Single-shot thermodynamics meets Crooks fluctuation theorem

Oscar Dahlsten
with Aaberg, Browne*, del Rio, Garner*, Egloff, Renner, Rieper, Yunger-Halpern & Vedral

*=here at conference

oscar.dahlsten@physics.ox.ac.uk
quantumoxford.com

UNIVERSITY OF OXFORD

h-bar kT, Grenoble, 30 Sep 2014

‘Single-shot thermodynamics’ meets ‘Crooks fluctuation theorem’

1. By *Crook’s fluctuation theorem* I mean

$$\frac{P_{fwd}(W)}{P_{rev}(-W)} = e^{\beta(W-\Delta F)}$$

   (I will define the terms later.)

2. Single-shot statistical mechanics is an approach to non-equilibrium statistical mechanics focussing on behaviour that is guaranteed to occur each realization (‘shot’) of an experiment, as opposed to average behaviour. It is inspired by single-shot information theory.

Key message: They are concerned with the same scenarios and quantities, so insights from one can be applied to the other.

Overview

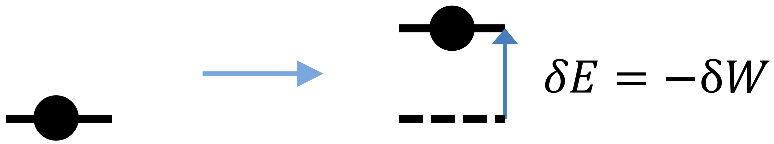
1. General intro to single-shot stat. mech.
2. General intro to Crooks fluctuation theorem.
   Key claim: W and P(W) used in the respective results refer to the same quantities in the game.
3. The setting for W and p(W): level-shifting work extraction protocols (games).
4. Crooks from this setting
5. Single-shot statements in this setting
6. Initial examples of advantages of link: theory and experiment.

General Intro to single-shot statistical mechanics

"Look, if you had one shot, or one opportunity
To seize everything you ever wanted.
One moment
Would you capture it or just let it slip?" Eminem

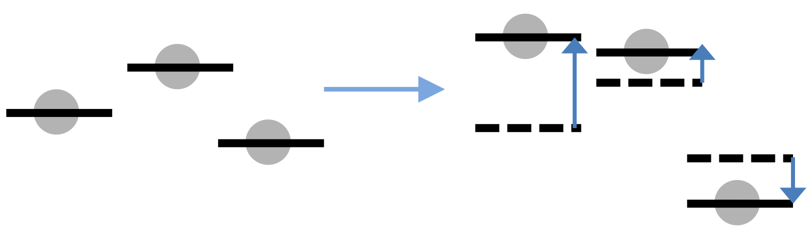
The amount of work out is not always certain.

- Consider a system in an energy level. Suppose you use a battery system to change that level by δE. The change in the battery systems energy is called work, $\delta W = -\delta E$.



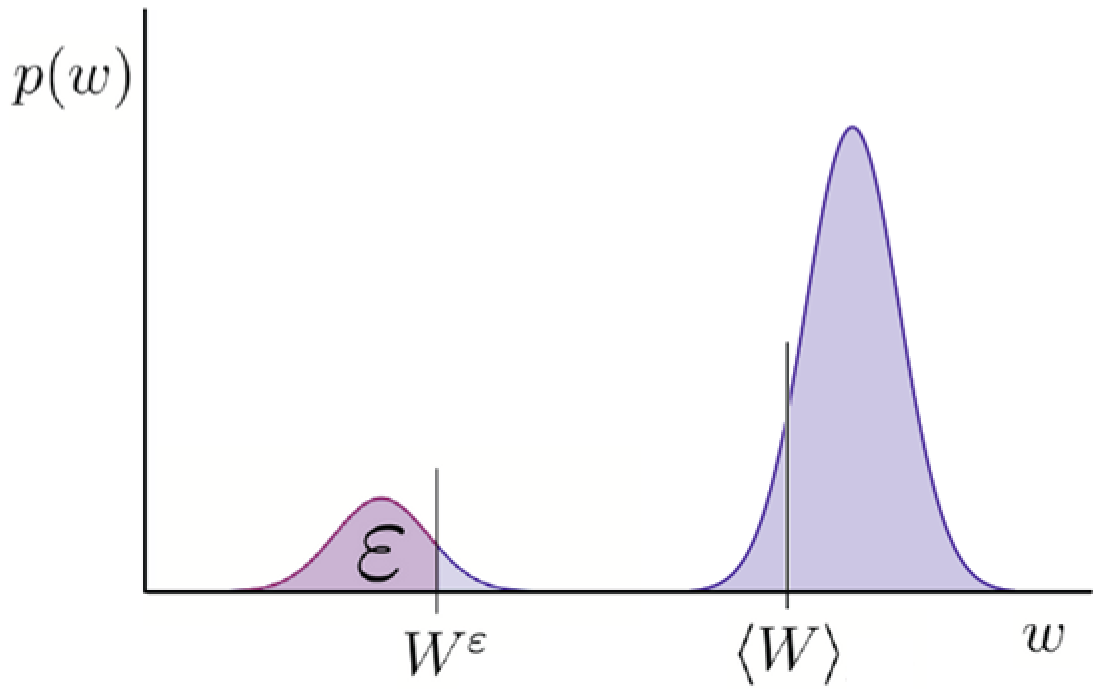
- Note also that shifting an unoccupied energy level costs no work.

- Now suppose there are several levels, and you do not know the which is occupied fully. As you lift/lower the i-th levels by dE_i, there is now a *probability distribution over work* $p(W)$.



Work guaranteed in every single shot

- Consider the guaranteed work for a given p(W): $W^{\varepsilon}(P(W))$.
- This is guaranteed (up to probability ε) to be extracted each and every run, every 'shot'. Its the sort of thing we study in **single-shot statistical mechanics**.

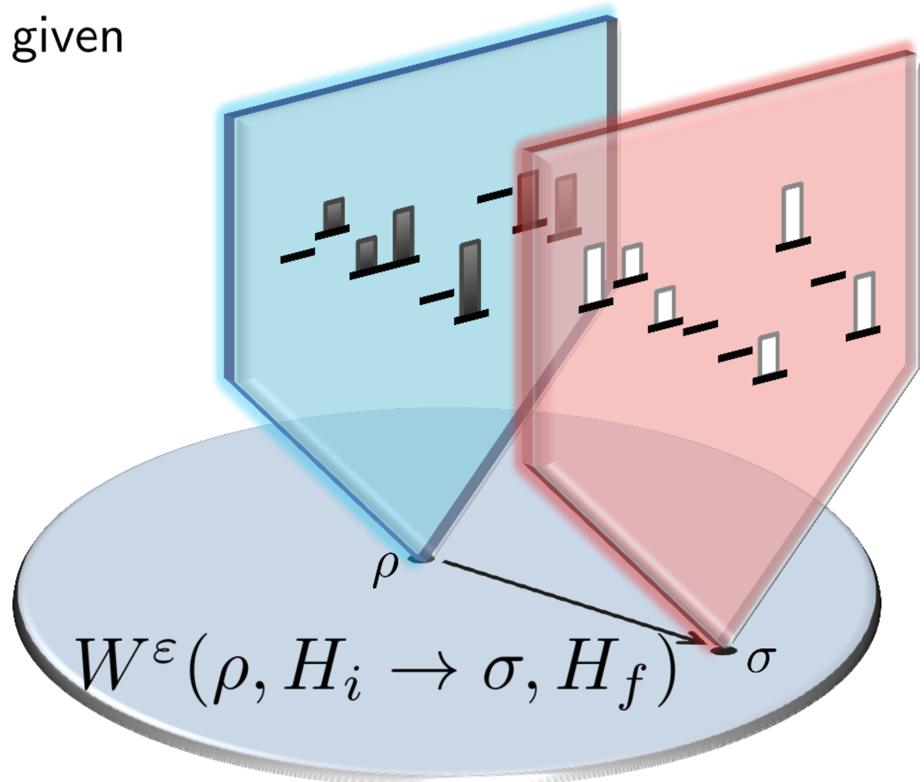


- Why care? Suppose your engine needs to provide a certain amount of work e.g. as an activation energy, or that it cannot dissipate more than a certain amount of heat without breaking something. Then guaranteed behaviour rather than average is what matters.

