|X| - H_{\max}^{\varepsilon}(X|B)$$

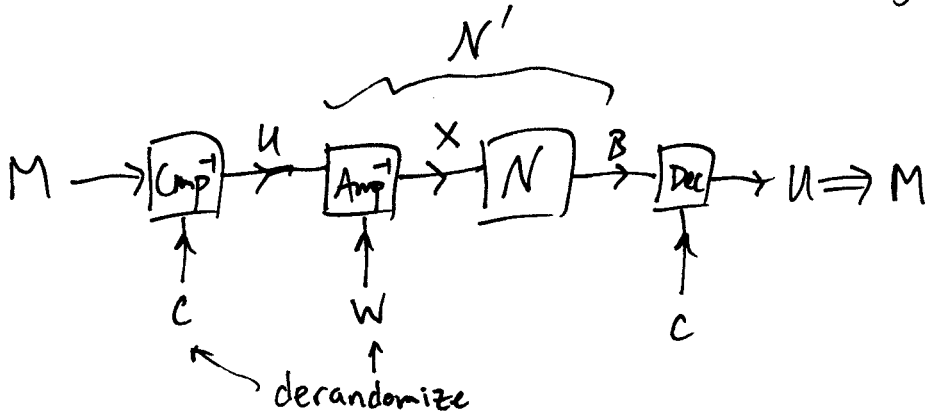
(c) derandomize C by taking best value

(2) that only gives "symmetric capacity".
for the actual capacity, prefix the channel with PA

Protocols 2



Again, assume Amp is linear. We don't need to worry about correlations.



rate: $\log|M| \approx \log|U| - H_{max}^{\varepsilon}(U|B)_{N'}$

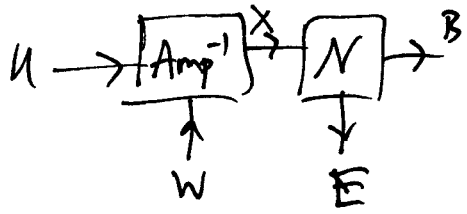
$\approx H_{min}^{\varepsilon}(X) - H_{max}^{\varepsilon}(U|B)$

$\gtrsim H_{min}^{\varepsilon}(X) - H_{max}^{\varepsilon}(X|B)$ U is a function of X via Amp (5.11)

can show converse (cf. Ligong/Renner)

Private Coding

Great thing is we can immediately do private coding



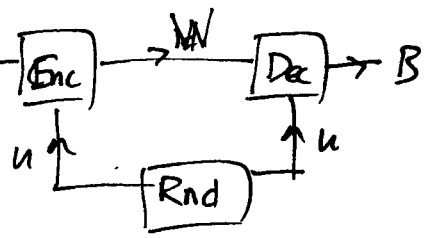
$\rho_{UE} \approx \pi_U \otimes \rho_E$

$\Rightarrow$ ~~rate~~

$\log|M| \gtrsim H_{min}^{\varepsilon}(X|E) - H_{max}^{\varepsilon}(X|B)$ Csiszar Körner

Protocols 3

Arb. Channel Simulation (for fixed state) $X \to [N] \to B \approx$



- specialize to classical to see the overview
- measurement compression can be done, too
- to get universal simulation, use post-selection (for iid)

from PA of $Y|X$: $P_{YX} \to P_{UWX} \to P_{W|UX}$

$(U, W) = \text{Amp}(Y)$

Enc: generate W from UX using $P_{W|UX}$

Dec: recreate Y from (U, W) $\quad$ Dec $= \text{Amp}^{-1}$

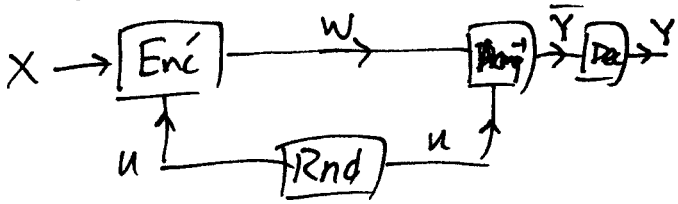
rate? $c = \log|W| = \log|Y| - \log|U| = \log|Y| - H_{\min}^{\varepsilon}(Y|X)$

$s = \log|U| = H_{\min}^{\varepsilon}(Y|X)$

---

improve by compressing Y first, to $\bar{Y}$

$\Rightarrow \log|\bar{Y}| = H_{\max}^{\varepsilon}(Y)$



rate of channel sim:

$c = \log|\bar{Y}| - H_{\min}^{\varepsilon}(\bar{Y}|X)$

$s = H_{\min}^{\varepsilon}(\bar{Y})$

$\Rightarrow c \approx H_{\max}^{\varepsilon}(Y) - H_{\min}^{\varepsilon}(Y|X)$

$s \approx H_{\min}^{\varepsilon}(Y)$

$H_{\min}^{\varepsilon}(\bar{Y}|X) \leq H_{\min}^{\varepsilon}(Y|X)$ $\bar{Y}$ function of Y

but $p_{\text{guess}}(\bar{Y}|X) \approx p_{\text{guess}}(Y|X)$

just guess $\bar{Y}$ and decompress

$\Rightarrow H_{\min}^{\varepsilon}(\bar{Y}|X) \approx H_{\min}^{\varepsilon}(Y|X)$

