Retrodictive thermodynamics and the "thermodynamic reverse bound"

Francesco Buscemi*

SUSTech-Nagoya workshop on Quantum Science, 24 Jun 2021

*www.quantumquia.com

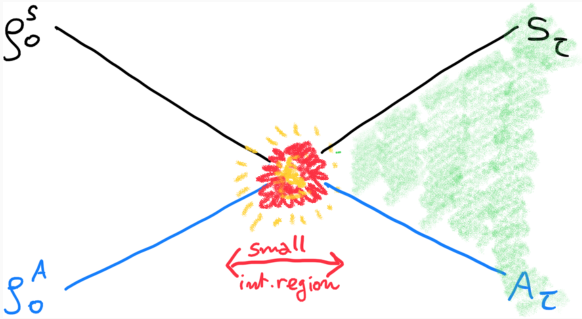
About this work

Four papers:

- with S. Das and M.M. Wilde. *Approximate reversibility in the context of entropy gain, information gain, and complete positivity*. Phys. Rev. A (2016). arXiv:1601.01207 [quant-ph]
- with N. Mitsui, D. Fujiwara, and M. Rotondo. *Thermodynamic reverse bounds for general open quantum processes*. Phys. Rev. A (2020). arXiv:2003.08548 [cond-mat.stat-mech]
- with V. Scarani. *Fluctuation relations from Bayesian retrodiction*. Phys. Rev. E (2021). arXiv:2009.02849 [quant-ph]
- with C.C. Aw and V. Scarani. *Fluctuation Theorems with Retrodiction rather than Reverse Processes*. arXiv:2106.08589 [cond-mat.stat-mech]

Open quantum systems thermodynamics

- system–ancilla initial factorization: $\rho_0^{SA} = \rho_0^S \otimes \rho_0^A$
- total Hamiltonian: $H^S(t) + H^A(t) + h^{SA}(t)$, for $0 \leq t \leq \tau$

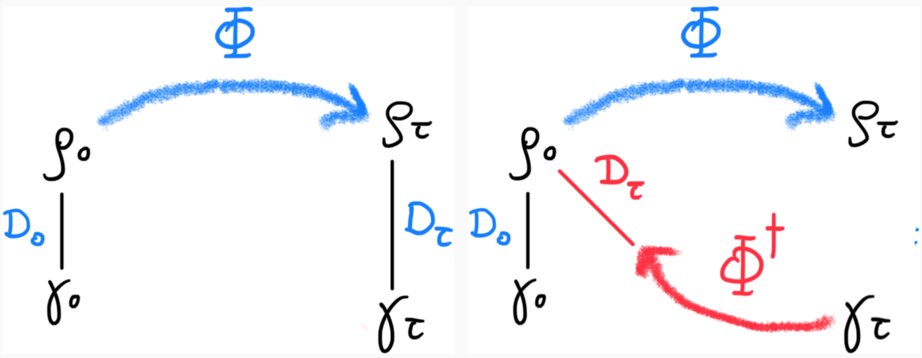


- $\implies \rho_\tau^S := \mathrm{Tr}_A \left\{ U^{SA}_{0\to\tau} \, (\rho_0^S \otimes \rho_0^A) \, (U^{SA}_{0\to\tau})^\dagger \right\} =: \Phi(\rho_0^S)$
- system's average energy change:
  $\Delta E \approx \mathrm{Tr}\{\rho_\tau^S \; H^S(\tau)\} - \mathrm{Tr}\{\rho_0^S \; H^S(0)\}$

Energy change as an information divergence

usual bound (γ is thermal state):

$$\beta(\Delta E - \Delta F) = \Delta D(\rho^S \| \gamma^S) + \Delta S \lesseqgtr \Delta D(\rho^S \| \gamma^S)$$



"reverse" bound:

$$\beta(\Delta E - \Delta F) \geq D(\rho_0^S \| \Phi^\dagger(\gamma_\tau^S)) - D(\rho_0^S \| \gamma_0^S)$$

