Memory in Processes II

Reading for this lecture:

CMR article RURO.

Memory in Processes II ...

Classes of Excess Entropy:

Finitary process: $\mathbf{E} < \infty$

Exponential or finite-length convergence

Infinitary process: $\mathbf{E} \to \infty$

Notable examples:

Finitary, finite-state: ∞-order Markov (Even Process)

Finitary, infinite-state: Simple Nonunifilar Source

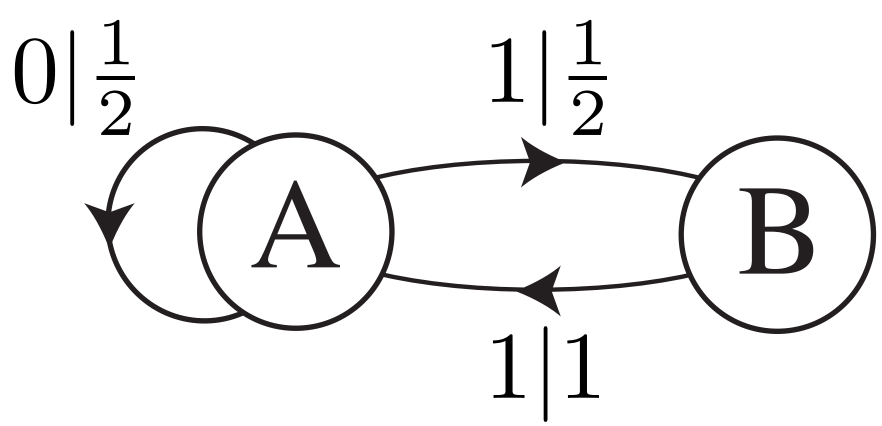
Infinitary, infinite-state: Topological complexity (Morse-Thue)

Memory in Processes II ...

Classes of Excess Entropy ...

