# Petz recovery, Jeffrey retrodiction, and von Neumann's "other" entropy

Francesco Buscemi* (Nagoya)

*www.quantumquia.com

Quantum Information and Mathematical Physics 2023, Budapest



## collaborators on this journey

- Clive Aw (CQT@NUS)
- Ge Bai (CQT@NUS)
- Kohtaro Kato (Nagoya)
- Teruaki Nagasawa (Nagoya)
- Arthur Parzygnat (MIT, formely Nagoya)
- Dominik Šafránek (IBS)
- Valerio Scarani (CQT@NUS)
- Joseph Schindler (UAB)

a growing list:
*The Observational Entropy Appreciation Club*
(www.observationalentropy.com)

# von Neumann's entropy

For $\varrho = \sum_{x=1}^{d} \lambda_x |\varphi_x\rangle\langle\varphi_x|$ $d$-dimensional density matrix ($\lambda_x \geq 0$, $\sum_x \lambda_x = 1$),

$$S(\varrho) := -\operatorname{Tr}[\varrho \log \varrho] = -\sum_{x=1}^{d} \lambda_x \log \lambda_x$$

with the convention $0 \log 0 := 0$.



Unfortunately though:

> *"The expressions for entropy given by the author [previously] are not applicable here in the way they were intended, as they were computed from the perspective of an observer who can carry out all measurements that are possible in principle—i.e., regardless of whether they are macroscopic [or not]."*
>
> von Neumann, 1929; transl. available in `arXiv:1003.2133`

# in formula:

**Theorem (least uncertainty)**

*For $\varrho$ density matrix, $\mathfrak{onb} = \{|\phi_i\rangle\}_i$ orthonormal basis, and $p_i = \langle\phi_i|\varrho|\phi_i\rangle$,*

$$S(\varrho) = \min_{\mathfrak{onb}} \left[ -\sum_i p_i \log p_i \right] .$$

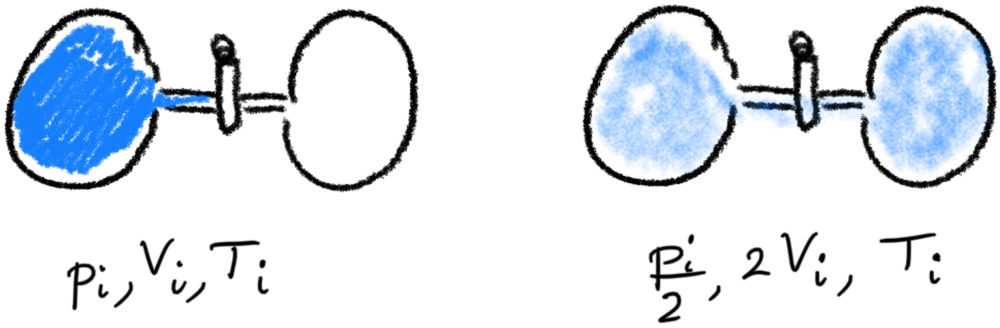
For a more general result, see [M. Dall'Arno and F.B., IEEE TIT, 65(4), 2018].



*"Although our entropy expression, as we saw, is completely analogous to the classical entropy, it is still surprising that it is invariant in the normal [Hamiltonian] evolution in time of the system, and only increases with measurements—in the classical theory (where the measurements in general played no role) it increased as a rule even with the ordinary mechanical evolution in time of the system. It is therefore necessary to clear up this apparently paradoxical situation."*

von Neumann, book (Math. Found. QM), 1932 (transl. 1955)

# the paradox: free expansion of an ideal gas




$$\Delta S(\text{universe}) = nR \log 2 > 0$$

$\Longrightarrow$ net entropy increase in an isolated system's evolution



# invariance of von Neumann entropy

Instead,

**Theorem**

*For any unitary operator $U$,*

$$S(\varrho) = S(U\varrho U^\dagger) ,$$

*for all density matrices $\varrho$.*

$\Longrightarrow$ the entropy increasing during a free expansion (isolated evolution, thus unitary) cannot be the von Neumann entropy

## von Neumann's insight (inspired by Szilard's)

> *"For a classical observer, who knows all coordinates and momenta, the entropy is constant. [...]*
> *The time variations of the entropy are then based on the fact that the observer does not know everything—that he cannot find out (measure) everything which is measurable in principle."*
>
> von Neumann, 1932 (transl. 1955)

Thus, von Neumann recognizes that thermodynamic entropy should be a quantity *relative* to the observer...