Reverse bound and approximate recoverability

$$\begin{aligned}\Delta S &= S(\Phi(\rho)) - S(\rho)\\ &= D(\rho\|\Phi^\dagger(\sigma)) - D(\Phi(\rho)\|\sigma) + \mathrm{Tr}[\rho\ (\underbrace{\ln \Phi^\dagger\sigma - \Phi^\dagger \ln\sigma}_{[\ln,\Phi^\dagger]\geq 0})]\\ &\geq D(\rho\|\Phi^\dagger(\sigma)) - D(\Phi(\rho)\|\sigma)\end{aligned}$$

- assumption: $\Phi^\dagger$ a positive, unit-preserving linear map (Choi, 1974)
- for $\sigma = \Phi(\rho)$:

$$S(\Phi(\rho)) - S(\rho) \geq D(\rho\|\Phi^\dagger \circ \Phi(\rho))\ ,$$

particularly meaningful for bistochastic channels

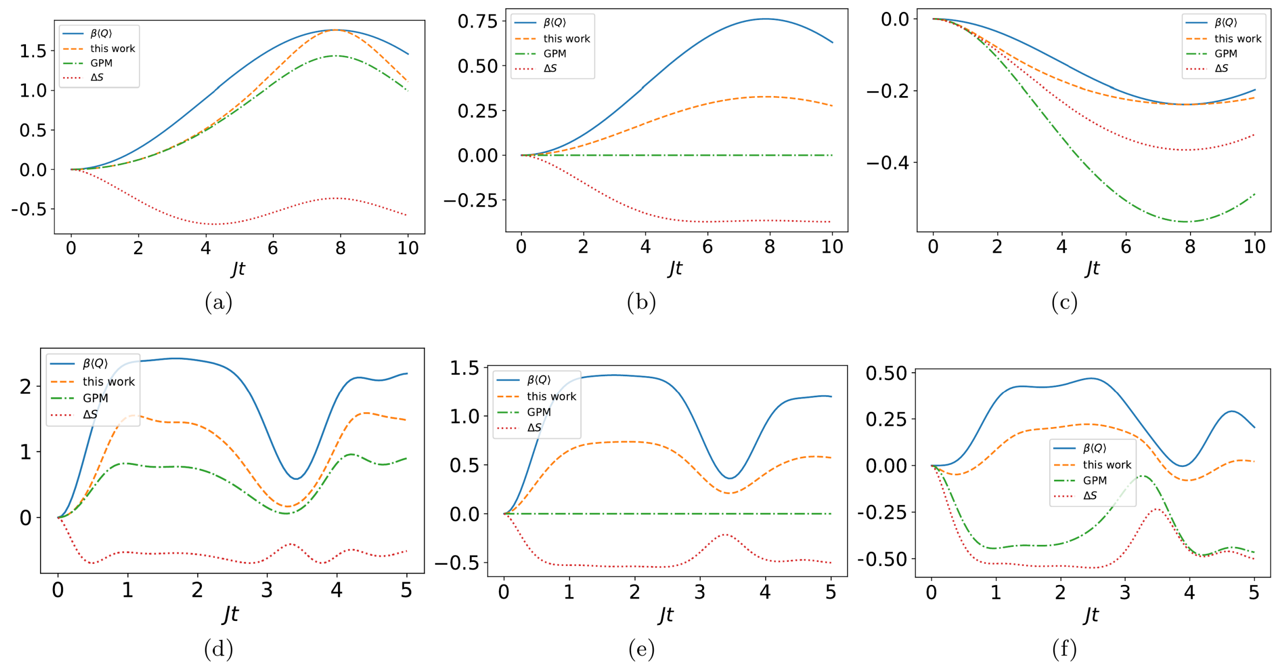
The “thermal pullback”

$$\begin{aligned}&\beta(\Delta E - \Delta F)\\ \geq& D(\rho_0^S\|\Phi^\dagger(\gamma_\tau^S)) - D(\rho_0^S\|\gamma_0^S)\\ =& D(\rho_0^S\|\tilde\gamma_0^S) - D(\rho_0^S\|\gamma_0^S) - \ln \mathrm{Tr}\{\Phi^\dagger(\gamma_\tau^S)\} \qquad \text{with } \tilde\gamma_0^S := \frac{\Phi^\dagger(\gamma_\tau^S)}{\mathrm{Tr}\{\Phi^\dagger(\gamma_\tau^S)\}}\end{aligned}$$

