Relations Between Work Extraction and Entropy Production of Information Driven FSMs

Neri Merhav

Department of Electrical Engineering

Technion—Israel Institute of Technology

Haifa 3200004, Israel

ITA 2017, San Diego, CA, February 2017

Information Thermodynamics – Some Background

Common theme: the presence of information “violates” the second law.

Systems with measurement and feedback control:

Systems controlled by parameters that depend on past (noisy) measurements.

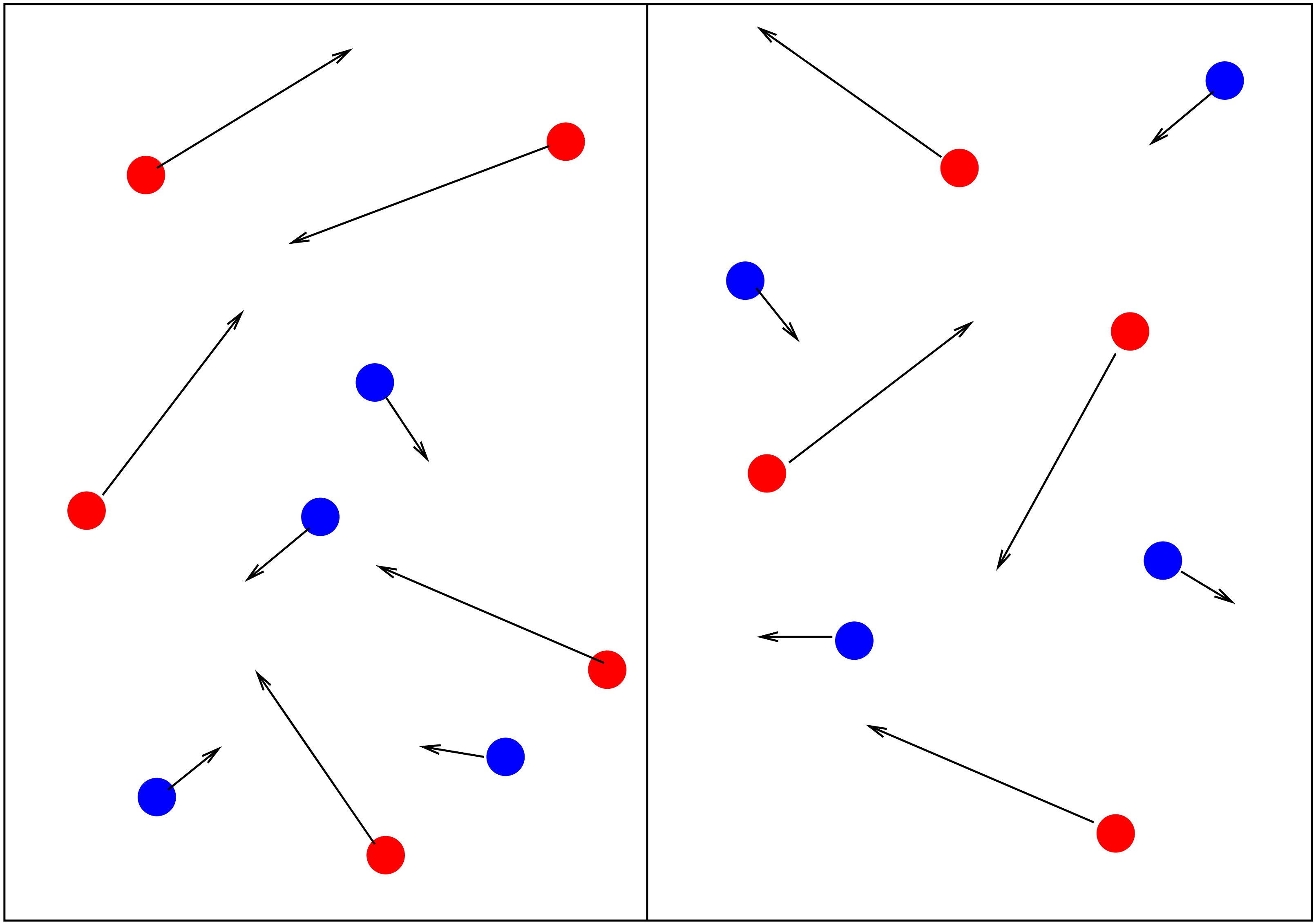
Barato & Seifert (2013, 2014); Deffner (2013); Deffner & Jarzynski (2013); Hoppenau & Angel (2014); Sagawa & Ueda (2011, 2012, 2013); Parrondo, Horowitz & Sagawa (2015); ...

