Improved Bounds on Entropy Production in Living Systems

Jonathan (Shao-Kai) Huang

February 14, 2025

Department of Physics, National Taiwan University

IMB, Academia Sinica

Table of Contents

1 Introduction

Overview

- **Living systems maintain or increase local order by working against the second law of thermodynamics**

- Optimization framework: thermodynamic infererence of improved lower bounds on entropy production rates

Deduction of thermodynamic quantities from partial observations.

- Non-zero bounds for entropy production by considering transition statistics beyond one-step dynamics

Present Problems

- Entropy production estimators fail when
  - Only a small set of observables is accessible
  - Nonequilibrium transport currents vanish

- Many degrees of freedom are typically hidden in experiments

- Sometimes trajectories *appear* symmetric in time

  ⇒ don’t obey detailed balance, no relative entropy!

Biological Examples

- Overview of biological data analyzed

- Bacterial flagellar motor
- Microtubule growth
- Calcium oscillation in human embryonic kidney cells

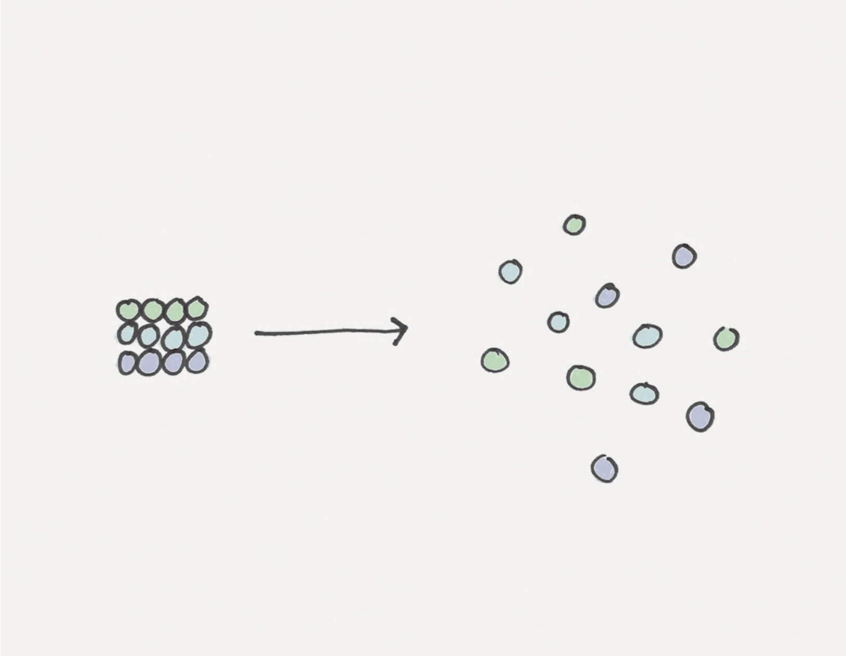
What is Entropy?

- Entropy as a measure of "disorder"
- Classically, entropy is a measure of the unavailability of a system's energy to do work

$$\Delta S = \frac{\Delta Q}{T}$$

- Entropy change can be experimentally measured

$$\underbrace{\Delta G}_{\text{Gibbs energy}} = \underbrace{\Delta U}_{\text{internal energy}} - T\Delta S$$



What is Entropy?

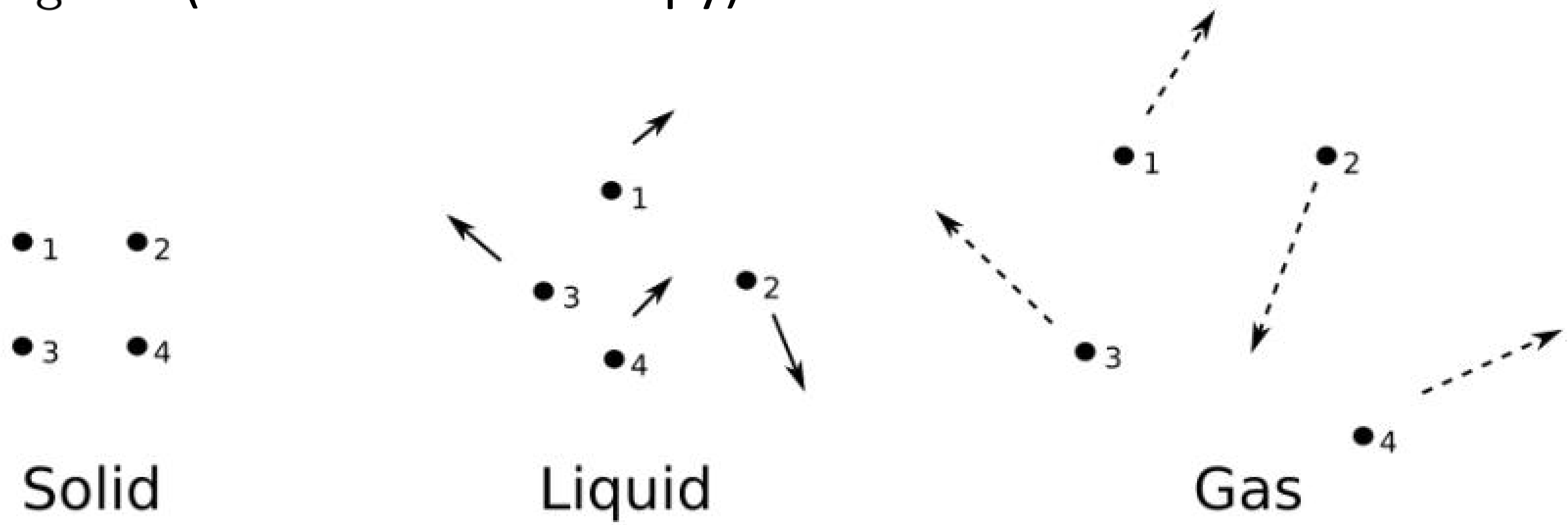
- Entropy as a probabilistic idea
- There seems to be more “disordered states” than there are “ordered states”

$S = k_\mathrm{B} \log \Omega$ (Boltzmann’s entropy)

Example (Phases of matter)

- More possible configurations in a gaseous system

⇒ gas has more entropy than solid



What is Entropy?

- As a number associated to a probability distribution over a finite sample space

$$H = -\sum_{i=1}^{N} p_i \log p_i$$

p_i is the probability of being in the state *i*

(Shannon entropy)

- Consistent with entropy in statistical physics (Cf. Boltzmann)

What is Entropy?