- the value $\mathrm{Tr}\{\Phi^\dagger(\gamma_\tau^S)\}$ is called *efficacy*: it often appears in fluctuation relations (e.g., Albash&al 2013, Goold&al 2015)
- the pullback mapping $x \to \frac{\Phi^\dagger(x)}{\mathrm{Tr}\{\Phi^\dagger(x)\}}$ is CPTP but (in general) nonlinear

Application: erasure processes

$\rho_0^S = \gamma_0^S$, $\Delta F = 0$, $\Delta E = \langle Q \rangle$ [see PRA 102, 032210 (2020)]

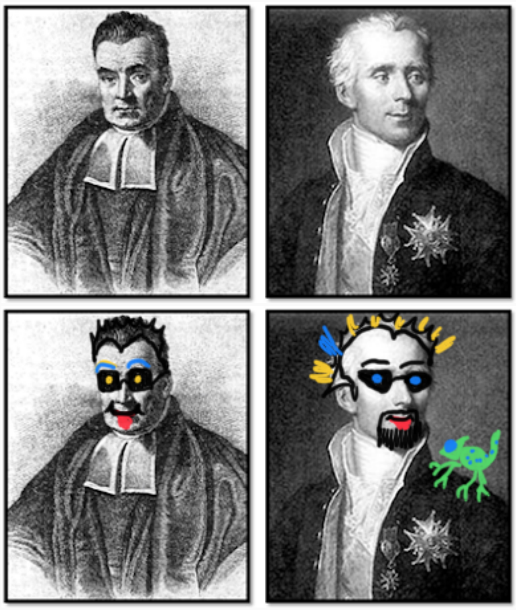


Does the pullback mapping

$$x \to \frac{\Phi^\dagger(x)}{\mathrm{Tr}\{\Phi^\dagger(x)\}}$$

remind us of anything?

The Bayes-Laplace Rule



Inverse Probability Formula

$$\underbrace{\mathcal{P}(H|D)}_{\text{inv. prob.}} \propto \underbrace{\mathcal{P}(D|H)}_{\text{likelihood}} \underbrace{\mathcal{P}(H)}_{\text{prior}}$$

where H is a hypothesis, D is the result of observation (i.e., evidence)

postmodern Bayesianism!

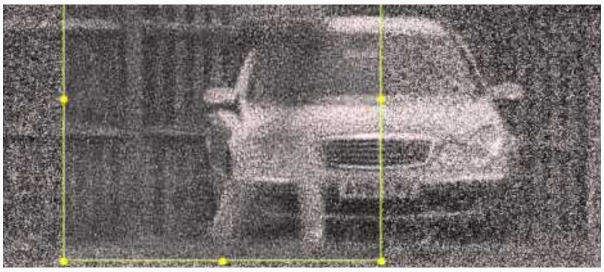
Meanings of the inverse probability

- it is the main *tool* of Bayesian statistics for problems like:
  - estimation (e.g.: how many red balls are in an urn?)
  - decision (e.g.: is ACME's stock a good investment? should I buy some?)
  - predictive inference (e.g.: weather forecasts)
  - retrodictive inference (e.g.: what kind of stellar event possibly caused the Crab Nebula?)
- it measures the degree of belief that a rational agent should have in one hypothesis, among other mutually exclusive ones, given the data

Noisy data and uncertain evidence

BUT! Bayes-Laplace Rule *does not* tell us how to update the prior in the face of uncertain data...

- suppose that a noisy observation suggests a probability distribution $\mathcal{Q}(D)$ for the data (e.g., the license plate no.)



- how should we update our prior $\mathcal{P}(H)$ given *uncertain evidence* in the from $\mathcal{Q}(D)$?

Jeffrey's rule of probability kinematics

Vanilla Bayes:

$$\mathcal{P}(H|D) = \mathcal{P}(D|H)\mathcal{P}(H)/\mathcal{P}(D)$$

Extended Bayes:

$$\mathcal{P}(H|\mathcal{Q}(D)) = ?$$

Jeffrey's conditioning* (1965)