Systems with information reservoirs:

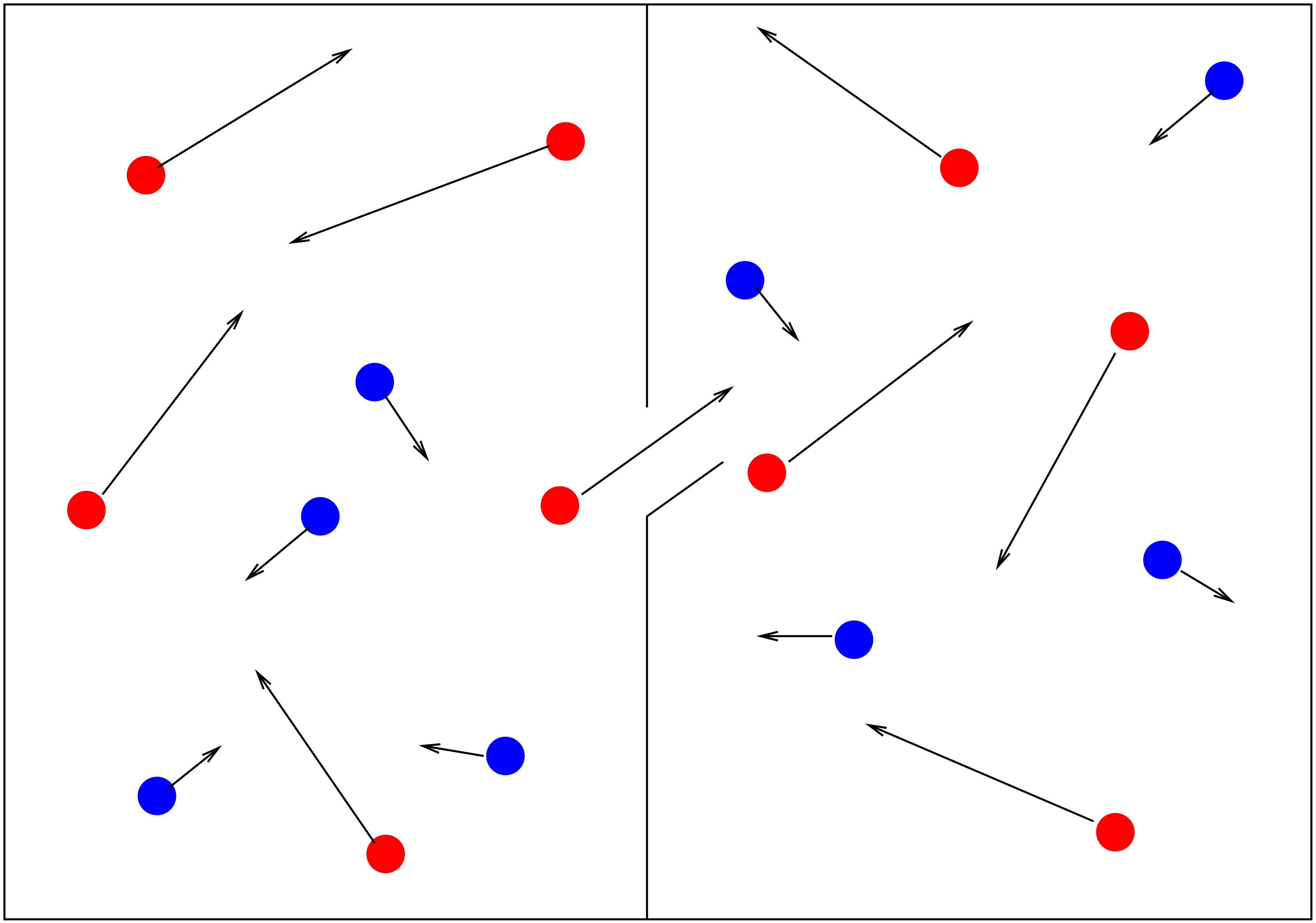
Systems that interact with an informational device (memory, digital tape, etc.).

Barato & Seifert (2014); Mandal & Jarzynski (2012, 2014); Boyd, Mandal & Crutchfield (2015); Mandal, Quan & Jarzynski (2013); Horowitz & Esposito (2014); ...

