Observing Microscopic Systems on a Macroscopic Scale: from Quantum Bayes' Rule to the Second Law

Francesco Buscemi

Department of Mathematical Informatics, Nagoya University

QMQI 2024, OIST, Japan
15 November 2024

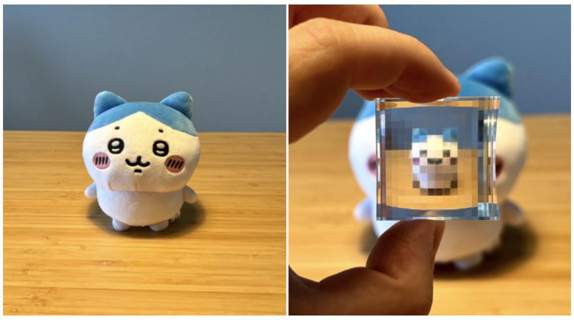
Abstract

In his 1932 book, von Neumann not only introduced the now familiar von Neumann entropy, but also discussed another entropic quantity that he called "macroscopic". He argued that this macroscopic entropy, rather than the von Neumann entropy, is the key measure for understanding thermodynamic systems. In this talk I will explore how macroscopic entropy, macroscopic states, and the emergence of the second law in isolated systems can be seen as consequences of a more general quantum Bayes' rule. This rule arises from a "minimum change" principle, just like the classical Bayes' rule, and recovers Petz's transpose map in several scenarios of physical interest. This is work done in collaboration with: Ge Bai, Kohtaro Kato, Teruaki Nagasawa, Valerio Scarani, and Eyuri Wakakuwa.

Observing Microscopic Systems on a Macroscopic Scale: from Quantum Bayes' Rule to the Second Law

Francesco Buscemi, Nagoya University

QMQI 2024, 11-15 November 2024
OIST, Okinawa



collaborators on this journey

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

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 ($\lambda_x \geqslant 0$, $\sum_x \lambda_x = 1$),

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

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

Unfortunately though:

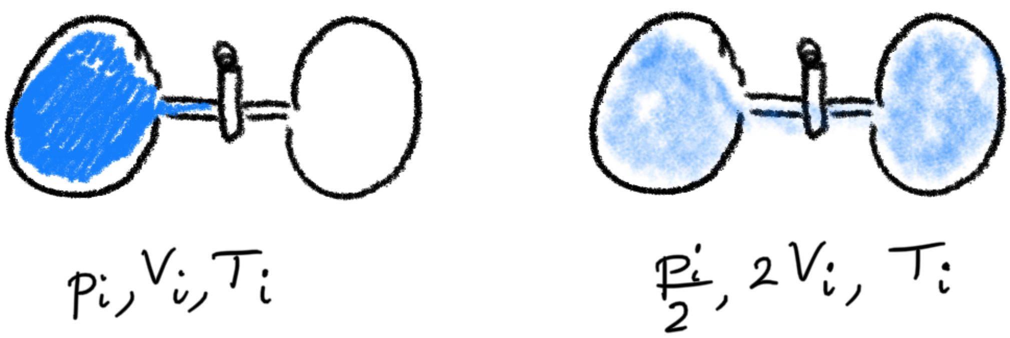
> *"The expressions for entropy given by the author [previously] are not applicable here in the way they were intended, as they were computed from the perspective of an observer who can carry out all measurements that are possible in principle—i.e., regardless of whether they are macroscopic [or not]."*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

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



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

$\Longrightarrow$ 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

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

$$S_{\mathbf{\Pi}}(\varrho) := -\sum_i p(i) \log \frac{p(i)}{\Omega(i)}$$

modern version: observational entropy

For

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

$$S_{\mathbf{P}}(\varrho) := -\sum_i p(i) \log \frac{p(i)}{V(i)}$$

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 $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n, \ldots\}$, 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.) μ 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 **P**) from which all macroscopic quantities can be simultaneously inferred

The meaning of ~~life~~ OE

Umegaki's quantum relative entropy

Definition

For density matrices ϱ, γ,

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

Useful properties:

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

In general, $S_{\mathbf{P}}(\varrho) = \log d - D(\mathcal{P}(\varrho) \| \mathcal{P}(u))$. By monotonicity, any postprocessing **P** $\to$ **Q** of the measurement outcomes lead to an increase of OE:

$$S_{\mathbf{Q}}(\varrho) \geqslant S_{\mathbf{P}}(\varrho) \ .$$

the fundamental bound (arXiv:2209.03803)

Theorem

For any d-dimensional density matrix ϱ and any POVM $\mathbf{P} = \{P_i\}_i$,

$$S(\tilde{\varrho}_{\mathbf{P}}) - S(\varrho) \geqslant S_{\mathbf{P}}(\varrho) - S(\varrho) \geqslant D(\varrho \| \tilde{\varrho}_{\mathbf{P}}) ,$$

where

$$\tilde{\varrho}_{\mathbf{P}} := \sum_i \mathrm{Tr}[\varrho \; P_i] \frac{P_i}{V_i} .$$

In particular, $\log d \geqslant S_{\mathbf{P}}(\varrho) \geqslant S(\varrho)$.

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

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

The macroscopic entropy tells us something about how much ϱ and $\tilde{\varrho}_{\mathbf{P}}$ "differ" from each other.

But what is the meaning of $\tilde{\varrho}_{\mathbf{P}}$?

Petz's transpose map

Petz (1986,1988)

Given a channel $\mathcal{E}$ and a prior state γ, the corresponding *transpose channel* is defined as

$$\mathcal{R}^{\gamma}_{\mathcal{E}}(\bullet) := \sqrt{\gamma}\mathcal{E}^{\dagger}\left[\frac{1}{\sqrt{\mathcal{E}(\gamma)}}\bullet\frac{1}{\sqrt{\mathcal{E}(\gamma)}}\right]\sqrt{\gamma}\,.$$

The "reconstructed" state

Defining the measurement channel $\mathcal{P}(\bullet) := \sum_i \mathrm{Tr}[P_i\ \bullet]\,|i\rangle\langle i|$, it turns out that

$$\tilde{\varrho}_{\mathbf{P}} = \mathcal{R}^{u}_{\mathcal{P}}\circ\mathcal{P}(\varrho) := \frac{1}{d}\mathcal{P}^{\dagger}[\mathcal{P}(u)^{-1/2}\ \mathcal{P}(\varrho)\ \mathcal{P}(u)^{-1/2}]$$

But what is the meaning of Petz's transpose map?

exact recovery

In general, for any channel $\mathcal{E}$ and any pair of states ϱ and γ, $D(\varrho\|\gamma) \geqslant D(\mathcal{E}(\varrho)\|\mathcal{E}(\gamma))$.

Question: for which triples $(\varrho, \gamma, \mathcal{E})$ does the equality $D(\varrho\|\gamma) = D(\mathcal{E}(\varrho)\|\mathcal{E}(\gamma))$ hold?

Petz (1986,1988)

Answer: if and only if $\mathcal{R}^\gamma_\mathcal{E} \circ \mathcal{E}(\varrho) = \varrho$. (The other equality $\mathcal{R}^\gamma_\mathcal{E} \circ \mathcal{E}(\gamma) = \gamma$ is satisfied *by construction*.)

But does Petz's transpose map also have a clear operational interpretation when $D(\varrho\|\gamma) > D(\mathcal{E}(\varrho)\|\mathcal{E}(\gamma))$?

Bayesian retrodiction

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

Theorem (Bayes–Jeffrey–Pearl retrodiction)

Given a channel $P(i|x)$ and a prior $\gamma(x)$, the result of a noisy observation $\sigma(i)$ is retrodicted to

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

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

When everything commutes, Petz's transpose map is coincides with the classical Bayes–Jeffrey–Pearl retrodiction rule.

But is this just a coincidence, or is there something deeper?

the principle of minimum change (arXiv:2410.00319)

"The updated belief should be consistent with the new observations, while deviating as little as possible from the initial belief."