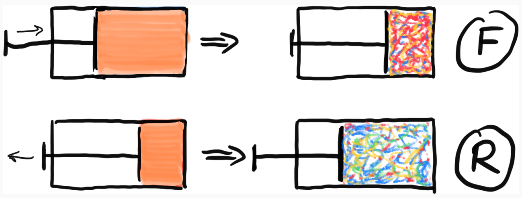
$$\mathcal{P}(H|\mathcal{Q}(D)) = \sum_D \underbrace{\mathcal{P}(H|D)}_{\text{inv. prob.}} \mathcal{Q}(D)$$

$$= \sum_D \frac{\mathcal{P}(D|H)\mathcal{P}(H)}{\sum_H \mathcal{P}(D|H)\mathcal{P}(H)} \mathcal{Q}(D)$$

* Jeffrey's rule was introduced *ad hoc*, but it can be proved from Bayes-Laplace Rule and Pearl's method of virtual evidence (1988)

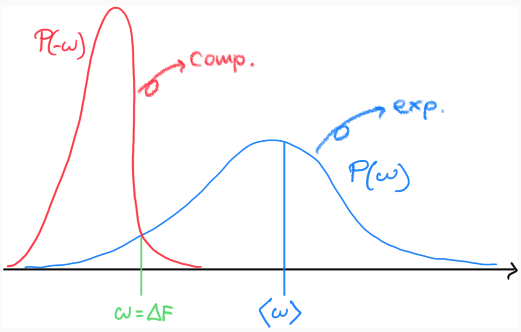
Reverse processes and fluctuation relations in thermodynamics

Reverse processes and the Second Law



Crooks' fluctation theorem (1999)

$$\frac{P_F(W)}{P_R(-W)} = e^{\beta(W-\Delta F)}$$



Crooks $\implies$ Jarzynski $\implies$ Clausius

Usual explanation

Crooks' theorem, and hence Jarzynski's relation, and hence the Second Law, all rely on two assumptions satisfied at equilibrium:

1. thermal equilibrium: initial distribution is $P(\xi) \propto e^{-\beta\epsilon(\xi)}$
2. microscopic reversibility (cf. *detailed balance*): molecular processes and their reverses occur at the same rate

So, do fluctuation relations (and the second law) require the existence of some microscopic "balancing mechanisms"?

A hint from Ed Jaynes

"To understand and like thermo we need to see it, not as an example of the n-body equations of motion, but as an example of the logic of ***scientific inference****."*

E.T. Jaynes (1984)

First idea: reverse process as Bayesian retrodiction

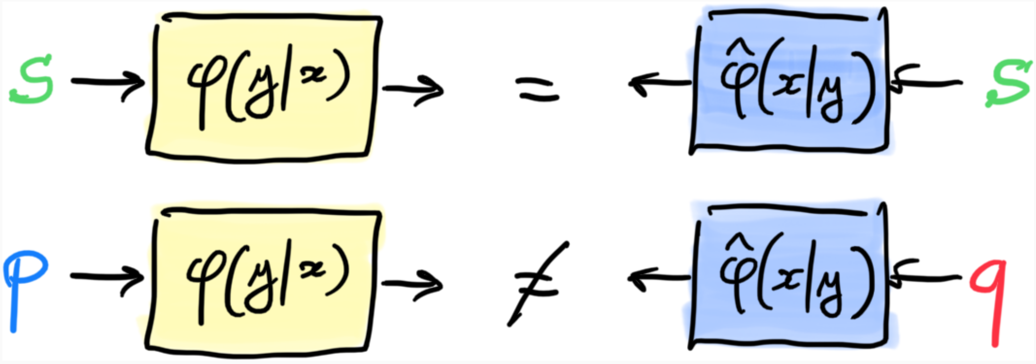
Retrodictive construction of the reverse process

- **physical setup**:
  - a stochastic transition rule: $\varphi(y|x)$
  - a steady (viz. invariant) state: $\sum_x \varphi(y|x)s(x) = s(y)$
- **Bayesian inversion at the steady state**:

$$s(y)\hat{\varphi}(x|y) := s(x)\varphi(y|x) \iff \frac{\varphi(y|x)}{\hat{\varphi}(x|y)} = \frac{s(y)}{s(x)}$$

- **two priors**:
  - predictor's prior: $p(x)$
  - retrodictor's prior $q(y)$
- **two processes**:
  - forward process (prediction): $\mathcal{P}_F(x, y) = \varphi(y|x)p(x)$
  - reverse process (retrodiction): $\mathcal{P}_R(x, y) = \hat{\varphi}(x|y)q(y)$

