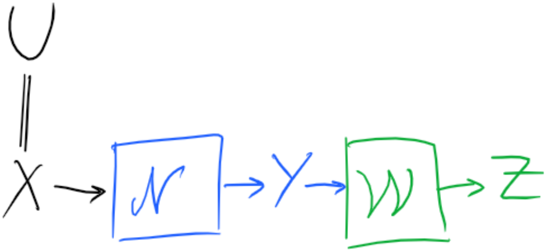
Reverse Data-Processing Theorems, Bayesian Structures, and the Flow of Information

Francesco Buscemi (Nagoya University)

Quantum Foundations Workshop
Dip. di Fisica, Università degli Studi di Pavia
21 June 2016

what is a data-processing theorem?

let $\mathcal{N} : \mathscr{X} \to \mathscr{Y}$ and $\mathcal{W} : \mathscr{Y} \to \mathscr{Z}$ be two noisy channels



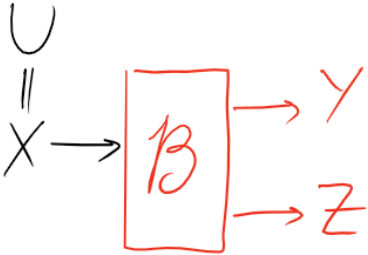
then, the joint distributions $(U, Y) \equiv (U, \mathcal{N}(X))$ and $(U, Z) \equiv (U, \mathcal{W}(Y))$ are such that

$$H(U|Y) \leqslant H(U|Z) \quad \text{for all initial} \quad (U, X)$$

notice that $H(U|Y) \leqslant H(U|Z) \iff H(U) - H(U|Y) \geqslant H(U) - H(U|Z) \iff I(U;Y) \geqslant I(U;Z)$

what is a reverse data-processing theorem?

let $\mathcal{B} : \mathscr{X} \to (\mathscr{Y}, \mathscr{Z})$ be a noisy broadcast channel



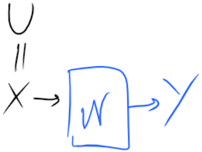
- ✔ assume that the final distribution $(U, Y, Z) \equiv (U, \mathcal{B}(X))$ is such that

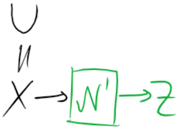
$$H(U|Y) \leqslant H(U|Z) \quad \text{for all initial} \quad (U, X)$$

- ✔ can we then conclude that there exists a noisy channel $\mathcal{W} : \mathscr{Y} \to \mathscr{Z}$ such that $(U, Z) = (U, \mathcal{W}(Y))$ for all initial (U, X)?
- ✔ no (Körner and Marton, 1977)

a useful hierarchy of conditions

Consider two noisy channels $\mathcal{N} : \mathscr{X} \to \mathscr{Y}$ and $\mathcal{N}' : \mathscr{X} \to \mathscr{Z}$





Körner and Marton (1977) introduce the following definitions:

- ✔ $\mathcal{N}$ is degradable into $\mathcal{N}'$ if there exists a noisy channel $\mathcal{W} : \mathscr{Y} \to \mathscr{Z}$ such that

$$\mathcal{N}' = \mathcal{W} \circ \mathcal{N}$$

- ✔ $\mathcal{N}$ is less noisy than $\mathcal{N}'$ if

$$H(U|Y) \leqslant H(U|Z) \quad \text{for all initial} \quad (U, X)$$

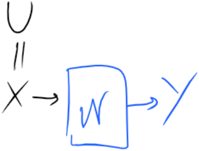
- ✔ $\mathcal{N}$ is more capable than $\mathcal{N}'$ if

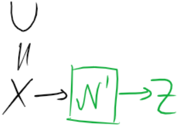
$$H(X|Y) \leqslant H(X|Z) \quad \text{for all initial} \quad X$$

- ✔ **fact**: degradable $\underset{\not\Leftarrow}{\Longrightarrow}$ less noisy $\underset{\not\Leftarrow}{\Longrightarrow}$ more capable

a reverse that works

Again, take two noisy channels $\mathcal{N}:\mathscr{X}\to\mathscr{Y}$ and $\mathcal{N}':\mathscr{X}\to\mathscr{Z}$





Question: when can we say that $\mathcal{N}$ is degradable into $\mathcal{N}'$, i.e., that there exists channel $\mathcal{W}$ such that $\mathcal{W}\circ\mathcal{N}=\mathcal{N}'$?