## von Neumann's proposal: macroscopic entropy

For

- $\varrho$ density matrix,
- $\mathfrak{P} = \{\Pi_i\}_i$ orthogonal resolution of identity,
- $p_i = \mathrm{Tr}[\varrho\, \Pi_i]$,
- $\Omega_i := \mathrm{Tr}[\Pi_i]$,

$$S_{\mathfrak{P}}(\varrho) := -\sum_i p_i \log \frac{p_i}{\Omega_i}$$

# modern generalization: observational entropy

For

- $\varrho$ density matrix,
- $\mathbf{P} = \{P_i\}_i$ POVM (i.e., $P_i \geq 0$, $\sum_i P_i = \mathbb{1}$),
- $p_i = \mathrm{Tr}[\varrho\ P_i]$,
- $V_i := \mathrm{Tr}[P_i]$,

$$S_{\mathbf{P}}(\varrho) := -\sum_i p_i \log \frac{p_i}{V_i}$$

References:


1. D. Šafránek, J.M. Deutsch, A. Aguirre. *Phys. Rev. A* **99**, 012103 (2019)
2. D. Šafránek, A. Aguirre, J. Schindler, J. M. Deutsch. *Found. Phys.* **51**, 101 (2021)




# "observational" = "of the observer"

- von Neumann defines a **macro-observer** as a collection of simultaneously measurable quantities $\{\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_n, \ldots\}$, where $\mathbf{Q}_n = \{Q_{x|n}\}_x$ are POVMs
- $\implies$ there exists one "mother" POVM $\mathbf{P} = \{P_i\}_i$ and a stochastic processing (i.e., cond. prob.) $\mu$ such that

$$Q_{x|n} = \sum_i \mu(x|n,i) P_i\ , \quad \forall x, n$$

- hence, a "macro-observer" is just a POVM, i.e., the mother POVM $\mathbf{P}$, from which all macroscopic measurements (i.e., coarse-grainings) can be simultaneously inferred by stochastic post-processing

Our aim ultimately is to extend OE's definition and scope.

So we need to understand what OE is and what it is all about.



# OE as "statistical deficiency"

## Umegaki's relative entropy

### Definition

For density matrices $\varrho, \gamma$,

$$D(\varrho\|\gamma) := \begin{cases} \mathrm{Tr}[\varrho(\log\varrho - \log\gamma)] \ , & \text{if } \operatorname{supp}\varrho \subseteq \operatorname{supp}\gamma \ , \\ +\infty \ , & \text{otherwise} \end{cases}$$

Useful properties:

- $D(A\|B) \geq 0$
- $S(\varrho) = \log d - D(\varrho\|u)$ where $u := d^{-1}\mathbb{1}$
- **monotonicity**: $D(\varrho\|\gamma) \geq D(\mathcal{N}(\varrho)\|\mathcal{N}(\gamma))$ for all channels (i.e., CPTP linear maps) $\mathcal{N}$ and all states $\varrho, \gamma$



## first interpretation

### Theorem

*Given a POVM* $\mathbf{P} = \{P_i\}_i$*, define the CPTP linear map* $\mathcal{P}(\bullet) := \sum_i \mathrm{Tr}[P_i \ \bullet] \, |i\rangle\langle i|$*. Then, for any state* $\varrho$*,*

$$\begin{aligned} \Sigma_{\mathbf{P}}(\varrho) :&= S_{\mathbf{P}}(\varrho) - S(\varrho) \\ &= D(\varrho\|u) - D(\mathcal{P}(\varrho)\|\mathcal{P}(u)) \\ &\geq 0 \ , \end{aligned}$$

*where* $u = d^{-1}\mathbb{1}$*. If* $\Sigma_{\mathbf{P}}(\varrho) = 0$*, the state* $\varrho$ *is said to be macroscopic for observer* $\mathbf{P}$*.*

Hence, the larger is the difference $\Sigma_{\mathbf{P}}(\varrho)$, the worse is $\mathbf{P}$ for distinguishing signal $\varrho$ from uniform $u$.

