Thermodynamic limits of information processing (in living systems)

Susanne Still
University of Hawaiʻi at Mānoa

www2.hawaii.edu/~sstill

Motivation (1)

- Information processing, learning and adaptation in living systems:

Are there simple physical principles?
What are the physical limits of information processing?

- Biological systems are adapted to stochastic environments.
- Efficiency arguments might lead to the discovery of principles.
- Candidates: **- Efficient use of energy**
  - Efficient modeling/information processing

(Nature computes! Living systems learn about their environment).

Proc Natl Acad Sci U S A. 2010 Jul 6;107(27):12329-34. Epub 2010 Jun 23.

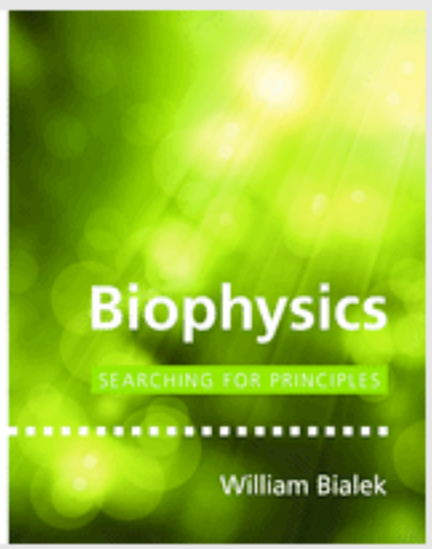
Metabolic cost as a unifying principle governing neuronal biophysics.

Hasenstaub A, Otte S, Callaway E, Sejnowski TJ.

Crick-Jacobs Center for Theoretical and Computational Biology, Salk Institute for Biological Studies, La Jolla, CA 92037, USA. andrea@salk.edu

Abstract

The brain contains an astonishing diversity of neurons, each expressing only one set of ion channels out of the billions of potential channel combinations. Simple organizing principles are required for us to make sense of this abundance of possibilities and wealth of related data. We suggest that energy minimization subject to functional constraints may be one such unifying principle. We compared the energy needed to produce action potentials singly and in trains for a wide range of channel densities and kinetic parameters and examined which combinations of parameters maximized spiking function while minimizing energetic cost. We confirmed these results for sodium channels using a dynamic current clamp in neocortical fast spiking interneurons. We find further evidence supporting this hypothesis in a wide range of other neurons from several species and conclude that the ion channels in these neurons minimize energy expenditure in their normal range of spiking.



Biophysics: Searching for Principles
William Bialek

William Bialek, Winner of the 2013 Swartz Prize for Theoretical and Computational Neuroscience, Society for Neuroscience

Efficient information processing as a first principle?

valueless category. Thus, separating predictive information from the background of non–predictive clutter is a formidable, and biologically relevant, challenge. Importantly, this very general task seems to contain within it, as special cases, problems ranging from signal processing to learning, problems that we usually think of as belonging to different levels of biological organization with very different mechanisms. Perhaps this is, after all, a path to the sort of general principle we are seeking.

Motivation (2)

- **Nature processes and transforms information...**
  - Interest in natural systems that:
    - run at finite rates (not infinitely slowly)
    - operate (arbitrarily far) away from thermodynamic equilibrium
- **fundamentally limited by the laws of physics!**

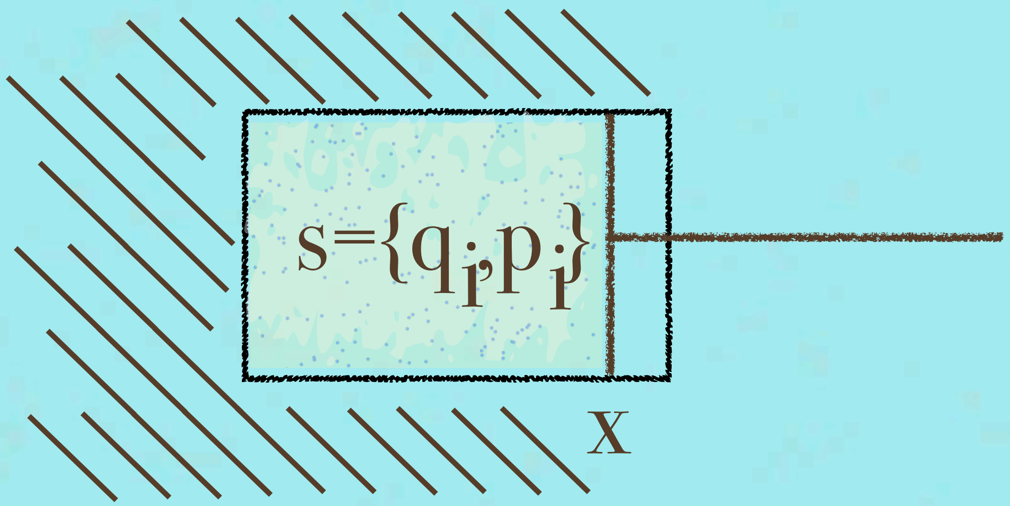
Information processing and dissipation of energy

Information erasure has to be paid for by heat production due to second law. (Szilard 1929, Landauer 1961)	Equilibrium thermodynamics Shannon information (Boltzmann, Shannon, Jaynes, ...)
Additional available free energy out of equilibrium (Shaw 1981, ...) => extract using measurement (Sagawa & Ueda 2010, ...)	Systems driven far from thermodynamic equilibrium (Schloegl, Schnakenberg, Prigogine, Fox, Jarzynski, Crooks, Searles, Evans, Esposito, Van den Broeck, Seifert, ...)
Instantaneous nonpredictive information is proportional to instantaneous dissipation due to change in environment. => Refinement of Landauer's argument	Driven systems, far from equilibrium, embedded in *stochastic* environments (Still, Sivak, Bell and Crooks PRL 2012)
This result holds also for quantum systems => new interpretation of quantum discord (Grimsmo PRA 2013)	Quantum Systems (Vedral, Renner, Wiesner, Grimsmo, ...)