- As a number associated to a probability distribution over a finite sample space

$$H = -\sum_{i=1}^{N} p_i \log p_i$$

(Shannon entropy)

p_i is the probability of being in the state *i*

- The most compressed amount of code needed to encompass the content of a signal (information)
- We can reinterpret “disorder” as “information content”

What is Entropy?

- An event E happens with probablity $\mathbb{P}(E)$
- $s(E)$ is the information content of E ⇒ function of $\mathbb{P}(E)$

1. Decreasing function of $\mathbb{P}(E)$
2. If events E_1 and E_2 are independent, then $s(E_1 \cap E_2) = s(E_1) + s(E_2)$

⇒ $s(E)$ is proportional to $-k \log \mathbb{P}(E), \quad k > 0$

Entropy is the expectation value of information content

Entropy in Biology

- Second Law of Thermodynamics:

  The entropy of the universe tends to a maximum.

- Biological organisms are highly ordered ⇒ nonequilibrium

  locally, entropy can be lowered by external action

  increase in environmental entropy ⇒ the Second Law

2 Literature Review

Upper Bound to Entropy Production Rates

- This paper gives a lower bound for entropy production
- An upper bound for steadystate has been derived:

$$a = \sum_{i<j} (\pi_i q_{ji} + \pi_j q_{ij}) \qquad R = \max_{i \neq j} \frac{q_{ij}}{q_{ji}} \geq 1$$

$$\Rightarrow \quad \sigma \leq a \, (\log R) \left(\frac{R-1}{R+1} \right)$$

(Ref.) Nishiyama, T.; Hasegawa, Y. (2023). *Upper bound for entropy production in Markov processes*.

Thermodynamic Uncertainty Relations

- Thermodynamic uncertainty relations (TURs) can constrain nonequilibrium fluctuations

$$\Sigma_\tau \geq 2k_B \frac{\langle J_\tau \rangle^2}{\mathrm{Var}(J_\tau)}$$

- Classic lower bound for entropy production rate

(Ref.) Horowitz, J. M.; Gingrich, T. R. (2020). *Thermodynamic uncertainty relations constrain non-equilibrium fluctuations*.

3 Theory

Entropy Production Rate

spends a fraction π_i
of the time in state *i*

- Systems in thermal equilibrium obey detailed balance: $\pi_i q_{ij} = \pi_j q_{ji}$

⇒ This is a reversible Markov process

rate of transition
from state *i* to *j*

- The entropy production rate in steadystate is

$$\sigma = \frac{k_B}{2} \sum_{i,j} (\pi_i q_{ij} - \pi_j q_{ji}) \log\left(\frac{\pi_i q_{ij}}{\pi_j q_{ji}}\right)$$

Detailed Balance

$$\sigma = \frac{k_B}{2} \sum_{i,j} (\pi_i q_{ij} - \pi_j q_{ji}) \log \left(\frac{\pi_i q_{ij}}{\pi_j q_{ji}} \right)$$

No entropy production when system is in equilibrium

Always positive

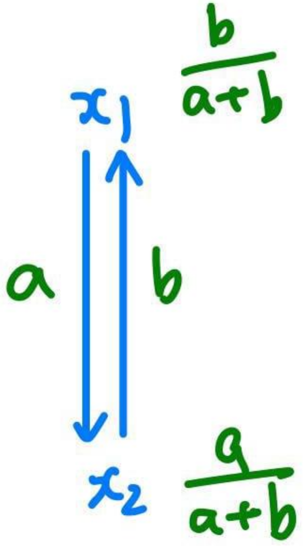
- Organisms operate far from equilibrium. Violations of detailed balance increase environment entropy

- Apply the formula? Not all microstates are accessible

⇒ Estimate σ from coarse-grained data

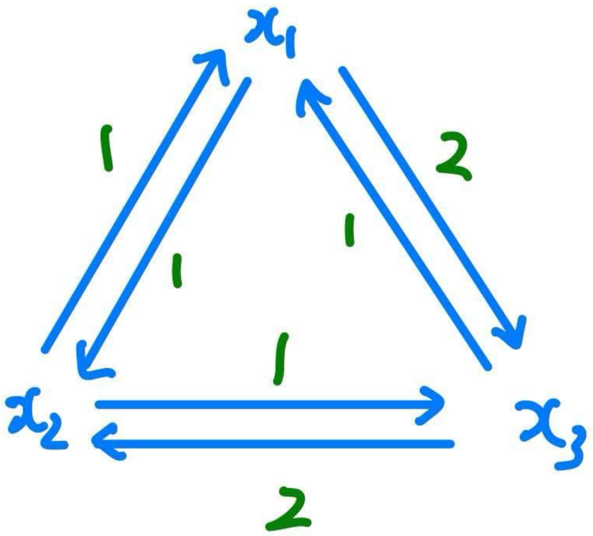
Detailed Balance

Example (Two nodes)



Equilibrium is guaranteed

Example (Three nodes)



Equilibrium is not guaranteed

Detail balance is a strong condition

Markov Processes

- Markov property (discrete): $X = (X_1, X_2, \dots, X_N, \dots)$

$$\mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n, \dots, X_1 = x_1) = \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n) \text{ for all } n \in \mathbb{N}.$$

- A continuous time Markov process a sequence of random variables with the Markov property

$$\mathbb{P}(X_t = j | X_0 = i) = P_{ij}(t) = (e^{Qt})_{ij}$$

Markov Processes

- History independence
- We may assume our biological system to be microscopically Markovian

⇒ Why do we emphasize “microscopic”?