Körner-Marton (1977)

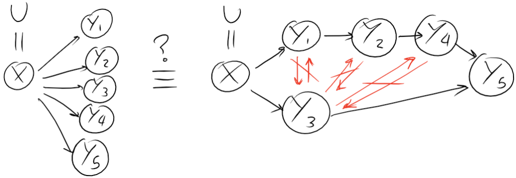
$$H(U|Y)\leqslant H(U|Z) \text{ for all initial } (U,X)$$
$$\not\Longrightarrow$$
$$\Longleftarrow$$
$$\exists\mathcal{W}:\ \mathcal{W}\circ\mathcal{N}=\mathcal{N}'$$

Theorem (2016)

$$H_{\min}(U|Y)\leqslant H_{\min}(U|Z) \text{ for all initial } (U,X)$$
$$\Longleftrightarrow$$
$$\exists\mathcal{W}:\ \mathcal{W}\circ\mathcal{N}=\mathcal{N}'$$

- ✔ $H_{\min}(U|Y)=-\log_2 P_{\text{guess}}(U|Y)=-\log_2\sum_y \max_u p(u,y)$
- ✔ it also holds in the quantum case
- ✔ it also holds approximately

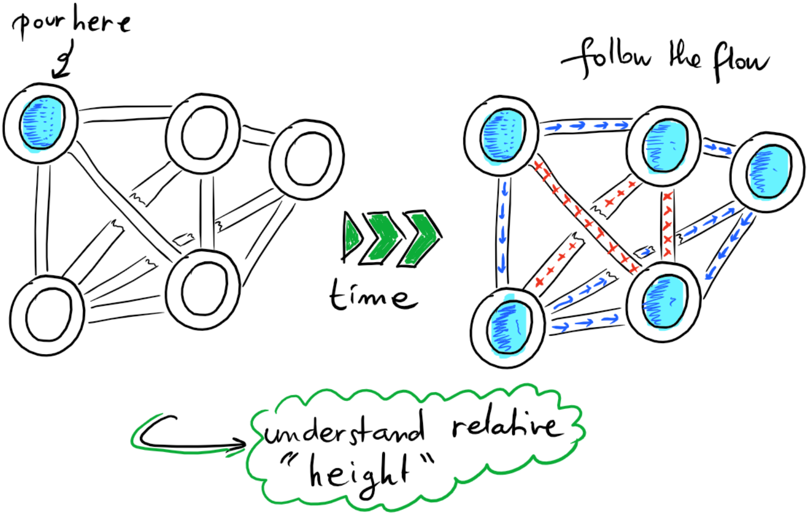
- suppose that, given a conditional probability $p(y_1, y_2, \cdots, y_N|x)$, we want to find stochastic dependencies between these variables



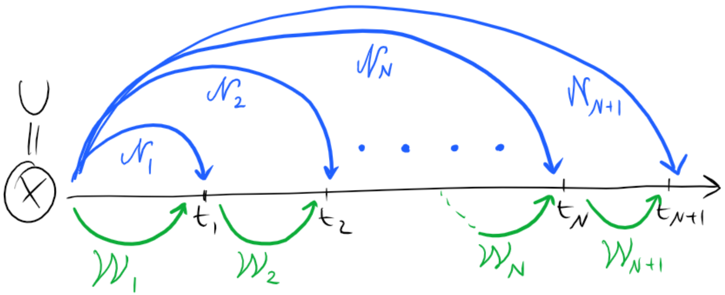
- reverse data-processing theorem: a path exists if and only if $H_{\min}$ never decreases (equivalently, no path exists if and only if $H_{\min}$ strictly decreases at some point for some initial conditions)
- stochastic dependencies follow the "flow" of $H_{\min}$
- partial ordering $\implies$ incomparable paths
- for example, the figure above is equivalent to the following entropic conditions:

$$\begin{cases} H_{\min}(U|Y_1) \leqslant H_{\min}(U|Y_2) \leqslant H_{\min}(U|Y_4) \leqslant H_{\min}(U|Y_5), & \text{for all initial } (U,X) \\ H_{\min}(U|Y_3) \leqslant H_{\min}(U|Y_5), & \text{for all initial } (U,X) \\ H_{\min}(U|Y_3) \lessgtr \{H_{\min}(U|Y_1), H_{\min}(U|Y_2), H_{\min}(U|Y_4)\}, & \text{for some initial } (U,X) \end{cases}$$