Equilibrium Thermodynamics (quick review)

- The state (variable s) of a system in thermodynamic equilibrium, given an environmental parameter, x.
  Best described by the Boltzmann distribution:

$$p_{\text{eq}}(s|x) = e^{-\beta[E(s,x)-F(x)]} \qquad \beta = \frac{1}{k_B T}$$

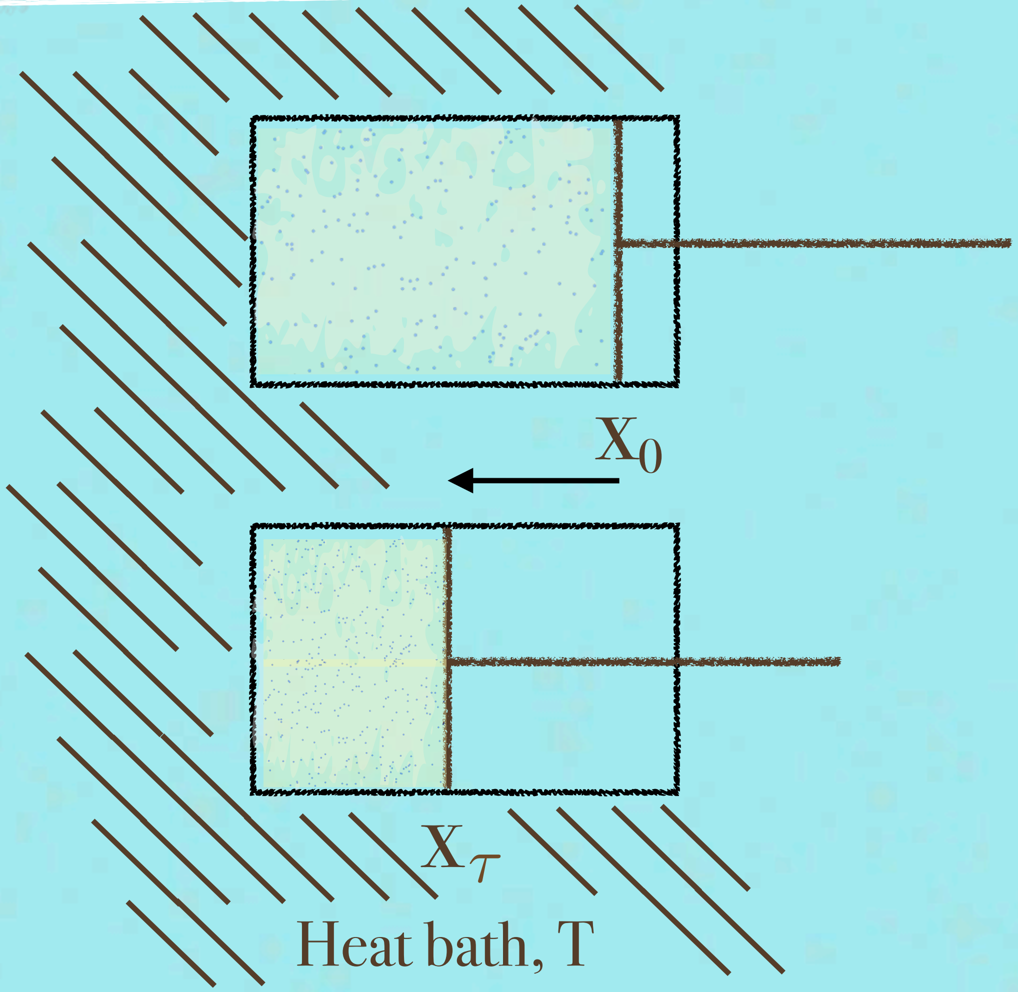


Example: Gas in box with piston. x = piston position, s = positions and momenta of molecules. Temperature T.

- This is the maximum entropy distribution, given the known quantities [here: the average energy $E = \langle E(s,x) \rangle$]. (Jaynes 1957)
- Entropy: $S = k_B H[p_{\text{eq}}] = -k_B \left\langle \ln[p_{\text{eq}}] \right\rangle_{p_{\text{eq}}}$
- (Equilibrium) Free Energy: $F = E - TS$

Work done in excess of free energy change

1. Start system in thermodynamic equilibrium:
   $p(s_0|x_0) = e^{-\beta[E(s_0,x_0)-F(x_0)]}$
2. Drive system, performing work, W.
   Contact to heat reservoir at temperature T.
   Heat flowing into gas = Q.
3. Let system relax back to equilibrium:
   $p(s_\tau|x_\tau) = e^{-\beta[E(s_\tau,x_\tau)-F(x_\tau)]}$



- Energy is conserved: $\Delta E = W + Q$
- Free energy change: $\Delta F = \Delta E - T\Delta S$
- Average work done in excess of free energy change is positive (second law): $\langle W_{\rm ex}\rangle = \langle W\rangle - \Delta F = -\langle Q\rangle + T\Delta S \geq 0$

$$\Delta E := E_\tau - E_0$$
$$\Delta F := F_\tau - F_0$$
$$\Delta S := S_\tau - S_0$$

Landauer's argument

Second law => $-\langle Q\rangle + T\Delta S \geq 0$

- Landauer (1961): What is the minimum amount of heat generated when one bit of information is reset, or "erased"? (Erasing information is equivalent to an entropy reduction.)
  (Shannon) Entropy (in bits): $H[p] = -\langle \log_2[p]\rangle_p$

