On characterizing classical and quantum entropy

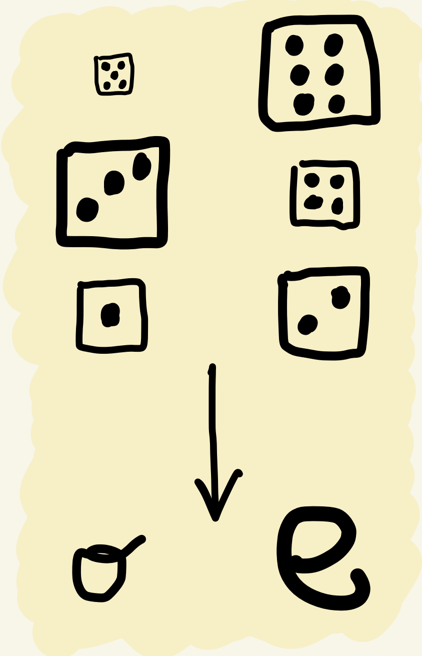
Arthur Parzygnat IHÉS

Categorical semantics of Entropy

CUNY Graduate Center ITS

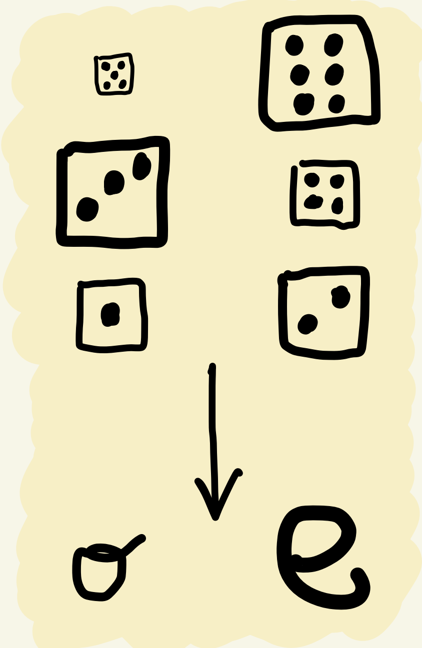
5/13/2022

Information Loss



Every (non-injective) function of probability spaces loses information.

Information Loss



$$\begin{array}{c} (X, p) \\ \Big\downarrow f \\ (Y, q) \end{array}$$

Every (non-injective) function of probability spaces loses information. The entropy difference

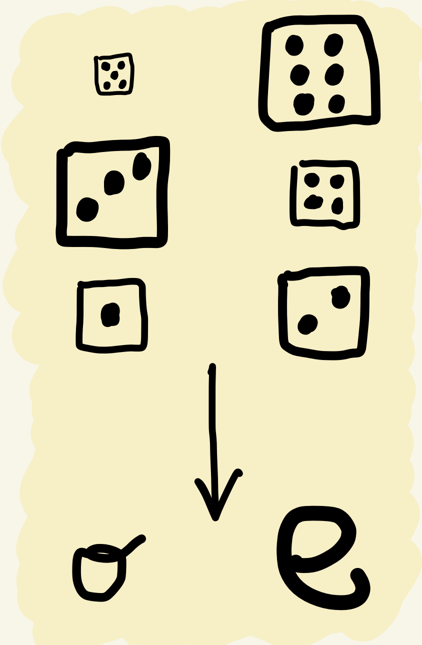
$$\Delta_H(f) := H(p) - H(q), \text{ where}$$

$$H(p) := -\sum_x p_x \log(p_x)$$

is the Shannon entropy, quantifies this.

Here, $q_y = \sum_{x \in f^{-1}(\{y\})} p_x$ is the pushforward of p along f, denoted $f \circ p$.

Information Loss



$$\begin{array}{c}(X,p)\\ \big\downarrow f\\ (Y,q)\end{array}$$

Here, $q_y = \sum_{x \in f^{-1}(\{y\})} p_x$ is the pushforward of p along f, denoted $f \circ p$.

Every (non-injective) function of probability spaces loses information. The entropy difference

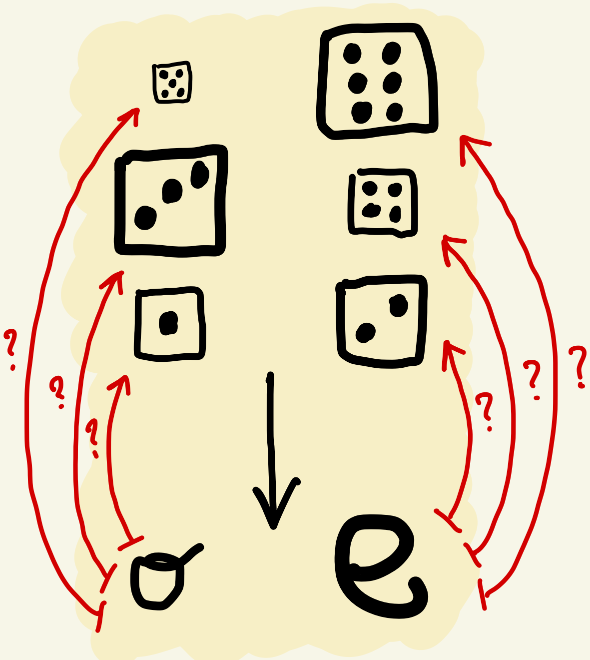
$$\Delta_H(f) := H(p) - H(q), \text{ where}$$

$$H(p) := -\sum_x p_x \log(p_x)$$

is the Shannon entropy, quantifies this.

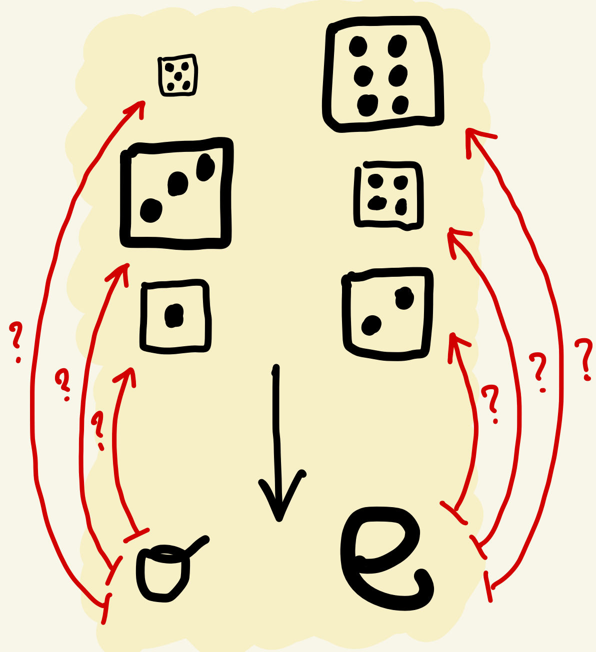
Baez, Fritz, Leinster characterized Δ_H as a continuous convex functor into $\mathbb{B}\mathbb{R}_{\geq 0}$ ($\mathbb{R}_{\geq 0}$ viewed as a one-object category) up to a constant ≥ 0.

Information Recovery



It is impossible to recover information in a deterministic manner!

Information Recovery



It is impossible to recover information in a deterministic manner!

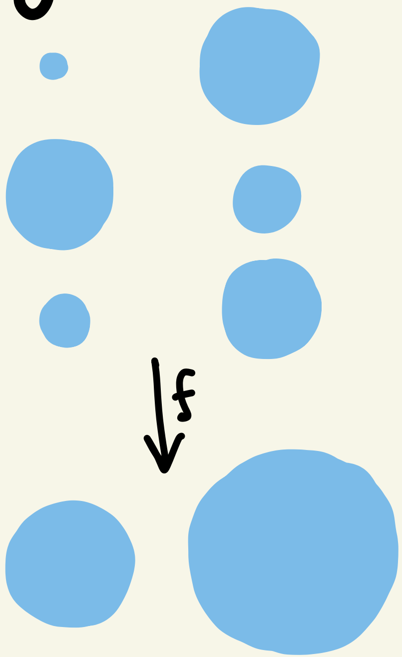
But we can try using a stochastic procedure $Y \overset{h}{\rightsquigarrow} X$.

Notation

h assigns a probability distribution h_y on X to each $y \in Y$.
The value on $x \in X$ is written h_{xy}.

