Single-shot thermo, lecture 5

18.11.2014

- There are now <u>lecture notes</u> online! (handwritten)
- How difficult was the exercise sheet?

2.6. Distillable nonuniformity and nonuniformity of formation

<u>Questions:</u> Given any state ρ, how much work can we extract from it?
How much work do we need to create a state ρ?

At this stage, "work" is only implicit (via Landauer erasure / Szilard engine). Later: model Hamiltonians explicitly.

<u>Questions for now:</u>

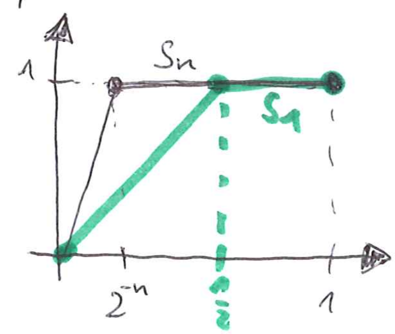
(i) How many pure bits can we extract from a given distribution p?

(ii) How many pure bits do we need to create p?

(i): we want $p \xrightarrow{\text{noisy}} s_n$, where $s_n = \underbrace{\begin{pmatrix}1\\0\end{pmatrix}^{\otimes n}}_{n \text{ pure bits}}$

with n as large as possible.

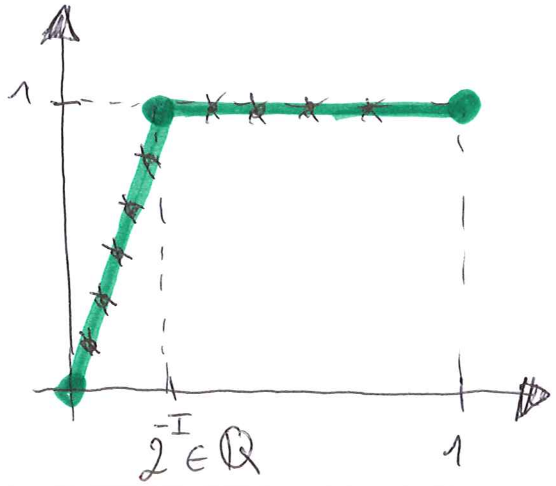
Lorenz curve of s_n:



(larger n $\Rightarrow$ more "valuable" as a resource)

Generalize to $n \notin \mathbb{N}$:

$$S_I := (\underbrace{\underbrace{\tfrac{1}{k}, \ldots, \tfrac{1}{k}}_{k}, 0, \ldots, 0}_{\ell}) \in \mathbb{R}^{\ell}$$



where $I = \log \frac{\ell}{k}$,

e.g. $S_{\log 7/3} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0, 0, 0\right)$.

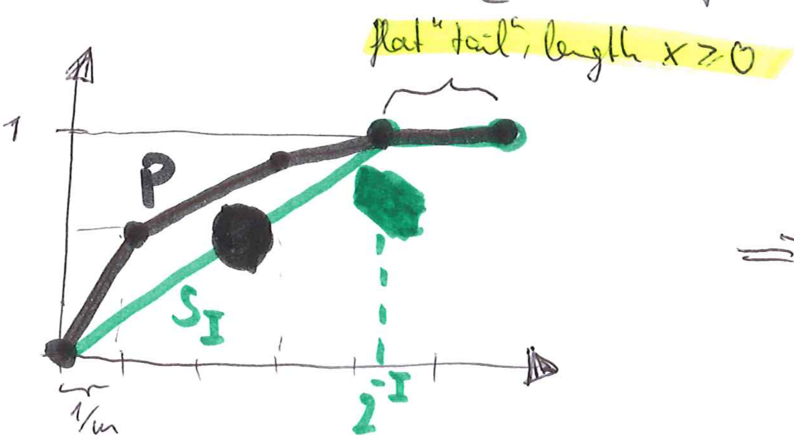
S_I is called a "sharp state".

Distillable nonuniformity:

$\Rightarrow$ We want $p \xrightarrow{\text{noisy}} S_I$, with I as large as possible.

(intuition: then we can extract exactly $I k_B T \ln 2$ of work in a Szilard engine).

$\Rightarrow$ Lorenz curve of S_I must fit below Lorenz curve of p:



$$\Rightarrow x \geq 1 - 2^{-I}$$

$x = 1 - \frac{1}{m} \cdot \#$ of nonzero entries of $p = 1 - \frac{1}{m} \operatorname{rank}(p)$

$\Rightarrow \frac{1}{m} \operatorname{rank}(p) \leq 2^{-I}$, $\quad I \leq \log m - \log \operatorname{rank}(p)$

With $H_0(p) := \log \operatorname{rank}(p)$, [redacted] writing $d_p := m$, we get

Lemma: $p \xrightarrow{\text{noisy}} S_I$ if and only if

$$I \leq \log d_p - H_0(p) .$$ [redacted]