What can we calculate in single-shot approach?

- Various papers have calculated the optimal guaranteed*) work for given initial and final energy levels, and initial and final states.

- The optimisation is over all Hamiltonian paths and thermalisations, for the given initial and final conditions.



*) Actually some are concerned with *deterministic* work, where one demands the work output distribution is fully concentrated around one value.

What can we calculate in single-shot approach?

More and more general expressions for the optimal ϵ-guaranteed work.

$$W^{\varepsilon}(\rho, H_i = 0 \to \gamma_T, H_f = 0) = (n - H^{\varepsilon}_{\max}(\rho))\, kT \ln 2$$

H_i: initial Hamiltonian, ρ: initial state; γ_T, thermal state, H_f: final Hamiltonian.

$$W^{\varepsilon}(\rho_{SM}, H_i = 0 \to |0\rangle\langle 0| \otimes \rho_M, H_i = 0) = H^{\varepsilon}_{\max}(S|M) kT \ln 2(+log....)$$

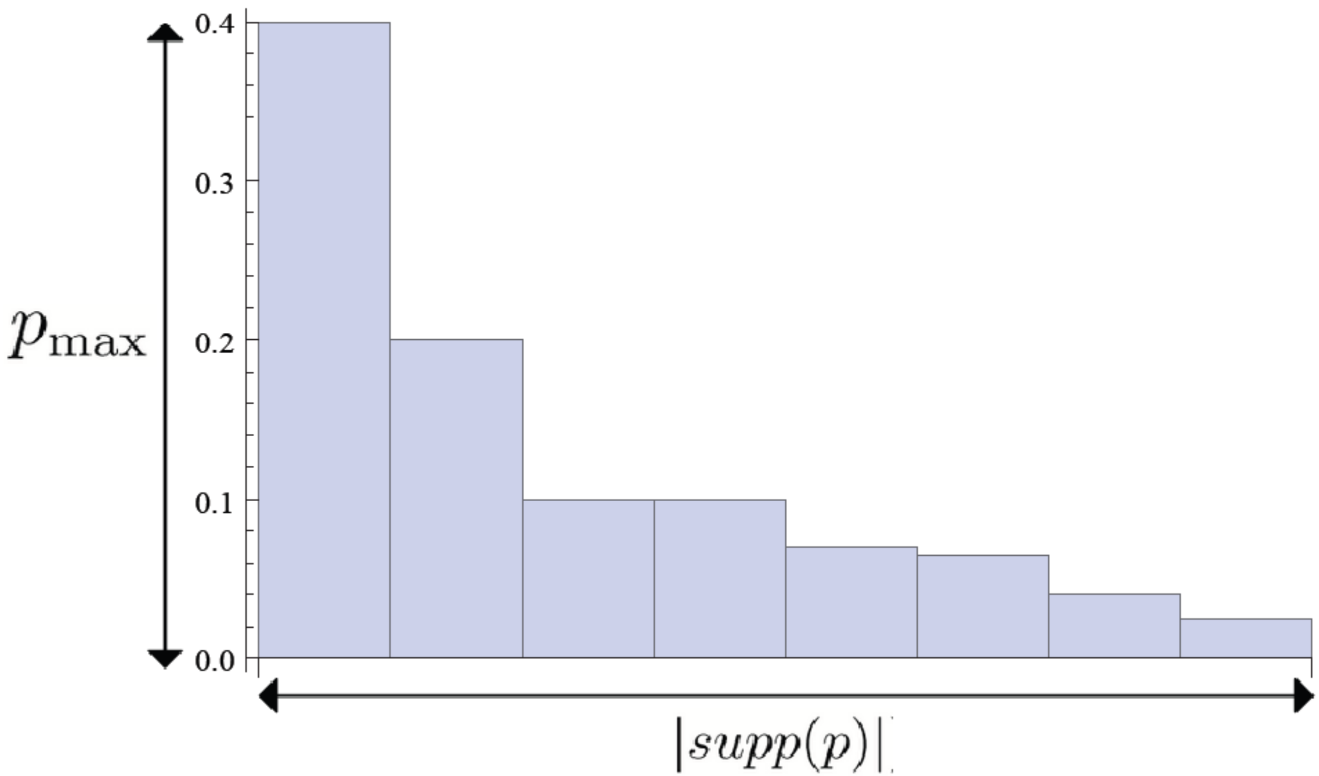
$$W^{\varepsilon}(\rho, H_i \to \gamma_T, H_i) = kT \ln(2) D^{\varepsilon}_0(\rho||\gamma_T)$$

$$W^{\varepsilon}(\rho, H_i \to \sigma, H_f) = kT \ln\left(M\left(\frac{G^T(\rho)}{1-\varepsilon}||G^T(\sigma)\right)\right)$$

What are these quantities on the RHS? They are quantities from (or given to) **single-shot information theory**.

Single-shot information theory entropies

- $H_{\max}(p) = \log(|supp(p)|)$ (and another single shot entropy $H_{\min} = \log(\frac{1}{p_{\max}})$).
- An important operational meaning of $H_{\max}$ is that it is the size a memory needs to have to reliable retain information. Not on average in some sense but in a given experiment. It is called a single-shot entropy.



Smoothing entropies

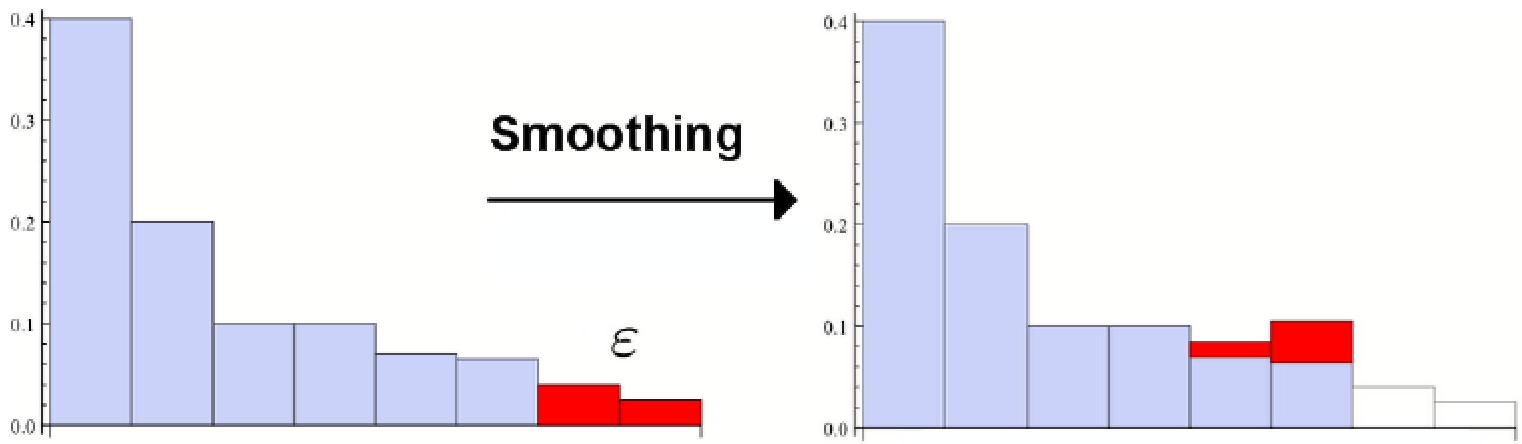
- Recall that there was also an ε on the entropies in the earlier slide, e.g.

$$W^{\varepsilon}(\rho, H_i = 0 \rightarrow \gamma_T, H_f = 0) = (n - H^{\varepsilon}_{\max}(\rho))\, kT \ln 2$$

- An important technique is to **smooth** a distribution before evaluating the entropy. This gives the **smooth entropy**

  $H^{\varepsilon}_{\max}(\rho) := \min H_{\max}(\rho') \,| d(\rho, \rho') \leq \varepsilon$

  If d(.,.) is the trace distance the smoothing looks like this.



