Reverse Bounds, Bayesian Retrodiction, and the Second Law of Thermodynamics

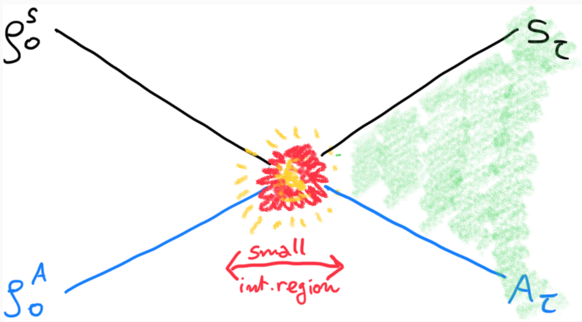
Francesco Buscemi*

NCTS Annual Theory Meeting, 17 Feb 2021

*www.quantumquia.com

Open quantum systems evolution

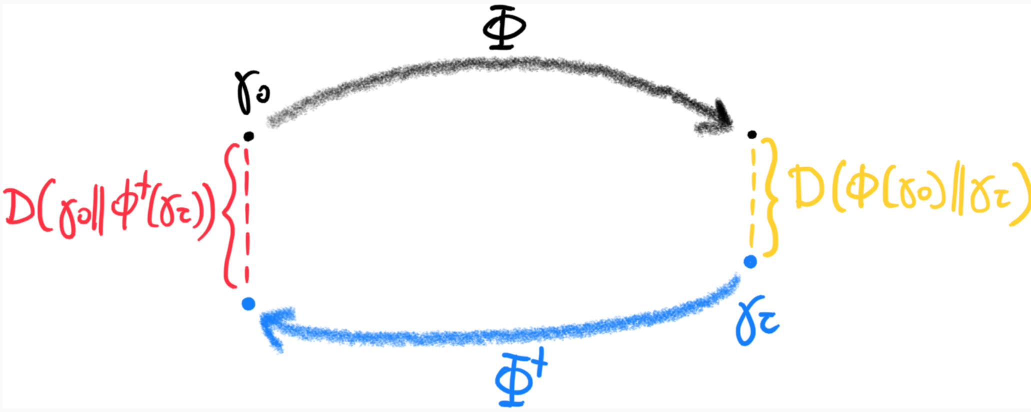
- system–ancilla initial factorization: $\rho_0^{SA} = \rho_0^S \otimes \rho_0^A$
- total Hamiltonian: $H^S(t) + H^A(t) + h^{SA}(t)$, for $0 \le t \le \tau$



- $\rightsquigarrow \rho_\tau^S := \mathrm{Tr}_A \left\{ U_{0\to\tau}^{SA} \, (\rho_0^S \otimes \rho_0^A) \, (U_{0\to\tau}^{SA})^\dagger \right\} =: \Phi(\rho_0^S)$
- 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

$$\beta(\Delta E - \Delta F) = \Delta S + D(\Phi(\gamma_0^S)\|\gamma_\tau^S) \lesseqqgtr D(\Phi(\gamma_0^S)\|\gamma_\tau^S)$$



$$\begin{aligned}\beta(\Delta E - \Delta F) &= \mathrm{Tr}\{\gamma_0^S\ [\ln \Phi^\dagger(\gamma_\tau^S) - \Phi^\dagger(\ln \gamma_\tau^S)]\} + D(\gamma_0^S\|\Phi^\dagger(\gamma_\tau^S)) \\ &\geq D(\gamma_0^S\|\Phi^\dagger(\gamma_\tau^S))\end{aligned}$$

The "thermal pullback"

$$\begin{aligned}\beta(\Delta E - \Delta F) &\geq D(\gamma_0^S\|\Phi^\dagger(\gamma_\tau^S)) \\ &= D\left(\gamma_0^S \,\middle\|\, \frac{\Phi^\dagger(\gamma_\tau^S)}{\mathrm{Tr}\{\Phi^\dagger(\gamma_\tau^S)\}}\right) - \ln \mathrm{Tr}\{\Phi^\dagger(\gamma_\tau^S)\} \\ &\geq -\ln \mathrm{Tr}\{\Phi^\dagger(\gamma_\tau^S)\}\end{aligned}$$

- 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

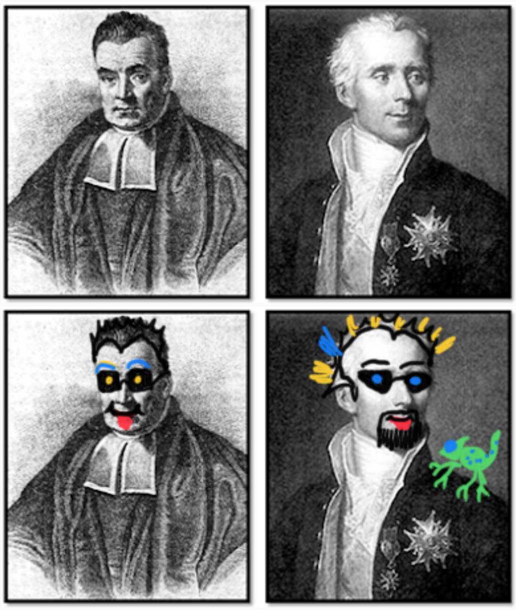
further discussions&applications in `arXiv:2003.08548`

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!