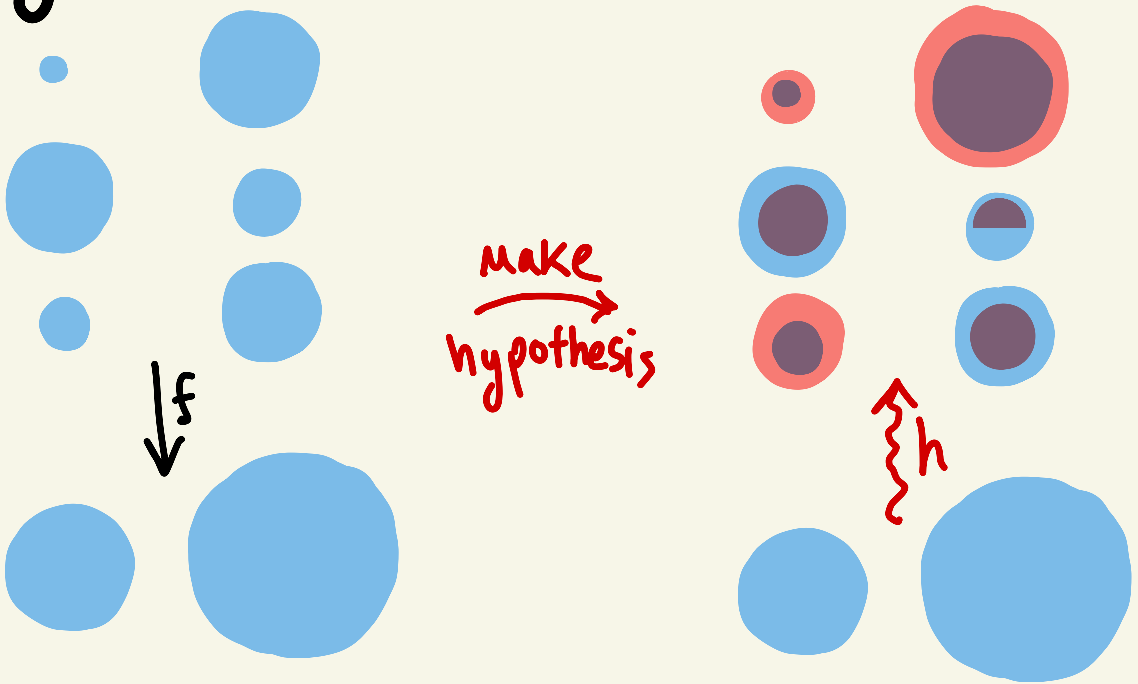
Compare to $\mathbb{P}(x \mid y)$ the conditional probability of x given y.

Hypotheses = stochastic sections



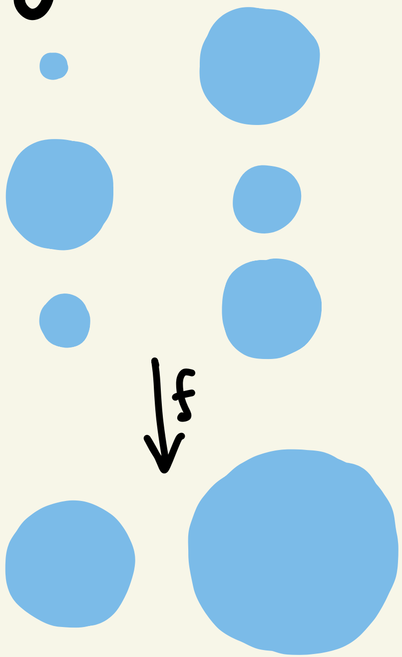
Water droplet picture on left due to Gromov

Hypotheses = stochastic sections

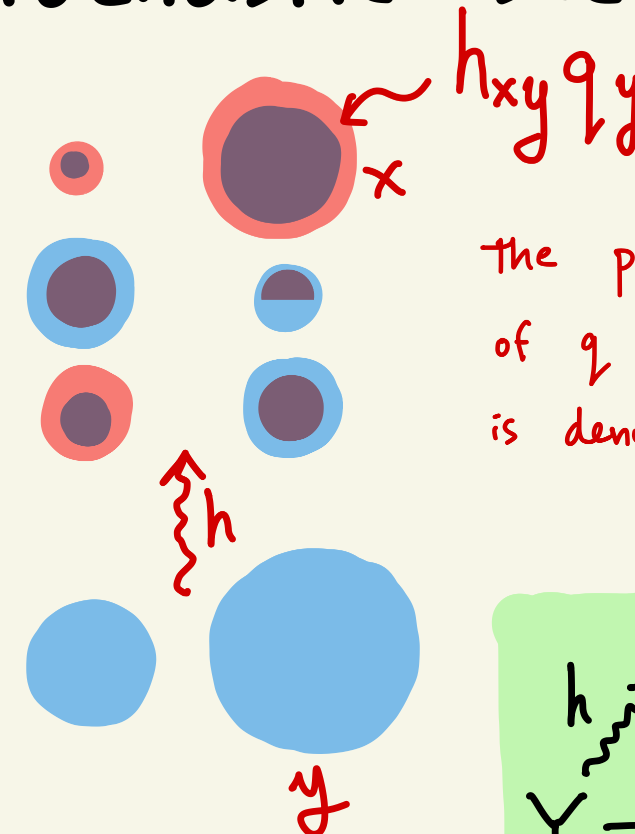


Water droplet picture on left due to Gromov

Hypotheses = stochastic sections



$\xrightarrow{\text{make}}$ hypothesis

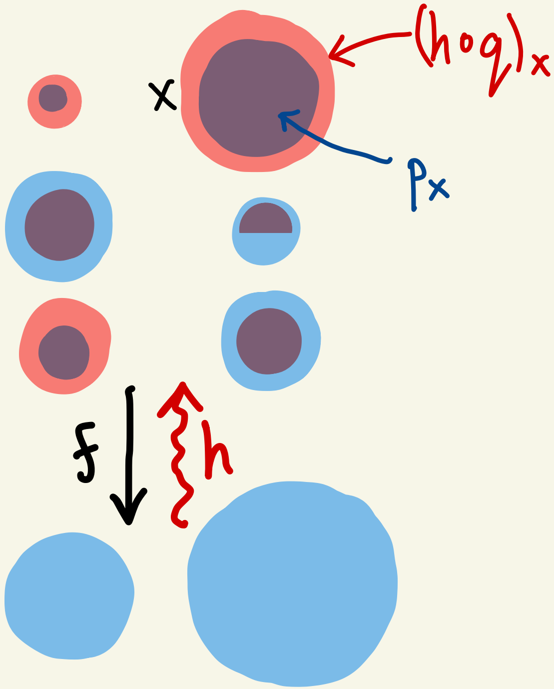


the pushforward of q along h is denoted $h \circ q$.

$$\begin{tikzcd} & X \arrow[dr, "f"] & \\ Y \arrow[ur, squiggly, "h"] \arrow[rr, "\mathrm{id}_Y"'] \arrow[rr, phantom, bend left=20, "="] & & Y \end{tikzcd}$$

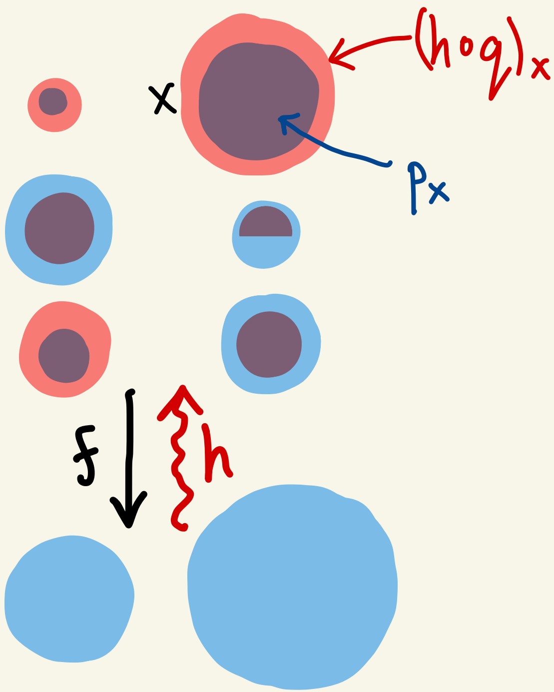
Water droplet picture on left due to Gromov

Relative Entropy of Recovery



But hypotheses are not always correct.

Relative Entropy of Recovery



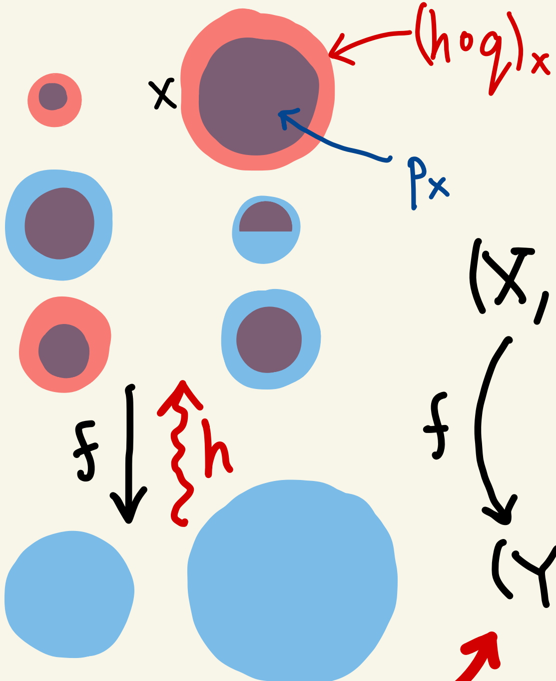
But hypotheses are not always correct. To quantify the accuracy of a hypothesis h, we use the relative entropy (of recovery)

$$S(p\,\|\,h\circ q) := \sum_x p_x \log\left(\frac{p_x}{(h\circ q)_x}\right)$$

$$= \sum_x p_x \log\left(\frac{p_x}{h_{x f(x)}\, q_{f(x)}}\right)$$

h is sometimes called a recovery map

Relative Entropy of Recovery



$$\begin{tikzcd} (X, p) \arrow[d, bend right, "f"'] \\ (Y, q) \arrow[u, squiggly, bend right, "h"'] \end{tikzcd}$$

But hypotheses are not always correct. To quantify the accuracy of a hypothesis h, we use the relative entropy (of recovery)

$$S(p \,\|\, h \circ q) := \sum_x p_x \log\left(\frac{p_x}{(h \circ q)_x}\right)$$

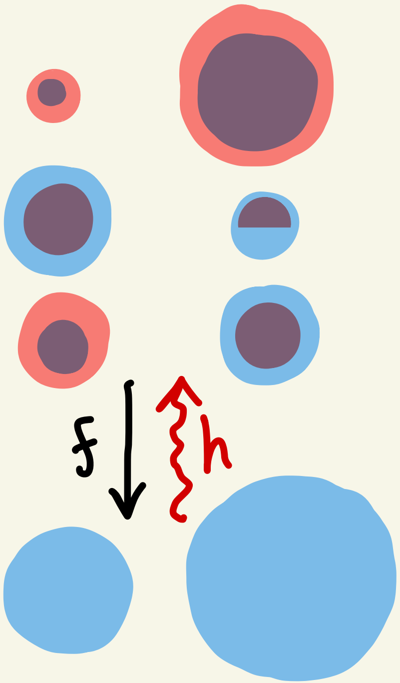
$$= \sum_x p_x \log\left(\frac{p_x}{h_{x f(x)}\, q_{f(x)}}\right)$$

h is sometimes called a recovery map

Warning: This notation is misleading.

we require $q = f \circ p$ but <u>NOT</u> $p = h \circ q$!