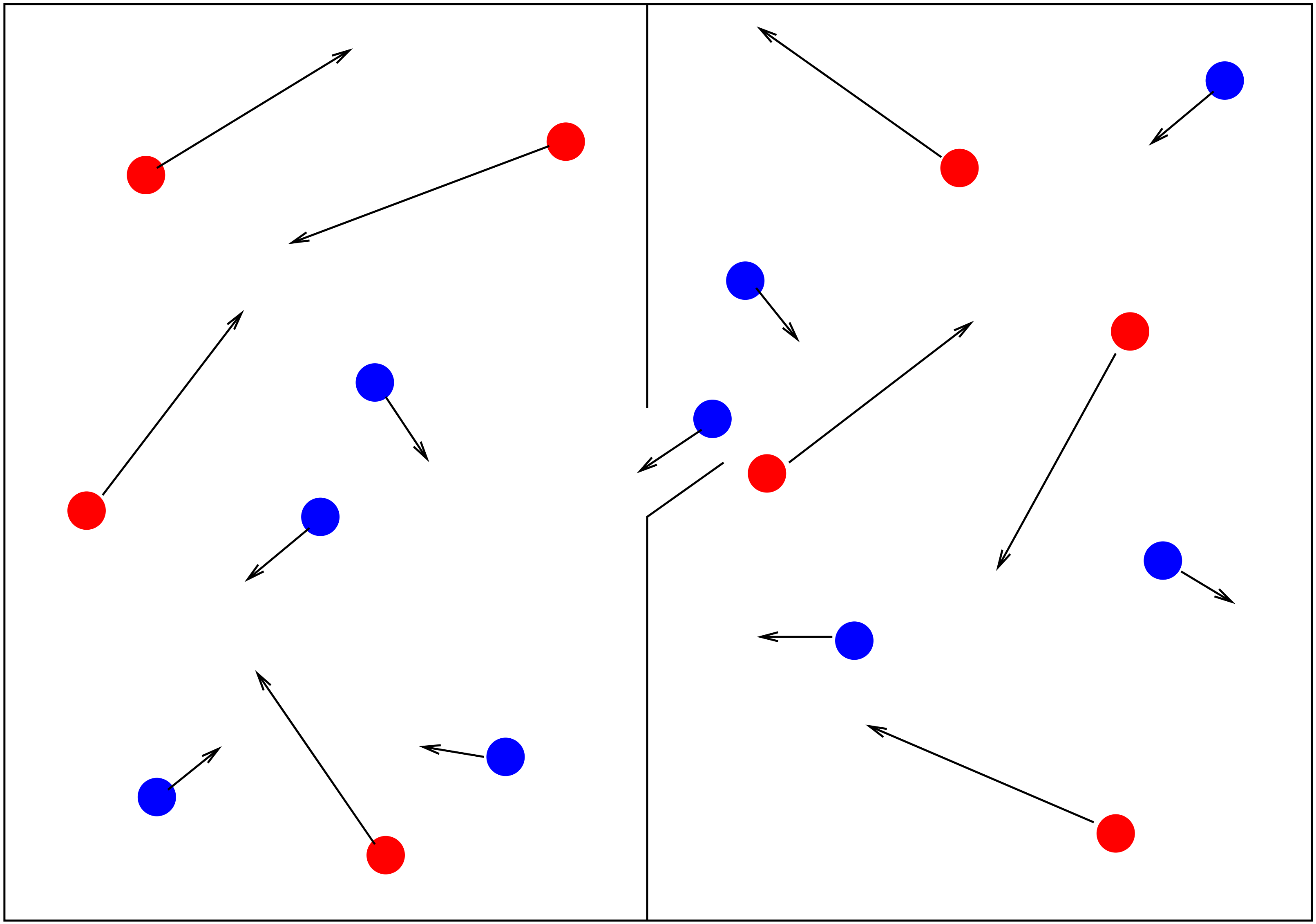
The Maxwell Demon (1867)



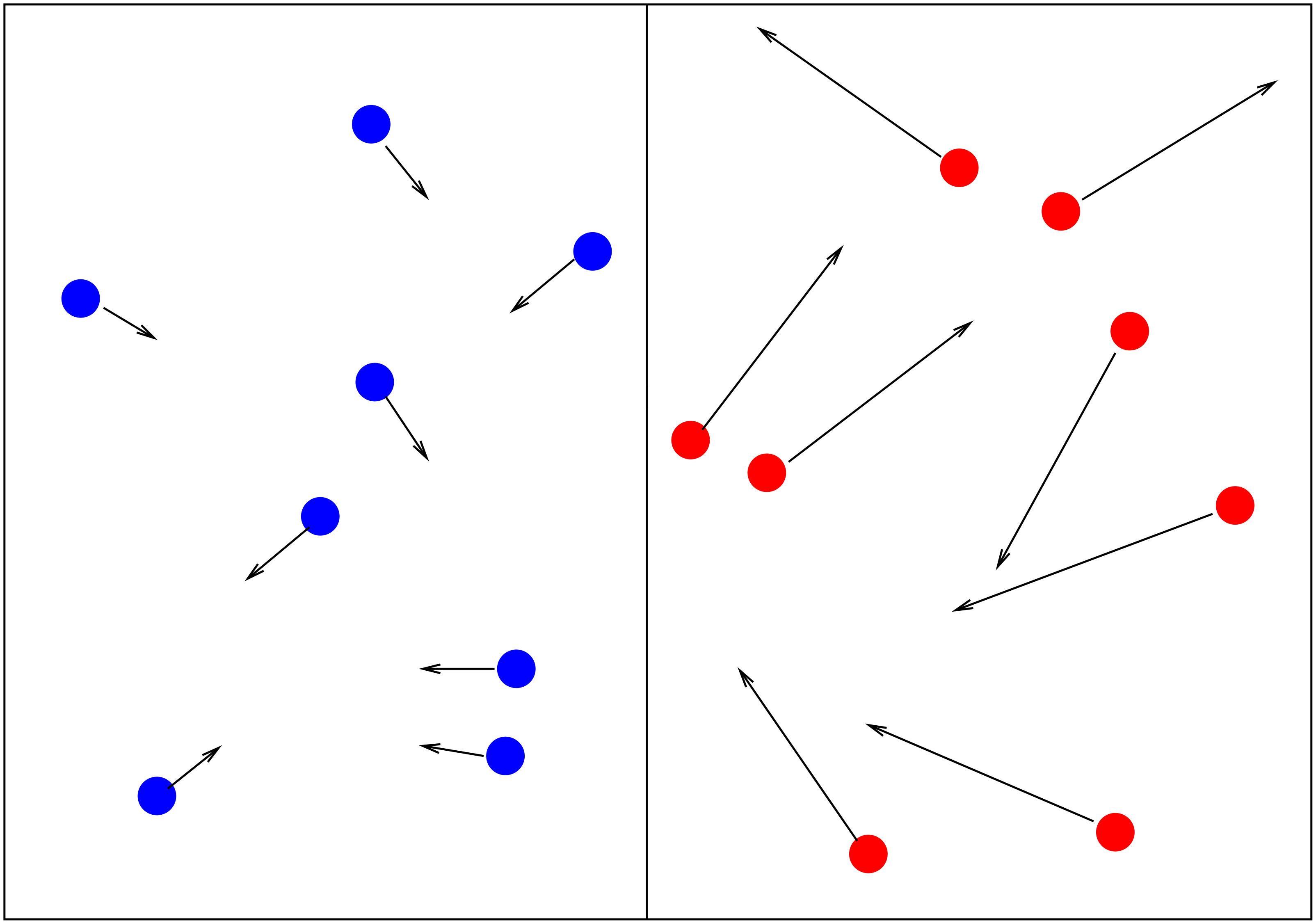
The Maxwell Demon (1867)