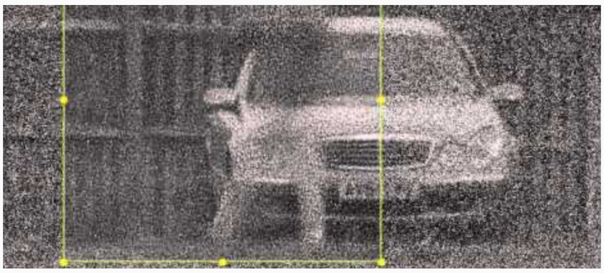
Meanings of the inverse probability

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

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

Jeffrey's conditioning* (1965)

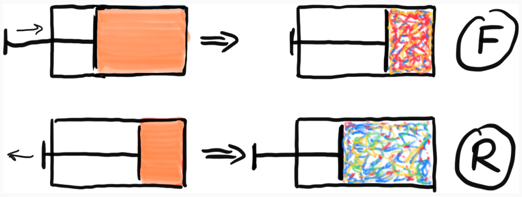
$$\mathcal{P}(H|\mathcal{Q}(D)) = \sum_D \underbrace{\mathcal{P}(H|D)}_{\text{inv. prob.}} \mathcal{Q}(D)$$

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

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

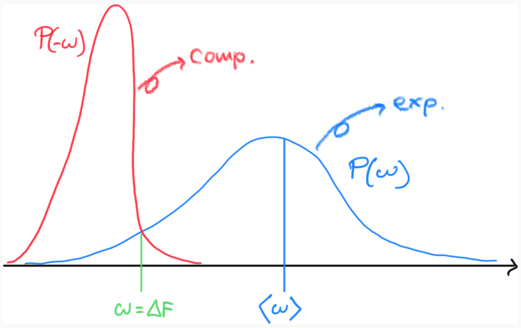
Reverse processes and fluctuation relations in thermodynamics

Reverse processes and the second law



Crooks' fluctuation theorem (1999)

$$\frac{\mathcal{P}_F(W)}{\mathcal{P}_R(-W)} = e^{\beta(W-\Delta F)}$$



$$\frac{\mathcal{P}_F(W)}{\mathcal{P}_R(-W)} = e^{\beta(W-\Delta F)} \implies \left\langle e^{-\beta(W-\Delta F)} \right\rangle_F = 1 \implies \langle W \rangle \geq \Delta F$$

What's behind this?

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

Do fluctuation relations (and the second law) rely on 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

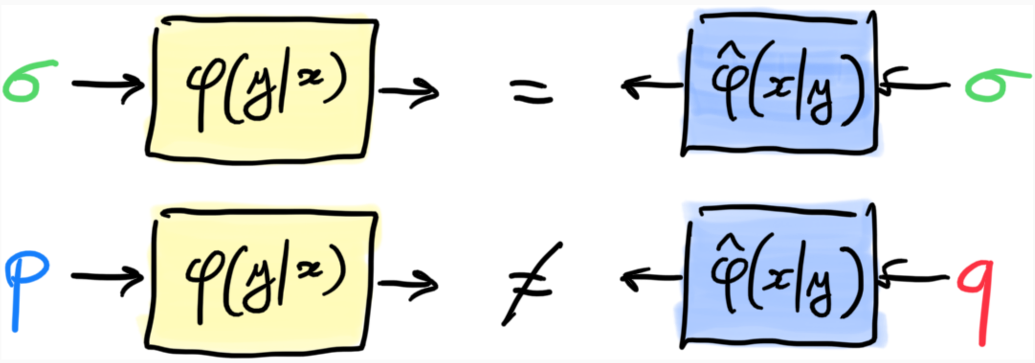
Construction of the reverse process

- starting point:
  - a stochastic transition rule: $\varphi(y|x)$
  - a steady (viz. invariant) state: $\sum_x \varphi(y|x)\sigma(x) = \sigma(y)$
- define reverse transition by Bayesian inversion at steady state:

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

- 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 steady state: prediction = retrodiction
- otherwise: asymmetry

Measures of statistical divergence

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

- f-divergences: $D_f(\mathcal{P}_F \| \mathcal{P}_R) := \sum \mathcal{P}_F(x,y) f(\frac{\mathcal{P}_F(x,y)}{\mathcal{P}_R(x,y)})$

  ⇝ $f(r) = \ln(r) \implies D_f$ is KL-divergence (viz. relative entropy)

  ⇝ $f(r) = r^\alpha$, $\alpha \neq 0 \implies D_f$ is a Hellinger-Rényi divergence