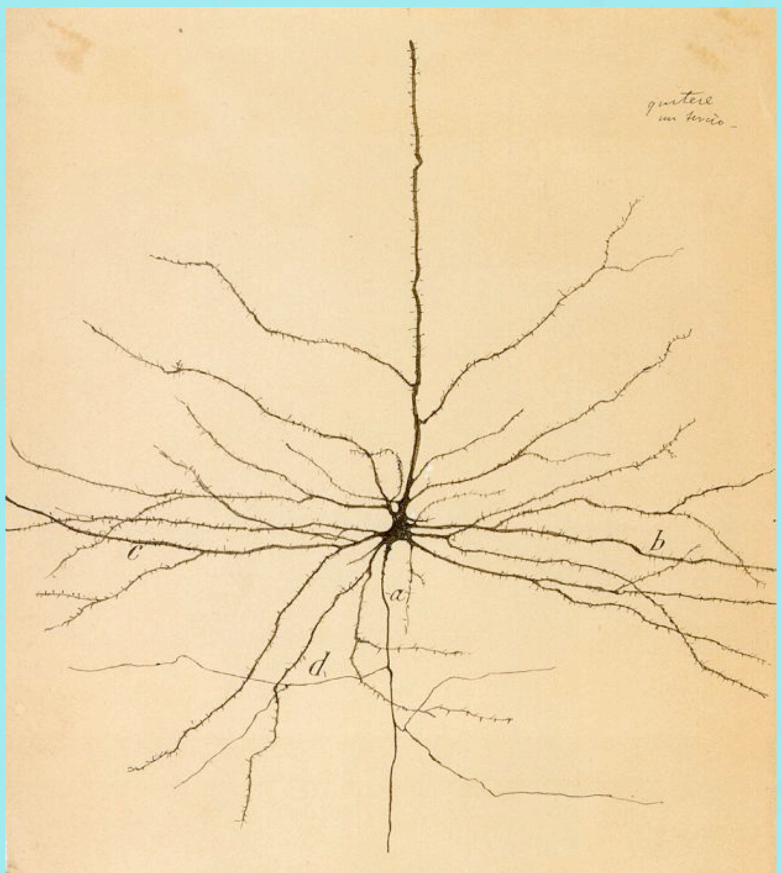
  Erasure $:= H_0 - H_\tau = -\Delta S/(k_B \ln(2))$

  => $-\langle Q\rangle \geq k_B T \ln(2)$

Heat produced when one bit of information is erased is at least kT ln(2)

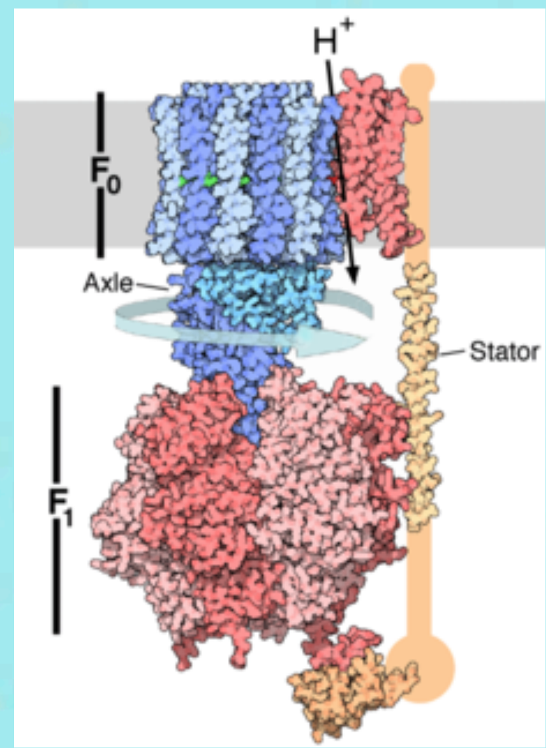
Living systems are not in equilibrium!

Neuron (nerve cell)



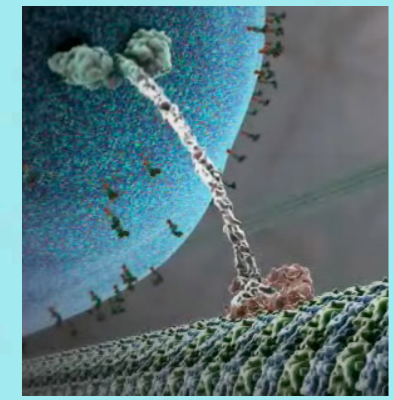
Rotary F_0-F_1 ATP synthase

(ATP - Adenosine triphosphate)

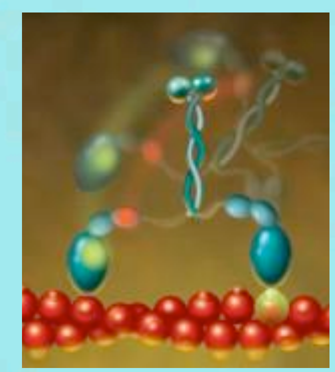


From: PDB, rcsb.org

kinesin



myosin



From: Inner Life of the Cell

Death = "the decay into thermodynamical equilibrium."
-Erwin Schrödinger, *What is Life?*

Far from thermodynamic equilibrium - what do we know?

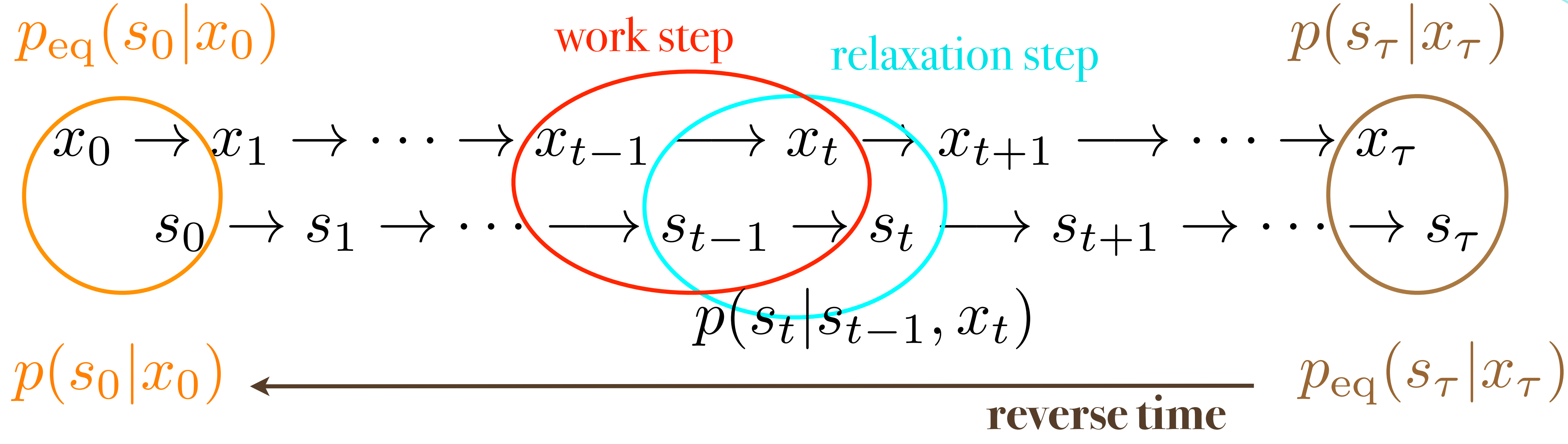
Jarzynski's work relation (PRL 1997):

$$\Delta F = -\frac{1}{\beta} \ln \left\langle e^{-\beta W} \right\rangle$$

- Jensen's inequality => $\langle W \rangle \geq \Delta F$
- Expand into sum of cumulants; Gaussian noise => fluctuation-dissipation relation: $\langle W \rangle - \Delta F = \dfrac{\sigma^2}{2k_B T}$

with: $\sigma^2 = \langle W^2 \rangle - \langle W \rangle^2$

Crooks' derivation of Jarzynski's relation:



1) define:

$$W = \sum_{t=0}^{\tau-1} \left(E(s_t, x_{t+1}) - E(s_t, x_t)\right)$$

$$Q = \sum_{t=1}^{\tau} \left(E(s_t, x_t) - E(s_{t-1}, x_t)\right)$$

$$W + Q = E(s_\tau, x_\tau) - E(s_0, x_0) = \Delta E$$

2) consider forward and reverse time protocol

=> Crooks equation:

$$\frac{P_F(\vec{s}|\vec{x})}{P_R(\vec{s}|\vec{x})} = \frac{p_{eq}(s_0|x_0)}{p_{eq}(s_\tau|x_\tau)} \prod_t \frac{p(s_t|s_{t-1}, x_t)}{p(s_{t-1}|s_t, x_t)}$$

assume detailed balance

$$= e^{\beta(\Delta E - \Delta F)} \prod_t \frac{p_{eq}(s_t|x_t)}{p_{eq}(s_{t-1}|x_t)}$$

$$= e^{\beta(Q+W-\Delta F)} e^{-\beta \sum_t (E(s_t,x_t) - E(s_{t-1},x_t))}$$

$$= e^{\beta(Q+W-\Delta F)} e^{-\beta Q}$$

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

3) easy proof:

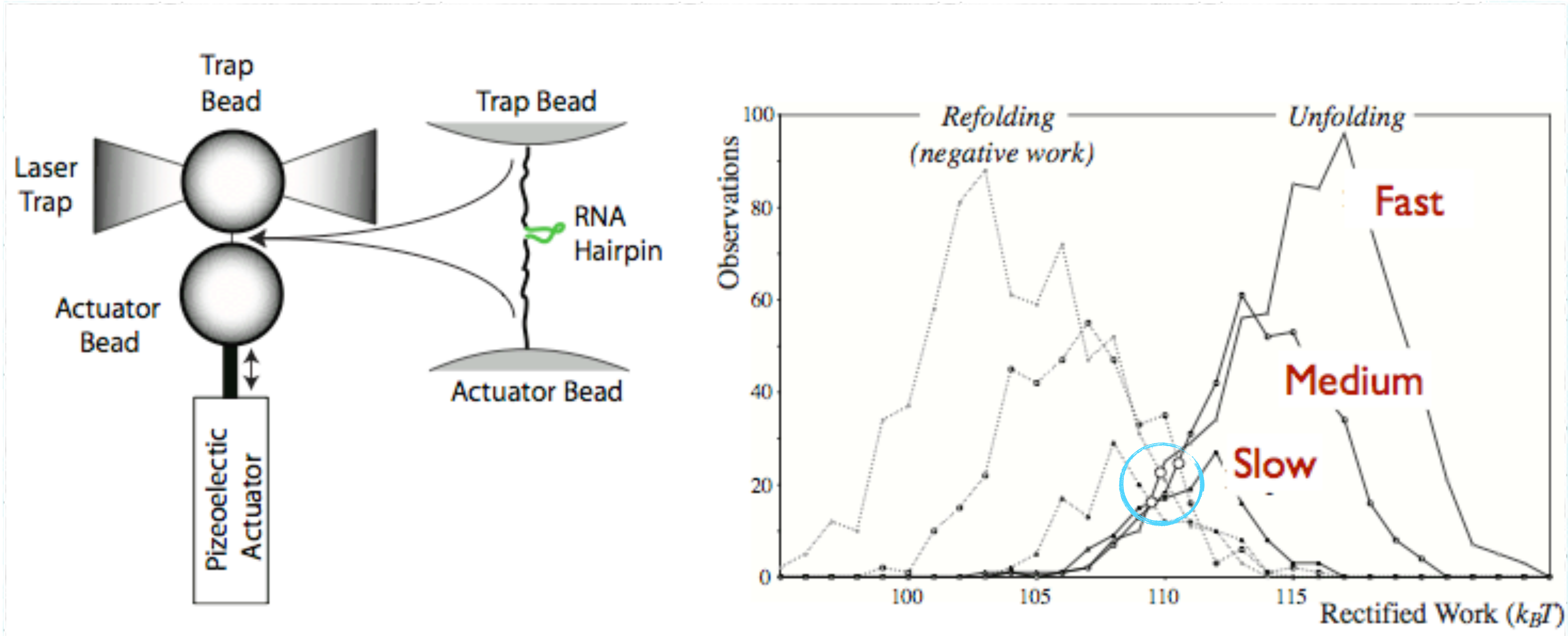
$$\langle e^{-\beta W} \rangle_{P_F} = \langle e^{-\beta W} \frac{P_F}{P_R} \rangle_{P_R}$$

$$= e^{-\beta \Delta F}$$

G. E. Crooks, J.Stat.Phys (1998)

=> Crooks' fluctuation theorem:

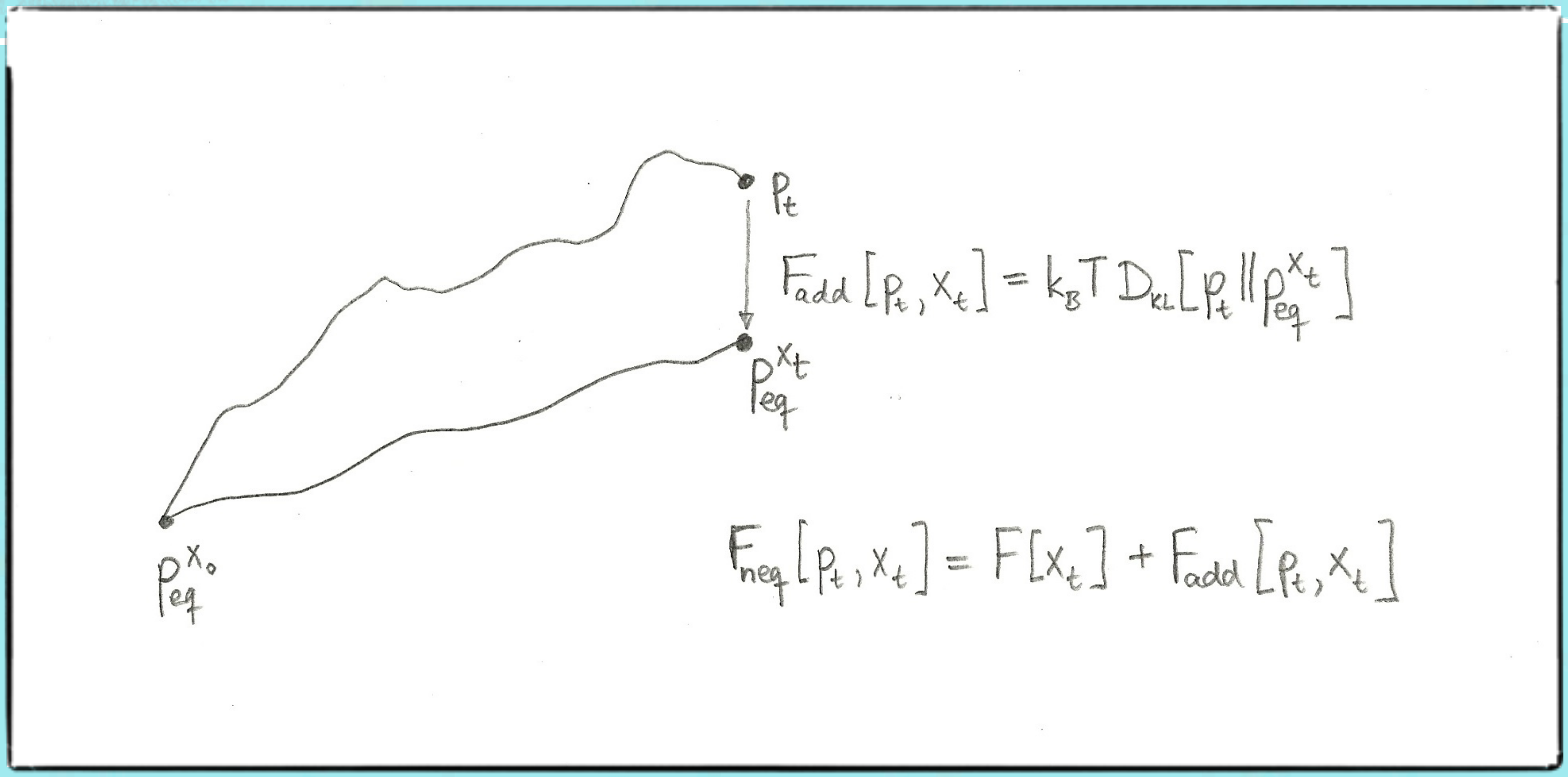
$$\frac{P_F(+W)}{P_R(-W)} = e^{\beta(W-\Delta F)}$$



D. Collin, F. Ritort, C. Jarzynski, S.B. Smith, I.Tinoco, C. Bustamante (2005)

- allows measurement of equilibrium free energy change in system that is far from equilibrium during the experiment, e.g. optical trap RNA hairpin unfolding (& refolding).

What if there is no relaxation back to equilibrium?



- Additional free energy compared to equilibrium $\propto D_{\rm kl}[p_t || p_{\rm eq}^{x_t}]$

Shaw (1981)

- Dissipation--work *irretrievably* lost:

$$W_{\rm diss} = W - \Delta F_{\rm neq}$$

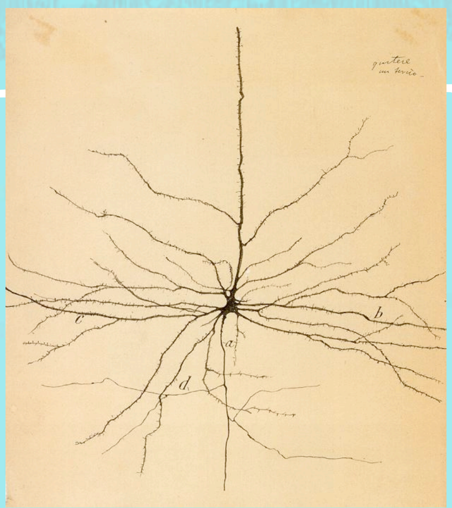
Esposito & Van den Broeck, EPL (2011); Still et al., PRL (2012)

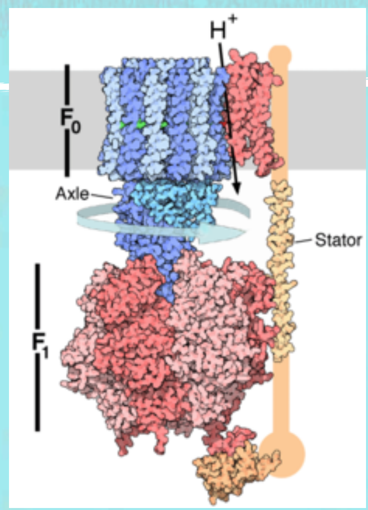
Relation to work lost in excess of equilibrium free energy change