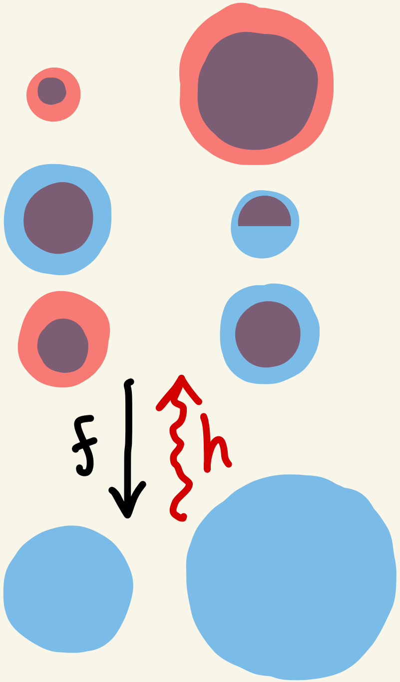
Relative Entropy & Optimal Hypotheses



The relative entropy satisfies

- $S(p \| h \circ q) \geq 0$

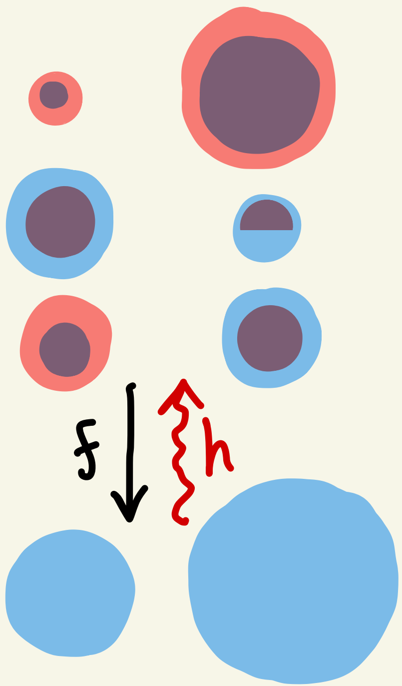
Relative Entropy & Optimal Hypotheses



The relative entropy satisfies

- $S(p \| h \circ q) \geq 0$
- $S(p \| h \circ q) = 0$ iff $h \circ q = p$

Relative Entropy & Optimal Hypotheses

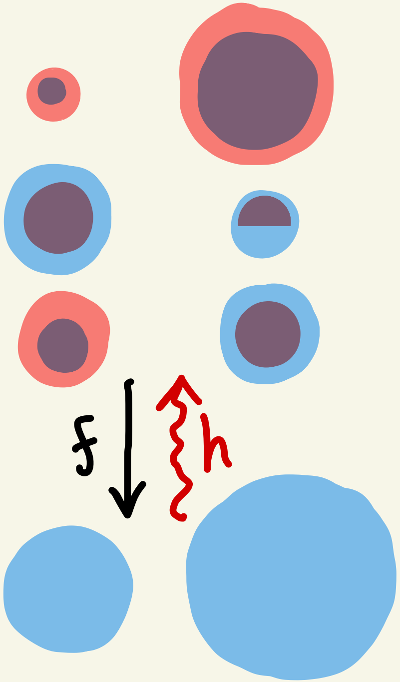


The relative entropy satisfies

- $S(p \| h \circ q) \geq 0$
- $S(p \| h \circ q) = 0$ iff $h \circ q = p$

An optimal hypothesis is a hypothesis h such that $h \circ q = p$.

Relative Entropy & Optimal Hypotheses



The relative entropy satisfies

- $S(p \| h \circ q) \geq 0$
- $S(p \| h \circ q) = 0$ iff $h \circ q = p$

An optimal hypothesis is a hypothesis h such that $h \circ q = p$.

<u>Fact</u> Optimal hypotheses always exist.

$$h_{xy} = \frac{\delta_{y f(x)} \, p_x}{q_y}, \quad \text{where } \delta_{x'x} = \begin{cases} 1 & \text{if } x' = x \\ 0 & \text{o.w.} \end{cases}$$

A Category of Hypotheses

The following category **FinStat** was constructed by Baez and Fritz.

A Category of Hypotheses

The following category FinStat was constructed by Baez and Fritz.

objects: pairs (X, p), (Y, q), (Z, r) finite probability spaces

A Category of Hypotheses

The following category **FinStat** was constructed by Baez and Fritz.

objects: pairs (X,p), (Y,q), (Z,r) finite probability spaces

morphisms:

$$\begin{tikzcd} (X,p) \arrow[r, bend left, "f"] & (Y,q) \arrow[l, bend left, squiggly, "h"] \end{tikzcd}$$

f = deterministic

$q = f \circ p$

$f \circ h = \mathrm{id}_Y$

(BUT NOT $p = h \circ q$)

A Category of Hypotheses

The following category **FinStat** was constructed by Baez and Fritz.

objects: pairs (X,p), (Y,q), (Z,r) finite probability spaces

morphisms:
$$\begin{tikzcd}
(X,p) \arrow[r, bend left, "f"] & (Y,q) \arrow[l, squiggly, bend left, "h"]
\end{tikzcd}$$

f = deterministic
$q = f \circ p$
$f \circ h = \mathrm{id}_Y$

(BUT $\underline{\text{NOT}}$ $p = h \circ q$)

composition:
$$\begin{tikzcd}
(X,p) \arrow[r, bend left, "f"] \arrow[rr, bend left=60, "g \circ f"] & (Y,q) \arrow[l, squiggly, bend left, "h"] \arrow[r, bend left, "g"] & (Z,r) \arrow[l, squiggly, bend left, "k"] \arrow[ll, squiggly, bend left=60, "h \circ k"]
\end{tikzcd}$$

$$(h \circ k)_{xz} := \sum_y h_{xy} k_{yz}$$

Chapman–Kolmogorov

Relative Entropy as a Functor

$$\mathrm{FinStat} \xrightarrow{RE} \mathbb{B}[0,\infty]$$

$\mathbb{B}[0,\infty]$ is the category with a single object and the morphism set is $[0,\infty]$ with addition as the composition.

Relative Entropy as a Functor

$$\mathrm{FinStat} \xrightarrow{RE} \mathbb{B}[0,\infty]$$

$$\begin{tikzcd} (X,p) \arrow[d, bend right, "f"'] \\ (Y,q) \arrow[u, bend right, squiggly, "h"'] \end{tikzcd} \longmapsto S(p \,\|\, h \circ q)$$

$\mathbb{B}[0,\infty]$ is the category with a single object and the morphism set is $[0,\infty]$ with addition as the composition.

Relative Entropy as a Functor

$$\mathrm{FinStat} \xrightarrow{RE} \mathbb{B}[0,\infty]$$

$$\begin{tikzcd} (X,p) \arrow[d, bend right, "f"'] \\ (Y,q) \arrow[u, squiggly, bend right, "h"'] \end{tikzcd} \longmapsto S(p \,\|\, h \circ q)$$

$\mathbb{B}[0,\infty]$ is the category with a single object and the morphism set is $[0,\infty]$ with addition as the composition.

is the unique lower semi-continuous convex functor that vanishes on the subcategory FP of optimal hypotheses (up to a non-negative constant).

This is the main theorem of Baez and Fritz.

Functoriality of Relative Entropy

$$
\begin{tikzcd}
(X,p) \arrow[r, bend left, "f"] \arrow[rr, bend left=60, "g \circ f"] & (Y,q) \arrow[l, squiggly, bend left, "h"] \arrow[r, bend left, "g"] & (Z,r) \arrow[l, squiggly, bend left, "k"] \arrow[ll, squiggly, bend left=60, "h \circ k"]
\end{tikzcd}
$$

$$S(p \,\|\, h \circ q) + S(q \,\|\, k \circ r) = S(p \,\|\, h \circ k \circ r)$$

Functoriality of Relative Entropy

$$
\begin{tikzcd}
(X,p) \arrow[r, bend left, "f"] \arrow[rr, bend left=50, "g\circ f"] & (Y,q) \arrow[l, squiggly, bend left, "h"] \arrow[r, bend left, "g"] & (Z,r) \arrow[l, squiggly, bend left, "k"] \arrow[ll, squiggly, bend left=50, "h\circ k"]
\end{tikzcd}
$$

$$S(p\,\|\,h\circ q) + S(q\,\|\,k\circ r) = S(p\,\|\,h\circ k\circ r)$$

Special case of functoriality 1:

$$
\begin{tikzcd}
(X,q) \arrow[r, bend left, "f"] \arrow[rr, bend left=50, "!_X"] & (Y,f\circ q) \arrow[l, squiggly, bend left, "g"] \arrow[r, bend left, "!_Y"] & (\bullet,1) \arrow[l, squiggly, bend left, "f\circ p"] \arrow[ll, squiggly, bend left=50, "p"]
\end{tikzcd}
$$

f ← fixed

g ← optimal hypothesis of (f,p)

p ← fixed

arbitrary → q

Functoriality of Relative Entropy

$$
\begin{tikzcd}
(X,p) \arrow[r, "f"] \arrow[rr, bend left=40, "g\circ f"] & (Y,q) \arrow[r, "g"] \arrow[l, squiggly, bend left, "h"] & (Z,r) \arrow[l, squiggly, bend left, "K"] \arrow[ll, squiggly, bend left=40, "h\circ K"]
\end{tikzcd}
$$

$$S(p \,\|\, h\circ q) + S(q \,\|\, K\circ r) = S(p \,\|\, h\circ K\circ r)$$

Special case of functoriality 1:

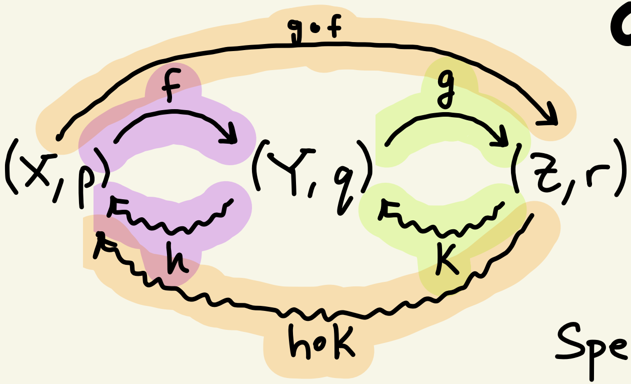
$$
\begin{tikzcd}
(X,q) \arrow[r, "f \leftarrow \text{fixed}"] \arrow[rr, bend left=40, "!_X"] & (Y,f\circ q) \arrow[r, "!_Y"] \arrow[l, squiggly, bend left, "g \leftarrow \text{optimal hypothesis of } (f,p)"] & (\bullet,1) \arrow[l, squiggly, bend left, "f\circ p"] \arrow[ll, squiggly, bend left=40, "p \leftarrow \text{fixed}"]
\end{tikzcd}
$$

arbitrary

$$S(q \,\|\, g\circ f\circ q) + S(f\circ q \,\|\, f\circ p) = S(q \,\|\, p) \quad \text{i.e.,}$$

$$S(q \,\|\, p) - S(f\circ q \,\|\, f\circ p) = S(q \,\|\, g\circ f\circ q)$$

Functoriality of Relative Entropy



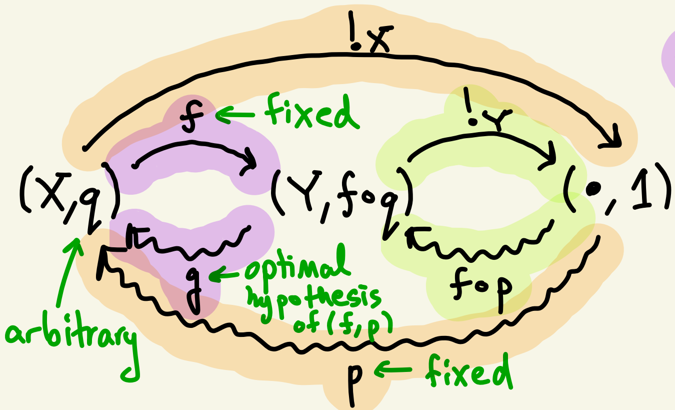
$$S(p \,\|\, h \circ q) + S(q \,\|\, K \circ r) = S(p \,\|\, h \circ K \circ r)$$

Special case of functoriality 1:

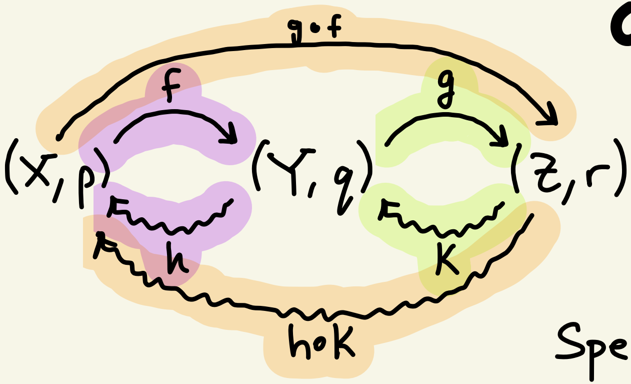
The conditional expectation property

$$S(q \,\|\, g \circ f \circ q) + S(f \circ q \,\|\, f \circ p) = S(q \,\|\, p) \quad \text{i.e.,}$$

$$S(q \,\|\, p) - S(f \circ q \,\|\, f \circ p) = S(q \,\|\, g \circ f \circ q)$$



Functoriality of Relative Entropy



$$S(p\,\|\,h\circ q) + S(q\,\|\,K\circ r) = S(p\,\|\,h\circ K\circ r)$$

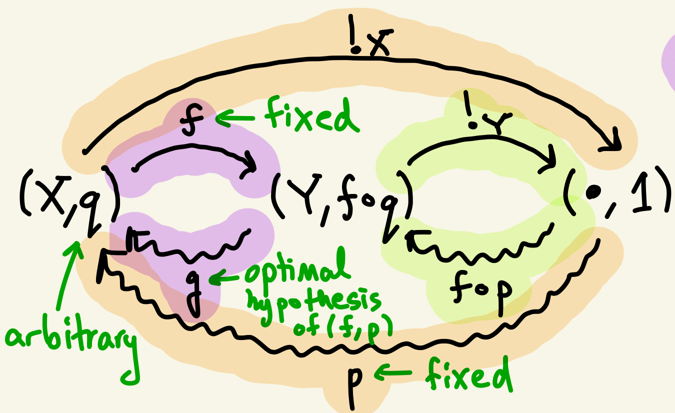
Special case of functoriality 1:

The conditional expectation property

$$S(q\,\|\,g\circ f\circ q) + S(f\circ q\,\|\,f\circ p) = S(q\,\|\,p) \quad \text{i.e.,}$$

$$\underbrace{S(q\,\|\,p) - S(f\circ q\,\|\,f\circ p)} = S(q\,\|\,g\circ f\circ q)$$

This term appears in the data-processing inequality (DPI). Since the RHS ≥ 0, this is an improvement of the DPI.



Functoriality of Relative Entropy

Special case of functoriality 2:

$$
\begin{tikzcd}
(X \times Y \times Z, p_{XYZ}) \arrow[r, bend left, "\pi^{XYZ}_{YZ}"] \arrow[rr, bend left=60, "\pi^{XYZ}_{Z}"] & (Y \times Z, p_{YZ}) \arrow[r, bend left, "\pi^{YZ}_{Z}"] \arrow[l, squiggly, bend left, "u^{YZ}_{XYZ}"] & (Z, p_Z) \arrow[l, squiggly, bend left, "u^{Z}_{YZ}"] \arrow[ll, squiggly, bend left=60, "u^{Z}_{XYZ}"]
\end{tikzcd}
$$

The π's are projections, while the u's are fiberwise uniform distributions.

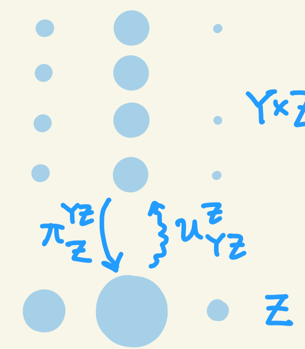
Functoriality of Relative Entropy

Special case of functoriality 2:

$$
\begin{tikzcd}[column sep=huge]
(X\times Y\times Z, p_{XYZ}) \arrow[r, bend left, "\pi^{XYZ}_{YZ}"] \arrow[rr, bend left=60, "\pi^{XYZ}_{Z}"] & (Y\times Z, p_{YZ}) \arrow[r, bend left, "\pi^{YZ}_{Z}"] \arrow[l, squiggly, bend left, "u^{YZ}_{XYZ}"] & (Z, p_Z) \arrow[l, squiggly, bend left, "u^{Z}_{YZ}"] \arrow[ll, squiggly, bend left=60, "u^{Z}_{XYZ}"]
\end{tikzcd}
$$

The π's are projections, while the u's are fiberwise uniform distributions.

For example



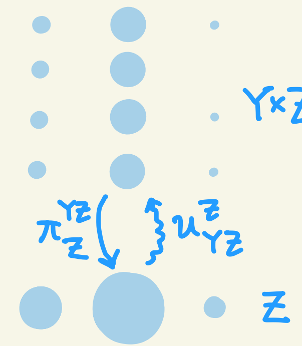
Functoriality of Relative Entropy

Special case of functoriality 2:

$$\begin{tikzcd}[column sep=huge]
(X\times Y\times Z, p_{XYZ}) \arrow[r, bend left, "\pi^{XYZ}_{YZ}"] \arrow[rr, bend left=50, "\pi^{XYZ}_{Z}"] & (Y\times Z, p_{YZ}) \arrow[l, squiggly, bend left, "u^{YZ}_{XYZ}"] \arrow[r, bend left, "\pi^{YZ}_{Z}"] & (Z, p_Z) \arrow[l, squiggly, bend left, "u^{Z}_{YZ}"] \arrow[ll, squiggly, bend left=50, "u^{Z}_{XYZ}"]
\end{tikzcd}$$

The π's are projections, while the u's are fiberwise uniform distributions.