## a stronger bound

**Theorem**

*For $d$-dimensional quantum system, density matrix $\varrho$, and POVM $\mathbf{P} = \{P_i\}_i$, the difference $\Sigma_{\mathbf{P}}(\varrho) = S_{\mathbf{P}}(\varrho) - S(\varrho)$ satisfies*

$$T\ln(d-1) + h(T) \geq \Sigma_{\mathbf{P}}(\varrho) \geq D(\varrho \| \tilde{\varrho}_{\mathbf{P}}) ,$$

*where*

- $\tilde{\varrho}_{\mathbf{P}} := (\mathcal{R}^u_{\mathcal{P}} \circ \mathcal{P})(\varrho) = \sum_i \mathrm{Tr}[\varrho \; P_i] \frac{P_i}{V_i}$
- $\mathcal{R}^u_{\mathcal{P}}(\cdot) := \frac{1}{d}\mathcal{P}^\dagger[\mathcal{P}(u)^{-1/2}(\cdot)\mathcal{P}(u)^{-1/2}]$
- $T := \frac{1}{2}\|\varrho - \tilde{\varrho}_{\mathbf{P}}\|_1$
- $h(x) := -x\ln x - (1-x)\ln(1-x)$

Hence, a state $\varrho$ is macroscopic for observer $\mathbf{P}$ if and only if $\varrho = \tilde{\varrho}_{\mathbf{P}}$.



## why OE resolves the paradox (and satisfies an $H$-theorem)

- start at $t = t_0$ from a macrostate $\varrho^{t_0} \in \mathfrak{M}(\mathbf{P}) := \{\text{macrostates of } \mathbf{P}\}$
- the system evolves unitarily, i.e., $\varrho^{t_0} \mapsto \varrho^{t_1} = U\varrho^{t_0}U^\dagger$; then,

$$\begin{aligned}
S_{\mathbf{P}}(\varrho^{t_1}) &= -\sum_i \mathrm{Tr}\big[P_i\;(U\varrho^{t_0}U^\dagger)\big]\log\frac{\mathrm{Tr}\big[P_i\;(U\varrho^{t_0}U^\dagger)\big]}{\mathrm{Tr}[P_i]} \\
&= -\sum_i \mathrm{Tr}\big[(U^\dagger P_i U)\;\varrho^{t_0}\big]\log\frac{\mathrm{Tr}\big[(U^\dagger P_i U)\;\varrho^{t_0}\big]}{\mathrm{Tr}[U^\dagger P_i U]} \\
&= S_{U^\dagger \mathbf{P} U}(\varrho^{t_0}) \\
&\geq S_{\mathbf{P}}(\varrho^{t_0}) = S(\varrho^{t_0}) = S(\varrho^{t_1})
\end{aligned}$$

- summarizing: in general, $S_{\mathbf{P}}(\varrho^{t_1}) \geq S_{\mathbf{P}}(\varrho^{t_0})$, with equality if and only if $\varrho^{t_1} \in \mathfrak{M}(\mathbf{P})$ too
- in words: the observational entropy of an isolated macroscopic state remains constant if and only if the state remains macroscopic, otherwise it will (generically) increase

If $\varrho$ is the microscopic (i.e., "true") state of the system...

...what does the corresponding macrostate $\tilde{\varrho}_{\mathbf{P}}$ represent exactly?



# OE as "irretrodictability"

## what is retrodiction?

With prior $\pi(H)$ and likelihood $P(D|H)$, when the observation returns a definite value $\bar{D}$, Bayes' update rule says that the posterior becomes $R_P^\pi(H|\bar{D}) \propto \pi(H)P(\bar{D}|H)$.

