# Macroscopic states and entropy: properties, meaning, and some recent results

Francesco Buscemi

Department of Mathematical Informatics, Nagoya University

QCMC 2024, IIT Madras, Chennai
30 August 2024

**Abstract**

In addition to the quantity now eponymously known as von Neumann entropy, in his 1932 book von Neumann also discusses another entropic quantity, which he calls “macroscopic”, and argues that it is the latter, and not the former, that is the relevant quantity to use in the analysis of thermodynamic systems. For a long time, however, von Neumann’s ”other” entropy was largely forgotten, appearing only sporadically in the literature, overshadowed by its more famous sibling. In this talk I will discuss a recent generalization of von Neumann’s macroscopic entropy, called “observational entropy”, focusing on its mathematical properties and logical interpretation, and presenting some recent results.

# Macroscopic states and entropy: interpretation, properties, and some recent results

Francesco Buscemi, Nagoya University

**QCMC 2024**, 26-30 August 2024
IIT Madras, Chennai



## 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 entropy

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

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

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



Unfortunately though:

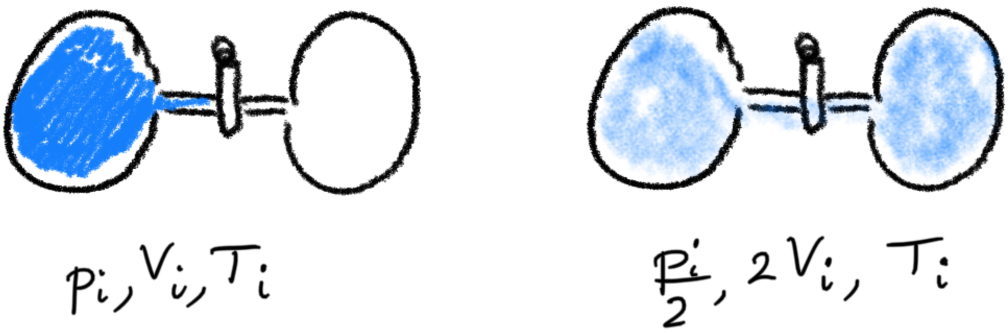
> *"The expressions for entropy given by the author [previously] are not applicable here in the way they were intended, as they were computed from the perspective of an observer who can carry out all measurements that are possible in principle—i.e., regardless of whether they are macroscopic [or not]."*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

And again:

> *"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




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

$\Longrightarrow$ there is net entropy increase in an isolated system's evolution, but von Neumann entropy remains constant

## 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)

von Neumann recognizes that thermodynamic entropy should be a quantity *relative* to the observer's knowledge



## von Neumann's proposal: macroscopic entropy

For

- $\varrho$ density matrix,
- $\mathbf{\Pi} = \{\Pi_i\}_i$ orthogonal resolution of identity (PVM),
- $p_i = \mathrm{Tr}[\varrho\, \Pi_i]$,
- $\Omega_i := \mathrm{Tr}[\Pi_i]$,

$$\boxed{S_{\mathbf{\Pi}}(\varrho) := -\sum_i p_i \log \frac{p_i}{\Omega_i}}$$

## modern version: observational entropy

For

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

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

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 $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n, \ldots\}$, where $\mathbf{X}_n = \{X_{j|n}\}_j$ are POVMs
- $\implies$ there exists one most-refined “parent” POVM $\mathbf{P} = \{P_i\}_i$ and a stochastic processing (i.e., cond. p.d.) $\mu$ such that

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

- hence, a macro-observer for von Neumann is “just” a POVM (i.e., the most-refined parent POVM $\mathbf{P}$) from which all macroscopic quantities can be simultaneously inferred