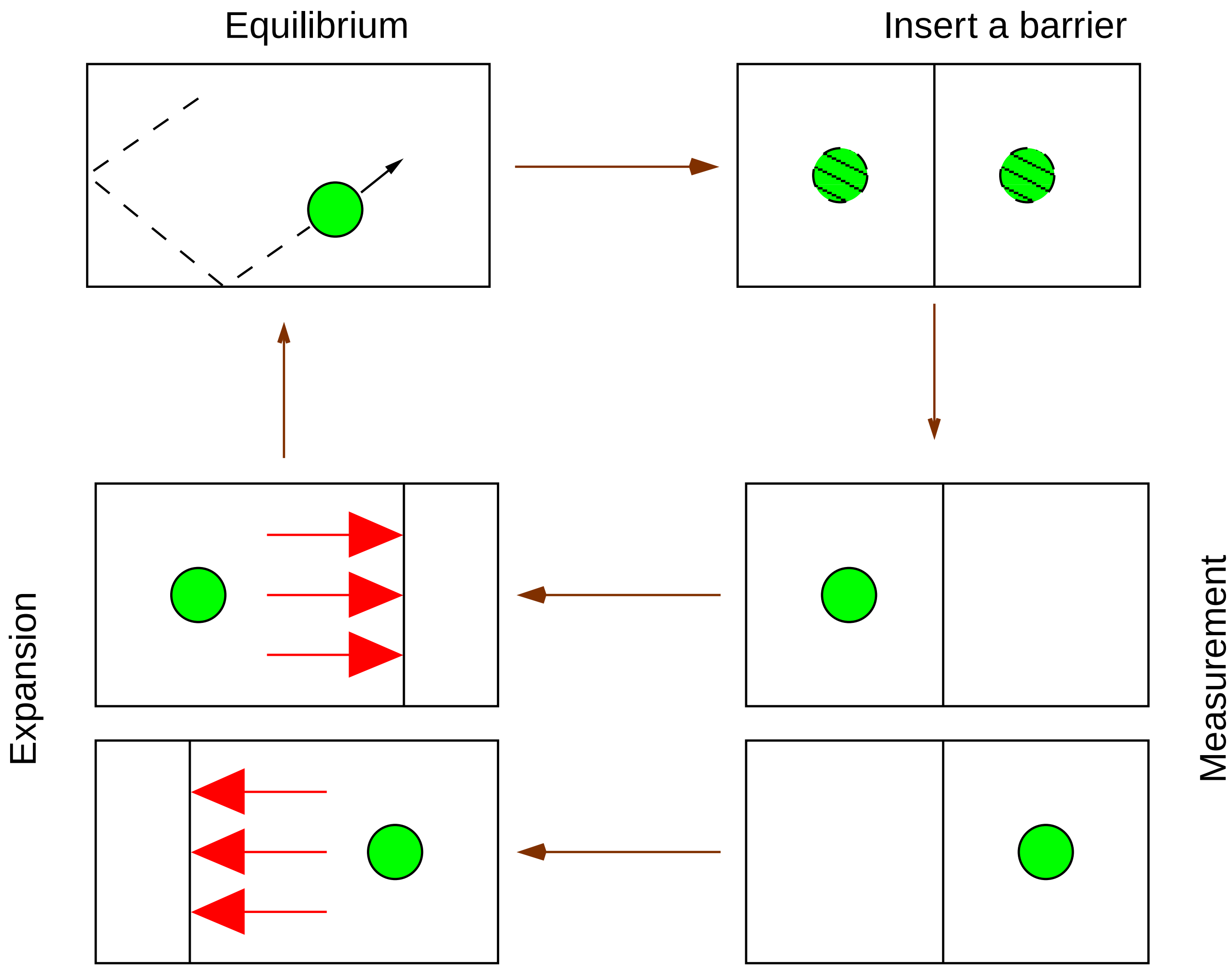
The Maxwell Demon (1867)