For example



$$S(p_{XYZ} \,\|\, u^{YZ}_{XYZ} \circ p_{YZ}) + S(p_{YZ} \,\|\, u^{Z}_{YZ} \circ p_Z) = S(p_{XYZ} \,\|\, u^{XYZ}_{Z} \circ p_Z)$$

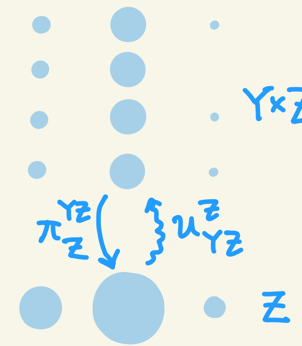
Functoriality of Relative Entropy

Special case of functoriality 2:

$$\begin{tikzcd}
(X\times Y\times Z,\, p_{XYZ}) \arrow[r, bend left, "\pi^{XYZ}_{YZ}"] \arrow[rr, bend left=50, "\pi^{XYZ}_{Z}"] & (Y\times Z,\, p_{YZ}) \arrow[l, squiggly, bend left, "u^{YZ}_{XYZ}"] \arrow[r, bend left, "\pi^{YZ}_{Z}"] & (Z,\, p_Z) \arrow[l, squiggly, bend left, "u^{Z}_{YZ}"] \arrow[ll, squiggly, bend left=50, "u^{Z}_{XYZ}"]
\end{tikzcd}$$

The π's are projections, while the u's are fiberwise uniform distributions.

For example



$$S(p_{XYZ} \,\|\, u^{YZ}_{XYZ}\circ p_{YZ}) + S(p_{YZ} \,\|\, u^{Z}_{YZ}\circ p_Z) = S(p_{XYZ} \,\|\, u^{XYZ}_{Z}\circ p_Z)$$

$$-H(X|YZ) + \log|X| \quad - H(Y|Z) + \log|Y| \quad = \quad -H(XY|Z) + \log|X\times Y|$$

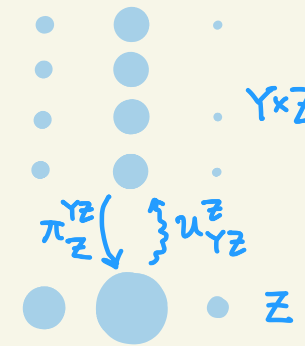
Functoriality of Relative Entropy

Special case of functoriality 2: chain rule for conditional entropy

$$\begin{tikzcd}
(X\times Y\times Z,\, p_{XYZ}) \arrow[r, bend left, "\pi^{XYZ}_{YZ}"] \arrow[rr, bend left=50, "\pi^{XYZ}_{Z}"] & (Y\times Z,\, p_{YZ}) \arrow[l, bend left, squiggly, "u^{YZ}_{XYZ}"] \arrow[r, bend left, "\pi^{YZ}_{Z}"] & (Z,\, p_Z) \arrow[l, bend left, squiggly, "u^{Z}_{YZ}"] \arrow[ll, bend left=50, squiggly, "u^{Z}_{XYZ}"]
\end{tikzcd}$$

The π's are projections, while the u's are fiberwise uniform distributions.

For example



$$S(p_{XYZ} \,\|\, u^{YZ}_{XYZ} \circ p_{YZ}) + S(p_{YZ} \,\|\, u^{Z}_{YZ} \circ p_Z) = S(p_{XYZ} \,\|\, u^{XYZ}_{Z} \circ p_Z)$$

$$-H(X|YZ) + \log|X| \quad -H(Y|Z) + \log|Y| \quad = \quad -H(XY|Z) + \log|X\times Y|$$

$$H(X|YZ) + H(Y|Z) = H(XY|Z)$$

[Ref Cover-Thomas Eqn (2.21)]

A stronger DPI

In our first example of functoriality, we found "given $X \xrightarrow{f} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the optimal hypothesis, then $S(q\|p) - S(f\circ q\|f\circ p) = S(q\|g\circ f\circ q)$ $\forall q$."

A stronger DPI

In our first example of functoriality, we found "given $X \xrightarrow{f} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the optimal hypothesis, then $S(q \| p) - S(f \circ q \| f \circ p) = S(q \| g \circ f \circ q)$ $\forall q$."

More generally, what can we say if f is stochastic?

A stronger DPI

In our first example of functoriality, we found "given $X \xrightarrow{f} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the optimal hypothesis, then $S(q\|p) - S(f\circ q\| f\circ p) = S(q\|g\circ f\circ q)$ $\forall q$."

More generally, what can we say if f is stochastic?

Li and Winter proved (in 2014!)

"given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the Bayesian inverse of (f,p), then $S(q\|p) - S(f\circ q\| f\circ p) \geq S(q\|g\circ f\circ q)$ $\forall q$."

A stronger DPI

In our first example of functoriality, we found
"given $X \xrightarrow{f} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the optimal hypothesis, then $S(q\|p) - S(f\circ q \| f\circ p) = S(q \| g\circ f\circ q)$ $\forall q$."
More generally, what can we say if f is stochastic?

Li and Winter proved (in 2014!)
"given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the Bayesian inverse of (f, p), then $\boxed{S(q\|p) - S(f\circ q \| f\circ p) \geq S(q \| g\circ f\circ q)}$ $\forall q$."
Note that this implies monotonicity of relative entropy

$$S(q\|p) - S(f\circ q \| f\circ p) \geq 0,$$

which is sometimes called the data-processing inequality (DPI).

Bayesian inverses

what's this?

"given $X \xrightarrow{f} Y$ and $\bullet \xrightarrow{p} X$, set $X \xleftarrow{g} Y$ to be the Bayesian inverse of (f, p), then $S(q\|p) - S(f\circ q \| f \circ p) \geq S(q \| g \circ f \circ q)$ $\forall q$."

Bayesian inverses

what's this?

"given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the Bayesian inverse of (f, p), then $S(q \| p) - S(f \circ q \| f \circ p) \geq S(q \| g \circ f \circ q)$ $\forall q$."

<u>Def'n</u> Given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, a Bayesian inverse of (f, p) is a stochastic map $X \overset{g}{\leftsquigarrow} Y$ s.t.

Bayesian inverses

what's this?

"given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the Bayesian inverse of (f, p), then $S(q \| p) - S(f \circ q \| f \circ p) \geq S(q \| g \circ f \circ q)$ $\forall q$."

<u>**Def'n**</u> Given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, a **Bayesian inverse of (f, p)** is a stochastic map $X \overset{g}{\leftsquigarrow} Y$ s.t.

$$\begin{tikzcd}
X \arrow[d, "\Delta_X"'] & \bullet \arrow[l, squiggly, "p"'] \arrow[r, squiggly, "f \circ p"] \arrow[d, phantom, "="] & Y \arrow[d, "\Delta_Y"] \\
X \times X \arrow[r, squiggly, "\mathrm{id}_X \times f"'] & X \times Y & Y \times Y \arrow[l, squiggly, "g \times \mathrm{id}_Y"]
\end{tikzcd}$$

Bayesian inverses

what's this?

"given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the Bayesian inverse of (f,p), then $S(q \| p) - S(f\circ q \| f\circ p) \geq S(q \| g\circ f\circ q)$ $\forall q$."

Def'n Given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, a Bayesian inverse of (f,p) is a stochastic map $X \overset{g}{\leftsquigarrow} Y$ s.t.

$$\begin{tikzcd}
X \arrow[d, "\Delta_X"'] & \bullet \arrow[l, squiggly, "p"'] \arrow[r, squiggly, "f\circ p"] \arrow[d, phantom, "="] & Y \arrow[d, "\Delta_Y"] \\
X\times X \arrow[r, squiggly, "\mathrm{id}_X\times f"'] & X\times Y & Y\times Y \arrow[l, squiggly, "g\times \mathrm{id}_Y"]
\end{tikzcd}$$

i.e., $f_{yx}\, p_x = g_{xy}\, (f\circ p)_y \qquad \forall x \in X, y \in Y$

Bayesian inverses

what's this?

"given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the Bayesian inverse of (f,p), then $S(q\|p) - S(f\circ q \| f\circ p) \geq S(q \| g\circ f\circ q)$ $\forall q$."

Def'n Given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, a Bayesian inverse of (f,p) is a stochastic map $X \overset{g}{\leftsquigarrow} Y$ s.t.

$$
\begin{tikzcd}
X \arrow[d, "\Delta_X"'] & \bullet \arrow[l, squiggly, "p"'] \arrow[r, squiggly, "f\circ p"] & Y \arrow[d, "\Delta_Y"] \\
X\times X \arrow[r, squiggly, "\mathrm{id}_X\times f"'] & X\times Y \arrow[u, phantom, "="] & Y\times Y \arrow[l, squiggly, "g\times \mathrm{id}_Y"]
\end{tikzcd}
$$

i.e., $f_{yx}\, p_x = g_{xy}\,(f\circ p)_y \quad \forall x\in X, y\in Y$

i.e., $\mathbb{P}(y|x)\,\mathbb{P}(x) = \mathbb{P}(x|y)\,\mathbb{P}(y) \quad \forall x\in X, y\in Y$

Bayesian inverses

what's this?

"given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the Bayesian inverse of (f,p), then $S(q \| p) - S(f \circ q \| f \circ p) \geq S(q \| g \circ f \circ q)$ $\forall q$."