- Interpretation: the entropy is effectively $H^{\varepsilon}_{\max}(\rho)$.

- One reason smoothing important: $\lim(n \rightarrow \infty, \varepsilon \rightarrow 0) H^{\varepsilon}_{\max}(\rho^{\otimes n}) = nH(\rho)$

Papers to date on this approach

Title	Authors	arXiv	Journal
Inadequacy of von Neumann entropy for characterizing extractable work	Dahlsten, Renner, Rieper, Vedral	0908.0424	NJP
Thermodynamic meaning of negative entropy	del Rio, Aaberg, Renner, Dahlsten, Vedral	1009.1630	Nature
Fundamental limitations for quantum and nano thermodynamics	Horodecki, Oppenheim	1111.3834	NCOMMS
Truly work-like work extraction	Aaberg	1110.6121	NCOMMS
The laws of thermodynamics beyond the von Neumann Regime	Egloff, Dahlsten, Renner and Vedral	1207.0434	
A quantitative Landauers Principle	Faist, Dupuis, Oppenheim, Renner	1211.1037v1	
The resource theory of informational non-equilibrium in thermodynamics	Gilad, Muller, Narasim-hachar, Spekkens, Yunger-Halpern	1309.6586v1	

Papers to date on this approach, cont'd

Title	Authors	arXiv	Journal
The second laws of quantum thermodynamics	Brandao, Horodecki, Ng, Oppenheim, Wehner	1305.5278	
Non-equilibrium Thermodynamics inspired by Modern Information Theory (a soft intro)	Dahlsten		Entropy, Maroney Ed.
Guaranteed energy efficient bit reset in finite time.	Browne, Garner, Dahlsten, Vedral	1311.7612	PRL
Unifying fluctuation theorems and single-shot statistical mechanics	Yunger-Halpern, Dahlsten, Garner, Vedral	1409.3878	

There are now more papers which I have not had a chance to read and add here

Quick recap of Crooks theorem

Crooks theorem on fluctuations of work

Crooks fluctuation theorem is that under certain assumptions

$$\frac{P_{\mathrm{fwd}}(W)}{P_{\mathrm{rev}}(-W)} = e^{\beta(W-\Delta F)},$$

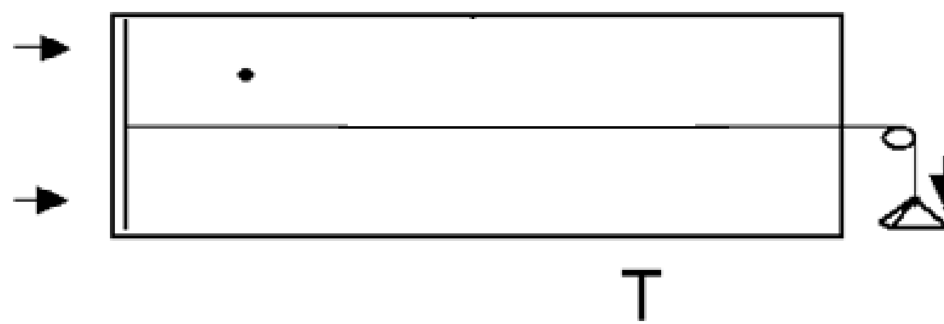
W: work output

$P_{\mathrm{fwd}}(W)$: probability distribution of work for a given path the Hamiltonian in one direction

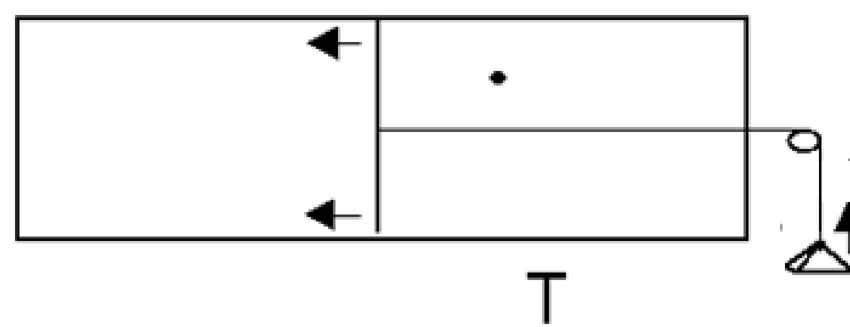
$P_{\mathrm{rev}}(W)$ for Hamiltonian taking the reverse path

ΔF: the free energy difference between two thermal states at the beginning and end points of the Hamiltonian path.