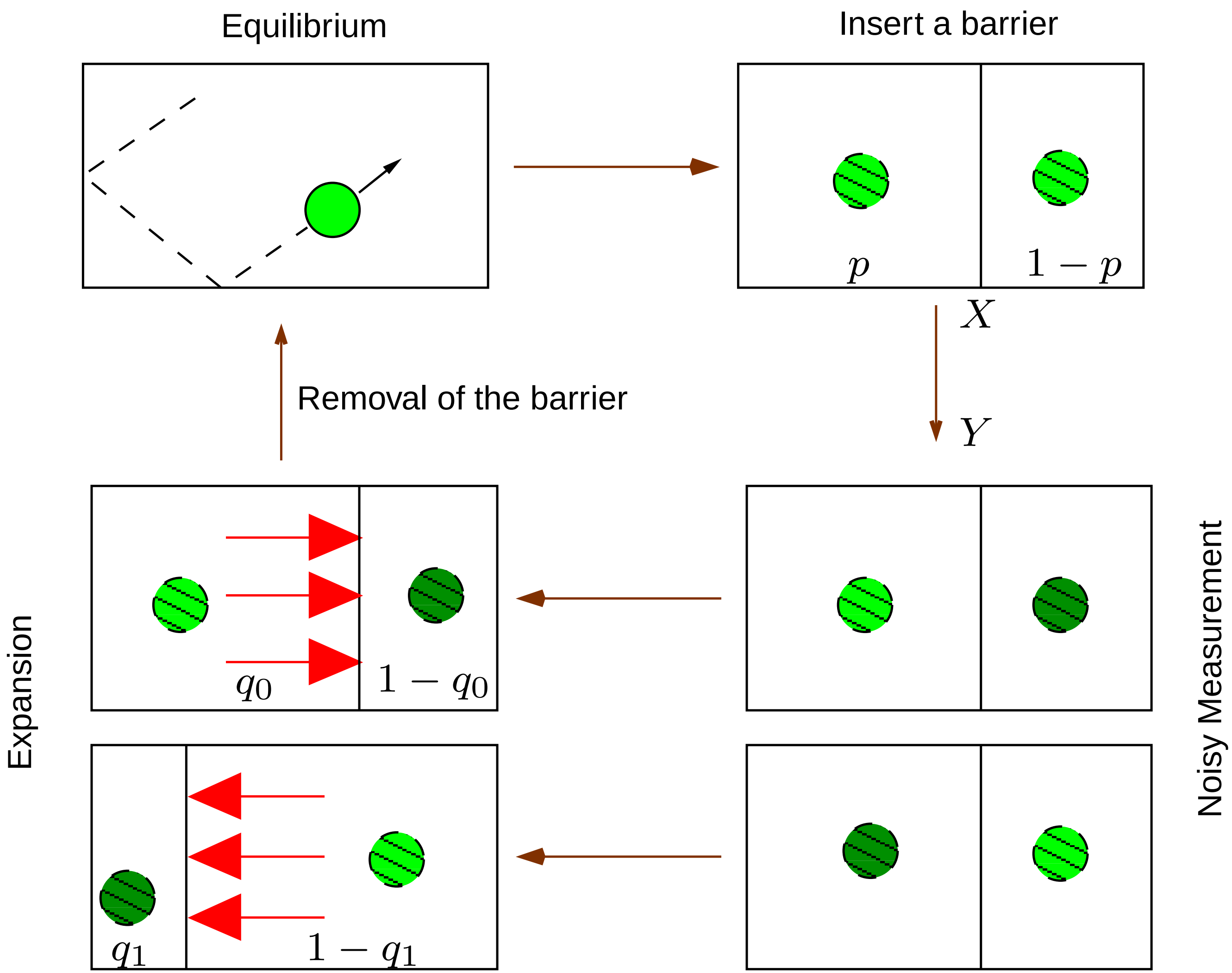
The Maxwell Demon (1867)



The Szilard Engine (1929)