But what if the observation is noisy and returns some p.d. $Q(D)$ instead?

**Theorem (Jeffrey, 1965)**

*Starting from a given prior $\pi(H)$ and a likelihood $P(D|H)$, the result of a noisy observation $Q(D)$ is retrodicted to*

$$\widetilde{Q}(H) := \sum_D R_P^\pi(H|D)Q(D) .$$

*The conventional Bayes' rule is recovered for $Q(D) = \delta_{D,\bar{D}}$.*



## what is irretrodictability?

- a channel $P(D|H)$ is given (objective)
- a predictor is given information about the hypothesis $H$ in the form of a p.d. $O(H)$
- the predictor's prediction (expectation) about the data distribution is modeled by $P_F(H,D) := O(H)P(D|H)$
- a retrodictor is given information about the data $D$ as $Q(D)$
- the retrodictor chooses a prior $\pi(H)$ and applies Jeffrey's rule to retrodict as $P_R^\pi(H,D) := Q(D)R_P^\pi(H|D)$

**Definition**

The model is retrodictable whenever

$$P_F(H,D) = P_F(D)R_P^\pi(H|D) .$$

In particular, the above implies $O(H) = \sum_D P_F(D)R_P^\pi(H|D)$.

## Jeffrey's retrodiction *versus* Petz's recovery

- retrodictor's prior $\pi(H) \longrightarrow$ state $\gamma$
- likelihood $P(D|H) \longrightarrow$ channel $\mathcal{N}$
- Bayes' inverse $R_P^\pi(H|D) \longrightarrow$ Petz recovery map $\mathcal{R}_\mathcal{N}^\gamma$
- add all sorts of possible "rotations" when noncommuting

That is: all rotated Petz maps are "quantum Jeffrey retrodictions".

Correspondingly, $\tilde{\varrho}_\mathbf{P} = \sum_i \frac{p_i}{V_i} P_i$ represents the state **retrodicted** from the observed outcomes statistics $p_i = \mathrm{Tr}[\varrho\ P_i]$, knowing the POVM $\{P_i\}$ and assuming the uniform prior $u$.

The connection with retrodiction can be made even more precise.



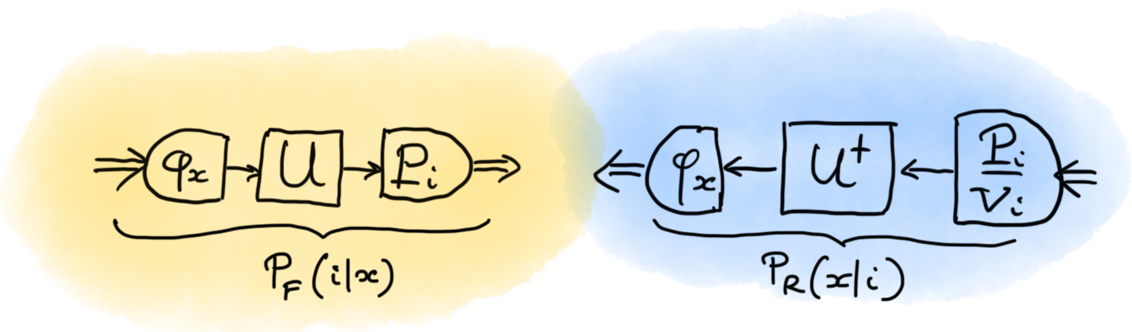
## second interpretation

**Theorem**

*Given a $d$-dimensional system, a density matrix $\varrho$ with diagonalization $\{\lambda_x, |\varphi_x\rangle\}_{x=1}^d$, a unitary operator $U$, and a POVM $\mathbf{P} = \{P_i\}_i$, let us define two joint probability distributions:*

$$P_F(x,i) := \lambda_x \underbrace{\mathrm{Tr}\left[U|\varphi_x\rangle\langle\varphi_x|U^\dagger\ P_i\right]}_{P_F(i|x)}, \qquad P_R^u(x,i) := P_F(i)\underbrace{\mathrm{Tr}\left[|\varphi_x\rangle\langle\varphi_x|\ \frac{U^\dagger P_i U}{V_i}\right]}_{P_R^u(x|i)}.$$

*Then,* $S_\mathbf{P}(U\varrho U^\dagger) - S(\varrho) = D(P_F\|P_R^u)$.