Even Process: After pair of 1s, coin flip

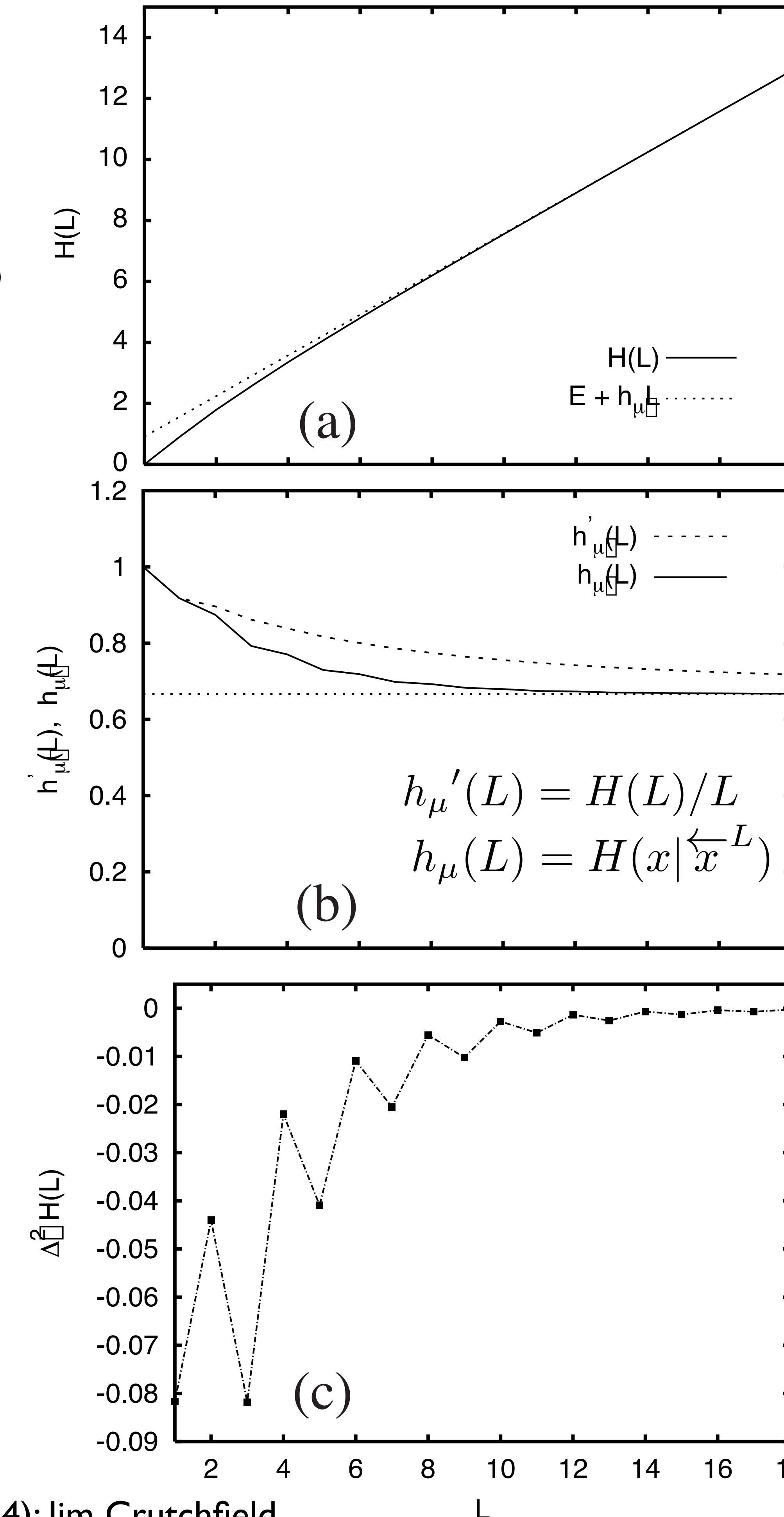
Presentation as Unifilar HMM



No finite-order Markov process exactly models the Even process.