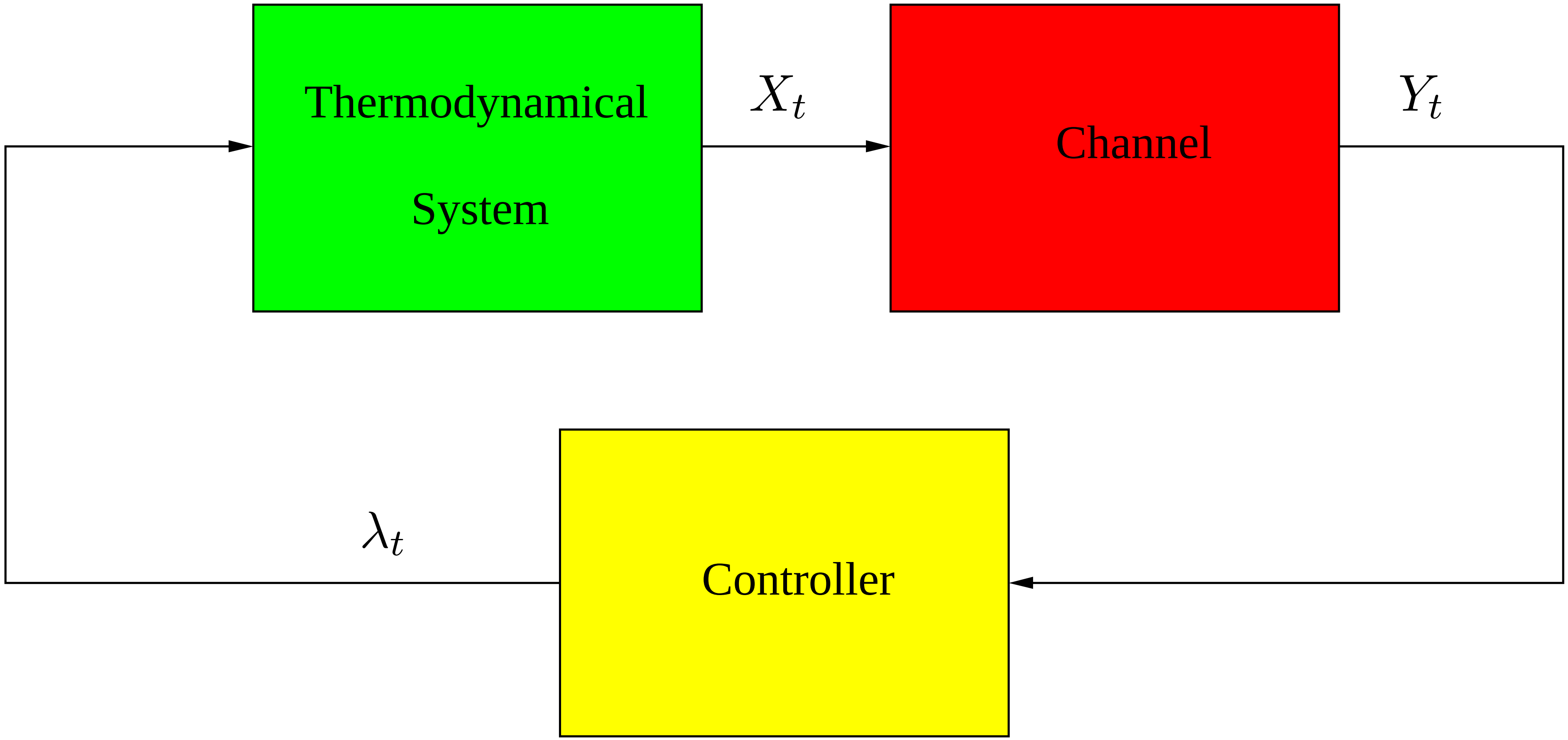
Generalized Szilard Engine (Sagawa & Ueda, 2011)



$\max_{q_0,q_1} W = kT \cdot I(X;Y)$

Vinkler, Permuter and Merhav (2014): relation to gambling.

System with Measurement and Feedback Control



Sagawa and Ueda (2008–2011) : Extracted work $\leq -\Delta F + kT \cdot nI(X;Y)$

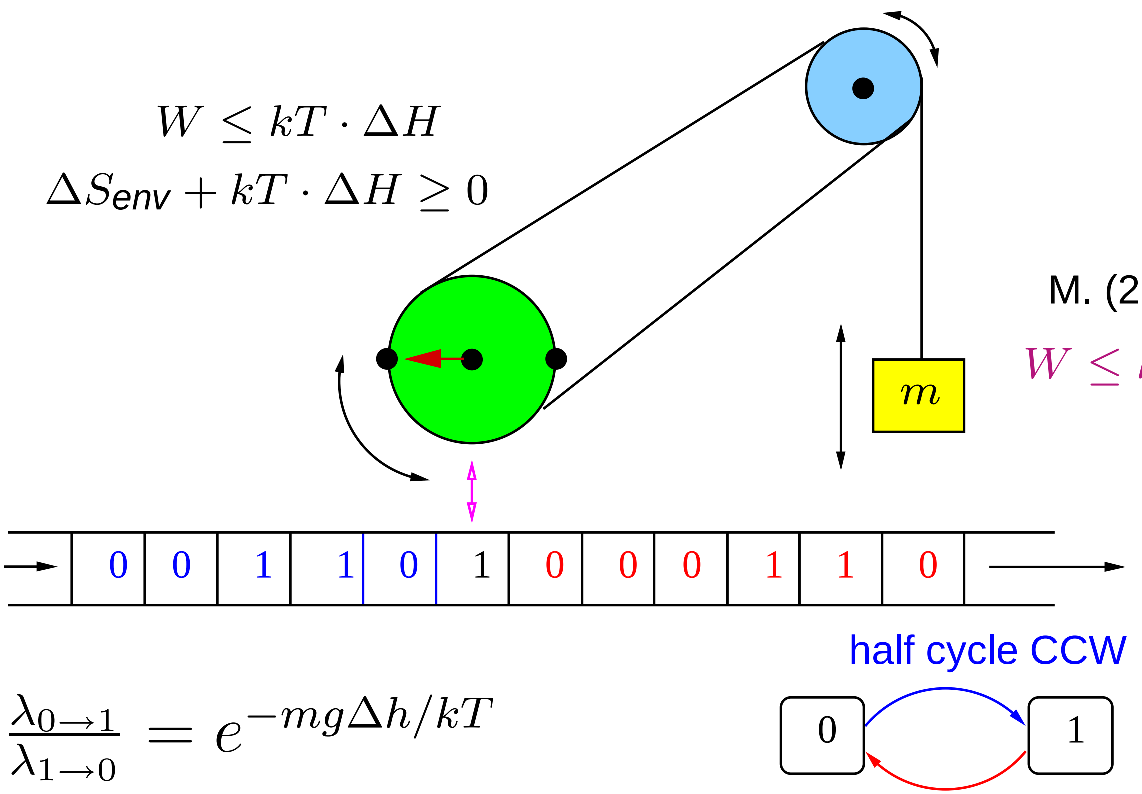
More generally, $nI(X;Y) \to I(X^n \to Y^n)$ = capacity of channel w. feedback.