# Properties of OE



## 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 $\varrho, \gamma$

## the fundamental bound

### 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* $\mathbf{P}$*.*



## a better bound (arXiv:2209.03803)

### 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*

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

**Remark.** It could be $[\varrho, \tilde{\varrho}_{\mathbf{P}}] \neq 0$.

**Remark.** The reconstructed state $\tilde{\varrho}_{\mathbf{P}}$ only depends on the observer's knowledge.

## coarse-grainings and macroscopic states (arXiv:2404.11985)

### Definition (coarse-grainings)

A POVM $\mathbf{Q} = \{Q_j\}_j$ is a coarse-graining of another POVM $\mathbf{P} = \{P_i\}_i$, denoted by $\mathbf{Q} \preceq \mathbf{P}$, whenever there exists a p.d. $p(j|i)$ such that $Q_j = \sum_i p(j|i) P_i$, for all $j$.

### Definition (macrostates)

Given a POVM $\mathbf{P} = \{P_i\}_i$, the set of states macroscopic w.r.t. $\mathbf{P}$ is $\mathfrak{M}(\mathbf{P}) = \{\varrho : \varrho = \tilde{\varrho}_{\mathbf{P}}\}$. These are the states that can be perfectly reconstructed from the observer's knowledge.

### Theorem ($\star$)

*A state $\varrho$ is in $\mathfrak{M}(\mathbf{P})$ if and only if there exists a PVM $\mathbf{\Pi} = \{\Pi_j\}_j$, with $\mathbf{\Pi} \preceq \mathbf{P}$, together with coefficients $c_j \geq 0$, such that $\varrho = \sum_j c_j \Pi_j$.*

**Remark.** $\varrho \in \mathfrak{M}(\mathbf{P}) \implies [\varrho, P_i] = 0$ for all $i$.



## how OE resolves the paradox

- suppose that $\mathfrak{M}(\mathbf{P}) \supsetneq \{u\}$ and let the initial state of the system at time $t = t_0$ be a macrostate $\varrho^{t_0} \neq u$
- the system evolves unitarily, i.e., $\varrho^{t_0} \mapsto \varrho^{t_1} = U \varrho^{t_0} U^\dagger$; thus,

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

- summarizing: in general, $S_{\mathbf{P}}(\varrho^{t_1}) \geq S_{\mathbf{P}}(\varrho^{t_0})$, with equality if and only if $U\varrho^{t_0}U^\dagger \in \mathfrak{M}(\mathbf{P})$
- Corollary of Theorem ($\star$): $\varrho^{t_1} \in \mathfrak{M}(\mathbf{P}) \implies [\varrho^{t_1}, P_i] = [U\varrho^{t_0}U^\dagger, P_i] = 0$ for all $i$
- hence, when the initial state is a macrostate $\varrho^{t_0} \neq u$, $S_{\mathbf{P}}(\varrho^{t_1}) > S_{\mathbf{P}}(\varrho^{t_0})$ *generically*

## an "H theorem" for OE (arXiv:2404.11985)

**Theorem**

*In a $d$-dimensional system, choose a state $\varrho$ and a POVM $\mathbf{P} = \{P_i\}_i$ with a finite number of outcomes. Choose also a (small) value $\delta > 0$. For a unitary operator $U$ sampled at random according to the Haar distribution, it holds:*

$$\mathbb{P}_H \left\{ \frac{S_{\mathbf{P}}(U\varrho U^\dagger)}{\log d} \leq (1-\delta) \right\} \leq \frac{4}{\kappa(\mathbf{P})} e^{-C\delta\kappa(\mathbf{P})^2 d \log d} ,$$

*where $\kappa(\mathbf{P}) = \min_i \mathrm{Tr}[P_i\, u]$ and $C \approx 0.0018$.*

**Remark.** A similar statement holds for unitaries sampled from an approximate 2-design.

$\Longrightarrow$ in the eyes of the observer, the state of a randomly evolving system **quickly** becomes **indistinguishable** from the maximally uniform one, regardless of the system's initial state.