Result for qc-channels

Given a qc-channel $\mathcal{P}(\bullet_{\text{in}}) = \sum_i \mathrm{Tr}[P_i \bullet] |i\rangle\langle i|_{\text{out}}$ and a prior state $\gamma > 0$ such that $\mathcal{P}(\gamma) > 0$, let $Q^\gamma_\mathcal{P} := \sum_i |i\rangle\langle i|_{\text{out}} \otimes \left(\sqrt{\gamma^T} P_i^T \sqrt{\gamma^T}\right)_{\text{in}}$. Then, given any observation result $\sigma(i)$, represented as $\sigma_{\text{out}} = \sum_i \sigma(i) |i\rangle\langle i|_{\text{out}}$, the optimization problem

$$\max_{\substack{Q \geqslant 0 \\ \mathrm{Tr}_{\text{in}}[Q] = \sigma_{\text{out}}}} F(Q^\gamma_\mathcal{P}, Q) ,$$

where $F(A, B) := \mathrm{Tr}\left[\sqrt{\sqrt{A} B \sqrt{A}}\right]$ is the (square-root) fidelity, has a unique solution $\tilde{Q}$, which in particular satisfies

$$\tilde{\sigma}_{\text{in}} := \mathrm{Tr}_{\text{out}}\left[\tilde{Q}^T\right] = \mathcal{R}^\gamma_\mathcal{P}(\sigma_{\text{out}}) = \sum_i \frac{\sqrt{\gamma} P_i \sqrt{\gamma}}{\mathrm{Tr}[P_i \, \gamma]} \sigma(i) .$$

According to the minimum change principle, Petz's transpose map is "the" quantum Bayes' rule...

...and thus $\tilde{\varrho}_{\mathbf{P}}$ is the quantum state to be retrodicted from the viewpoint of the macroscopic observer.

macroscopic states

Definition (macroscopic states)

Given a POVM $\mathbf{P} = \{P_i\}_i$ and a prior γ, the set of states macroscopic w.r.t. $\mathbf{P}$ and γ is $\mathfrak{M}^{\gamma}(\mathbf{P}) = \{\varrho : \varrho = (\mathcal{R}^{\gamma}_{\mathcal{P}} \circ \mathcal{P})(\varrho)\}$. These are the states that can be *perfectly retrodicted* from the observer's knowledge.

Theorem ($\star$)

A state ϱ is in $\mathfrak{M}^{\gamma}(\mathbf{P})$ if and only if there exists a PVM $\mathbf{\Pi} = \{\Pi_j\}_j$, with $\Pi_j = \sum_i \mu(j|i) P_i$, such that $[\Pi_i, \gamma] = 0$, together with coefficients $c_j \geqslant 0$, such that $\varrho = \sum_j c_j \Pi_j \gamma$.

Remark. The prior state is always macroscopic: $\gamma \in \mathfrak{M}^{\gamma}(\mathbf{P})$ for all POVMs $\mathbf{P}$.

Remark. For uniform prior, i.e., $\gamma = u$, $\varrho \in \mathfrak{M}^{u}(\mathbf{P}) \implies [\varrho, P_i] = 0$ for all i. (In general, it may be $[\gamma, P_i] \neq 0$.)

resolving the paradox of entropy increase in closed systems

- 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,

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

- summarizing: in general, $S_{\mathbf{P}}(\varrho^{t_1}) \geqslant S_{\mathbf{P}}(\varrho^{t_0})$, 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 ϱ 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:

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

where $\kappa(\mathbf{P}) = \min_i \operatorname{Tr}[P_i \ u]$ 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.

parenthesis: Watanabe's contention

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

1. The second law is not about the arrow of time, but rather it is about the arrow of inference.
2. The "mysterious" coarse-graining operation that appears in Gibbs' proof of the second law is nothing but Bayesian retrodiction.

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.* Quantum, vol. 8, 1524 (2024).

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).

6. G. Bai, F. Buscemi, and V. Scarani, *Quantum Bayes' rule and Petz transpose map from the minimal change principle.* Preprint arXiv:2410.00319 (2024).