Physical Systems with an Informational Device:

Mandal & Jarzynski (2012): system converting thermal fluctuations to work while writing info.

$W \leq kT \cdot \Delta H$

$\Delta S_{\text{env}} + kT \cdot \Delta H \geq 0$



M. (2015):

$W \leq kT \cdot F[\text{LZ complexity}]$

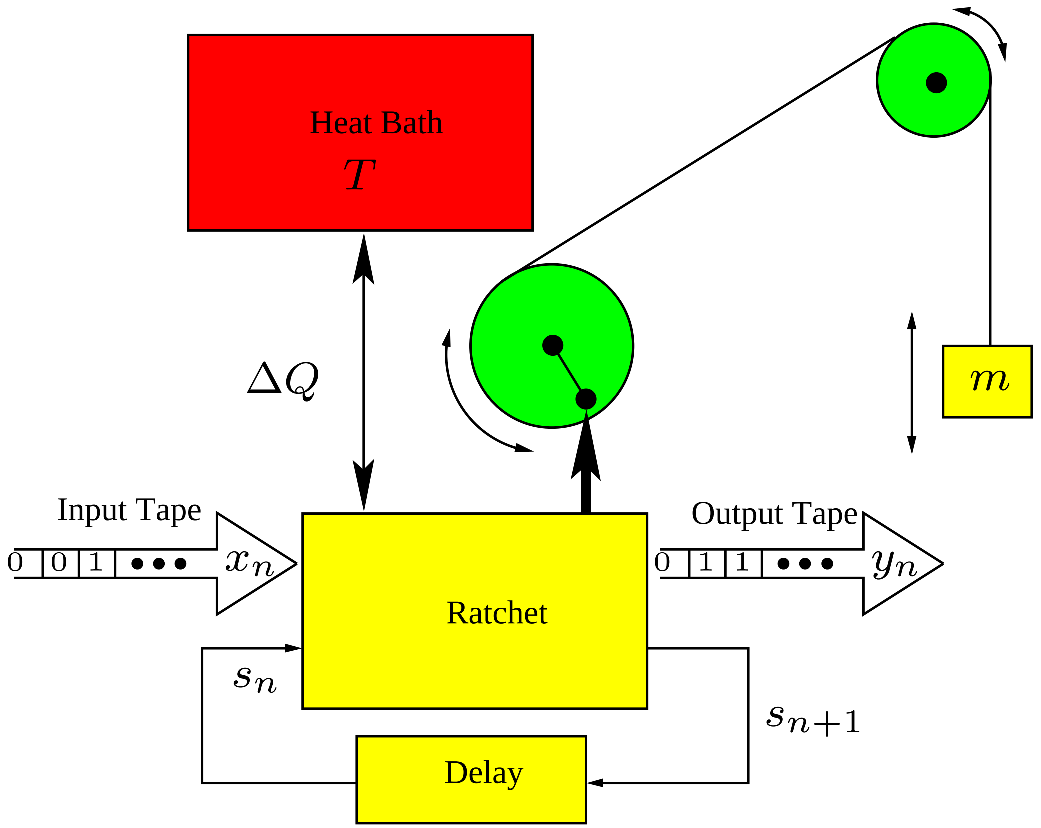
$$\frac{\lambda_{0\to1}}{\lambda_{1\to0}} = e^{-mg\Delta h/kT}$$

More generally, Deffner and Jarzynski (2013):

- A system (device) + a heat bath (heat reservoir),
- An information reservoir, e.g., a memory device with N bits (2^N states).

An extended 2nd law: $\Delta S_{\text{dev}} + \Delta S_{\text{heat-res}} + \Delta S_{\text{info-res}} \geq 0$.

More General System Model



- $\{x_n\}$ – finite-alphabet input process. Each symbol interacts for τ seconds.
- In the n-th cycle: $\{(\xi_t, \sigma_t),\ n\tau \leq t < (n+1)\tau\}$ = Markov jump process.
- $(\xi_{n\tau}, \sigma_{n\tau}) = (x_n, s_n)$; $(y_n, s_{n+1}) = (\xi_{(n+1)\tau-}, \sigma_{(n+1)\tau-})$.
- Energy $E(x,s) = mg\Delta h(x,s)$; stat. dist. $\propto e^{-E(x,s)/kT}$ (detailed balance).