Def'n Given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, a Bayesian inverse of (f,p) is a stochastic map $X \overset{g}{\leftsquigarrow} Y$ s.t.

$$\begin{tikzcd}
X \arrow[d, "\Delta_X"'] & \bullet \arrow[l, squiggly, "p"'] \arrow[r, squiggly, "f\circ p"] \arrow[d, phantom, "="] & Y \arrow[d, "\Delta_Y"] \\
X \times X \arrow[r, squiggly, "\mathrm{id}_X \times f"'] & X \times Y & Y \times Y \arrow[l, squiggly, "g \times \mathrm{id}_Y"]
\end{tikzcd}$$

i.e., $f_{yx} p_x = g_{xy} (f \circ p)_y \quad \forall x \in X, y \in Y$

i.e., $\mathbb{P}(y|x)\mathbb{P}(x) = \mathbb{P}(x|y)\mathbb{P}(y) \quad \forall x \in X, y \in Y$

Theorem
- If f is deterministic and g is a Bayesian inverse of (f,p), then g is an optimal hypothesis.

Bayesian inverses

what's this?

"given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, set $X \overset{g}{\leftsquigarrow} Y$ to be the Bayesian inverse of (f,p), then $S(q \| p) - S(f \circ q \| f \circ p) \geq S(q \| g \circ f \circ q)$ $\forall q$."

Def'n Given $X \overset{f}{\rightsquigarrow} Y$ and $\bullet \overset{p}{\rightsquigarrow} X$, a Bayesian inverse of (f,p) is a stochastic map $X \overset{g}{\leftsquigarrow} Y$ s.t.

$$\begin{tikzcd} X \arrow[d, "\Delta_X"'] & \bullet \arrow[l, squiggly, "p"'] \arrow[r, squiggly, "f\circ p"] \arrow[d, phantom, "="] & Y \arrow[d, "\Delta_Y"] \\ X\times X \arrow[r, squiggly, "\mathrm{id}_X \times f"'] & X \times Y & Y \times Y \arrow[l, squiggly, "g \times \mathrm{id}_Y"] \end{tikzcd}$$

i.e., $f_{yx} p_x = g_{xy} (f \circ p)_y$ $\quad \forall x \in X, y \in Y$

i.e., $\mathbb{P}(y|x)\mathbb{P}(x) = \mathbb{P}(x|y)\mathbb{P}(y)$ $\quad \forall x \in X, y \in Y$

Theorem
- If f is deterministic and g is a Bayesian inverse of (f,p), then g is an optimal hypothesis.
- If g is an optimal hypothesis, then g is a Bayesian inverse and f is necessarily deterministic.

Enough classical stuff!

Li and Winter are quantum information theorists!

Why is any of this useful?

Enough classical stuff!

Li and Winter are quantum information theorists!

Why is any of this useful?

Some goals/achievements since Li-Winter's work:

Enough classical stuff!

Li and Winter are quantum information theorists!

Why is any of this useful?

Some goals/achievements since Li-Winter's work:

- improve the DPI to get tighter bounds

Enough classical stuff!

Li and Winter are quantum information theorists!

Why is any of this useful?

Some goals/achievements since Li-Winter's work:

- improve the DPI to get tighter bounds
- extend the 2nd law of thermodynamics beyond equilibrium

Enough classical stuff!

Li and Winter are quantum information theorists!

Why is any of this useful?

Some goals/achievements since Li-Winter's work:

- improve the DPI to get tighter bounds
- extend the 2nd law of thermodynamics beyond equilibrium
- discover alternatives to error-correcting codes

Enough classical stuff!

Li and Winter are quantum information theorists!

Why is any of this useful?

Some goals/achievements since Li-Winter's work:

- improve the DPI to get tighter bounds
- extend the 2^{nd} law of thermodynamics beyond equilibrium
- discover alternatives to error-correcting codes
- understand entanglement-wedge reconstruction in AdS/CFT

Enough classical stuff!

Li and Winter are quantum information theorists!

Why is any of this useful?

Some goals/achievements since Li-Winter's work:

- improve the DPI to get tighter bounds
- extend the 2^{nd} law of thermodynamics beyond equilibrium
- discover alternatives to error-correcting codes
- understand entanglement-wedge reconstruction in AdS/CFT
- apply quantum info ideas to the renormalization group

Enough classical stuff!

Li and Winter are quantum information theorists!

Why is any of this useful?

Some goals/achievements since Li-Winter's work:

- improve the DPI to get tighter bounds
- extend the 2^{nd} law of thermodynamics beyond equilibrium
- discover alternatives to error-correcting codes
- understand entanglement-wedge reconstruction in AdS/CFT
- apply quantum info ideas to the renormalization group
- adjust Hawking's calculation to prove information conservation in black hole evaporation (formerly "the information paradox")

Classical $\longleftrightarrow$ Quantum^{op}

space X

*-algebra of functions
$\mathbb{C}^X := \{X \xrightarrow{\varphi} \mathbb{C}\}$

*-algebra $\mathcal{A}$

Classical $\longrightarrow$ Quantum^{op}

space X

$*$-algebra of functions $\mathbb{C}^X := \{X \xrightarrow{\varphi} \mathbb{C}\}$

$*$-algebra $\mathcal{A}$

probability $\bullet \overset{p}{\rightsquigarrow} X$

expectation value $\mathbb{C}^X \overset{\mathbb{E}_p[\cdot]}{\rightsquigarrow} \mathbb{C}$

$\varphi \longmapsto \mathbb{E}_p[\varphi] := \sum_x p_x \varphi(x)$

state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$

Classical $\longrightarrow$ Quantum$^{\text{op}}$

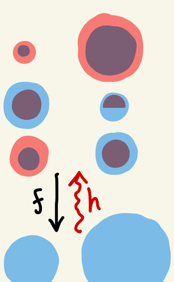
space X	$*$-algebra of functions $\mathbb{C}^X := \{X \xrightarrow{\varphi} \mathbb{C}\}$	$*$-algebra $\mathcal{A}$
probability $\bullet \overset{p}{\rightsquigarrow} X$	expectation value $\mathbb{C}^X \overset{\mathbb{E}_p[\cdot]}{\rightsquigarrow} \mathbb{C}$, $\varphi \longmapsto \mathbb{E}_p[\varphi] := \sum_x p_x \varphi(x)$	state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$
stochastic map $X \overset{f}{\rightsquigarrow} Y$	$\mathbb{C}^Y \rightsquigarrow \mathbb{C}^X$, $\psi \longmapsto \left(x \longmapsto \mathbb{E}_{f_x}[\psi]\right)$	(unital) positive map $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$

Classical $\longrightarrow$ Quantum^{op}

space X	*-algebra of functions $\mathbb{C}^X := \{X \xrightarrow{\varphi} \mathbb{C}\}$	*-algebra $\mathcal{A}$
probability $\bullet \overset{p}{\rightsquigarrow} X$	expectation value $\mathbb{C}^X \overset{\mathbb{E}_p[\cdot]}{\rightsquigarrow} \mathbb{C}$ $\varphi \longmapsto \mathbb{E}_p[\varphi] := \sum_x p_x \varphi(x)$	state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$
stochastic map $X \overset{f}{\rightsquigarrow} Y$	$\mathbb{C}^Y \rightsquigarrow \mathbb{C}^X$ $\psi \longmapsto (x \longmapsto \mathbb{E}_{f_x}[\psi])$	(unital) positive map $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$
deterministic map $X \xrightarrow{f} Y$	$\mathbb{C}^Y \longrightarrow \mathbb{C}^X$ $\psi \longmapsto \psi \circ f$ (pullback)	(unital) *-homomorphism $\mathcal{B} \xrightarrow{F} \mathcal{A}$

Classical $\longrightarrow$ Quantum^{op}

space X	*-algebra of functions $\mathbb{C}^X := \{X \xrightarrow{\varphi} \mathbb{C}\}$	*-algebra $\mathcal{A}$
probability $\bullet \overset{p}{\rightsquigarrow} X$	expectation value $\mathbb{C}^X \overset{\mathbb{E}_p[\cdot]}{\rightsquigarrow} \mathbb{C}$, $\varphi \longmapsto \mathbb{E}_p[\varphi] := \sum_x p_x \varphi(x)$	state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$
stochastic map $X \overset{f}{\rightsquigarrow} Y$	$\mathbb{C}^Y \rightsquigarrow \mathbb{C}^X$, $\psi \longmapsto (x \longmapsto \mathbb{E}_{f_x}[\psi])$	(unital) positive map $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$
deterministic map $X \xrightarrow{f} Y$	$\mathbb{C}^Y \longrightarrow \mathbb{C}^X$, $\psi \longmapsto \psi \circ f$ (pullback)	(unital) *-homomorphism $\mathcal{B} \xrightarrow{F} \mathcal{A}$



hypothesis

$$\begin{tikzcd} (X, p) \arrow[d, bend right, "f"'] \\ (Y, q) \arrow[u, squiggly, bend right, "h"'] \end{tikzcd}$$

$$\begin{tikzcd} (\mathbb{C}^X, \mathbb{E}_p) \arrow[d, squiggly, bend left, "H"] \\ (\mathbb{C}^Y, \mathbb{E}_q) \arrow[u, bend left, "F"] \end{tikzcd}$$

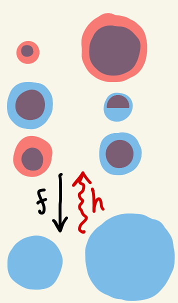
$$\begin{tikzcd} (\mathcal{A}, \omega) \arrow[d, squiggly, bend left, "H"] \\ (\mathcal{B}, \xi) \arrow[u, bend left, "F"] \end{tikzcd}$$