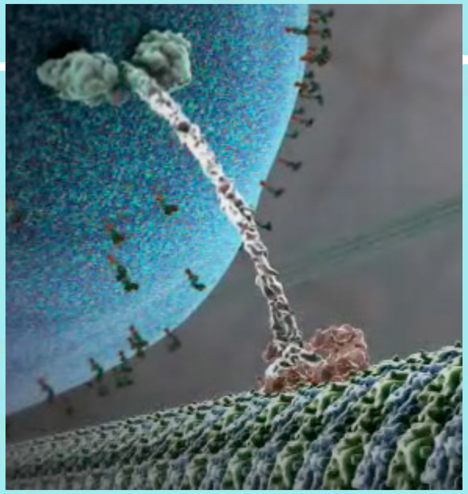
- Dissipation: $$W_{\text{diss}} = W - \Delta F_{\text{neq}} = W_{\text{ex}} - \Delta F_{\text{add}}$$
  - is non-negative
- Therefore: $W_{\text{ex}} \geq \Delta F_{\text{add}}$
  - — this motivates control using measurement of the system; information could be used to lower dissipation

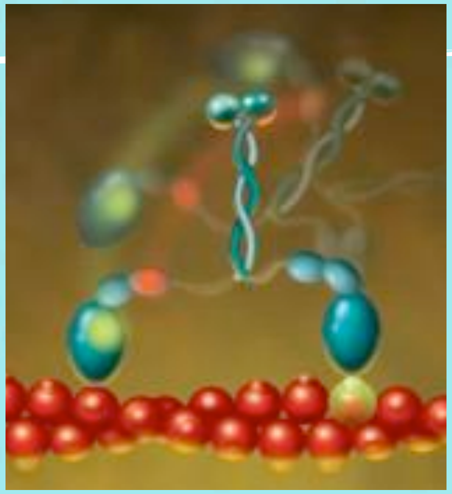
Esposito&VandenBroeck, EPL (2011); Sagawa&Ueda,PRL (2012); Still et al., PRL (2012)

Living systems are embedded in *stochastic* environments









- Dynamics of a stochastic environment governed by $p(x_0, \ldots, x_\tau)$
  - p is not necessarily "known"/given
  - system could be adapted to a certain type of environment

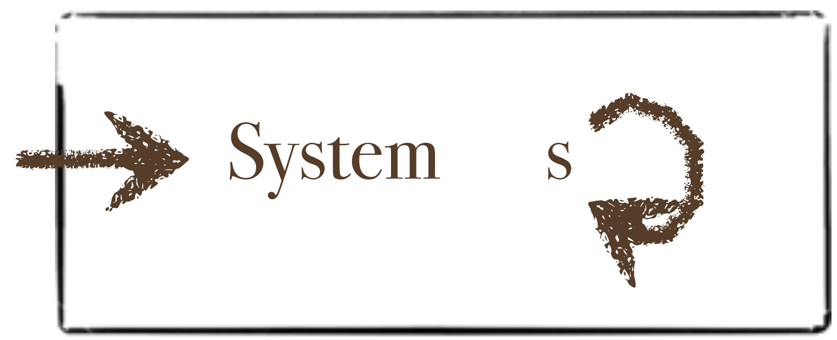
Setup:

Environmental variables x(t) —drive→ System [state s(t)]

Heat bath, T

Physical systems process information!

Environment -> external drive/data x



Heat bath, T

Environment: $x_0 \rightarrow x_1 \rightarrow \cdots \rightarrow x_{t-1} \longrightarrow x_t \rightarrow x_{t+1} \longrightarrow \cdots \rightarrow x_\tau$

(Learning) System: $s_0 \rightarrow s_1 \rightarrow \cdots \longrightarrow s_{t-1} \rightarrow s_t \longrightarrow s_{t+1} \rightarrow \cdots \rightarrow s_\tau$

$p(s_t|s_{t-1}, x_t)$

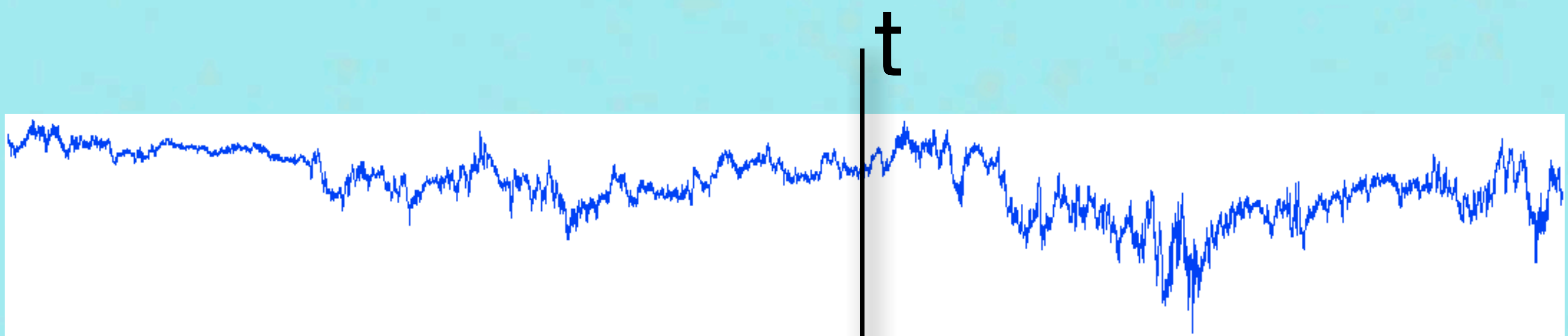
System computes by changing its state in response to changes in the environment.

Dynamics of the system encode **model of environment**.

Still EPL (2009); Still et al. PRL (2012); Still Entropy (in press)

Simple framework

- Signal:



past $\overleftarrow{x}_t$ future $\overrightarrow{x}_t$

- Predictive information (Shannon mutual information btw. past and future):