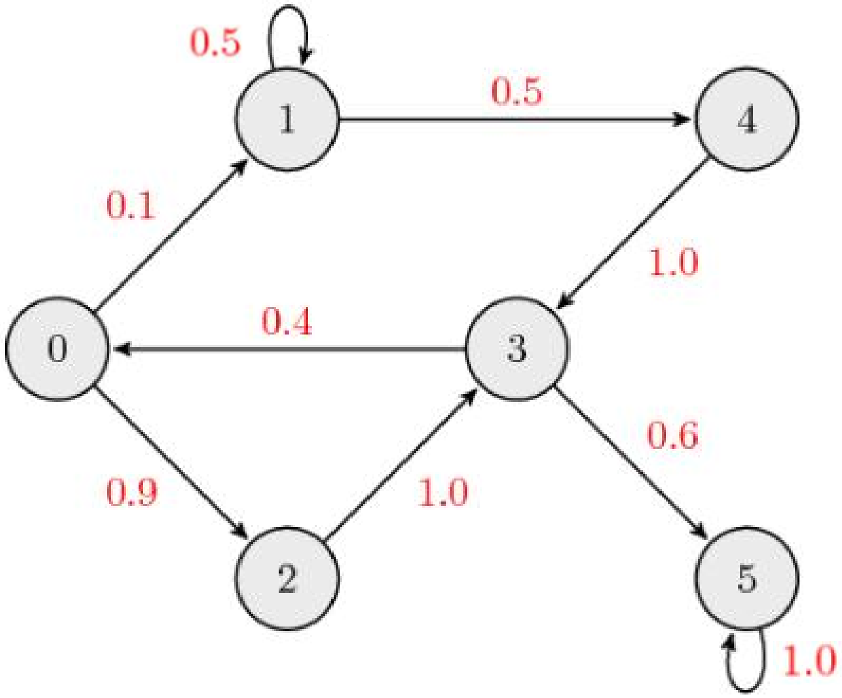
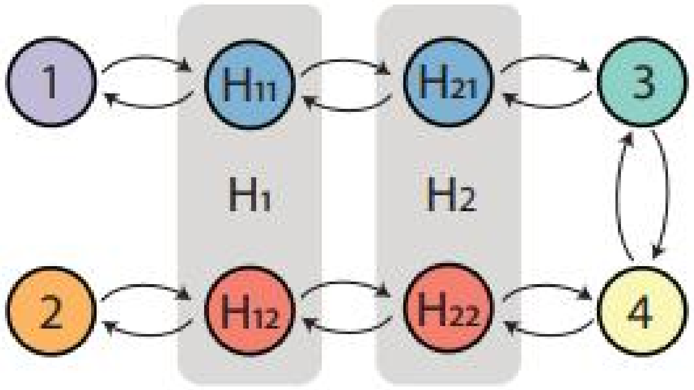


FIG. Example of a Markov chain

An Example



- System currently in macrostate H_1

  The probabilities of jumping to 1 or 2 are both non-zero

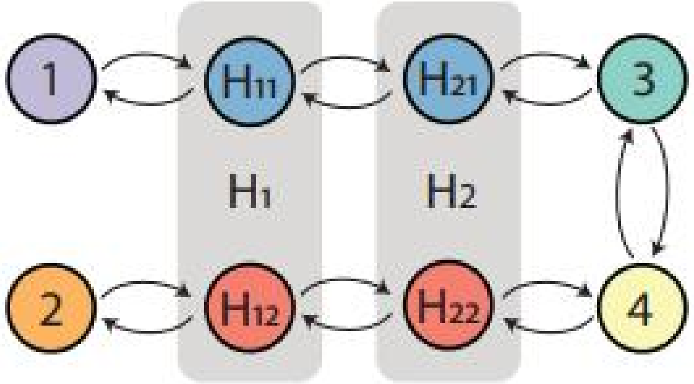
- Previous observed macrostate was 1

  The probability of transitioning to 2 is zero

- Coarse-grained observables may be non-Markovian!

  ⇒ History dependence emerges from partial observation

An Example



- Arbitrary trajectory $(3, H_2, H_1, H_2, \ldots, H_1)$

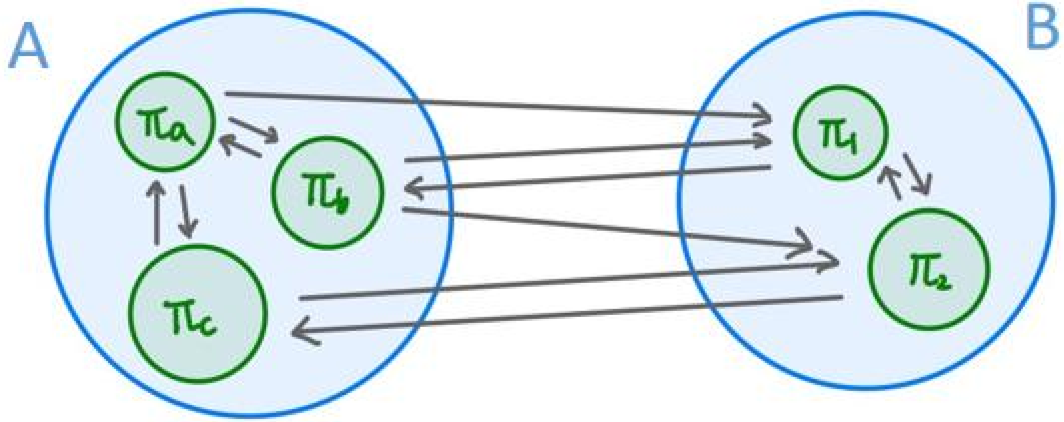
  The probability of transitioning to 2 is zero

- Truncate the trajectory to length N : $(H_2, H_1, H_2, \ldots, H_1)$

  Can the process transition to 2? It shouldn't! We lose information

Markovian Systems

- No finite set of observables guaranteed to describe the system ⇒ Construct an estimator that preserves the statistics
- Observable: $\hat{n}_{AB} = \sum_{i \in A} \sum_{j \in B} \pi_i q_{ij}$



Example calculation:

$$A = \{a, b, c\}, \quad B = \{1, 2\}.$$

$$\hat{n}_{AB} = \sum_{i=a,b,c} \sum_{j=1,2} \pi_i q_{ij}$$
$$= \pi_a q_{a1} + \pi_b q_{b1} + \pi_b q_{b2} + \pi_c q_{c2}$$

As as Optimization Problem

set of all possible macroscopic observables for our system

The best possible estimator is

$$\sigma(\mathcal{S}) \geq \min\{\sigma(\mathcal{R}) | \mathcal{O}(\mathcal{R}) = \mathcal{O}(\mathcal{S})\}.$$

$$\Downarrow$$

This is not practical

Let’s look at a subset of the observable space

$$\sigma(\mathcal{S}) \geq \min\{\sigma(\mathcal{R}) | \mathcal{O}_k(\mathcal{R}) = \mathcal{O}_k(\mathcal{S})\},$$