$\xi = \omega \circ F$

$H \circ F = id_{\mathcal{B}}$

Classical $\longrightarrow$ Quantum^{op}

Classical		Quantum
space X	$*$-algebra of functions $\mathbb{C}^X := \{X \xrightarrow{\varphi} \mathbb{C}\}$	$*$-algebra $\mathcal{A}$
probability $\bullet \overset{p}{\rightsquigarrow} X$	expectation value $\mathbb{C}^X \overset{\mathbb{E}_p[\cdot]}{\rightsquigarrow} \mathbb{C}$, $\varphi \longmapsto \mathbb{E}_p[\varphi] := \sum_x p_x \varphi(x)$	state $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$
stochastic map $X \overset{f}{\rightsquigarrow} Y$	$\mathbb{C}^Y \rightsquigarrow \mathbb{C}^X$, $\psi \longmapsto (x \longmapsto \mathbb{E}_{f_x}[\psi])$	(unital) positive map $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$
deterministic map $X \xrightarrow{f} Y$	$\mathbb{C}^Y \longrightarrow \mathbb{C}^X$, $\psi \longmapsto \psi \circ f$ (pullback)	(unital) $*$-homomorphism $\mathcal{B} \xrightarrow{F} \mathcal{A}$



hypothesis

$$\begin{tikzcd} (X,p) \arrow[d, bend right, "f"'] \\ (Y,q) \arrow[u, squiggly, bend right, "h"'] \end{tikzcd}$$

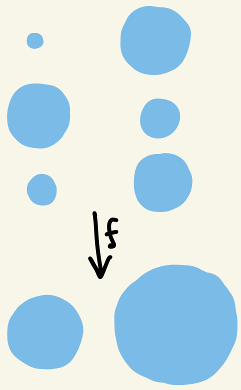
$$\begin{tikzcd} (\mathbb{C}^X, \mathbb{E}_p) \arrow[d, squiggly, bend left, "H"] \\ (\mathbb{C}^Y, \mathbb{E}_q) \arrow[u, bend left, "F"] \end{tikzcd}$$

$$\begin{tikzcd} (\mathcal{A}, \omega) \arrow[d, squiggly, bend left, "H"] \\ (\mathcal{B}, \xi) \arrow[u, bend left, "F"] \end{tikzcd}$$

$\xi = \omega \circ F$

$H \circ F = \mathrm{id}_{\mathcal{B}}$

Subtle differences in quantum



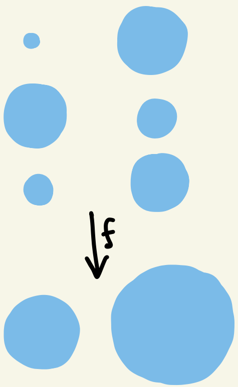
Classically, info loss

$$\Delta_H(f) = H(p) - H(q)$$

associated to $(X,p) \xrightarrow{f} (Y,q)$

is always non-negative.

Subtle differences in quantum



Classically, info loss

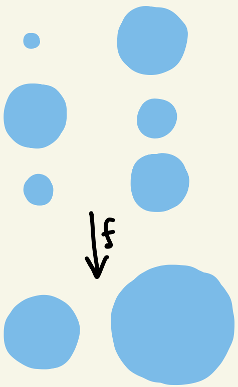
$$\Delta_H(f) = H(p) - H(q)$$

associated to $(X,p) \xrightarrow{f} (Y,q)$

is always non-negative.

Baez-Fritz-Leinster's proof relied this!

Subtle differences in quantum



Classically, info loss

$\Delta_H(f) = H(p) - H(q)$

associated to $(X,p) \xrightarrow{f} (Y,q)$

is always non-negative.

In quantum,

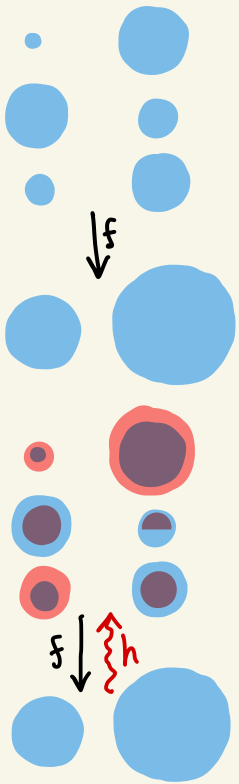
$\Delta_H(F) = H(\omega) - H(\xi)$

associated to $(\mathcal{B},\xi) \xrightarrow{F} (\mathcal{A},\omega)$

need not have a definite sign.

Baez-Fritz-Leinster's proof relied this!

Subtle differences in quantum



Classically, info loss

$\Delta_H(f) = H(p) - H(q)$

associated to $(X,p) \xrightarrow{f} (Y,q)$

is always non-negative.

In quantum,

$\Delta_H(F) = H(\omega) - H(\xi)$

associated to $(\mathcal{B},\xi) \xrightarrow{F} (\mathcal{A},\omega)$

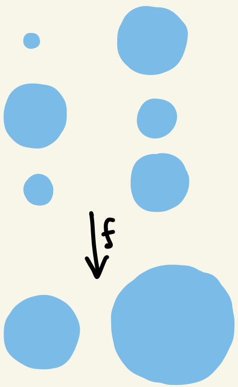
need not have a definite sign.

Baez-Fritz-Leinster's proof relied this!

Classically, optimal

hypotheses always exist.

Subtle differences in quantum



Classically, info loss

$\Delta_H(f) = H(p) - H(q)$

associated to $(X,p) \xrightarrow{f} (Y,q)$

is always non-negative.

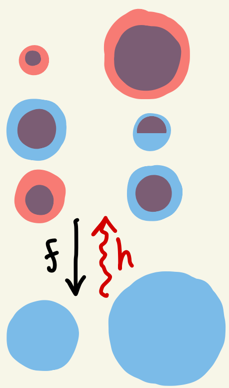
In quantum,

$\Delta_H(F) = H(\omega) - H(\xi)$

associated to $(\mathcal{B},\xi) \xrightarrow{F} (\mathcal{A},\omega)$

need not have a definite sign.

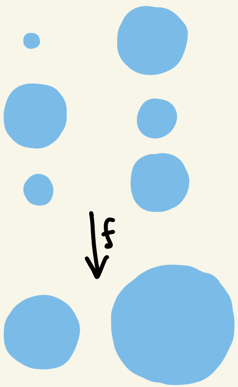
Baez-Fritz-Leinster's proof relied this!

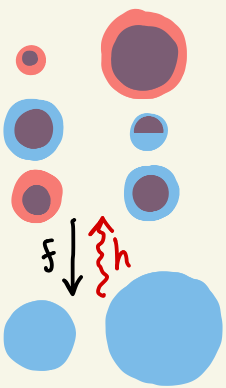


Classically, optimal hypotheses always exist.

In quantum, optimal hypotheses don't always exist.

Subtle differences in quantum





Classically, info loss

$$\Delta_H(f) = H(p) - H(q)$$

associated to $(X,p) \xrightarrow{f} (Y,q)$ is always non-negative.

In quantum,

$$\Delta_H(F) = H(\omega) - H(\xi)$$

associated to $(\mathcal{B},\xi) \xrightarrow{F} (\mathcal{A},\omega)$ need not have a definite sign.

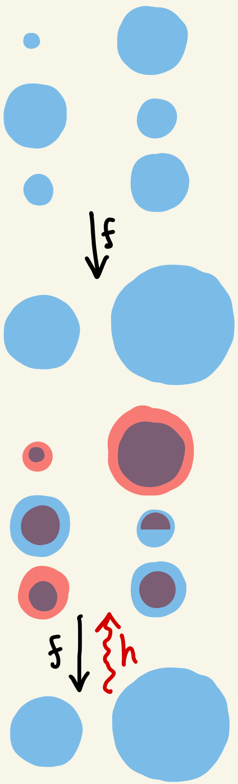
Baez-Fritz-Leinster's proof relied this!

Classically, optimal hypotheses always exist.

In quantum, optimal hypotheses don't always exist.

<u>Fact</u>: Given $(\mathcal{B},\xi) \xrightarrow{F} (\mathcal{A},\omega)$, if an optimal hypothesis exists, then $\Delta_H(F) \geq 0$.

Subtle differences in quantum



Classically, info loss

$\Delta_H(f) = H(p) - H(q)$

associated to $(X,p) \xrightarrow{f} (Y,q)$

is always non-negative.

In quantum,

$\Delta_H(F) = H(\omega) - H(\xi)$

associated to $(\mathcal{B},\xi) \xrightarrow{F} (\mathcal{A},\omega)$

need not have a definite sign.

Baez-Fritz-Leinster's proof relied this!

Classically, optimal hypotheses always exist.

In quantum, optimal hypotheses don't always exist.

<u>Fact</u>: Given $(\mathcal{B},\xi) \xrightarrow{F} (\mathcal{A},\omega)$, if an optimal hypothesis exists, then $\Delta_H(F) \geq 0$.

See my talk @ N. Yanofsky's seminar

This was one of the key observations in extending Baez-Fritz-Leinster's Theorem to quantum.