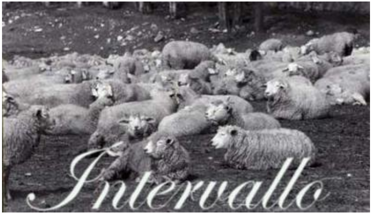
a picture (information as a “fluid”)



Again: while height is a total ordering, the info-ordering is only a partial ordering



- divisibility: $\mathcal{N}_i = \mathcal{W}_i \circ \mathcal{W}_{i-1} \circ \cdots \circ \mathcal{W}_1$
- a dynamical map is divisible if and only if the sequence $\{H_{\min}(U|Y_i)\}_{i\geqslant 1}$ is non-decreasing for all initial (U, X)
- namely, divisibility is equivalent to "no $H_{\min}$ backflow"
- the same insight holds also in the quantum case (Buscemi, Datta; PRA 2016) and approximately (Buscemi, Prob. Inf. Trans. 2016; Jencova, ISIT 2016)



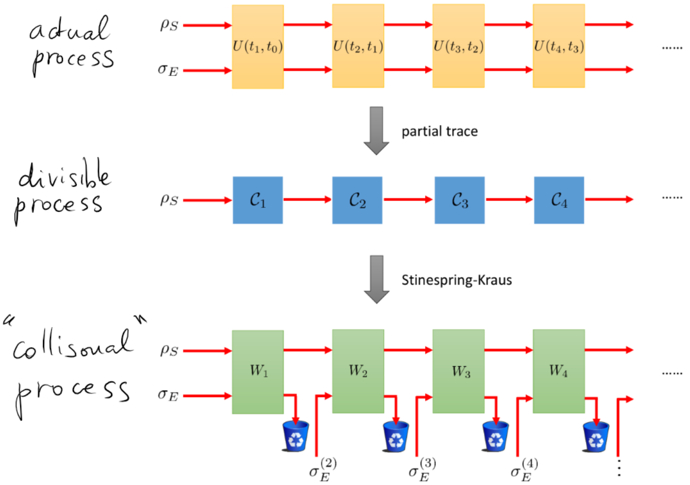
Reverse Data-Processing Theorems

- ✔ Data-Processing as a principle
- ✔ $H_{\min}(U|X_1) \leqslant H_{\min}(U|X_2)$, for U, X_1, X_2 random variables
- ✔ Equivalent to existence of a memoryless process
- ✔ No information backflows

Lieb and Yngvasson (1999)

- ✔ Second Law as a principle
- ✔ $S(X_1) \leqslant S(X_2)$, for X_1, X_2 thermodynamical equilibrium states
- ✔ Equivalent to existence of an adiabatic process
- ✔ No heat flows

S: system, E: environment, $S+E$: conservative

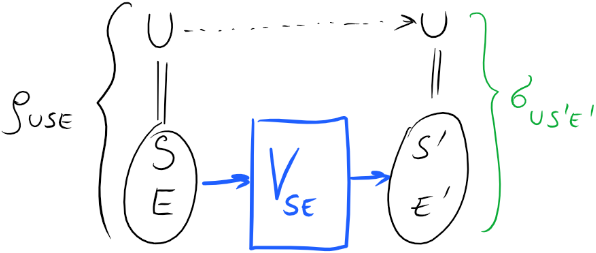


However, the given process need not be collisional to be divisible, i.e., there are system-environment correlations that do not break divisibility. **Question**: how to characterize such correlations?

- ✔ experimentally, the initial factorization condition is an approximation (and strong-coupling regimes are of interest)
- ✔ what happens to the reduced dynamics in the presence of initial system-environment correlations?
- ✔ Pechukas (PRL, 1994): "Here we show that complete positivity is an artifact of product initial conditions. In general, reduced dynamics need not be CP"
- ✔ Lindblad (J. Phys. A, 1995) and Alicki (PRL, 1995, comment to Pechukas)
- ✔ Rodriguez-Rosario, Modi, Kuah, Sudarshan (2008): null discord $\Longrightarrow$ divisibility
- ✔ Shabani–Lidar (2009): null discord $\Longleftrightarrow$ divisibility (Erratum 2016)
- ✔ Brodutch, Datta, Modi, Rivas, Rodriguez-Rosario (2013): divisibility $\not\Longrightarrow$ null discord
- ✔ FB (2014): what follows