Unique stationary distribution, use *n* as observable

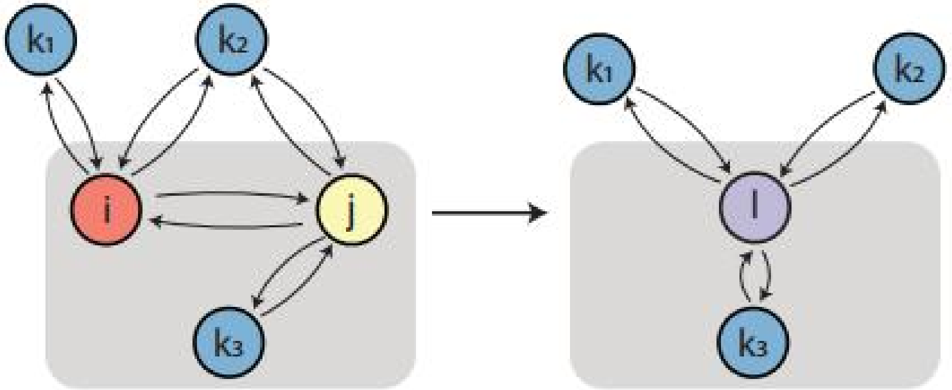
$$\sigma = \frac{k_B}{2} \sum_{i \neq j} (n_{ij} - n_{ji}) \log \left(\frac{n_{ij}}{n_{ji}} \right)$$

One-Step Estimator

- We can simplify the internal topology without affecting $\mathcal{O}_1$ observables

$$(n_{ik} - n_{ki}) \log \left(\frac{n_{ik}}{n_{ki}} \right) + (n_{jk} - n_{kj}) \log \left(\frac{n_{jk}}{n_{kj}} \right)$$

$$\mapsto \; ((n_{ik} + n_{jk}) - (n_{ki} + n_{kj})) \log \left(\frac{n_{ik} + n_{jk}}{n_{ki} + n_{kj}} \right)$$



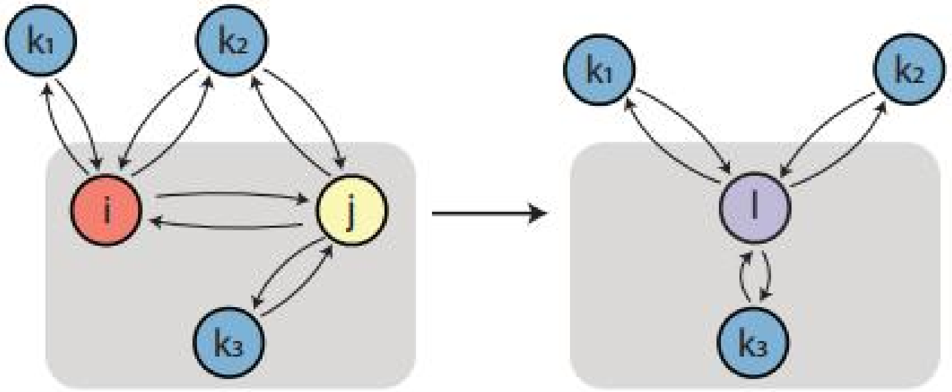
$$\mathcal{O}_1 = \{n_{IJ} | I, J \text{ macrostates}\}$$

One-Step Estimator

- Convexity + Jensen’s inequality

  ⇒ Entropy production rate decreases

- Markovian system with no hidden states and the same statistics (observables)



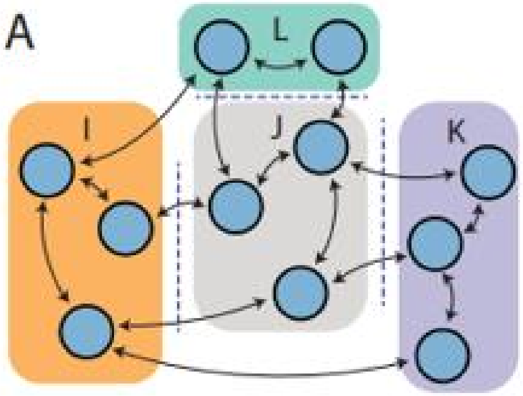
$$\mathcal{O}_1 = \{n_{IJ} | I, J \text{ macrostates}\}$$

Two-Step Estimator

$$\mathcal{O}_2 = \{n_{UV}, n_{UVW} | U, V, W \text{ macrostates}\}$$

- We can obtain a bound for each macrostate, consistent with $\mathcal{O}_2$ observables

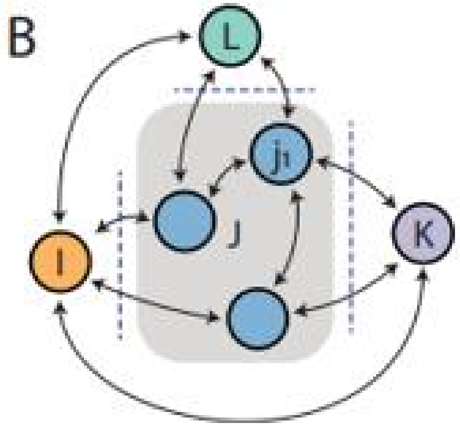
- Example: find the minimal EP rate across the edges (*J*, *I*), (*J*, *K*), and (*J*, *L*)



Example network

$\Longrightarrow$

Apply σ_1 to simplify *I*, *K*, *L*

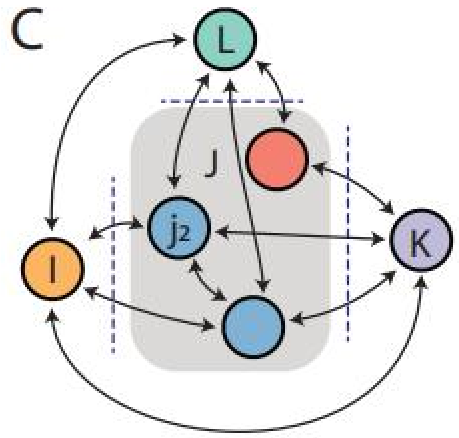


$$\sigma(j) = \sum_{i \in \mathcal{K}} (n_{ij} - n_{ji}) \log\left(\frac{n_{ij}}{n_{ji}}\right)$$