Quantum Relative Entropy

Defining NCFinStat (non commutative) via

Quantum Relative Entropy

Defining **NCFinStat** (non commutative) via

objects: pairs $(\mathcal{A},\omega)$, $(\mathcal{B},\xi)$, $(\mathcal{C},\zeta)$ of non-commutative probability spaces

Quantum Relative Entropy

Defining **NCFinStat** (non commutative) via

objects: pairs $(\mathcal{A},\omega)$, $(\mathcal{B},\xi)$, $(\mathcal{C},\zeta)$ of non-commutative probability spaces

morphisms:
$$\begin{tikzcd} (\mathcal{A},\omega) \arrow[r, bend right, squiggly, "H"'] & (\mathcal{B},\xi) \arrow[l, bend right, "F"'] \end{tikzcd}$$

F = deterministic

$\xi = \omega \circ F$

$H \circ F = \mathrm{id}_{\mathcal{B}}$

Quantum Relative Entropy

Defining **NCFinStat** (non commutative) via

objects: pairs $(\mathcal{A},\omega), (\mathcal{B},\xi), (\mathcal{C},\zeta)$ of non-commutative probability spaces

morphisms:
$$\begin{tikzcd} (\mathcal{A},\omega) \arrow[r, bend right, squiggly, "H"'] & (\mathcal{B},\xi) \arrow[l, bend right, "F"'] \end{tikzcd}$$

F = deterministic

$\xi = \omega \circ F$

$H \circ F = \mathrm{id}_{\mathcal{B}}$

$\mathbf{NCFinStat} \xrightarrow{RE} \mathbb{B}[0,\infty]$ also defines a convex functor

$$\begin{tikzcd} (\mathcal{A},\omega) \arrow[d, squiggly, bend left, "H"] \\ (\mathcal{B},\xi) \arrow[u, bend left, "F"] \end{tikzcd} \longmapsto S(\omega \| \xi \circ H)$$

Quantum Relative Entropy

Defining **NCFinStat** (non commutative) via

objects: pairs $(\mathcal{A},\omega)$, $(\mathcal{B},\xi)$, $(\mathcal{C},\zeta)$ of non-commutative probability spaces

morphisms:
$$\begin{tikzcd}(\mathcal{A},\omega) \arrow[r, bend right, squiggly, "H"'] & (\mathcal{B},\xi) \arrow[l, bend right, "F"']\end{tikzcd}$$

F = deterministic

$\xi = \omega \circ F$

$H \circ F = \mathrm{id}_{\mathcal{B}}$

$$\mathbf{NCFinStat} \xrightarrow{RE} \mathbb{B}[0,\infty]$$

$$\begin{tikzcd}(\mathcal{A},\omega) \arrow[d, squiggly, bend left, "H"] \\ (\mathcal{B},\xi) \arrow[u, "F"]\end{tikzcd} \longmapsto S(\omega \| \xi \circ H)$$

also defines a convex functor

(at least, I checked this when the states satisfy appropriate support conditions).

Quantum Relative Entropy

Defining **NCFinStat** (non commutative) via

objects: pairs $(\mathcal{A},\omega)$, $(\mathcal{B},\xi)$, $(\mathcal{C},\zeta)$ of non-commutative probability spaces

morphisms:
$$\begin{tikzcd} (\mathcal{A},\omega) \arrow[r, rightsquigarrow, bend right, "H"'] & (\mathcal{B},\xi) \arrow[l, bend right, "F"'] \end{tikzcd}$$

F = deterministic

$\xi = \omega \circ F$

$H \circ F = \mathrm{id}_{\mathcal{B}}$

$$\mathbf{NCFinStat} \xrightarrow{RE} \mathbb{B}[0,\infty]$$

$$\begin{tikzcd} (\mathcal{A},\omega) \arrow[d, rightsquigarrow, "H"] \\ (\mathcal{B},\xi) \arrow[u, "F"] \end{tikzcd} \longmapsto S(\omega \| \xi \circ H)$$

also defines a convex functor (at least, I checked this when the states satisfy appropriate support conditions).

In progress:
1) functoriality for all states
2) lower semi-continuity
3) characterization

Properties of RE

Functoriality specializes to

Properties of RE

Functoriality specializes to

- The conditional expectation property (originally due to Petz)

Properties of RE

Functoriality specializes to

- The conditional expectation property (originally due to Petz)
- The chain rule for quantum conditional entropy

Properties of RE

Functoriality specializes to

- The conditional expectation property (originally due to Petz)
- The chain rule for quantum conditional entropy

But, the Li-Winter strengthened DPI fails. Namely,

"Given $B \overset{F}{\rightsquigarrow} A$ and $A \overset{\omega}{\rightsquigarrow} \mathbb{C}$, there exists a

$B \overset{G}{\leftsquigarrow} A$ s.t. $S(\xi \| \omega) - S(\xi \circ F \| \omega \circ F) \geq S(\xi \| \xi \circ F \circ G) \ \forall \xi$"

is false.

Properties of RE

Functoriality specializes to

- The conditional expectation property (originally due to Petz)
- The chain rule for quantum conditional entropy

But, the Li-Winter strengthened DPI fails. Namely,

"Given $\mathcal{B} \overset{F}{\rightsquigarrow} \mathcal{A}$ and $\mathcal{A} \overset{\omega}{\rightsquigarrow} \mathbb{C}$, there exists a

$\mathcal{B} \overset{G}{\leftsquigarrow} \mathcal{A}$ s.t. $S(\xi \| \omega) - S(\xi \circ F \| \omega \circ F) \geq S(\xi \| \xi \circ F \circ G)$ $\forall \xi$"

is false. Variants and weaker versions have been found.

Properties of RE

Functoriality specializes to

- The conditional expectation property (originally due to Petz)
- The chain rule for quantum conditional entropy

But, the Li-Winter strengthened DPI fails. Namely,

"Given $\mathcal{B} \xrightarrow{F} \mathcal{A}$ and $\mathcal{A} \xrightarrow{\omega} \mathbb{C}$, there exists a $\mathcal{B} \xleftarrow{G} \mathcal{A}$ s.t. $S(\xi \| \omega) - S(\xi \circ F \| \omega \circ F) \geq S(\xi \| \xi \circ F \circ G) \; \forall \xi$"

is false. Variants and weaker versions have been found.

In progress: Is it true when a Bayesian inverse exists?

Such a map is defined by

$$\begin{tikzcd} \mathcal{A} \arrow[r, squiggly, "\omega"] & \mathbb{C} \arrow[d, phantom, "="] & \mathcal{B} \arrow[l, squiggly, "\omega \circ F"'] \\ \mathcal{A} \otimes \mathcal{A} \arrow[u, squiggly, "\text{mult}"'] & \mathcal{A} \otimes \mathcal{B} \arrow[l, squiggly, "\mathrm{id}_{\mathcal{A}} \otimes F"] \arrow[r, squiggly, "G \otimes \mathrm{id}_{\mathcal{B}}"'] & \mathcal{B} \otimes \mathcal{B} \arrow[u, squiggly, "\text{mult}"] \end{tikzcd}$$

Categorical approaches to entropy

Classical

Quantum

Categorical approaches to entropy

Classical

Shannon entropy
Baez–Fritz–Leinster

Quantum

Categorical approaches to entropy

Classical	Quantum
Shannon entropy	von Neumann entropy
Baez–Fritz–Leinster	Parzygnat

Categorical approaches to entropy

Classical

Shannon entropy
Baez – Fritz – Leinster

Conditional entropy
Fullwood – Parzygnat

Quantum

von Neumann entropy
Parzygnat

Categorical approaches to entropy

Classical	Quantum
Shannon entropy Baez–Fritz–Leinster	von Neumann entropy Parzygnat
Conditional entropy Fullwood–Parzygnat	Conditional entropy In progress

Categorical approaches to entropy

Classical

Shannon entropy
Baez–Fritz–Leinster

Conditional entropy
Fullwood–Parzygnat

Relative entropy
Baez–Fritz
Gagné–Panangaden

Quantum

von Neumann entropy
Parzygnat

Conditional entropy
In progress

Categorical approaches to entropy

Classical	Quantum
Shannon entropy Baez–Fritz–Leinster	von Neumann entropy Parzygnat
Conditional entropy Fullwood–Parzygnat	Conditional entropy In progress
Relative entropy Baez–Fritz Gagné–Panangaden	Relative entropy In progress, but partial results

Infinite-dim'l generalizations unknown

Categorical approaches to entropy

Classical

Shannon entropy
Baez–Fritz–Leinster

Conditional entropy
Fullwood–Parzygnat

Relative entropy
Baez–Fritz
Gagné–Panangaden

Quantum

von Neumann entropy
Parzygnat

Conditional entropy
In progress

Relative entropy
In progress, but partial results

Also homological,
operadic,
topos-theoretic
characterizations of some of these exist.

Infinite-dim'l generalizations unknown

Categorical approaches to entropy

Classical	Quantum
Shannon entropy Baez–Fritz–Leinster	von Neumann entropy Parzygnat
Conditional entropy Fullwood–Parzygnat	Conditional entropy In progress
Relative entropy Baez–Fritz Gagné–Panangaden	Relative entropy In progress, but partial results

Infinite-dim'l generalizations unknown

Also homological,
operadic,
topos-theoretic
characterizations of some of these exist.

what connects all of these different approaches?

Thank you!

References

J. Baez, T. Fritz, T. Leinster arXiv: 1106.1791 [cs.IT] ← Shannon entropy

J. Baez and T. Fritz arXiv: 1402.3067 [cs.IT] ← Relative entropy

K. Li and A. Winter arXiv: 1410.4184 [quant-ph] ← Data-processing & recovery

A. Parzygnat and B. Russo arXiv: 1907.09689 [quant-ph] ← Disintegrations

A. Parzygnat arXiv: 2001.08375 [quant-ph] ← Quantum Markov cats

A. Parzygnat and B. Russo arXiv: 2005.03886 [quant-ph] ← Quantum Bayesian inverses

A. Parzygnat arXiv: 2009.07125 [quant-ph] ← von Neumann entropy

A. Parzygnat arXiv: 2105.04059 [quant-ph] ← Relative entropy