But,

$$\mathbf{E} \approx 0.902 \text{ bits}$$



Memory in Processes II ...

Classes of Excess Entropy ...

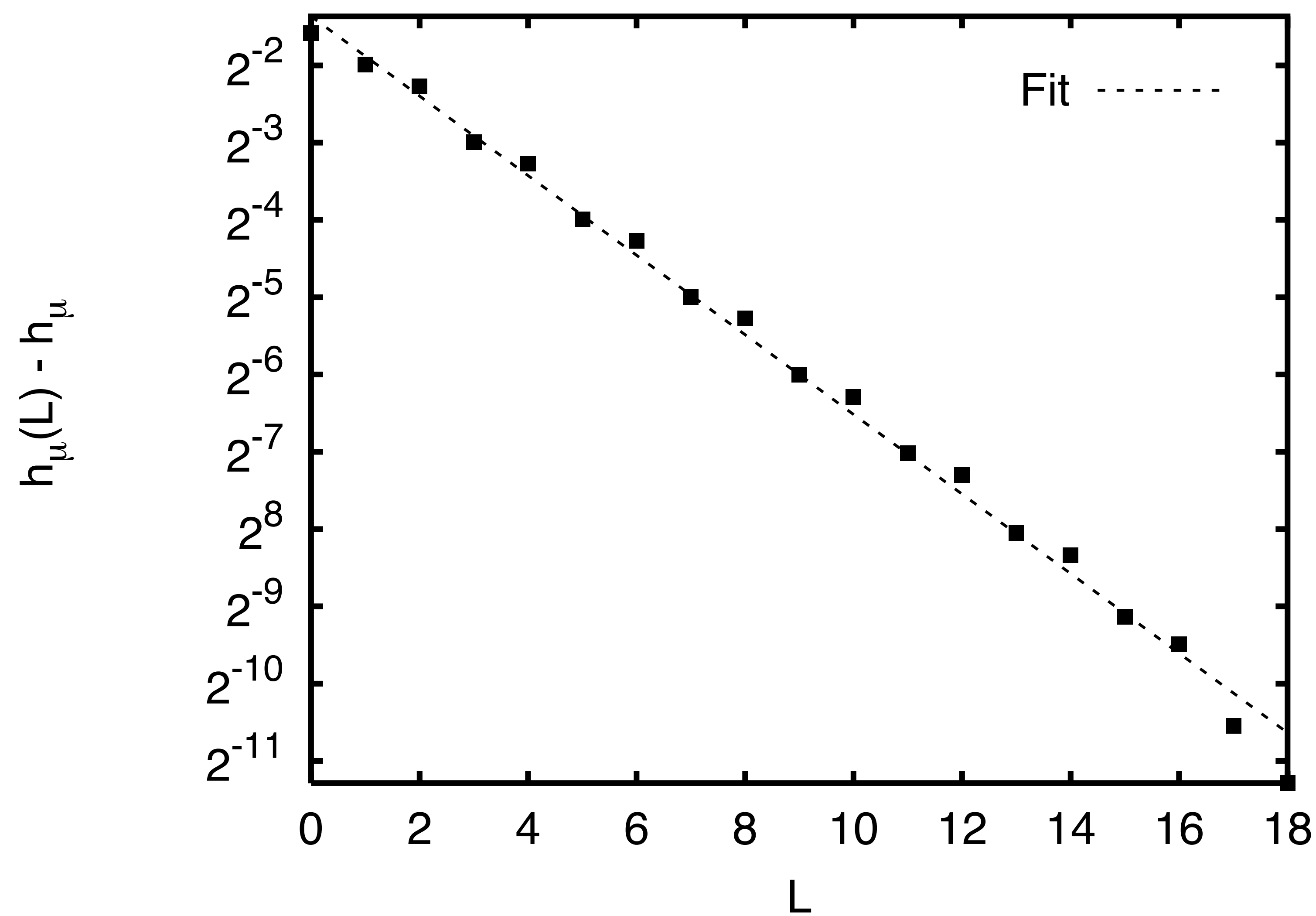
Even Process ...