# Interpretation of OE

## retrodiction

- consider a discrete noisy channel $P : \mathcal{X} \to \mathcal{I}$ and a prior $\gamma(x)$ on the input
- when the receiver reads a definite value $i_0$, Bayes' update rule says that their posterior should be updated to $R_P^\gamma(x|i_0) \propto \gamma(x)P(i_0|x)$
- but what if the observation is noisy and returns some p.d. $\sigma(i)$ instead?

**Theorem (Jeffrey, 1965)**

*Starting from a given prior $\gamma(x)$ and a likelihood $P(i|x)$, the result of a noisy observation $\sigma(i)$ is retrodicted to*

$$\widetilde{\sigma}(x) := \sum_i R_P^\gamma(x|i)\sigma(i) \ .$$

*The conventional Bayes' rule is recovered for $\sigma(i) = \delta_{i,i_0}$.*



## irretrodictability

- a channel $P(i|x)$ is given (objective)
- the predictor has their prior $\pi(x)$
- the retrodictor has their prior $\gamma(x)$
- the predictor's expected data distribution is $P_F(x,i) := \pi(x)P(i|x)$
- the retrodicted channel is $R^\gamma(x|i) = \frac{1}{[P\gamma](i)}\gamma(x)P(i|x)$

**Definition**

The triple $(\pi, P, \gamma)$ is retrodictable whenever

$$P_F(x,i) = P_F(i)R^\gamma(x|i) =: P_R^\gamma(x,i) \ .$$

More generally, the degree of irretrodictability is given by

$$D(P_F \| P_R^\gamma) = D(\pi\|\gamma) - D(P\pi\|P\gamma) \ .$$

**Remark.** The above implies $D(\pi\|\gamma) - D([P\pi]\|[P\gamma]) \geq D(\pi\|[R^\gamma \circ P]\pi)$.

# OE as irretrodictability

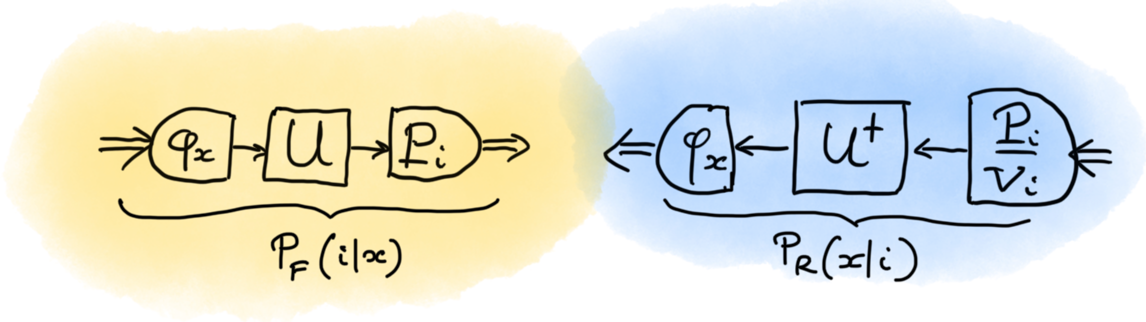
## Theorem

*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 $\mathbf{P} = \{P_i\}_i$,*

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

*where*

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






# parenthesis: Watanabe's contention

*"The phenomenological onewayness of temporal developments in physics is due to irretrodictability, and not due to irreversibility."* Satosi Watanabe (1965)

The reconstructed (i.e., retrodicted) state $\tilde{\varrho}_{\mathbf{P}}$ is exactly the coarse-grained state that appears in the Gibbsian "proof" of the second law.

# Extension to general priors



## two aspects of OE

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?
(especially relevant in thermodynamic situations, or for $\infty$-dim systems)

## first aspect: statistical deficiency

$\Sigma_{\mathbf{P}}(\varrho) = S_{\mathbf{P}}(\varrho) - S(\varrho) = D(\varrho\|u) - D(\mathcal{P}(\varrho)\|\mathcal{P}(u)) \rightsquigarrow$ replace $u$ with $\gamma$

### Definition

Given a POVM $\mathbf{P} = \{P_i\}_i$ and a prior state $\gamma > 0$, the set of macroscopic states is $\mathfrak{M}(\mathbf{P}, \gamma) := \{\varrho : D(\varrho\|\gamma) - D(\mathcal{P}(\varrho)\|\mathcal{P}(\gamma)) = 0\}$.