That is, the "distillable nonuniformity" is

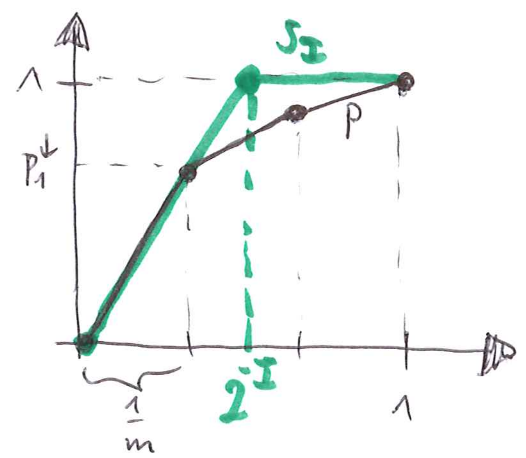
$$\boxed{I_0(p) := \log d_p - H_0(p) .}$$

Interpretation: This is the amount of work that we can extract from p in a Szilard engine with absolute certainty (not just "on average").

Nonuniformity of formation:

We want $S_I \xrightarrow{\text{noisy}} p$, with I as small as possible.

$\Rightarrow$ Lorenz curve of p must fit below Lorenz curve of S_I:



$\Rightarrow$ initial slope of L_{S_I} $\geq$ initial slope of L_p.

$$\Rightarrow \frac{1}{2^{-I}} \geq \frac{p_1^{\downarrow}}{1/m}$$

$$2^I \geq m \max_i p_i$$ [redacted]

$H_\infty(p) := -\log \max_i p_i \Rightarrow$

Lemma: $S_I \xrightarrow{\text{noisy}} p$ if and only if

$$I \geq \log d_p - H_\infty(p)$$

That is, the "nonuniformity of formation" is

$$\boxed{I_\infty(p) := \log d_p - H_\infty(p)}.$$

Interpretation: This is the amount of work that we have to invest to create p perfectly (exactly / with zero probability of failure) from "nothing" (i.e. from maximally mixed states), by

- using Landauer erasure to prepare pure bits;
- using noisy operations to prepare p from those.

It is easy to show that

$$0 \leq H_\infty(p) \leq H_0(p) \leq \log d_p$$

$$\Rightarrow \quad 0 \leq I_0(p) \leq I_\infty(p) \leq \log d_p$$

$\Rightarrow$ In general, it costs more work to create a state than what we can gain / extract from that state.

What about "failure probability" etc.? $\to$ next step.

2.7. Closeness of (quantum) states: the trace distance

Def.: For two density matrices $\rho, \sigma \in B(\mathbb{C}^n)$, we define their trace distance as

$$D(\rho,\sigma) := \frac{1}{2} \operatorname{tr} |\rho - \sigma|$$

This is a special case of the "Schatten p-norm"

$$\|M\|_p := \left[\mathrm{tr}(|M|^p)\right]^{1/p} \quad [redacted] \quad (p \geq 1);$$

$$D(\rho,\sigma) = \frac{1}{2}\|\rho-\sigma\|_1 .$$

This is a distance [redacted] measure:

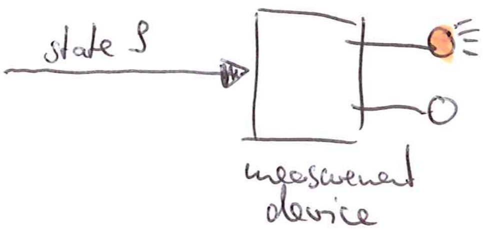
$$0 \leq D(\rho,\sigma) \leq 1,$$

$$D(\rho,\sigma) = 0 \iff \rho = \sigma,$$

$$D(\rho,\sigma) = 1 \iff \rho \text{ and } \sigma \text{ are orthogonal, i.e. } \rho\sigma = 0$$

$$D(\rho,\tau) \; [redacted] \leq D(\rho,\sigma) + D(\sigma,\tau)$$

Operational meaning as <u>distinguishability</u> of ρ and σ:
Two-outcome projective measurement:



$\mathrm{Prob}(\text{1st outcome}) = \mathrm{tr}(\rho P)$
$\quad " \quad (\text{2nd} \quad " \quad) = \mathrm{tr}(\rho P^\perp),$
$\quad P^\perp = \mathbb{1} - P.$

<u>Lemma</u>:
$$D(\rho,\sigma) = \max_{P \text{ orthogonal projector}} \left| \mathrm{tr}(\rho P) - \mathrm{tr}(\sigma P) \right|$$
$$= \max_{P \text{ orth. proj.}} \mathrm{tr}\left[P(\rho-\sigma)\right]$$
$$= \max_{0 \leq E \leq \mathbb{1}} \mathrm{tr}\left[E(\rho-\sigma)\right].$$