∞-order Markov process.

But, still exponential entropy-rate decay:

$$h_\mu(L) - h_\mu \propto 2^{-\gamma L}$$

$$\gamma \approx \frac{1}{2}$$



Memory in Processes II ...

Classes of Excess Entropy ...

Morse-Thue Process:

Support is a context-free language
Generated by Logistic map at onset of chaos

Production rules:

$$\sigma(0) = 01$$

$$\sigma(1) = 10$$

For example:

$$\sigma^5(1) = 10010110011010010110100110010110$$

Aperiodic, infinite memory, predictable!

Memory in Processes II ...

Classes of Excess Entropy ...

Exact entropy-rate approximates:

$$h_\mu(1) = 1$$

$$h_\mu(2) = \log_2 3 - \tfrac{2}{3}$$

$$h_\mu(3) = \tfrac{2}{3}$$

$$h_\mu(L) = \begin{cases} 4/(3 \cdot 2^k), \\ \qquad \text{if } \; 2^k + 1 \leq L - 1 \leq 3 \cdot 2^{k-1} \\ 2/(3 \cdot 2^k), \\ \qquad \text{if } \; 3 \cdot 2^{k-1} + 1 \leq L - 1 \leq 2^{k+1} \end{cases}$$

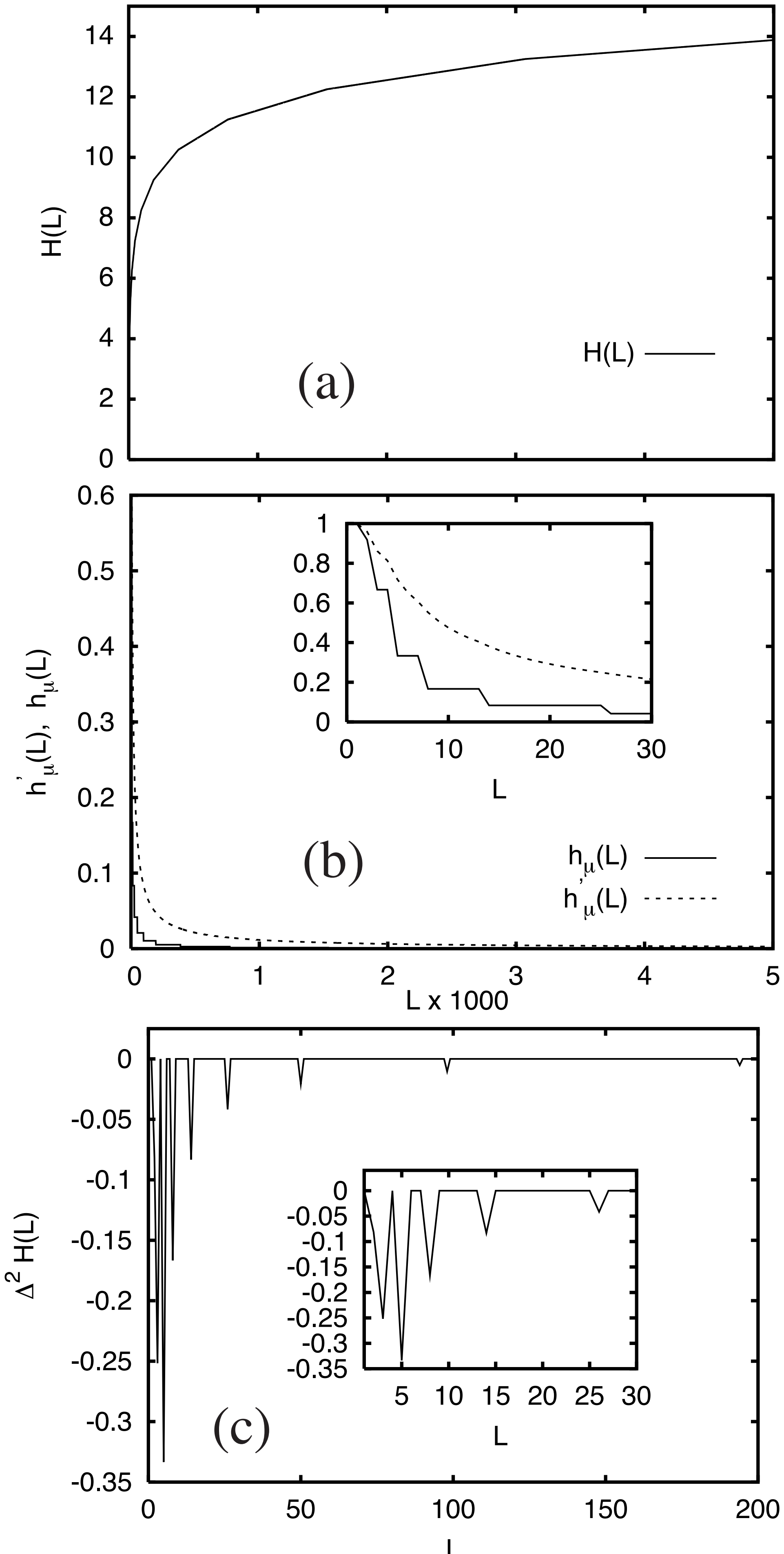
Slow entropy convergence (power-law):

$$h_\mu(L) \propto \frac{1}{L}$$

Entropy-rate vanishes:

$$h_\mu = 0 \text{ bits per symbol}$$

$$H(L) \propto \log_2(L)$$



Memory in Processes II ...

Classes of Excess Entropy ...