FWD Process

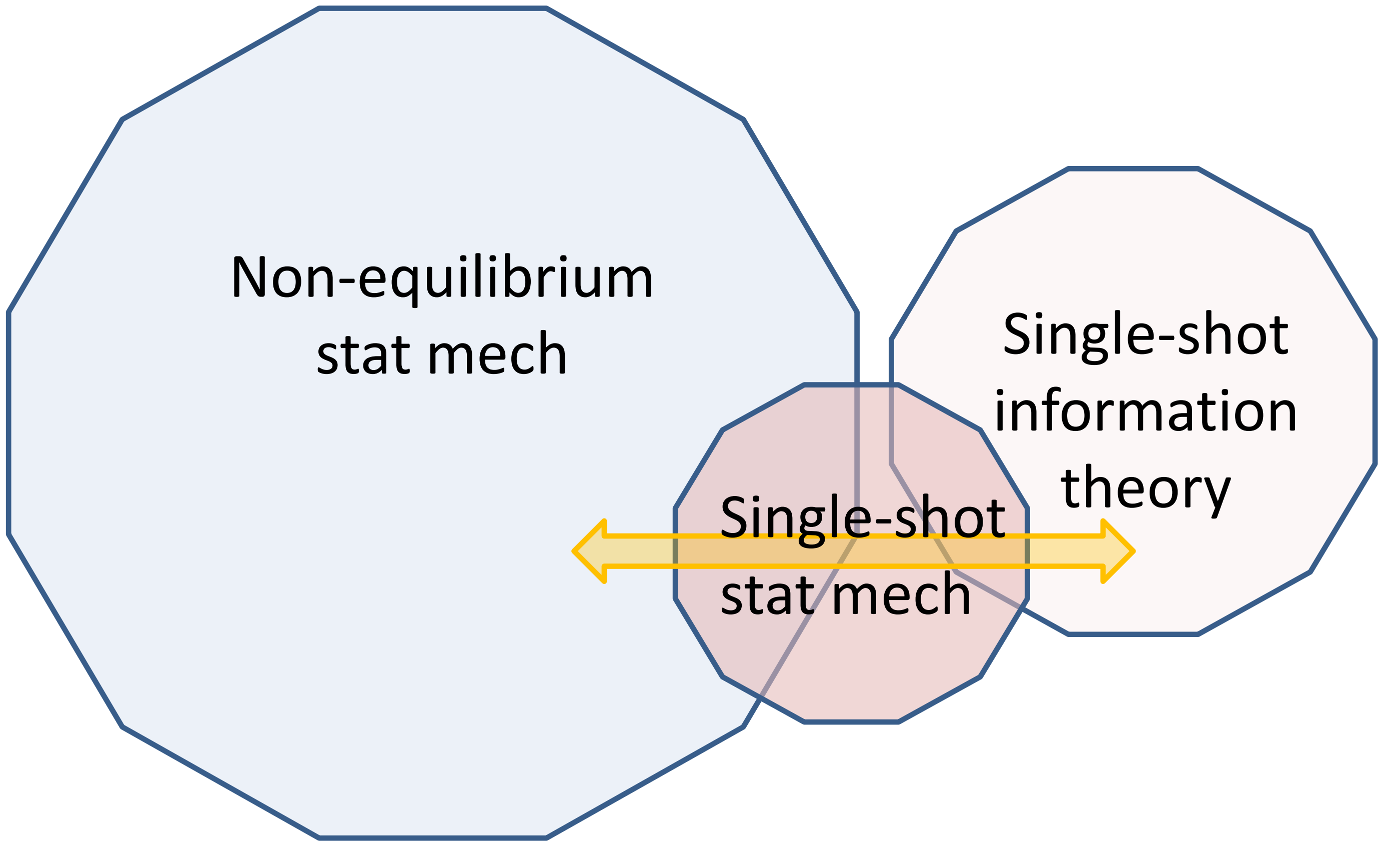


REV Process

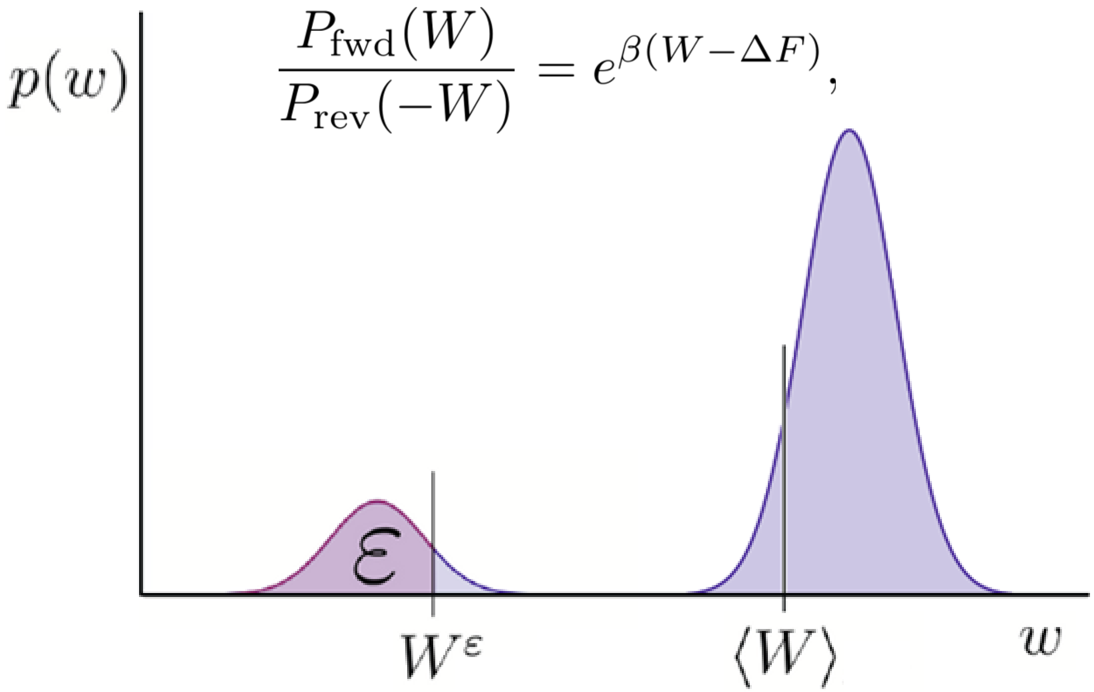


Why look for a link between the two?

- Single-shot approach can benefit from the wide range of theoretical and experimental papers relating to fluctuation theorems.
- Fluctuation results can be linked to modern (I mean single-shot) information theory as well as the move away from entropy to majorisation as the central object of thermodynamics.



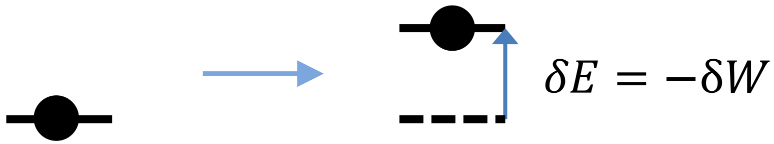
Key link: p(w) in single-shot papers represents same physical thing as that in Crooks scenario.



- I will now explain a work extraction model before showing how Crooks and single-shot results both apply in that model.

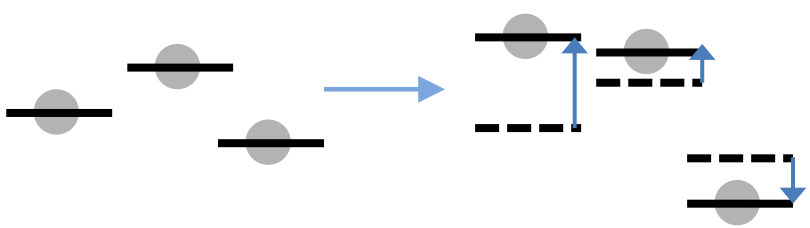
The setting for W, p(W) -in pictures

- Consider a system in an energy level. Suppose you use a battery system to change that level by δE. The change in the battery systems energy is called work, $\delta W = -\delta E$.



- Note also that shifting an unoccupied energy level costs no work.

- Now suppose there are several levels, and you do not know the which is occupied fully. As you lift/lower the i-th levels by dE_i, there is now a *probability distribution over work* $p(W)$.



- There are also thermalisations, hopping the system between levels, doing no work.

The setting for W, p(W)- in writing

Hamiltonian function of λ_m, m is time step. System undergoes two alternating changes:

1. *Hamiltonian changes*. Take $\lambda_m \to \lambda_{m+1}$, energy level invariant*, i.e. $|i_n, \lambda_m\rangle \to |i_n, \lambda_{m+1}\rangle$. This costs work $\delta W = E(|i_n, \lambda_{m+1}\rangle - E(|i_n, \lambda_m\rangle)$.
2. *Thermalisations*. No work cost: $\delta W = 0$. They hop system around levels, without altering λ. Assume thermal detailed balance respected:

$$\frac{p(|i_n, \lambda_m\rangle \to |i_{n+1}, \lambda_m\rangle)}{p(|i_{n+1}, \lambda_m\rangle \to |i_n, \lambda_m\rangle} = \exp -\beta \left(E(|i_{n+1}, \lambda_m\rangle - E|i_n, \lambda_m\rangle\right).$$

System then evolves according to a (random) *trajectory*:

$$|i_0, \lambda_0\rangle \to |i_0, \lambda_1\rangle \to |i_1, \lambda_1\rangle \to |i_{f-1}, \lambda_f\rangle ... |i_f, \lambda_f\rangle.$$

In the forwards (reverse) direction the initial state is thermal w.r.t. the Hamiltonian associated with λ_0 (λ_f).

*) Quantumly superpositions may arise, see later slide.