A picture



- at the steady state: prediction = retrodiction
- otherwise: asymmetry (irreversibility, *irretrodictability*)

Measures of statistical divergence

Second idea: fluctuation relations as measures of *statistical divergence* between $\mathcal{P}_F(x,y)$ and $\mathcal{P}_R(x,y)$

- relative entropy:

$$D(\mathcal{P}_F \| \mathcal{P}_R) := \left\langle -\ln \frac{\mathcal{P}_R(x,y)}{\mathcal{P}_F(x,y)} \right\rangle_F =: \langle -\ln r(x,y) \rangle_F$$

$\rightsquigarrow$ more generally, one can use $D_f(\mathcal{P}_R \| \mathcal{P}_F) := \langle f(r(x,y)) \rangle_F$

- introduce probability density functions

$\rightsquigarrow$ $\Omega(x,y) := f(r(x,y))$ (f-entropy production)

$\rightsquigarrow$ $\mu_F(\omega) := \sum_{x,y} \delta[\omega - \Omega(x,y)]\ \mathcal{P}_F(x,y)$

$\rightsquigarrow$ $\mu_R(\omega) := \sum_{x,y} \delta[\omega - \Omega(x,y)]\ \mathcal{P}_R(x,y)$

$$\Longrightarrow \langle \omega \rangle_F = D_f(\mathcal{P}_R \| \mathcal{P}_F)$$

From f-divergences to f-fluctuation theorems

- for $f : \mathbb{R}^+ \to \mathbb{R}$ invertible

f-Fluctuation Theorem

$$\mu_R(\omega) = f^{-1}(\omega)\mu_F(\omega) \quad \Longrightarrow \quad \langle f^{-1}(\omega)\rangle_F = 1$$

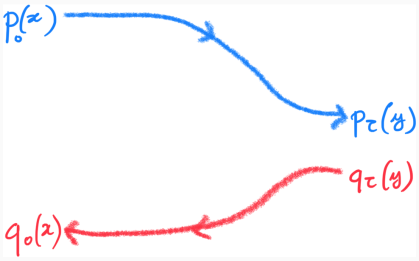
$\rightsquigarrow$ for $f(u) = -\ln u$, we have $f^{-1}(v) = e^{-v}$, that is

$$\frac{\mu_F(\omega)}{\mu_R(\omega)} = e^{\omega} \quad \Longrightarrow \quad \left\langle e^{-\omega}\right\rangle_F = 1$$

further discussions in arXiv:2009.02849 and arXiv:2106.08589

Examples of known results recovered as retrodiction

Example: driven closed system evolution



- driving protocol: $H(0) \to H(t) \to H(\tau)$
- $H(0) = \sum_x \epsilon_x \pi_x$, $H(\tau) = \sum_y \eta_y \pi'_y$
- $\varphi(y|x) = \delta_{y,y(x)}$, i.e., one-to-one
- $s(x) = d^{-1} \implies \varphi(y|x) = \hat{\varphi}(x|y)$
- $p_0(x) = e^{\beta(F-\epsilon_x)}$, $q_\tau(y) = e^{\beta(F'-\eta_y)}$

In this case, for the choice $f(u) = -\ln u$,

$$\Omega(x,y) = \ln \frac{\mathcal{P}_F(x,y)}{\mathcal{P}_R(x,y)} = \ln \frac{s(y)p(x)}{s(x)q(y)} = \ln \frac{p(x)}{q(y)}$$
$$= \beta(F - \epsilon_x + F' + \eta_y) = \beta(W - \Delta F)$$

$$\implies \frac{\mu_F(W)}{\mu_R(W)} = e^{\beta(W-\Delta F)}$$

Example: nonequilibrium steady states

- stochastic process $\varphi(y|x)$ with non-thermal steady state $s(x)$
- thermal equilibrium priors: $p(x) = q(x) \propto e^{-\beta\epsilon_x}$
- fluctuation variable:
  $\omega = \ln \frac{\mathcal{P}_F(x,y)}{\mathcal{P}_R(x,y)} = \ln \frac{p(x)}{q(y)} \frac{s(y)}{s(x)} = \beta(\epsilon_y - \epsilon_x) + (\ln s(y) - \ln s(x))$
- *nonequilibrium potential*: $V(x) := -\frac{1}{\beta} \ln s(x)$ (e.g., Manzano&al 2015)
- $\left\langle e^{\beta(\Delta E - \Delta V)} \right\rangle_F = 1$, but $\left\langle e^{\beta \Delta E} \right\rangle_F =$ "efficacy"
- $\implies$ nonequilibrium potentials (usually introduced *ad hoc*) are understood here as remnants of Bayesian inversion

But why known relations are compatible with Bayesian inversion?