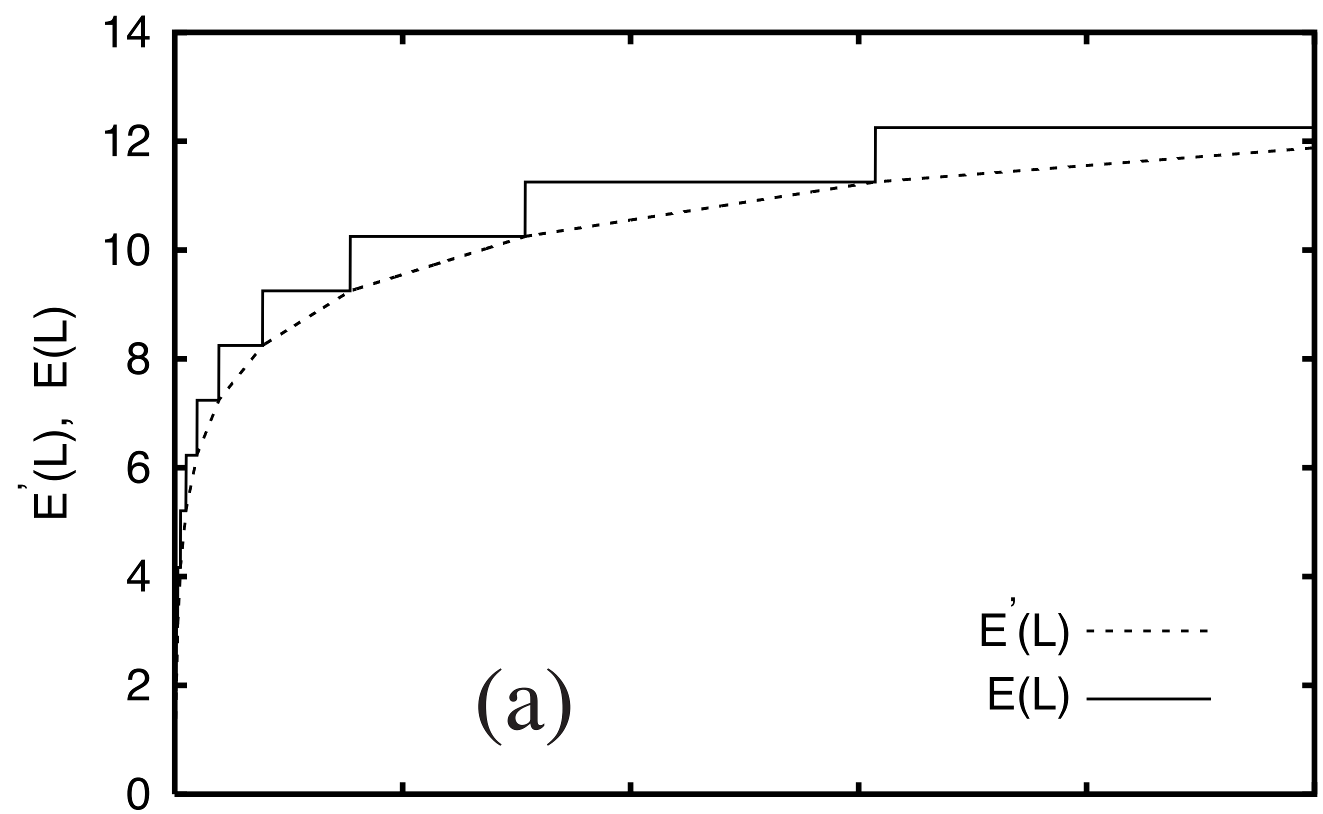
Excess entropy diverges:

Arbitrarily long-range correlations

(e.g., critical phenomena at phase transitions)

Infinitary Process!