## irretrodictability

Hence, for $P_F(x,i) = \lambda_x P_F(i|x)$ and $P^u_R(x,i) = P_F(i) P^u_R(x|i)$ given above:

- $P_F$ corresponds to the prediction $\boldsymbol{\lambda} \to \bullet$: the inference about $i$
- $P^u_R$ corresponds to the retrodiction $\bullet \leftarrow \boldsymbol{p}$: the inference about $x$, done from the uniform prior on $x$ and the predicted distribution on $i$

Hence, the larger is the difference $S_{\mathbf{P}}(U\varrho U^\dagger) - S(\varrho)$, the more irretrodictable the forward process is.



## parenthesis: Watanabe's contention

*"The phenomenological oneway-ness of temporal developments in physics is due to irretrodictability, and not due to irreversibility."*

Satosi Watanabe (1965)

## intermediate summary

The difference $\Sigma_{\mathbf{P}}(\varrho) = S_{\mathbf{P}}(\varrho) - S(\varrho)$ admits two forms:

- as deficiency, i.e., $\Sigma_{\mathbf{P}}(\varrho) = D(\varrho\|u) - D(\mathcal{P}(\varrho)\|\mathcal{P}(u))$
- as irretrodictability, i.e., $\Sigma_{\mathbf{P}}(\varrho) = D(P_F\|P_R^u)$

In both, the uniform prior is assumed.

Can we generalize the discussion to an arbitrary prior?



# Generalizing the prior

## simple case: the retrodictor's prior commutes with $\varrho$

Suppose that the retrodictor's uniform prior $u$ is replaced with another state $\gamma$, but such that $[\varrho, \gamma] = 0$, i.e., predictor's and retrodictor's priors commute.

Then, everything goes through: if we define

$$S^{\text{clax}}_{\mathbf{P},\gamma}(\varrho) := -\operatorname{Tr}[\varrho \ \log \gamma] + \sum_i p_i \log \frac{p_i}{q_i} \ ,$$

with $p_i := \operatorname{Tr}[\varrho \ P_i]$ and $q_i := \operatorname{Tr}[\gamma \ P_i]$, it is easy to check that

$$S^{\text{clax}}_{\mathbf{P},\gamma}(\varrho) - S(\varrho) = D(\varrho\|\gamma) - D(\mathcal{P}(\varrho)\|\mathcal{P}(\gamma)) = D(P_F\|P_R^\gamma) \ ,$$

with $P_F(x,i) = \lambda_x \langle\varphi_x|P_i|\varphi_x\rangle$ and $P_R^\gamma(x,i) = P_F(i)\frac{\gamma_x\langle\varphi_x|P_i|\varphi_x\rangle}{q_i}$.

But what if $[\varrho, \gamma] \neq 0$?



## a first candidate

A generalized deficiency-like definition is easy.

Even if $[\varrho, \gamma] \neq 0$, maintain $\Sigma^{(1)}_{\mathbf{P},\gamma} = D(\varrho\|\gamma) - D(\mathcal{P}(\varrho)\|\mathcal{P}(\gamma))$, that is, define $S^{(1)}_{\mathbf{P},\gamma}(\varrho) := -\operatorname{Tr}[\varrho \ \log \gamma] - D(\mathcal{P}(\varrho)\|\mathcal{P}(\gamma))$.

Instead, if $[\varrho, \gamma] \neq 0$, a generalized retrodiction-like definition is difficult, since these is no straightforward generalization of the joint input-output distribution for a quantum channel.

# Choi-like joint input-output representations

Arbitrarily fix o.n.b. $\{|i\rangle\}_{i=1}^{d}$ for input and $\{|\tilde{k}\rangle\}_{k=1}^{d'}$ for output.