The entropy production along edges connected to *j*

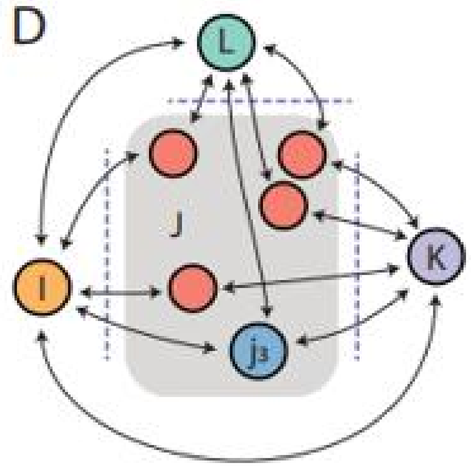
Two-Step Estimator

$\Longrightarrow$



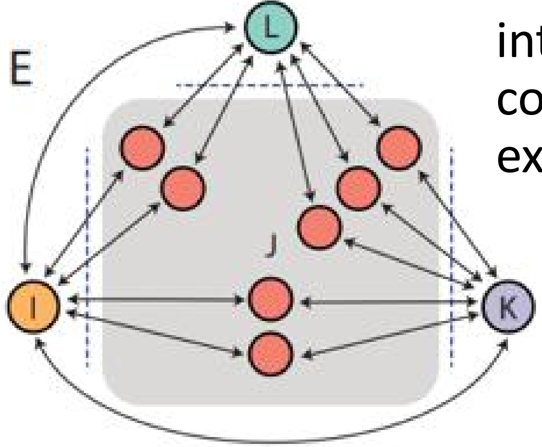
Rerouting the network
does not increase
entropy production

$\Longrightarrow$



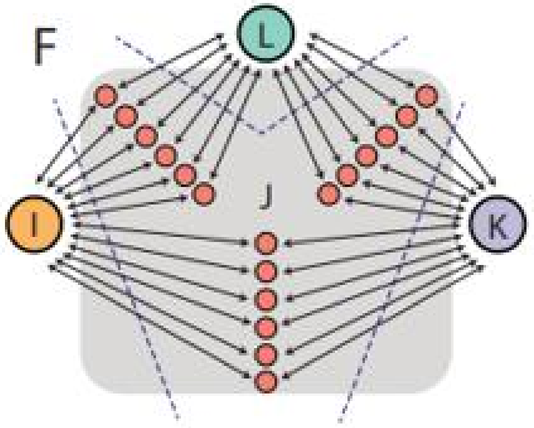
Iteration

$\Longrightarrow$



Canonical form: each internal state in *J* is connected only to two external macrostates

$\Downarrow$



It is always possible to optimize over 6 internal states connecting any two macrostates

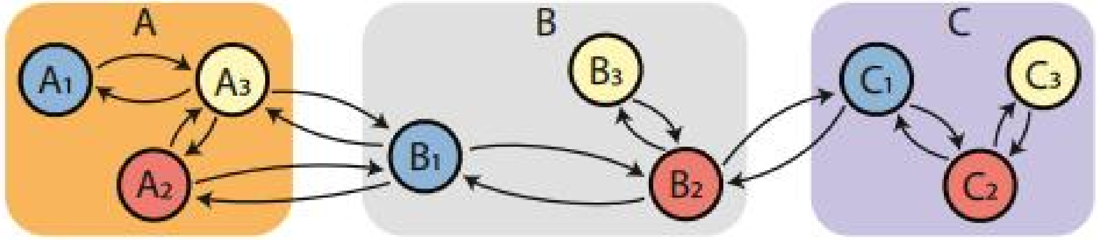
- Summing over all macrostates gives the $\mathcal{O}_2$ estimator

$$\sigma_2 = \frac{1}{2} \sum_J \sigma_2(J) \leq \sigma$$

Two-Step Estimator

$$\sigma_2 = \frac{1}{2}\sum_J \sigma_2(J) \leq \sigma$$

- For a simple system A-B-C, $\sigma_2(B)$ is the optimal bound
- Less computation required



A three state topology with no “loops”