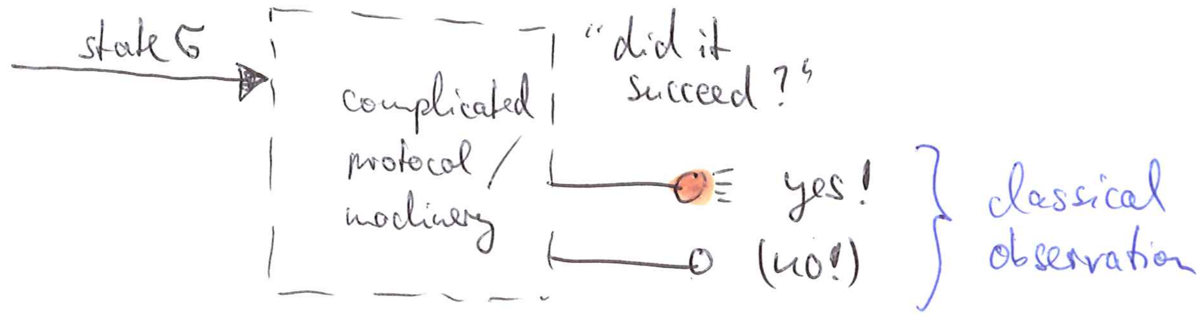
Discuss meaning.

Quantum operations $\mathcal{N}$ (e.g. open-system evolutions) only ever <u>decrease</u> the distinguishability:

$$D(\mathcal{N}(\rho), \mathcal{N}(\sigma)) \leq D(\rho,\sigma).$$

$$D(U\rho U^\dagger, U\sigma U^\dagger) = D(\rho,\sigma).$$
unitary evolution preserves trace distance.

What about "success probability"?



There is some operation element $0 \le E \le \mathbb{1}$ such that $\text{Prob}(\text{success}) = \text{tr}(E\sigma)$.

If protocol <u>always</u> succeeds on state ρ, i.e. $\text{Prob}(\text{success} \mid \rho) = 1$, but we have instead $\rho' \neq \rho$ then

$$\text{Prob}(\text{success} \mid \rho') \ge 1 - D(\rho, \rho').$$

Classical analog: the variation distance.

$[\rho, \sigma] = 0 \Rightarrow \rho = \begin{pmatrix} p_1 & & \\ & \ddots & \\ & & p_n \end{pmatrix}$, $\sigma = \begin{pmatrix} q_1 & & \\ & \ddots & \\ & & q_n \end{pmatrix}$ in some basis

$$\Rightarrow D(\rho, \sigma) = \frac{1}{2} \text{tr} |\rho - \sigma| = \frac{1}{2} \text{tr} \left| \begin{pmatrix} p_1 - q_1 & & \\ & \ddots & \\ & & p_n - q_n \end{pmatrix} \right| = \frac{1}{2} \text{tr} \begin{pmatrix} |p_1 - q_1| & & \\ & \ddots & \\ & & |p_n - q_n| \end{pmatrix}$$

$$= \frac{1}{2} \sum_{i=1}^{n} |p_i - q_i|$$

$\Rightarrow$ For prob. vectors $p, q \in \mathbb{R}^n$,

$$\boxed{D(p,q) := \frac{1}{2} \sum_{i=1}^{n} |p_i - q_i|} = \frac{1}{2} \|p - q\|_1 \text{, special case}$$

of ℓ_p-norm $\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$.

<u>Lemma</u>: $D(p,q) = \max_{S \subseteq \{1,..,n\}} |p(S) - q(S)| = \max_{S \subseteq \{1,..,n\}} (p(S) - q(S))$

where [redacted] $p(S) := \sum_{i \in S} p_i$.

Same "distinguishability" interpretation as in the quantum case.

2.8. Approximate state formation and distillation of pure bits

Write $P \xrightarrow{\varepsilon\text{-noisy}} Q$ if there exists a prob. vector $\tilde{Q}$ with $D(Q, \tilde{Q}) \leq \varepsilon$ such that $P \xrightarrow{\text{noisy}} \tilde{Q}$.

Consistency of quantum and classical case:

<u>Lemma</u>: For quantum states ρ, σ, there exists a noisy (quantum) operation $\mathcal{N}$ and another state $\tilde{\sigma}$ with $D(\sigma, \tilde{\sigma}) \leq \varepsilon$ and $\mathcal{N}(\rho) = \tilde{\sigma}$ if and only if $\lambda(\rho) \xrightarrow{\varepsilon\text{-noisy}} \lambda(\sigma)$.

Proof not here.

<u>Definition</u> ("smooth entropies"):

For a prob. vector p and $\varepsilon \geq 0$, define

$$H_\infty^\varepsilon(p) := \max_{p' : D(p,p') \leq \varepsilon} H_\infty(p'), \quad \text{and}$$

$$H_0^\varepsilon(p) := \min_{p' : D(p,p') \leq \varepsilon} H_0(p').$$

Smooth entropies have <u>many</u> applications in classical and quantum information theory.