Forward process $\varrho_{\rm in} \longmapsto \mathcal{P}(\varrho)_{\rm out}$:

- Choi operator: $C_{\mathcal{P}} := \sum_{i,j=1}^{d} \mathcal{P}(|i\rangle\langle j|) \otimes |i\rangle\langle j| = \sum_k |\tilde{k}\rangle\langle\tilde{k}| \otimes P_k^T$
- define $Q_F := (\mathbb{1}_{\rm out} \otimes \sqrt{\varrho^T})\, C_{\mathcal{P}}\, (\mathbb{1}_{\rm out} \otimes \sqrt{\varrho^T})$
- then $\mathrm{Tr}_{\rm out}[Q_F] = \varrho^T$ and $\mathrm{Tr}_{\rm in}[Q_F] = \mathcal{P}(\varrho)$

Reverse process $\sigma_{\rm out} \longmapsto \mathcal{R}_{\mathcal{P}}^{\gamma}(\sigma)_{\rm in}$:

- Choi operator: $C_{\mathcal{R}_{\mathcal{P}}^{\gamma}} := \sum_{k,\ell=1}^{d'} |\tilde{k}\rangle\langle\tilde{\ell}| \otimes \mathcal{R}_{\mathcal{P}}^{\gamma}(|\tilde{k}\rangle\langle\tilde{\ell}|)$
- it holds that $C_{\mathcal{R}_{\mathcal{P}}^{\gamma}}^T = (\mathcal{P}(\gamma)^{-1/2} \otimes \sqrt{\gamma^T})\, C_{\mathcal{P}}\, (\mathcal{P}(\gamma)^{-1/2} \otimes \sqrt{\gamma^T})$
- define $Q_R^{\gamma} := (\sqrt{\sigma} \otimes \mathbb{1}_{\rm in})\, C_{\mathcal{R}_{\mathcal{P}}^{\gamma}}^T\, (\sqrt{\sigma} \otimes \mathbb{1}_{\rm in})$
- then $\mathrm{Tr}_{\rm out}[Q_R^{\gamma}] = (\mathcal{R}_{\mathcal{P}}^{\gamma}(\sigma))^T$ and $\mathrm{Tr}_{\rm in}[Q_R^{\gamma}] = \sigma$



# a second candidate

Having the Choi-like representations $Q_F$ and $Q_R^{\gamma}$, we define

$$\Sigma_{\mathbf{P},\gamma}^{(2)}(\varrho) := D(Q_F \| Q_R^{\gamma}) ,$$

where we put $\sigma \equiv \mathcal{P}(\varrho)$.

But we face a dilemma, because $\Sigma_{\mathbf{P},\gamma}^{(1)}(\varrho) \neq \Sigma_{\mathbf{P},\gamma}^{(2)}(\varrho)$, i.e.,

$$D(\varrho\|\gamma) - D(\mathcal{P}(\varrho)\|\mathcal{P}(\gamma)) \neq D(Q_F\|Q_R^{\gamma}) .$$

(Proof by explicit numerical counterexamples).

Can we save goat and cabbages?

## how to save goat and cabbages



## how to save goat and cabbages

- instead of Umegaki's, use Belavkin–Staszewski's: $D_{BS}(\varrho\|\gamma) := \mathrm{Tr}[\varrho \log \varrho\gamma^{-1}]$ (assume $\gamma > 0$)
- instead of $Q_F$ use ${}^tQ_F := \sqrt{C_\mathcal{P}}\,(\mathbb{1}_{\text{out}} \otimes \varrho^T)\,\sqrt{C_\mathcal{P}}$
- instead of $Q_R^\gamma$ use ${}^tQ_R^\gamma := \sqrt{C_\mathcal{P}}\,(\mathcal{P}(\gamma)^{-1/2}\mathcal{P}(\varrho)\mathcal{P}(\gamma)^{-1/2} \otimes \gamma^T)\,\sqrt{C_\mathcal{P}}$

then:

$$D_{BS}(\varrho\|\gamma) - D(\mathcal{P}(\varrho)\|\mathcal{P}(\gamma)) = D_{BS}({}^tQ_F\|{}^tQ_R^\gamma)$$