### Theorem

*A state $\varrho$ is in $\mathfrak{M}(\mathbf{P}, \gamma)$ if and only if there exists a coarse-graining $\mathbf{\Pi} \preceq \mathbf{P}$ such that*

1. $\mathbf{\Pi} = \{\Pi_j\}_j$ *is a PVM;*
2. $[\Pi_j, \gamma] = 0$ *for all $j$;*
3. $\varrho = \sum_j c_j \gamma \Pi_j$ *for some $c_j \geq 0$.*

**Remark.** Now, for a general prior, $\varrho \in \mathfrak{M}(\mathbf{P}, \gamma) \;\;\not\Rightarrow\;\; [P_i, \varrho] = 0$: indeed, $\gamma \in \mathfrak{M}(\mathbf{P}, \gamma)$, regardless of the POVM $\mathbf{P}$.



## simple case: semi-classical case

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.

- define

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

- then

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

  with

  - $P(i|x) = \langle\varphi_x|P_i|\varphi_x\rangle$
  - $P_F(x, i) = \lambda_x P(i|x)$
  - $P_R^\gamma(x, i) = P_F(i) R^\gamma(x|i)$.

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

## a first candidate

A generalized deficiency-like definition is easy.

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

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 $\varrho_{\text{in}} \longmapsto \mathcal{P}(\varrho)_{\text{out}}$:

- 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 $Q_F := (\mathbb{1}_{\text{out}} \otimes \sqrt{\varrho^T})\; C_{\mathcal{P}}\; (\mathbb{1}_{\text{out}} \otimes \sqrt{\varrho^T})$
- then $\operatorname{Tr}_{\text{out}}[Q_F] = \varrho^T$ and $\operatorname{Tr}_{\text{in}}[Q_F] = \mathcal{P}(\varrho)$

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

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

## a second candidate

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

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

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

But we face a dilemma, because $\Sigma^{(1)}_{\mathbf{P},\gamma}(\varrho) \neq \Sigma^{(2)}_{\mathbf{P},\gamma}(\varrho)$, 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 ${}^tQ_F := \sqrt{C_{\mathcal{P}}}\ (\mathbb{1}_{\mathrm{out}} \otimes \varrho^T)\ \sqrt{C_{\mathcal{P}}}$
- 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:

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



# Conclusions

## take-home messages

When the use of von Neumann entropy in thermodynamics is problematic, try consider observational entropy (OE) instead, because:

1. OE has a fully operational/inferential definition
2. OE fits nicely within recent developments in quantum mathematical statistics (e.g., approximate Petz recovery, strengthened monotonicity bounds, etc.)
3. OE simplifies a number of conceptual issues within the foundations of statistical mechanics

The End: Thank You!

# References


1. F. Buscemi and V. Scarani, *Fluctuation theorems from Bayesian retrodiction.* Physical Review E, vol. 103, 052111 (2021).

2. C.C. Aw, F. Buscemi, and V. Scarani, *Fluctuation theorems with retrodiction rather than reverse processes.* AVS Quantum Science, vol. 3, 045601 (2021).

3. F. Buscemi, J. Schindler, and D. Šafránek, *Observational entropy, coarse-grained states, and the Petz recovery map: information-theoretic properties and bounds.* New Journal of Physics, vol. 25, 053002 (2023).

4. G. Bai, D. Šafránek, J. Schindler, F. Buscemi, and V. Scarani, *Observational entropy with general quantum priors.* Preprint arXiv:2308.08763 (2023).

5. T. Nagasawa, K. Kato, E. Wakakuwa, and F. Buscemi, *On the generic increase of observational entropy in isolated systems.* Preprint arXiv:2404.11985 (2024).