Recent Related Work + This Work

- Boyd, Mandal and Crutchfield (2015, 2016):

$$\lim_{N\to\infty} \overline{W}_N \leq kT[\overline{H}(Y) - \overline{H}(X)].$$

- Several open issues (if not gaps..) in the proof:
  - 2nd law using Shannon entropy (true only under some conditions).
  - Confusing Shannon entropy with thermodynamic entropy.
  - Assuming independence between different parts of the system.
- In this work:
  - Simple (and rigorous) approach; mild asssumptions.
  - Exact results for every number of cycles, N.
  - Bounds – simple to calculate and potentially tight.
  - The state plays a role.

The Basic Idea

Using the inequality [Cover & Thomas, 2006]:

$$D(P_\tau \| P_{\text{eq}}) \le D(P_0 \| P_{\text{eq}}),$$

with

$$P_{\text{eq}}(x, s) = \frac{e^{-E(x,s)/kT}}{Z},$$

we readily obtain

$$\mathbf{E}\{\Delta W_n\} \equiv \mathbf{E}\{E(Y_n, S_{n+1}) - E(X_n, S_n)]\} \le kT[H(Y_n, S_{n+1}) - H(X_n, S_n)].$$

The Basic Idea (Cont'd)

Therefore,

$$
\begin{aligned}
\overline{W}_N &\leq kT \cdot \sum_{n=1}^{N} [H(Y_n, S_{n+1}) - H(X_n, S_n)] \\
&= kT \cdot \sum_{n=1}^{N} [H(Y_n | S_{n+1}) - H(X_n | S_n)] + kT \cdot \sum_{n=1}^{N} [H(S_{n+1}) - H(S_n)] \\
&= kT \cdot \sum_{n=1}^{N} [H(Y_n | S_{n+1}) - H(X_n | S_n)] + kT[H(S_{N+1}) - H(S_1)]
\end{aligned}
$$

First term $=$ input/output entropy production given the past (via the state);

Second term $=$ entropy production of the state.

Tightness

If P_0 is very close to P_{eq} in the sense that

$$P_0(\xi,\sigma) = P_{\text{eq}}(\xi,\sigma)[1+\epsilon(\xi,\sigma)], \qquad \epsilon = \max_{\xi,\sigma} |\epsilon(\xi,\sigma)| \ll 1$$

and if P_τ is similarly within $1 \pm \epsilon^2$ away from P_{eq}, then the bound is tight in the sense that

$$\frac{\mathbf{E}\{\Delta W_n\}}{kT[H(Y_n,S_{n+1}) - H(X_n,S_n)]} = 1 - O(\epsilon).$$

If P_0 and P_{eq} are far apart, one can make many intermediate steps,

$$P_0 \rightarrow P_1 \rightarrow \ldots \rightarrow P_L \rightarrow P_{\text{eq}}$$

in between and still saturate the bound up to a term of $O(1/L)$. This requires a cascade of L systems.

Memoryless/Markov Inputs

For a memoryless input, $H(X_n|S_n) = H(X_n)$, and

$$\frac{\overline{W}_N}{kT} \leq \sum_{n=1}^{N}[H(Y_n) - H(X_n)] + H(S_{N+1}) - H(S_1) - \sum_{n=1}^{N} I(S_{n+1};Y_n)$$

Conclusion: memoryless inputs are best processed by memorlyess machines.

For a Markov input:

- $\{(X_n, S_n)\}$ is Markov as well.
- If it has a stationary distribution $P_\infty(x, s)$, then

$$P(x, s, y, s') = P_\infty(x, s)P(y, s'|x, s)$$

  completely dictates the steady–state work and the entropy production.
- Easier to calculate than $kT[\overline{H}(Y) - \overline{H}(X)]$, as $\{Y_n\}$ is an HMM.

Conditional Entropy Bounds

For a given n, let $V_n = f_n(X^{n-1}, Y^{n-1}, S^n)$, then

$$\overline{W}_N \leq kT \sum_{n=1}^{N} [H(Y_n, S_{n+1}|V_n) - H(X_n, S_n|V_n)],$$

with the freedom of choosing f_n, but the best choice is $V_n = \emptyset$......
However, for $V - n = (X^{n-1}, Y^{n-1})$, we can obtain a lower bound on the output entropy:

$$H(Y^N) \geq H(X^N) + \frac{\overline{W}_N}{kT} - N \ln |\mathcal{S}|.$$

If $\{X_n\}$ is memoryless/Markov, then $\{Y_n\}$ is an HMM, but the lower bound is easy to calculate.

Summary

- Simple relations between work and Δ-entropy (generalized 2nd law).
- The bounds are (potentially) tight.
- Memory of the past – via the state; a state entropy–production term.
- Relatively easy to calculate (examples in the paper).
- Lower bound on the entropy of HMM.
- Generalization – monotonicity of the f–divergence, $D_f(P_\tau \| P_{\text{eq}})$.