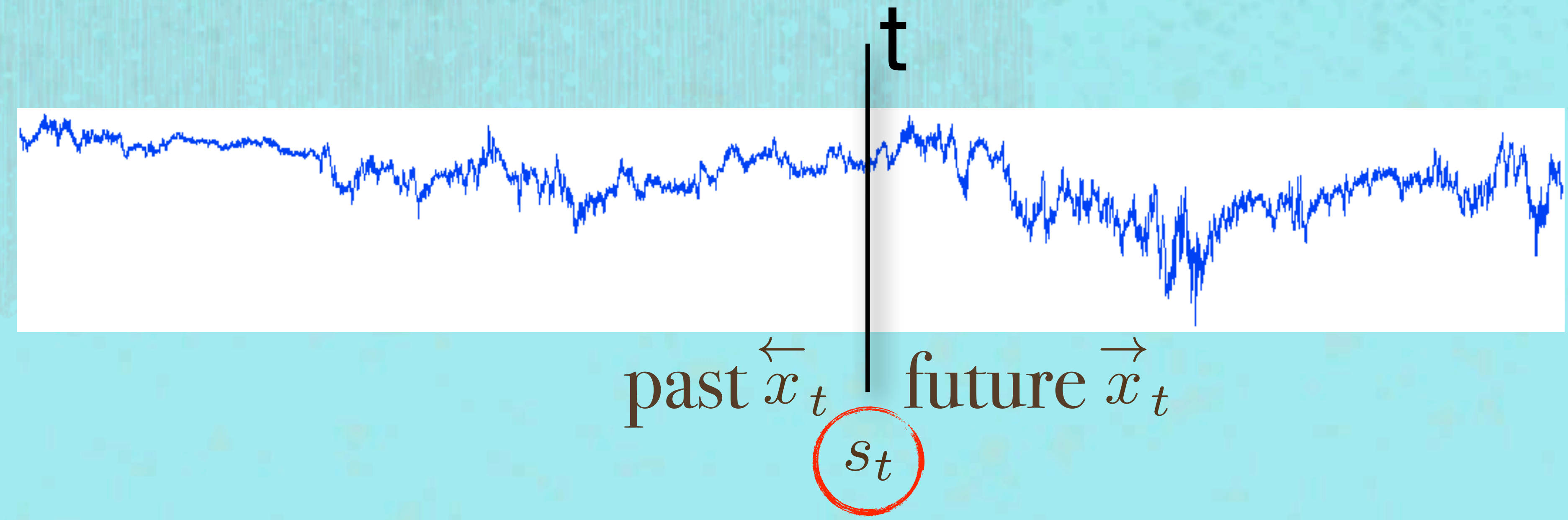
e.g. Shaw (1981)

$$I[\overleftarrow{x}_t; \overrightarrow{x}_t] := \left\langle \log \left[\frac{p(\overrightarrow{x}_t \mid \overleftarrow{x}_t)}{p(\overrightarrow{x}_t)} \right] \right\rangle_{p(\overleftarrow{x}_t; \overrightarrow{x}_t)} = H[\overrightarrow{x}_t] - H[\overrightarrow{x}_t \mid \overleftarrow{x}_t]$$

Predictive information in a nonequilibrium critical model

Martin Tchernookov, Ilya Nemenman (2013) J. Stat. Phys. 153:442–459

Abstract We propose predictive information, that is, information between a long past of duration T and the entire infinitely long future of a time series, as a general order parameter to study phase transitions in physical systems independently of the underlying dynamics. It can be used, in particular, to study nonequilibrium transitions and other exotic transitions, where a simpler order parameter cannot be identified using traditional symmetry arguments. As an example, we calculate predictive information for a stochastic nonequilibrium dynamics problem that forms an absorbing state under a continuous change of a parameter. The information at the transition point diverges as $\propto \log T$, and we calculate the expression for a smooth crossover to $\propto T^0$ away from the transition.



- **Representation** s_t of the signal (implemented by a **system**); induces a model that:
  - shares information with the past (memory): $I[s_t;\overleftarrow{x}_t] = H[\overleftarrow{x}_t] - H[\overleftarrow{x}_t|s_t]$
  - might have some predictive power: $I[s_t;\overrightarrow{x}_t] = H[\overrightarrow{x}_t] - H[\overrightarrow{x}_t|s_t]$
- **What constitutes a good model?**
  Large predictive power at lowest possible complexity!
  - Information theoretic approach: want large predictive power for given memory.
  - **Then nonpredictive information** $I[s_t;\overleftarrow{x}_t] - I[s_t;\overrightarrow{x}_t]$ **measures information processing inefficiency.**

Still EPL (2009); Still et al. PRL (2012); Still Entropy (in press)

Explore relationship between

information processing inefficiency and
thermodynamic inefficiency

What are the thermodynamic limits of information processing?

Have specified:

- measure information processing inefficiency by *nonpredictive information*
- measure thermodynamic inefficiency in systems driven *arbitrarily far* from thermodynamic equilibrium by dissipation based on *non-equilibrium free energy*

Stochastic thermodynamics of learning machines

$$x_0 \to x_1 \to \cdots \to x_{t-1} \longrightarrow x_t \to x_{t+1} \longrightarrow \cdots \to x_\tau$$

$$s_0 \to s_1 \to \cdots \longrightarrow s_{t-1} \to s_t \longrightarrow s_{t+1} \to \cdots \to s_\tau$$

Dissipation due to change in environment is proportional to instantaneous nonpredictive information:

$$I[S_t; X_t] - I[S_t; X_{t+1}] = H[S_t|X_{t+1}] - H[S_t|X_t]$$

$$= \beta\Big[\underbrace{\langle E(s_t, x_{t+1})\rangle - \langle E(s_t, x_t)\rangle}_{\text{av. work}} - \underbrace{\big(\langle F_{\text{neq}}[p_t(s|x_{t+1})]\rangle - \langle F_{\text{neq}}[p_t(s|x_t)]\rangle\big)}_{\text{av. non-eq. free energy change}}\Big]$$

$$I[S_t; X_t] - I[S_t; X_{t+1}] = \beta \langle W_{\text{diss}}(x_t \to x_{t+1})\rangle$$

Still, Sivak, Bell & Crooks PRL (2012)

Quantum systems

- **Result carries over!** (Grimsmo, PRA 2013)
- Provides a new interpretation of quantum discord as: "the thermodynamic inefficiency of the most energetically efficient classical approximation of a quantum memory".
- Discord "quantifies the contribution to the lost work coming from quantum correlations".
- Possible quantum advantage in terms of lost work (when environmental signal is non-classical).

Landauer refined (classical)

- **New bound on excess work:** $$\beta \langle W_{\rm ex} \rangle \geq I_{\rm nonpred}$$

  ($I_{\rm nonpred}$ is the *total* instataneous nonpredictive information, summed over time.)