Why one might think of Crooks theorem here

$$\begin{aligned} p(traj) &= p(|i_0, \lambda_0\rangle, |i_1, \lambda_1\rangle, ..., |i_f, \lambda_f\rangle) \\ &= p(|i_0, \lambda_0\rangle) p(|i_0, \lambda_1\rangle \to |i_1, \lambda_1\rangle) ... p(|i_{f-1}, \lambda_f\rangle \to |i_f, \lambda_f\rangle)); \end{aligned}$$

$$\begin{aligned} p(traj - inv) &= p(|i_f, \lambda_f\rangle, |i_{f-1}, \lambda_{f-1}\rangle, ..., |i_0, \lambda_0\rangle) \\ &= p(|i_f, \lambda_f\rangle) p(|i_f, \lambda_f\rangle \to |i_{f-1}, \lambda_f\rangle) ... p(|i_1, \lambda_1\rangle \to |i_0, \lambda_1\rangle)); \end{aligned}$$

Hmmm, will many of these things cancel or become something neat if we take the ratio?
Especially if we assume the thermalisations respect detailed balance:

$$\frac{p(|i_0, \lambda\rangle \to |i_1, \lambda\rangle)}{p(|i_1, \lambda\rangle \to |i_0, \lambda\rangle)} = \exp(-\beta(E(i_0) - E(i_1)))$$

Crooks holds in this set-up (1/2)

$$
\begin{aligned}
&\frac{p(traj)}{p(traj-inv)} = \\
= \quad & \frac{p(|i_0,\lambda_0\rangle)}{p(|i_f,\lambda_f\rangle)} \frac{p(|i_0,\lambda_1\rangle \to |i_1,\lambda_1\rangle) ... p(|i_{f-1},\lambda_f\rangle \to |i_f,\lambda_f\rangle))}{p(|i_f,\lambda_f\rangle \to |i_{f-1},\lambda_f\rangle) ... p(|i_1,\lambda_1\rangle \to |i_0,\lambda_1\rangle))} \\
= \quad & \frac{p(|i_0,\lambda_0\rangle)}{p(|i_f,\lambda_f\rangle)} \frac{p(|i_0,\lambda_1\rangle \to |i_1,\lambda_1\rangle)}{p(|i_1,\lambda_1\rangle \to |i_0,\lambda_1\rangle)} ... \frac{p(|i_{f-1},\lambda_f\rangle \to |i_f,\lambda_f\rangle)}{p(|i_f,\lambda_f\rangle \to |i_{f-1},\lambda_f\rangle)} \\
= \quad & \frac{p(|i_0,\lambda_0\rangle)}{p(|i_f,\lambda_f\rangle)} \exp -\beta \left(E(|i_1,\lambda_1\rangle - E|i_0,\lambda_1\rangle\right) ... \exp -\beta \left(E(|i_f,\lambda_f\rangle - E|i_{f-1},\lambda_f\rangle\right) \\
= \quad & \frac{\exp -\beta E(|i_0,\lambda_0\rangle)/Z_0}{\exp -\beta E(|i_f,\lambda_f\rangle)/Z_f} \exp -\beta \left(E(|i_1,\lambda_1\rangle - E|i_0,\lambda_1\rangle\right) ... \exp -\beta \left(E(|i_f,\lambda_f\rangle - E|i_{f-1},\lambda_f\right. \\
= \quad & \frac{Z_f}{Z_0} \exp -\beta \left([E(|i_0,\lambda_0\rangle) - E|i_0,\lambda_1\rangle] + [E(|i_1,\lambda_1\rangle - E(|i_1,\lambda_2\rangle] ... [E(|i_f,\lambda_{f-1}\rangle) - E|i_f,\lambda_f\right. \\
= \quad & \frac{Z_f}{Z_0} \exp(\beta W)
\end{aligned}
$$

where W denotes the work of a given trajectory.

Crooks holds in this set-up (2/2)

- To recover standard formulation of Crooks with LHS: $p_{fwd}(W)/p_{rev}(-W)$.
- Let $\mathcal{W}$ be set of trajectories with work gain W.
- Note that the set of reverse trajectories with work -W are precisely the inverse trajectories of those in $\mathcal{W}$:

$$
\begin{aligned}
\frac{p(W)}{p(-W)} &= \frac{\sum_{traj \in \mathcal{W}} p(traj)}{\sum_{traj \in \mathcal{W}} p(traj-inv)} \\
&= \frac{\sum_{traj \in \mathcal{W}} p(traj)}{\sum_{traj \in \mathcal{W}} p(traj) \frac{Z_0}{Z_f} \exp(-\beta W)} \\
&= \frac{Z_f}{Z_0} \exp(\beta W).
\end{aligned}
$$

Quantum extension of model: superposition problem

Say system in $|i_n, \lambda_n\rangle$. Suddenly $\lambda_n \to \lambda_{n+1}$. State is now $|i_n, \lambda_n\rangle$ which is not necessarily an eigenstate for λ_{n+1}, it is in general a *superposition* of the new energy eigenstates.

Simple example: $|i_0, \lambda_0\rangle = |0\rangle$.
$H(\lambda_0) = 0|0\rangle\langle 0| + 0|1\rangle\langle 1| \to H(\lambda_1) = |+\rangle\langle +| - |-\rangle\langle -|$

Suddenly $\lambda_0 \to \lambda_1$: the state $|0\rangle$ has no time to evolve. We no longer have an energy eigenstate, but a superposition of energy eigenstates

$$|0\rangle = |+\rangle + |-\rangle = |i_0, \lambda_1\rangle + |i_1, \lambda_1\rangle$$

It is ambiguous how to define the work of trajectories involving such states, but there is a trick...

(notation: $|+\rangle = |0\rangle + |1\rangle$, $|-\rangle = |0\rangle - |1\rangle$)

Quantum extension of model, Quan-Dong style

Quan-Dong trick to deal with superpositions: measure energy at end of each Hamiltonian change. This gives a new type of hop with transition probabilities:

$$|\langle i'_n, \lambda_{n+1}|U|i_n, \lambda_n\rangle|^2.$$

The reverse process is carefully defined so that the probability of the reverse jump in the reverse process equals that of the forwards jump above.

Thus $\frac{p(traj)}{p(traj-inv)}$ is unaffected by this modification.

The work cost can be different as $\delta W := E(|i'_n, \lambda_{n+1}\rangle - E(|i_n, \lambda_n\rangle)$. Same argument carries through though, so that

$$\frac{p(traj)}{p(traj - inv)} = \exp(-\beta((F_f - F_0) - W)),$$

and the last summation over trajectories also follows through, so Crooks holds also after this quantum modification.