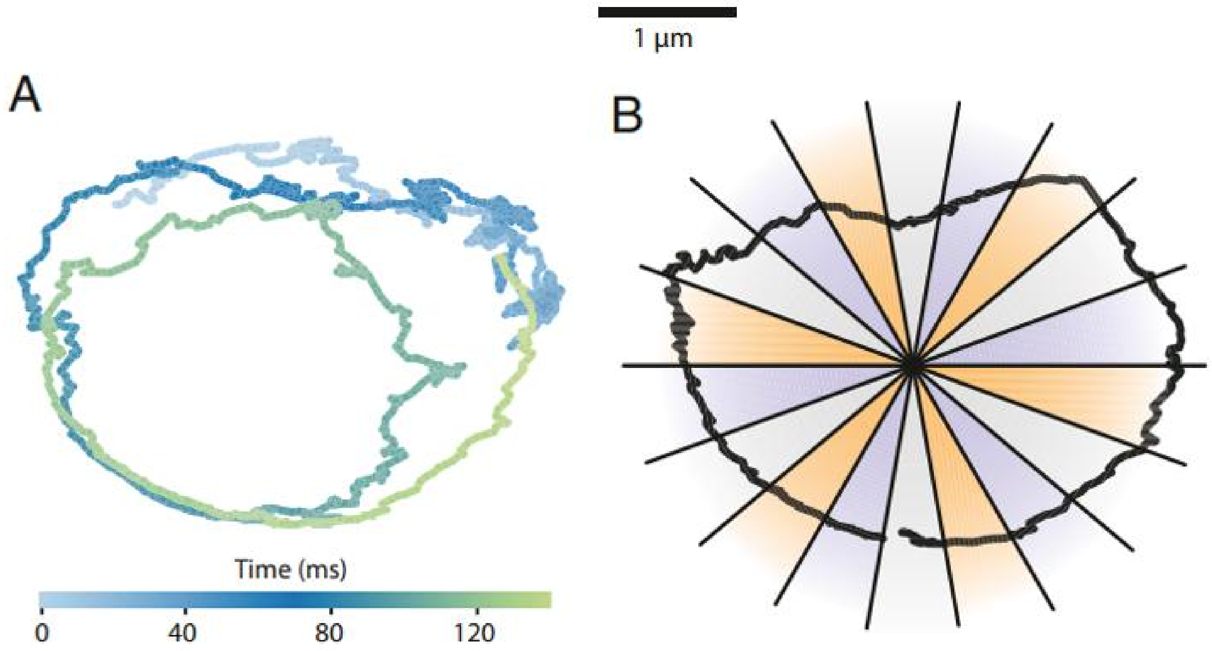
3 Results

Ex 1. Microbial Flagellar Motor

- Each flagellum is driven by a motor

- Stochastic direction switching

  ⇒ May not violate time-irreversibility



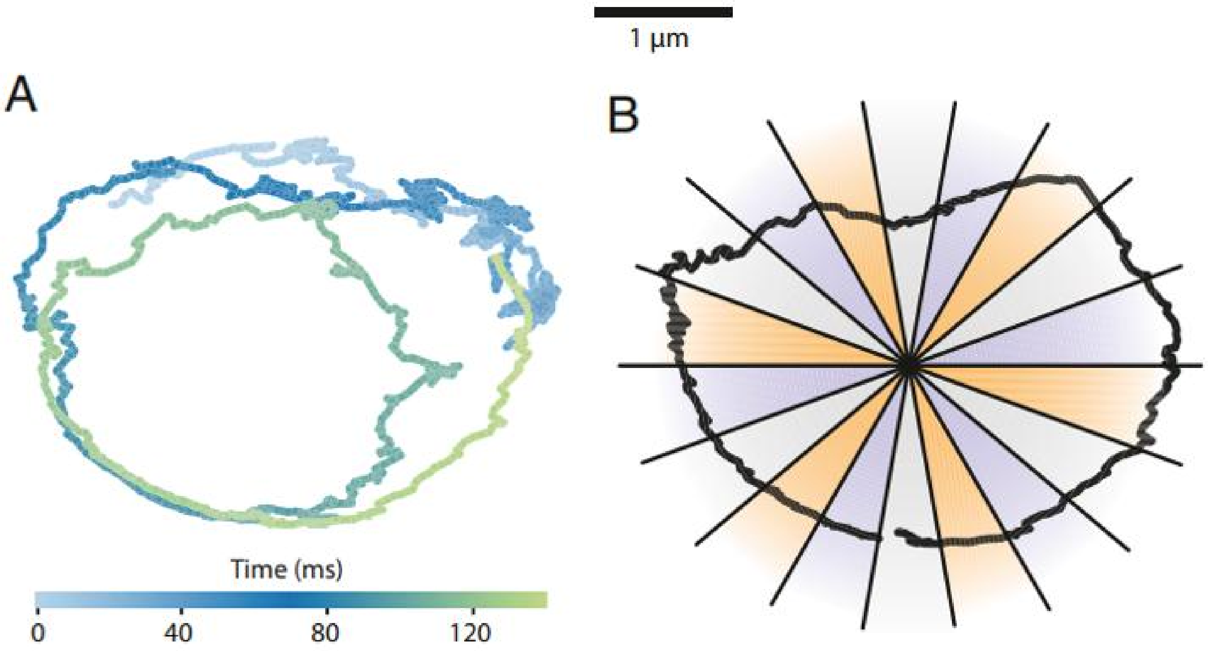
This example exhibits small net fluxes

Ex 1. Microbial Flagellar Motor

- Coarse-grain space into 3 regions, in total $3N$ segments ($N = 6$)

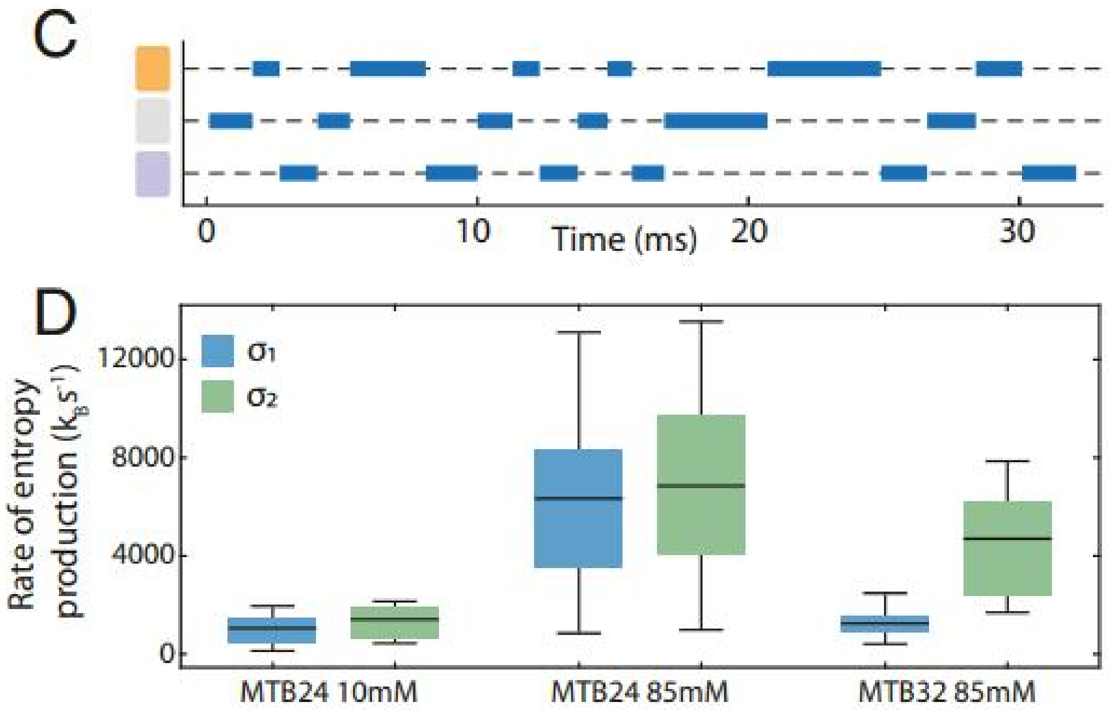
⇒ The σ_2 gives a nontrivial lower bound

- The quantity σT is interpreted as the free energy consumption rate



This example exhibits small net fluxes

Ex 1. Microbial Flagellar Motor



Compare: specimen with single turning direction and stochastically switching direction