- Therefore: $$-\beta Q \geq +\mathcal{I}_{\rm erasure} + I_{\rm nonpred}$$

Landauer's bound augmented by nonpredictive information (which is a signature of the *dynamics* of the driven system)

Still, Sivak, Bell & Crooks PRL (2012)

Punchline

Dissipation of energy is fundamentally related to information processing inefficiency via nonpredictive information

Provocative statements:

- Living systems might predict their environment in order to operate at maximal thermodynamic efficiency.
- Predictive inference--not only a higher cognitive function? Perhaps implemented on small scales, such as bio-molecular machines?
- Is prediction/predictive inference a signature of life?

Looking back at predictive inference

- **What constitutes a good model?**
- **Large predictive power at lowest possible complexity!**
- Information theoretic approach:
  - Measure model complexity by memory (coding rate): information retained about the past $I[s_t; \overleftarrow{x}_t] = H[\overleftarrow{x}_t] - H[\overleftarrow{x}_t | s_t]$
  - Measure predictive power by information about the future

$$I[s_t; \overrightarrow{x}_t] = H[\overrightarrow{x}_t] - H[\overrightarrow{x}_t | s_t]$$

  - Want **large predictive power at fixed memory! => optimization problem**

maximize $\left(I[s_t; \overrightarrow{x}_t] - \lambda I[s_t; \overleftarrow{x}_t] \right)$

e.g. Still EPL (2009); Still Entropy (in press)

=> Optimal prediction machines! - Theorems and Algorithms -

- **Dynamical learning system** finds asymptotically (Still, in press) the ϵ-**machine** (Crutchfield and Young,1989), a deterministic hidden Markov model that is maximally predictive. All characteristics of the underlying process can be computed from the ϵ-machine, in many cases analytically: entropy rate, predictive/stored information,... (Crutchfield and colleagues,1989 onward).
- Used in **batch mode** (Still and Crutchfield , 2007) it is an instantiation of **Information Bottleneck Method** (Tishby, Pereira and Bialek,1999) which, in turn, is a special case of **rate-distortion theory** (Shannon, 1948). Family of optimal solutions: maximally predictive models at fixed memory. More efficient models are infeasible! Iterative algorithm (compare to Blahut/Arimoto ,1972).
- As the constraint on model complexity is relaxed, batch method finds (Still, Crutchfield and Ellison, 2010) the "**causal state partition**" (Crutchfield and Young,1989), that is **minimal and unique sufficient statistics** (Shalizi and Crutchfield, 2001).
- **Gaussian assumptions** -> method is related (Creutzig and Sprekeler, 2008) to **slow feature analysis** (Wiskott and Sejnowski, 2002) in batch, and (Creutzig and Globerson and Tishby, 2009) to **canonical correlation analysis** in dynamic learning mode.

What about feedback (interaction)?

- A system's actions influence the environment and thus the system's own future input.
- Optimal prediction needs to include this effect!
- Same framework extended to feedback situation =>

1. Extension of causal state partition and ϵ -machine
2. **Optimal action policies have to balance exploration and control!**

Still EPL (2009)

- To do: thermodynamics for coupled systems with interaction (work in progress with Gavin Crooks and David Sivak)

What about quantum predictive systems?

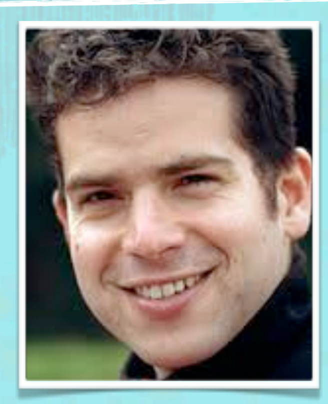
* Optimal quantum predictive systems in the sense of predictive inference outlined here
* Are there intrinsic quantum advantages?

(work in progress with Arne Grimsmo)

Collaborators

(noneq. thermodynamics)

Gavin E. Crooks
(LBL)



David A. Sivak
(LBL, now UCSF)

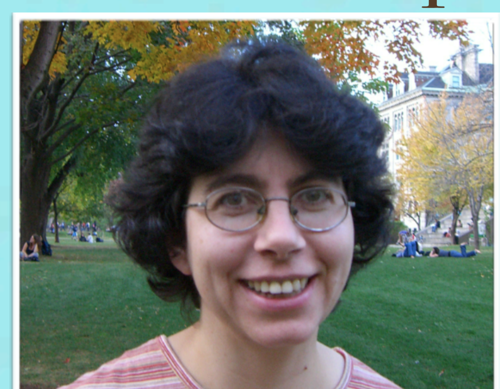
Anthony J. Bell

(RCTW, Berkeley)

(learning theory)

Bill Bialek
(Princeton)

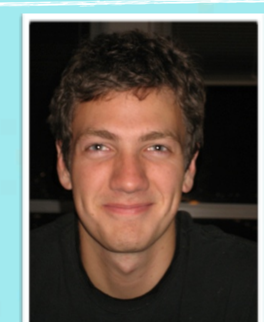
Doina Precup (McGill)



Jim Crutchfield
(UC Davis)

- (work in progress)
  Arne Grimsmo (NTNU/UoA)
  Giacomo Indiveri (ETH/UNIZ)

- (w.i.p. w/students)
  Lisa Miller (UHM ICS)
  Lane MacIntosh (UHM Math, now Stanford)
  Pradis Niknejadi (UHM Physics)
  Sarah Marzen (Berkeley, Physics)



Our papers

S.Still. Information Bottleneck approach to predictive inference. Entropy (in press).

S. Still, D. A. Sivak, A. J. Bell and G. E. Crooks. **The thermodynamics of prediction. Physical Review Letters 109, 120604 (2012)**

S. Still and D. Precup. An information-theoretic approach to curiosity-driven reinforcement learning. Theory in Biosciences 131 (3) pp. 139-148 (2012)

S. Still, J. P. Crutchfield, and C. Ellison. Optimal causal inference: estimating stored information and approximating causal architecture. Chaos 20, 037111 (2010)

S. Still. **Information-theoretic approach to interactive learning. EPL 85, 28005 (2009)**

S. Still and J. P. Crutchfield. Structure or Noise? arxiv:0708.0654 (2007)

S. Still and W. Bialek. How many clusters? An information theoretic perspective. Neural Computation16, pp. 2483-2506 (2004)

www2.hawaii.edu/~sstill