Single-shot statistical mechanics also uses this setting

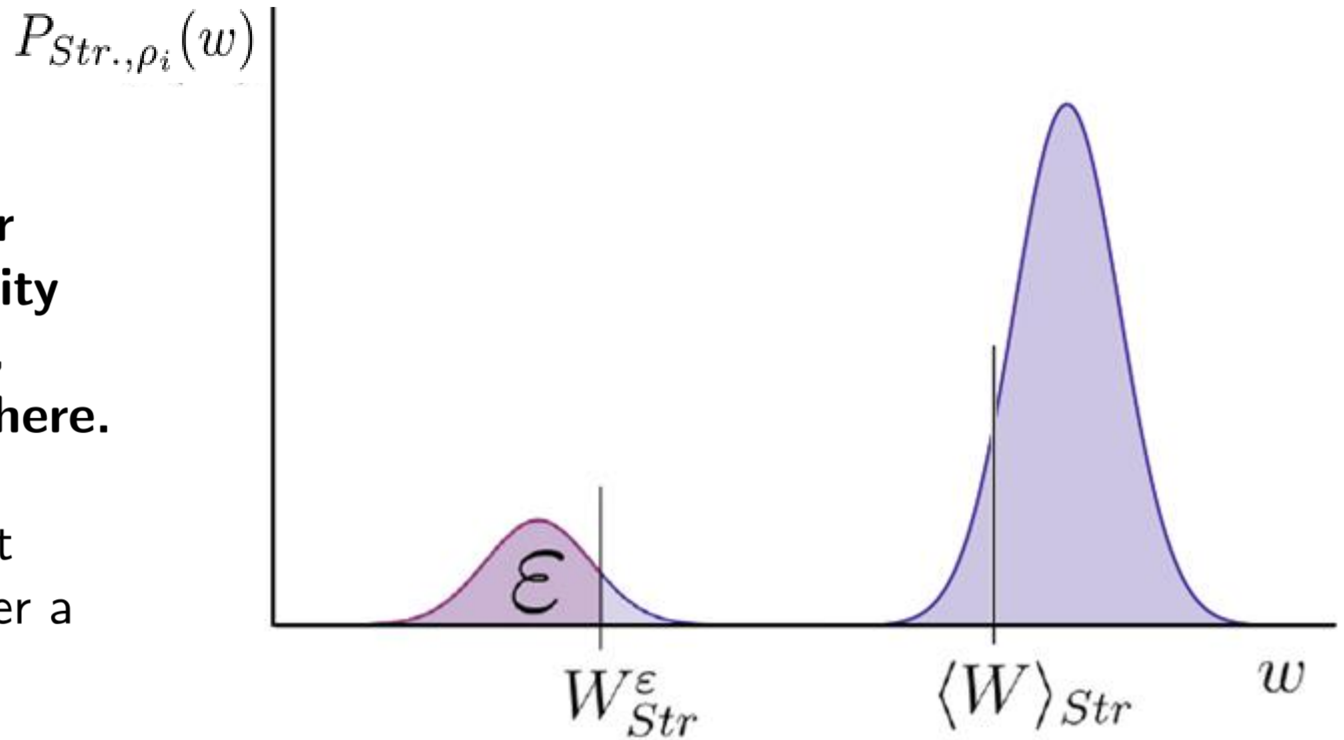
· The setting just described is also that used* for single-shot paper by Egloff et. al. *(setting similar to [Aberg], [delRio et. al], all leaning on [Alicki and Horodecki$^{\otimes 3}$]).*

· In that language there is an agent engaged in a work-extraction game.

· The agent chooses the Hamiltonian path, and which interactions with the heat-bath to include when. This choice is called the agent's strategy Str.

· **The strategy Str and initial state ρ_i together induce a work probability distribution $P_{Str.,\rho_i}(w)$, precisely as described here.**

*) detailed balance is not always required but rather a weaker condition.



Single-shot statistical mechanics in this setting

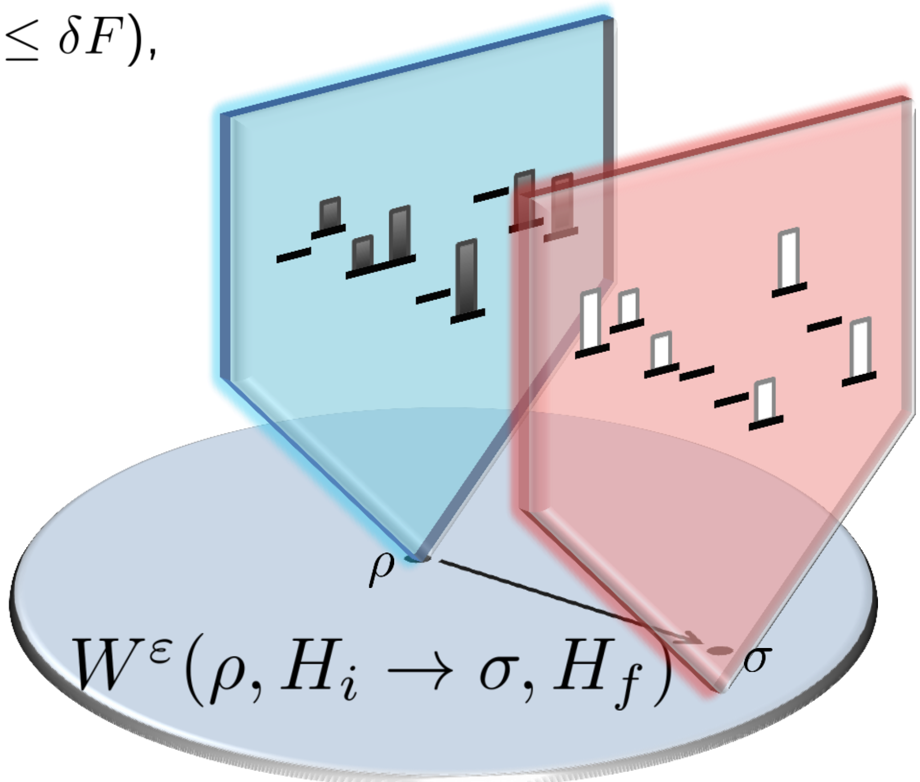
Recall that I flashed several equations earlier for W^ε, the optimal guaranteed work:

$$W^\varepsilon(\rho, H_i \to \sigma, H_f) = \max_{\{Str.\}} W^\varepsilon_{Str.}(\rho, H_i \to \sigma, H_f), e.g.$$

$$W^\varepsilon(\rho, H_i = 0 \to \gamma_T, H_f = 0) = (n - H^\varepsilon_{\max}(\rho))\, kT \ln 2$$

(This is single-shot analogue of $\langle W_{out} \rangle \leq \delta F$),

Those expressions apply here too, in this setting where Crooks also holds.



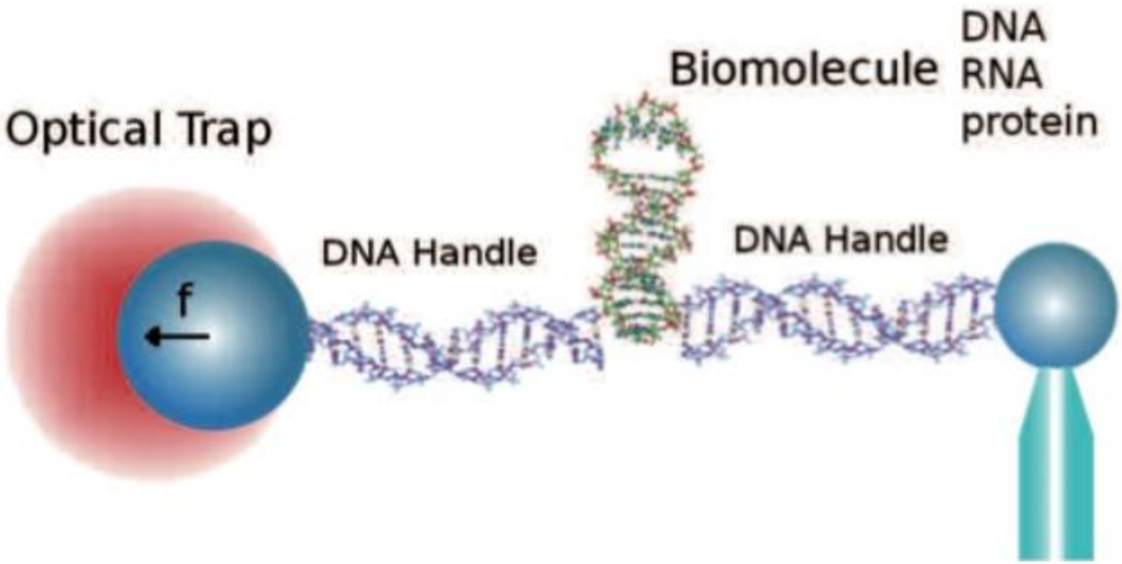
Early advantages of the links between the two approaches