$$\mathbf{E} \rightarrow \infty$$

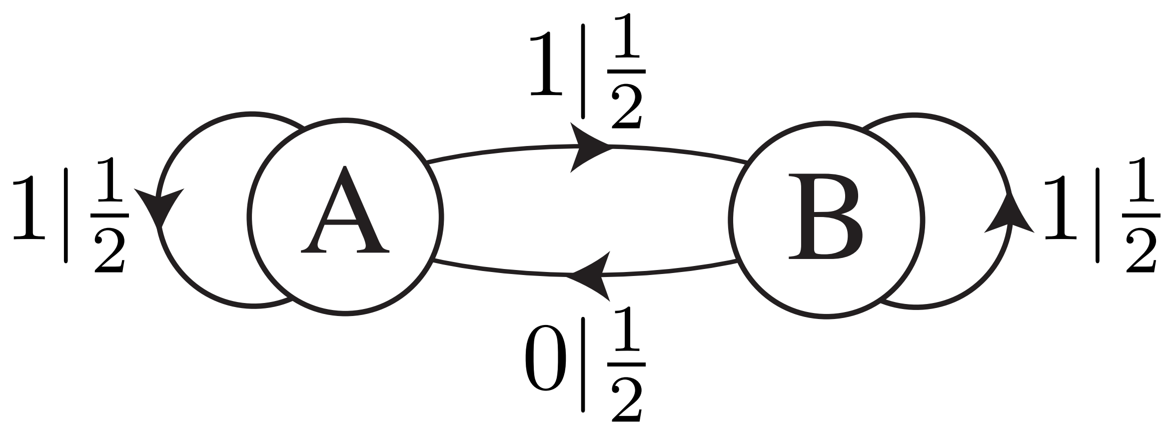


$$\mathbf{E}'(L) = I(\overleftarrow{X}^{L/2}; \overrightarrow{X}^{L/2})$$

$$\mathbf{E}(L) = H(L) - h_\mu L$$

Classes of Excess Entropy ...

Simple Nonunifilar Source:



What is its entropy rate?

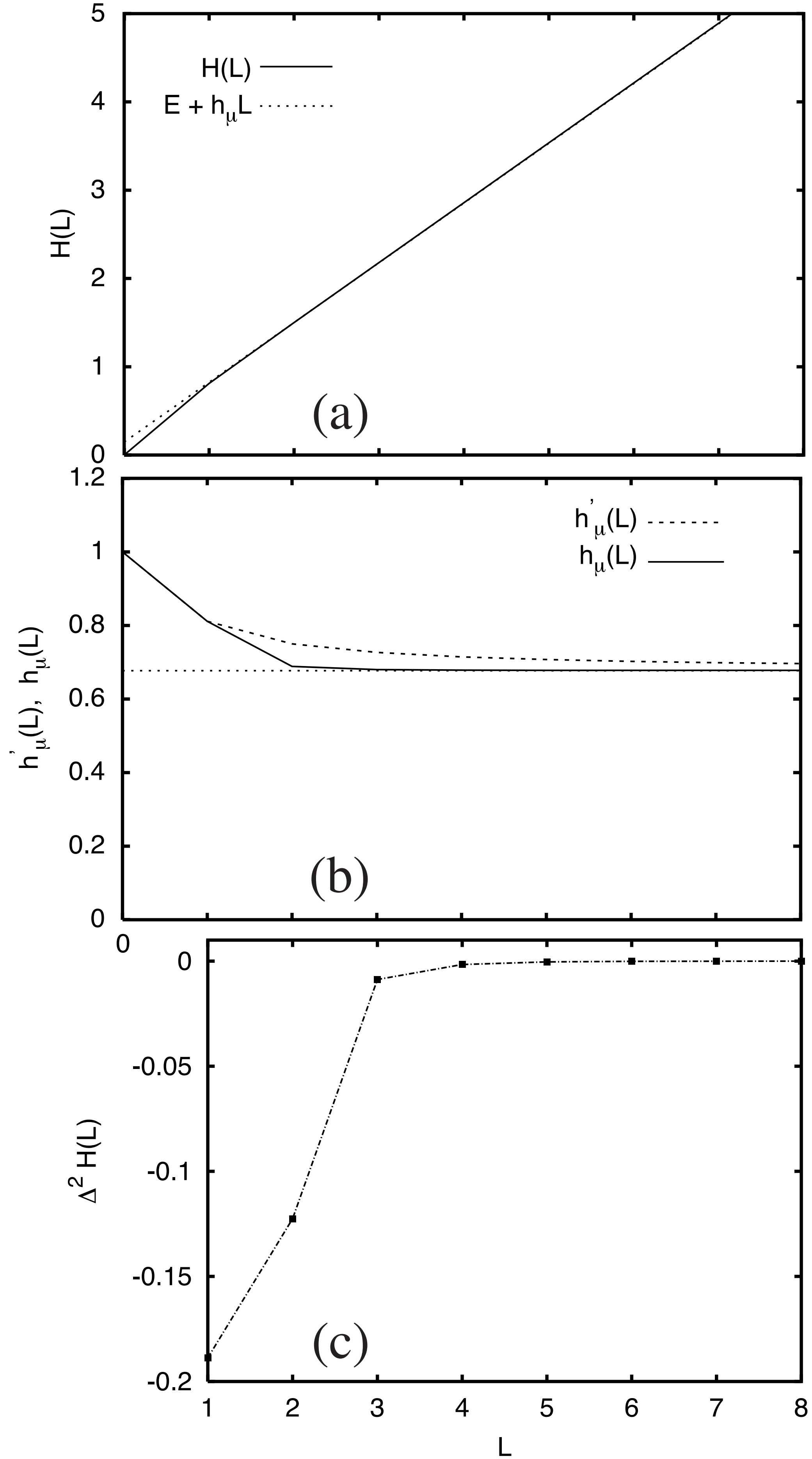
Recall: Cannot use nonunifilar presentation to answer.