Ex 2. Microtubule Growth

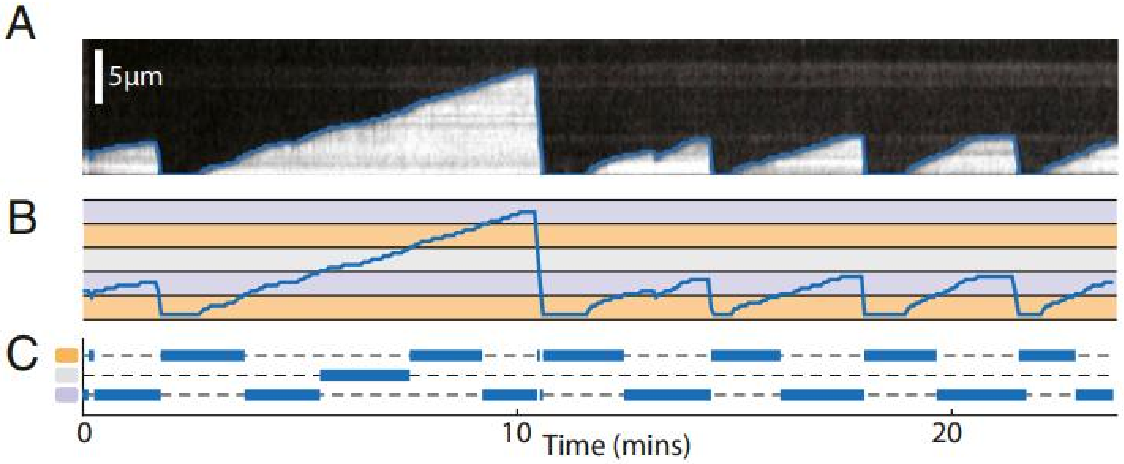
Recap: periods of steady growth before stochastically swicthing into rapid shrinking

Ex 2. Microtubule Growth

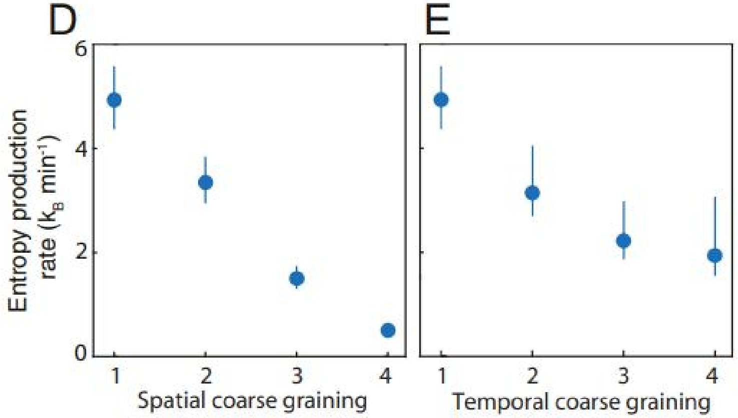
- Length oscillates around its mean value ⇒ no net flux

- Coarse-grain space modulo 3 and apply σ_2
  nontrivial bound of entropy production

- Higher-resolution experiments promise improved bounds



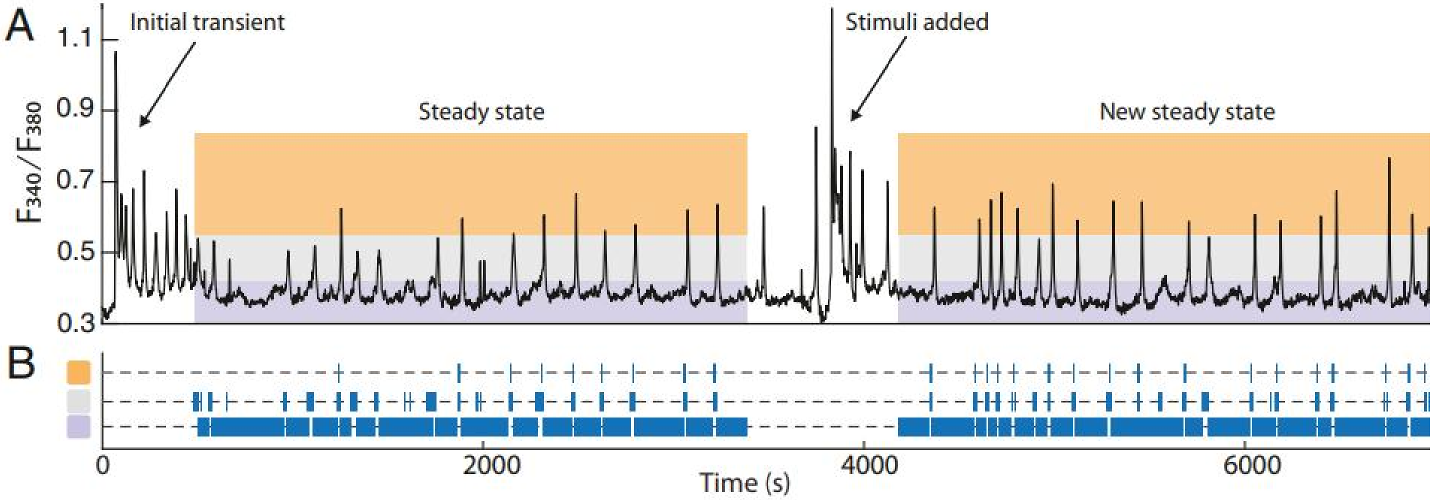
(A) Kymograph from experimental observations and (B)(C) corase-grained data



(D) Coarse-graining over space and (E) coarse-graining over time

Ex 3. Calcium Oscillation

- A cell responses to external stimuli via intracellular signaling



Manipulation of the concentration of calcium ions within the cytosol of human embryonic kidney cells

- Coarse-graining (modulo 3) followed by optimization
- The entropy production increases after stimulation with carbachol

$$4\, k_\mathrm{B} \cdot \mathrm{min}^{-1} \longrightarrow 8\, k_\mathrm{B} \cdot \mathrm{min}^{-1}$$

Relative Entropy

- Entropy production rates depend on the relative entropy between forward and backward trajectories

$$H(\mathbf{P},\mathbf{Q}) = \sum_{i\in X} P_i \log\left(\frac{P_i}{Q_i}\right)$$

$$e_p \overset{\text{def}}{=} \lim_{n\to+\infty} \frac{1}{n} H\left(\mathbf{P}|_{\mathcal{F}_0^n}, \mathbf{P}^-|_{\mathcal{F}_0^n}\right)$$

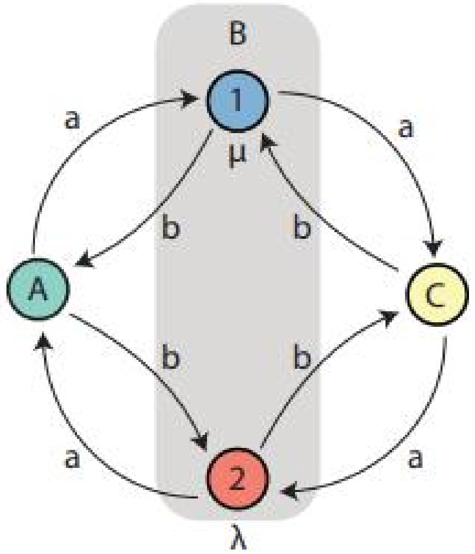
forward Markov chain

backward (time-reversed)
Markov chain

- Sometimes relative entropy appears to be zero!