Source: http://arxiv.org/abs/1409.3878
Unification of fluctuation theorems and one-shot statistical mechanics
Yunger-Halpern, Garner, Dahlsten, Vedral

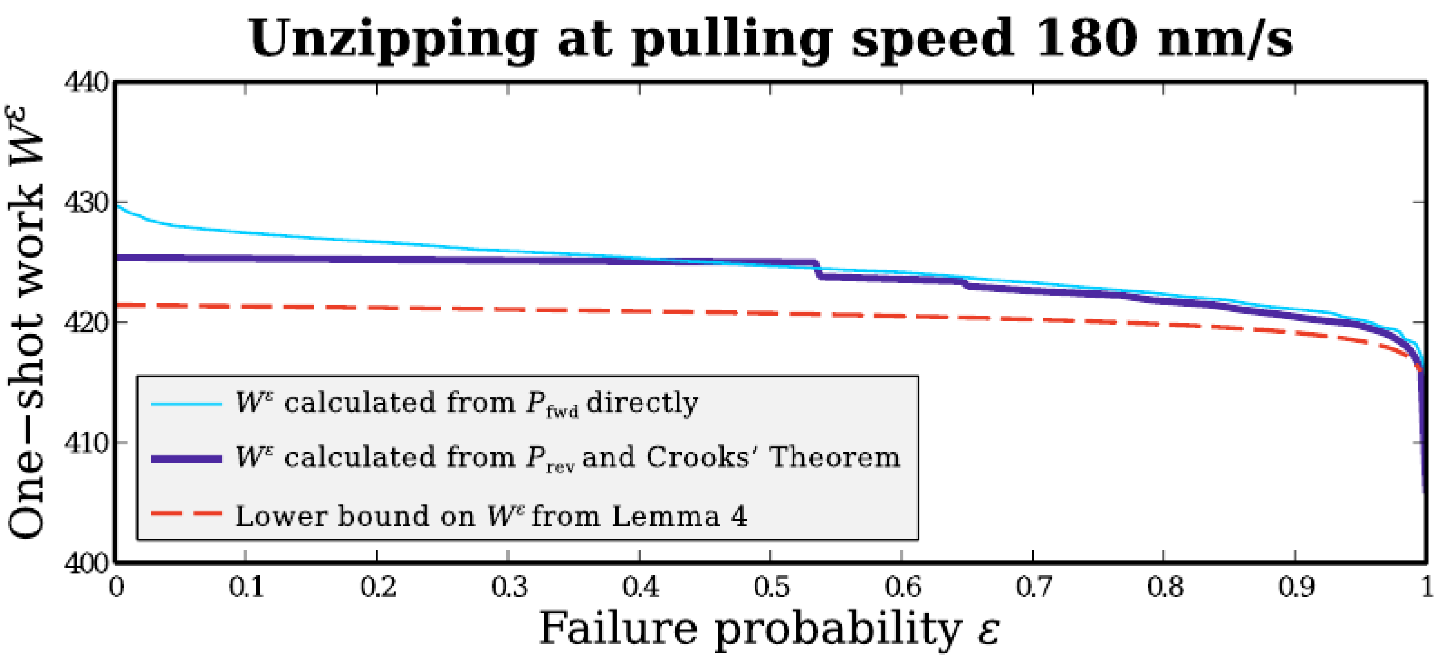
— Dangers lurked even in the bathroom.

Nicole Y-H Caltech Blog:
http://quantumfrontiers.com/2014/09/16/the-experimentalist-next-door/

1. Connecting single-shot stuff to experiment



Image& experimental data: Alemani&Ritort



2.Single-shot information theory combined with Crooks

Known: Crooks implies

$$\frac{1}{\beta}D(P_{fwd}(W)||P_{rev}(-W)) = \langle W\rangle_{fwd} - \Delta F \equiv \langle W_{diss}\rangle$$

where D is the standard relative entropy. In single-shot information theory this tends to get replaced with D_0 and D_∞, two different relative Renyi entropies (In this notation $D = D_1$). We shall here need:

$$D_\infty(P(x)||Q(x)) \equiv \log\{\min \lambda \in \mathbb{R} : P(x) \leq \lambda Q(x) \ \forall x\}.$$

New: Crooks also gives

$$\frac{1}{\beta}D_\infty(P_{fwd}(W)||||P_{rev}(-W)) = W_{\max} - \Delta F,$$

where at worst a finite-time forwards trial costs work $W_{\max}$. Moreover

$$\frac{1}{\beta}D_\infty(P_{\text{rev}}(-W)||P_{\text{fwd}}(W)) = \Delta F - W_{\min}.$$

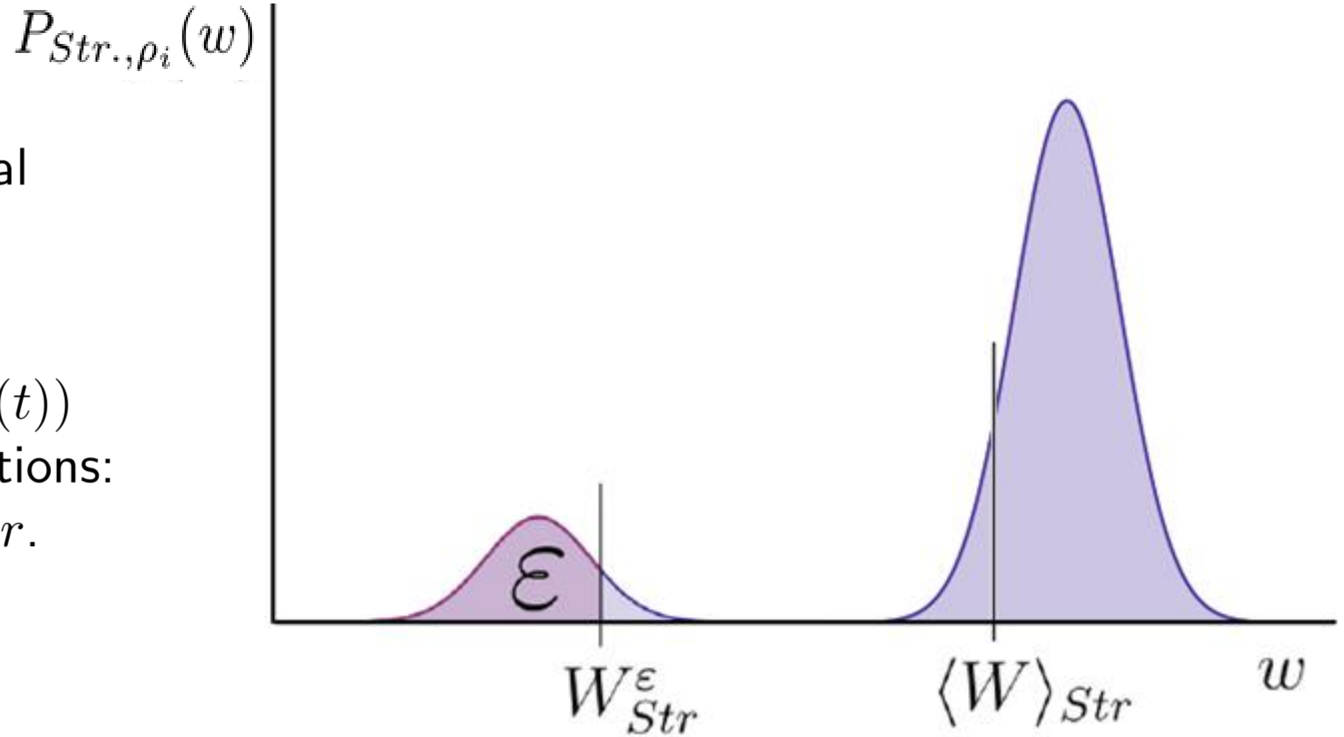
where at worst, a finite-time reverse trial outputs work $W_{\min}$

Summary

- Single-shot statistical mechanics and Crooks fluctuation theorem deal with same scenarios.