Is that a necessity?

Sketch argument

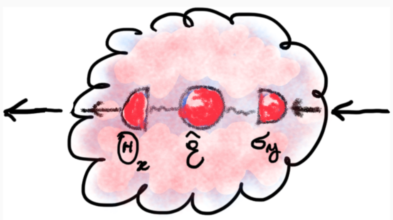
- $D(\mathcal{P}_F \| \mathcal{P}_R) = \left\langle \ln \frac{\mathcal{P}_F(x,y)}{\mathcal{P}_R(x,y)} \right\rangle_F$
- let us impose that the fluctuation variable is **local**:
  $\ln \frac{\mathcal{P}_F(x,y)}{\mathcal{P}_R(x,y)} = \Omega(x,y) \stackrel{!}{=} G'(y) - G(x)$
  - $\implies \quad \frac{\mathcal{P}_F(y|x)}{\mathcal{P}_R(x|y)} = \frac{H'(y)}{H(x)}$
  - $\implies \quad H(x)\mathcal{P}_F(y|x) = H'(y)\mathcal{P}_R(x|y)$
  - sum over $x \implies \quad H'(y) = \sum_x H(x)\mathcal{P}_F(y|x)$
- $\implies \quad \mathcal{P}_R(x|y) = \frac{1}{\sum_x H(x)\mathcal{P}_F(y|x)} H(x)\mathcal{P}_F(y|x)$

Hence, a Bayesian inverse-like form for the reverse process is **inevitable** if we want the fluctuating variable to have a local form!

Finally, what about the quantum case?

Quantum retrodiction and the Petz map

- assume $\varphi(y|x) = \mathrm{Tr}[\Pi_y \; \mathcal{E}(\rho_x)]$
- let $s(x)$ be invariant distribution
- according to the formalism of *quantum retrodiction*:
  - $\Sigma := \sum_x s(x)\rho_x$
  - $\hat{\rho}_y := \frac{1}{s(y)}\sqrt{\mathcal{E}(\Sigma)}\Pi_y\sqrt{\mathcal{E}(\Sigma)}$
  - $\hat{\Pi}_x := s(x)\frac{1}{\sqrt{\Sigma}}\rho_x\frac{1}{\sqrt{\Sigma}}$
  - $\hat{\mathcal{E}}(\cdot) := \sqrt{\Sigma}\left\{\mathcal{E}^\dagger\left[\frac{1}{\sqrt{\mathcal{E}(\Sigma)}}(\cdot)\frac{1}{\sqrt{\mathcal{E}(\Sigma)}}\right]\right\}\sqrt{\Sigma}$



- Bayesian inversion works seamlessly $\hat{\varphi}(x|y) = \mathrm{Tr}[\hat{\Pi}_x \; \hat{\mathcal{E}}(\hat{\rho}_y)]$

Some remarks

- the Petz recovery map reduces to Bayes–Laplace rule when operators commute
- to a unique Bayes–Laplace rule there correspond infinite possible Petz maps ("rotated" Petz maps)
- retrodiction (both classical and quantum) depends on the choice of reference prior
- exceptions are unitary (i.e., "bilateral deterministic") channels, for which:
  1. there is a unique Petz reverse (the retrodiction is independent of the choice of prior, and all rotated Petz maps coincide)
  2. retrodiction and (linear) inversion coincide

Conclusion

Final messages

1. retrodiction (viz. pullback) can be a useful concept in quantum thermodynamics and approximate recoverability
2. predictive and retrodictive inference provide the logical foundations of fluctuation theorems
3. the Second Law is special among the laws of physics, because it is in fact a law of logic
4. a clear distinction between mechanical *(ir)reversibility* and logical *(ir)retrodictability* avoids unnecessary paradoxes
5. quantum retrodiction and quantum fluctuation relations follow seamlessly using Petz recovery map

thank you