- introduce probability density functions

  ⇝ $\mu_F^f(u) := \sum_{x,y} \delta[u - f(r(x,y))]\ \mathcal{P}_F(x,y)$

  ⇝ $\mu_R^f(u) := \sum_{x,y} \delta[u - f(\frac{1}{r(x,y)})]\ \mathcal{P}_R(x,y)$

From f-divergences to f-fluctuation theorems

- for $f : \mathbb{R}^+ \to \mathbb{R}$ smooth and invertible, define $g := f \circ \frac{1}{x} \circ f^{-1}$

  ⇝ $f(r) = \ln(r) \implies g(r) = -r$

  ⇝ $f(r) = r^\alpha \implies g(r) = \frac{1}{r}$

f-Fluctuation Theorem

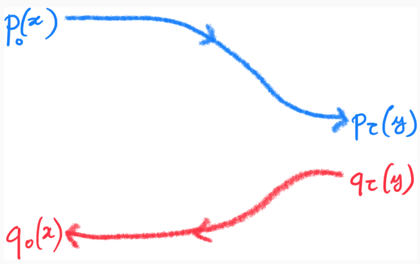
$$\frac{\mu_F^f(u)}{\mu_R^f(g(u))} = \frac{|g'(u)|}{f^{-1}(g(u))} \qquad \implies \qquad \left\langle f^{-1}(g(u)) \right\rangle_F = 1$$

⇝ for $f = \ln$, we have $\frac{\mu_F(u)}{\mu_R(-u)} = e^u$ and $\langle e^{-u} \rangle_F = 1$

further discussions in `arXiv:2009.02849`

Examples

Example: driven Hamiltonian 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
- $\sigma(x) = d^{-1} \implies \varphi(y|x) = \hat{\varphi}(x|y)$
- $p_0(x) = e^{\beta(F-\epsilon_x)}$, $q_\tau(y) = e^{\beta(F'-\eta_y)}$

In this case, for the choice $f(r) = \ln r$ (viz. $g(r) = -r$),

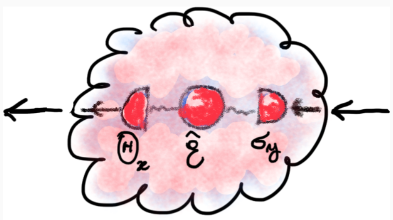
$$u(x,y) = \ln \frac{\mathcal{P}_F(x,y)}{\mathcal{P}_R(x,y)} = \ln \frac{\sigma(y)p(x)}{\sigma(x)q(y)} = \ln \frac{p(x)}{q(y)}$$
$$= \beta(F - \epsilon_x + F' + \eta_y) = \beta(W - \Delta F)$$

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

Example: nonequilibrium steady states

- stochastic process $\varphi(y|x)$ with non-thermal steady state $\sigma(x)$
- thermal equilibrium priors: $p(x) = q(x) \propto e^{-\beta\epsilon_x}$
- fluctuation variable:
  $u = \ln \frac{\mathcal{P}_F(x,y)}{\mathcal{P}_R(x,y)} = \ln \frac{p(x)}{q(y)} \frac{\sigma(y)}{\sigma(x)} = \beta(\epsilon_y - \epsilon_x) + (\ln \sigma(y) - \ln \sigma(x))$
- *nonequilibrium potential*: $V(x) := -\ln \sigma(x)$ (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

Example: quantum processes



- assume $\varphi(y|x) = \mathrm{Tr}[\Pi_y \ \mathcal{E}(\rho_x)]$
- according to the formalism of *quantum retrodiction*:
  - $\Sigma := \sum_x \sigma(x)\rho_x$
  - $\hat{\rho}_y := \frac{1}{\sigma(y)} \sqrt{\mathcal{E}(\Sigma)} \Pi_y \sqrt{\mathcal{E}(\Sigma)}$
  - $\hat{\Pi}_x := \sigma(x) \frac{1}{\sqrt{\Sigma}} \rho_x \frac{1}{\sqrt{\Sigma}}$
  - $\hat{\mathcal{E}}(\cdot) := \sqrt{\Sigma} \left\{ \mathcal{E}^\dagger \left[\frac{1}{\sqrt{\mathcal{E}(\Sigma)}} (\cdot) \frac{1}{\sqrt{\mathcal{E}(\Sigma)}} \right] \right\} \sqrt{\Sigma}$
- Bayesian inversion carries through directly $\hat{\varphi}(x|y) = \mathrm{Tr}[\hat{\Pi}_x \ \hat{\mathcal{E}}(\hat{\rho}_y)]$

Conclusions

Summary

- role of retrodiction (viz. Jeffrey conditioning) in thermodynamics and statistical mechanics
- reverse process not as *physical time-reversal*, but as retrodiction
- fluctuation relations (FRs) as quantitative measures (f-divergences) of asymmetry between prediction and retrodiction
- FRs not from *complex microscopic balancing mechanisms*, but from consistent inference (viz. Bayes-Laplace rule)
- logical origin of the perceived "one-wayness" of time

thank you