- The agent chooses the Hamiltonian path $H(\lambda(t))$ and thermalising operations: the agent's strategy Str.

- Str and initial state ρ_i together induce a work probability distribution $P_{Str,\rho_i}(w)$.



- Single-shot focus: $W^{\varepsilon}(\rho, H_i \to \sigma, H_f) = \max_{\{Str.\}} W^{\varepsilon}_{Str.}(\rho, H_i \to \sigma, H_f)$

- As both approaches deal with same scenario one can inter-mingle theoretical and experimental results, connecting single-shot stuff with experiments, and fluctuation stuff with single-shot information theory.

Outlook

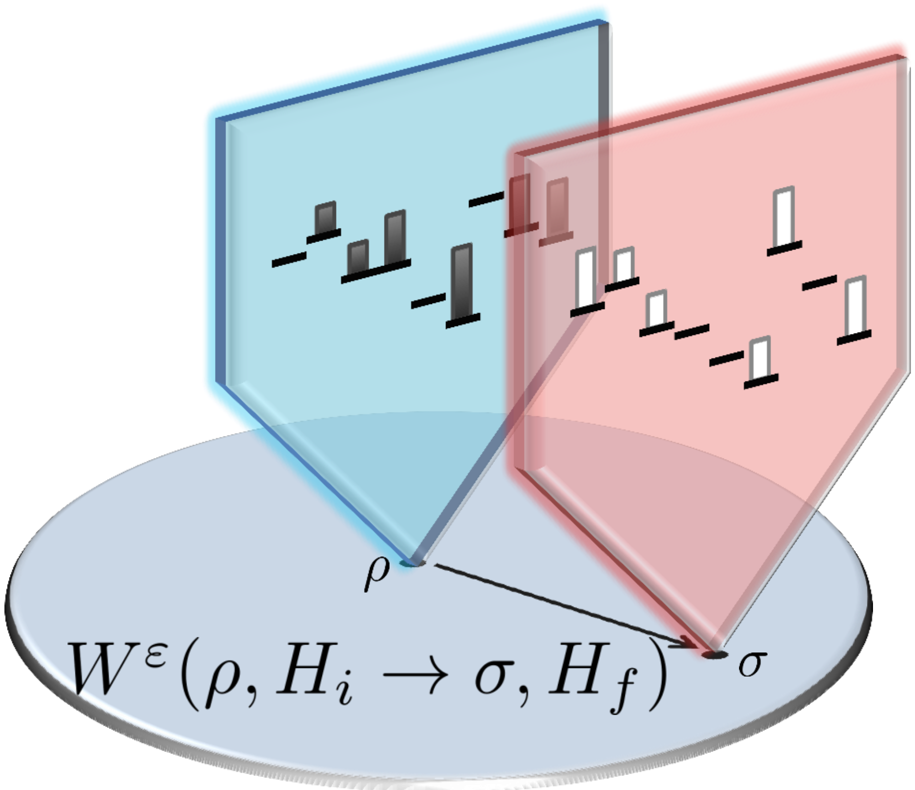
- Go beyond Crooks to entropy production relations in making the connection.
- Go beyond $H(\lambda)$-setting in making the connection, prefer battery included explicitly.
- Can these theoretical ideas contribute to new technology, perhaps involving optomechanical systems like those here in Grenoble?
- Energy and work as concepts sit oddly in quantum information theory, is there a 'better' way of defining them.
- Single-shot statistical mechanics more generally ends up replacing entropy differences with majorisation relations, I would say

> In 20 years it will be standard to be doing thermodynamics using majorisation rather than entropy

How does this tie into fluctuation relations?

Thank you,
Especially Alexia and Maxime
for extremely interesting workshop

UNIVERSITY OF OXFORD



0.1 Conditional entropy, relative entropy

The expressions for the *conditional* single-shot entropies are considerably more intimidating and arbitrary-looking at first sight. We now jump straight to the quantum case. A helpful way to see where the conditional entropies come from is to follow Datta [**?**] and define them via the relative Renyi entropy $S_\alpha(\rho||\sigma)$. It has often been argued in the context of the Shannon/von Neumann entropy that relative entropy,

$$S_1(\rho||\sigma) := Tr(\rho \log \rho) - Tr(\rho \log \sigma)$$

is a 'parent-quantity', in that $S_1(\rho) = -S_1(\rho||\mathbb{1})$ and conditional entropy (defined for the von Neumann entropy $S = S_1$ via $S_1(A|B) := S_1(AB) - S_1(B)$) can be written as

$$S_1(A|B) = -S_1(\rho_{AB}||\mathbb{1} \otimes \rho_B).$$

Datta notes that the relative Renyi entropies are parent-quantities in the same way. The definition of $D_0^\varepsilon(.||.)$ (called $D_{\min}$ in [**?**]) is as follows:

$$D_0(\rho||\sigma) := -\log Tr(\Pi_\rho \sigma),$$

where Π_ρ is the projector onto the support of ρ. The smooth version is defined as

$$D_0^\varepsilon(\rho||\sigma) := \sup_{\bar{\rho} \in B^\varepsilon(\rho)} D_0(\bar{\rho}||\sigma),$$

where $B^\varepsilon(\rho)$ is the set of states within ε trace distance of ρ.

If we now demand, in analogy with the case of the von Neumann entropies, that

$$S_0(A|B) = -S_0(\rho_{AB}||\mathbb{1} \otimes \rho_B)$$

we recover one definition of the conditional max entropy:

$$S_0(A|B) = -\log Tr((\Pi_{AB})(\mathbb{1} \otimes \rho_B)).$$

We showed that the optimal guaranteed work is determined by an expression involving a measure of how much more mixed one state ρ is than another, σ. We call this the relative mixedness and write it as $M(\rho||\sigma)$.

Definition 1. *Consider two probability distributions $\lambda(x)$ and $\mu(x)$ defined over $x \in \mathbb{R}^{(\geq 0)}$. Let $\lambda(x)\!\downarrow$ and $\mu(x)\!\downarrow$ denote these distributions after a (measure-preserving) rearrangement so that they are in descending order. Let the cumulative distribution function associated with a function γ be denoted as $\mathcal{F}_\gamma(x) := \int_0^x dx' \gamma(x')$. Then the relative mixedness of $\lambda(x)$ and $\mu(x)$ is defined as*

$$\mathsf{M}(\lambda||\mu) := \max m \, s.t. \, \mathcal{F}_{\lambda\downarrow}\left(\frac{x}{m}\right) \geq \mathcal{F}_{\mu\downarrow}(x) \;\; \forall x,$$

where $m \in \mathbb{R}$. In words: the relative mixedness of λ and μ is the maximal amount by which one can stretch $\lambda\!\downarrow$ under the condition that its integral upper bounds the integral of $\mu\!\downarrow$ at all points.

By the definition of majorisation, if and only if $M \geq 1$ does (the spectrum of) ρ majorise σ, $\rho \succ \sigma$. The actual number M can thus be viewed as putting a number to how much ρ majorises σ.