An Example

Brownian clock on four microstates {1, 2, A, C}



States 1 and 2 are not distinguished by observation

$\mathbb{P}(\bar{Y}) = (1/4)a^{n-k}b^k$

- Consider the trajectory $X = (A, B, C, B, C, B)$

  The trajectories $Y = (A, 1, C, 2, C, 2)$ and $\bar{Y} = (A, 2, C, 1, C, 1)$ cannot be distinguished

  $\mathbb{P}(Y) = (1/4)a^k b^{n-k}$

- Sum over all *Y*

  Forward and backward trajectories are equally probable...

  Partial observation causes us to think the system is time-symmetric

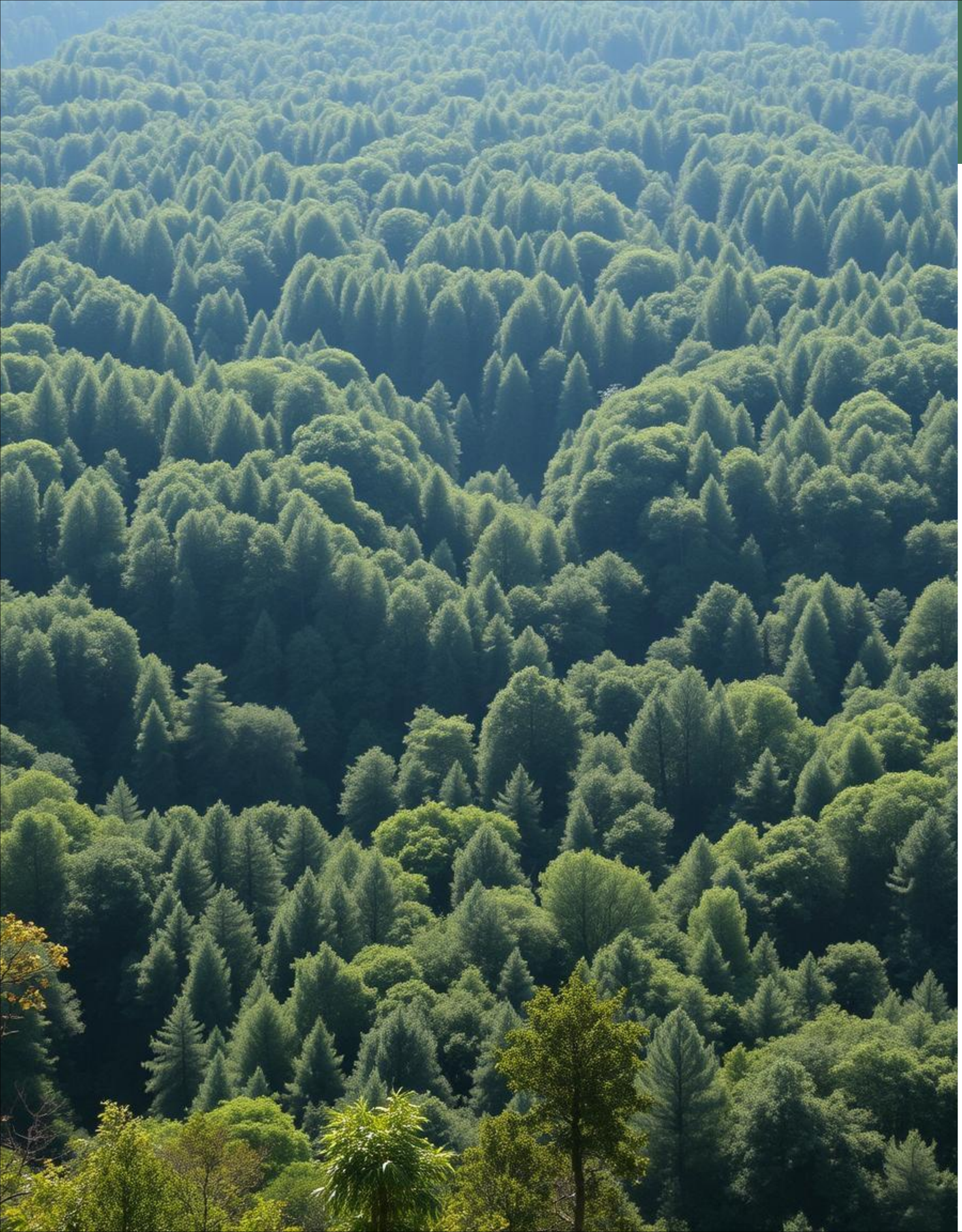
4 Concluding Remarks

Outlook

- Apply coarse graining to a small network ⇒ Larger networks
- Increased data resolution and trajectory length will lead to better bounds

  ⇒ imaging annd other experimental techniques

- Quantify tradeoffs between the faithful execution of a biological function and the actual energy expended



Conclusion

- Living systems expend entropy to maintain biological functions
- Ability to infer entropy production rates from partial experimental observations using optimization

  E.g. bacterial motors, microtubules, and calcium oscillations
- Tighter bounds on entropy production when systems appear time symmetric

References

D. J. Skinner and J. Dunkel.
Improved bounds on entropy production in living systems.
Proceedings of the National Academy of Sciences, 118(18):e2024300118, 2021.

Graphs and markov chains.
https://courses.grainger.illinois.edu/cs357/sp2021/notes/ ref-13-markov.html.
Accessed: February 11, 2025.

References

D. Jiang, M. Qian, and M. Qian.
Mathematical theory of nonequilibrium steady states: On the frontier of probability and dynamical systems.
2004.

D. Berry.
Cell organelles: Cytoskeleton (2014).
https://www.youtube.com/watch?v=tO-W8mvBa78.
Accessed: February 11, 2025.