guiding idea: data-processing as a principle

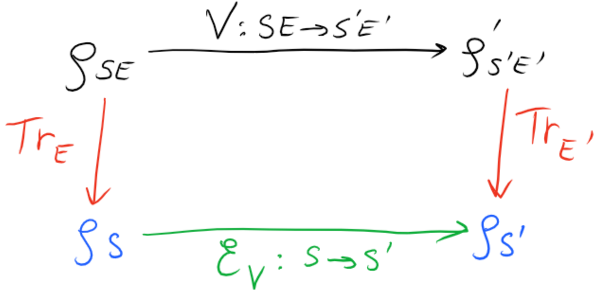
- ✔ system and environment start in a correlated state: they have to be considered as one composite system
- ✔ reference–system–environment: $\mathcal{H}_U \otimes \mathcal{H}_S \otimes \mathcal{H}_E$
- ✔ the initial condition is given as a tripartite density operator ρ_{USE}



- ✔ the state ρ_{USE} is "Markovian" if $H(U|S)_\rho \leqslant H(U|S')_\sigma$ for all isometries V
- ✔ **first fact**: ρ_{USE} Markovian $\iff$ $I(U;E|S)_\rho = 0$
- ✔ **second fact**: ρ_{USE} Markovian $\iff$ $\forall V_{SE}, \exists \mathcal{E}_S : \sigma_{US'} = (\mathrm{id}_U \otimes \mathcal{E}_S)(\rho_{US})$
- ✔ therefore, Markovian $\iff$ no backflows of information

conventional approach

- ✔ no reference system; instead, given is a *family* $\mathscr{S} = \{\rho_{SE} : \rho_{SE} \in \mathscr{S}\}$ of possible initial joint system–environment states



- ✔ in the above diagram, ρ_{SE} is a generic element of $\mathscr{S}$ (i.e., the condition must hold for all $\rho_{SE} \in \mathscr{S}$)
- ✔ if the above diagram holds for all isometries V, we say that the family $\mathscr{S}$ is *Markovian*
- ✔ example: the initial factorization condition, i.e., $\mathscr{S} = \{\rho_S \otimes \bar{\xi}_E : \rho_S \text{ any state of } S\}$ for some fixed environment state $\bar{\xi}_E$, defines a Markovian family

relations

- ✔ how to connect the tripartite with the bipartite scenario?
- ✔ steering: $\rho_{SE}^{\Pi} = \frac{\mathrm{Tr}_U[\rho_{USE}\ (\Pi_U \otimes \mathbb{1}_{SE})]}{\mathrm{Tr}[\rho_{USE}\ (\Pi_U \otimes \mathbb{1}_{SE})]}$ for some $\Pi > 0$
- ✔ we say that the family $\mathscr{S}$ is steerable if there exists a ρ_{USE} such that:
  1. for all $\rho_{SE} \in \mathscr{S}$ there exists a $\Pi_U > 0$ such that the above holds, and
  2. for all $\Pi_U > 0$, the steered state ρ_{SE}^{Π} is in $\mathscr{S}$
- ✔ example: if $\mathscr{S}$ is a polytope (e.g., bipartite cq-states) then it's steerable
- ✔ example: the family $\mathscr{S} = \{\rho_S \otimes \bar{\xi}_E\}$ corresponding to the initial factorization condition is steerable
- ✔ **main fact**: bipartite family $\mathscr{S}$ is steerable and Markovian $\iff$ it can be steered from Markovian ρ_{USE}, i.e., $I(U;E|S)_\rho = 0$
- ✔ **first corollary**: all previous cases (they all happen to be steerable Markovian families)
- ✔ many other more general constructions are possible too $\implies$ no direct connection between "strength/character of initial correlations" and "existence of CPTP reduced dynamics"
- ✔ **second corollary**: assume $\mathscr{S}$ Markovian and $\mathrm{Tr}_E[\mathscr{S}]$ complete (contains all states on $\mathcal{H}_S$) $\implies$ initial factorization condition
- ✔ **extra goody**: we can apply all the tools recently developed for approximate recoverability (family is approximately steerable, family is approximately Markov, tripartite state is approximately Markov, etc)

summary

- ✔ data-processing theorem: if there is a process, information always decreases
- ✔ reverse data-processing theorem: if information *always* decreases, then there exists a process
- ✔ data-processing inequality as a "physical principle": the flow of information determines the evolution
- ✔ no backflows of information + completeness $\iff$ initial factorization condition
- ✔ analogy with strong (i.e., necessary and sufficient) second law-like statements (e.g., Lieb–Yngvasson formulation of adiabaticity)
- ✔ work in progress: applications to generalized resource theories

è finita la comedia