Protocols 4

in the quantum world PA $\Longleftrightarrow$ IR !

it's all about the uncertainty principle

use von Neumann $\quad H(X_A|B) + H(Z_A|E) \geq \log|X|$

Start with a pure state (actually any state) $|\psi\rangle_{ABE}$

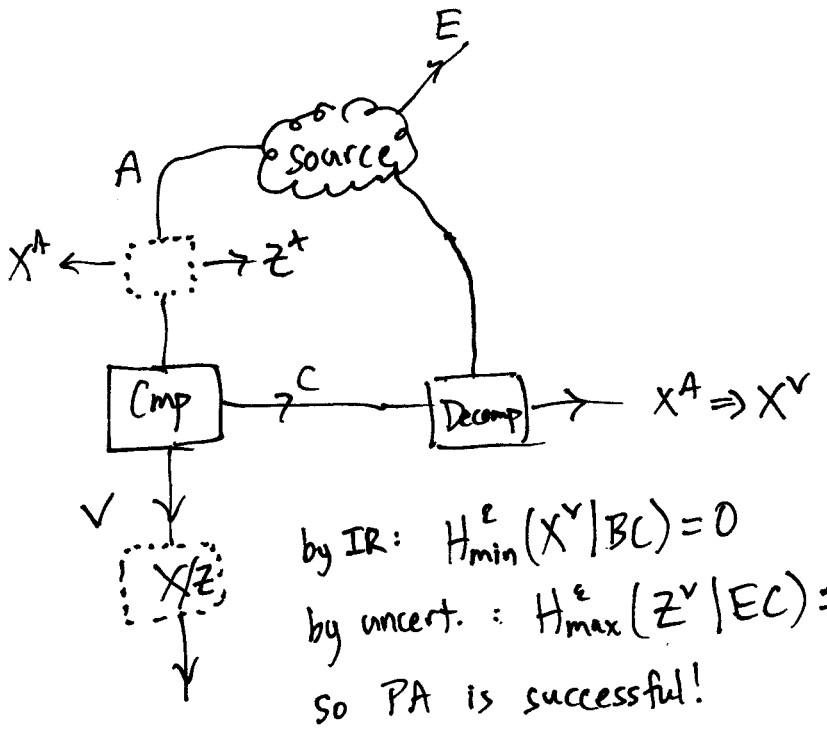
X arises from measurement in a given basis

Z arises from using the conjugate basis.

$$|z\rangle = \frac{1}{\sqrt{d}} \sum_x e^{2\pi i/d\, xz} |x\rangle$$

IR $\Rightarrow$ PA

using a linear compressor, a map on X can also be viewed as a map on Z



by IR: $H^\varepsilon_{min}(X^V|BC) = 0$

by uncert.: $H^\varepsilon_{max}(Z^V|EC) \simeq \log d$

So PA is successful!

rate: $\log|V| = \log d - \log|C| = \log d - H^\varepsilon_{max}(X^A|B)$

(but $\leq H^\varepsilon_{min}(Z^A|E)$)

PA ⇒ IR

seems like it won't work, we need a certainty principle

⇒ we have that for certain states!

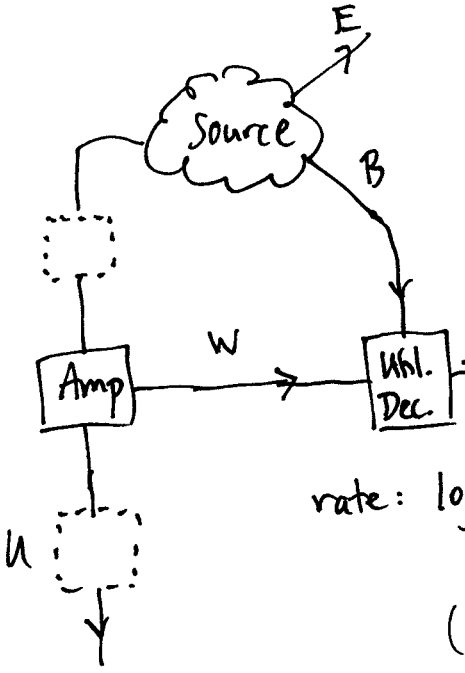
recall: $H(X^A|B) + H(Z^A|E) \geq \log d$

↳ but if state pure and $H(X^A|E) = 0$ or $H(Z^A|B) = 0$

then equality

we could have assumed $H(Z^A|B) = 0$ in (IR $\overset{\Rightarrow PA}{\not\Rightarrow}$ ~~$H(X^A|B) = 0$~~)

now assume $H(X^A|E) = 0$ for PA ⇒ IR



by PA $H^\varepsilon_{max}(Z^U|E) = 0$

by same ⇒ $H^\varepsilon_{min}(X^U|WB) = 0$

can also say $H^\varepsilon_{max}(Z^U|EX^W) = \log d$

and therefore $H^\varepsilon_{min}(X^U|X^W B) = 0$

and hence $H^\varepsilon_{min}(X^A|BX^W) = 0$

rate: $\log|W| = \log d - \log|U|$

$= \log d - H^\varepsilon_{min}(X^A|E)$

(but $\leq H^\varepsilon_{max}(X^A|B)$)

(10b)

(10a)