Memory in Processes II ...

Classes of Excess Entropy ...

Simple Nonunifilar Source ...

Entropy curves



Memory in Processes II ...

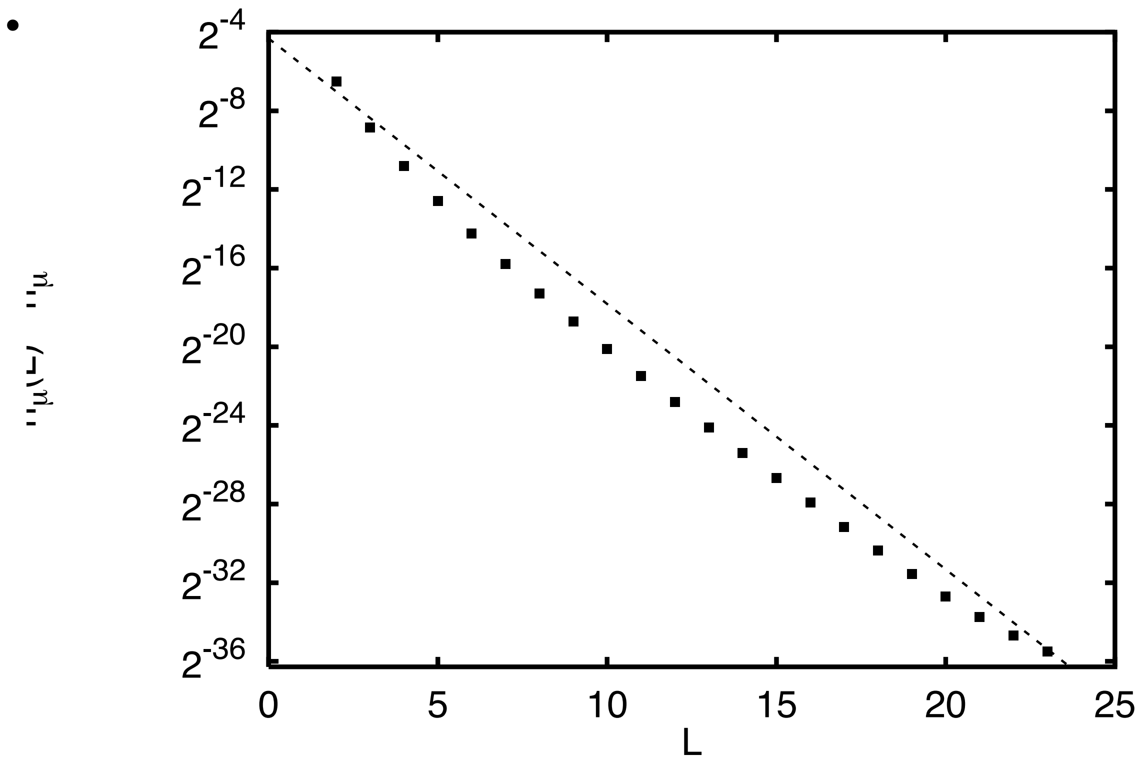
Classes of Excess Entropy ...

Simple Nonunifilar Source ...

∞-order Markov process.

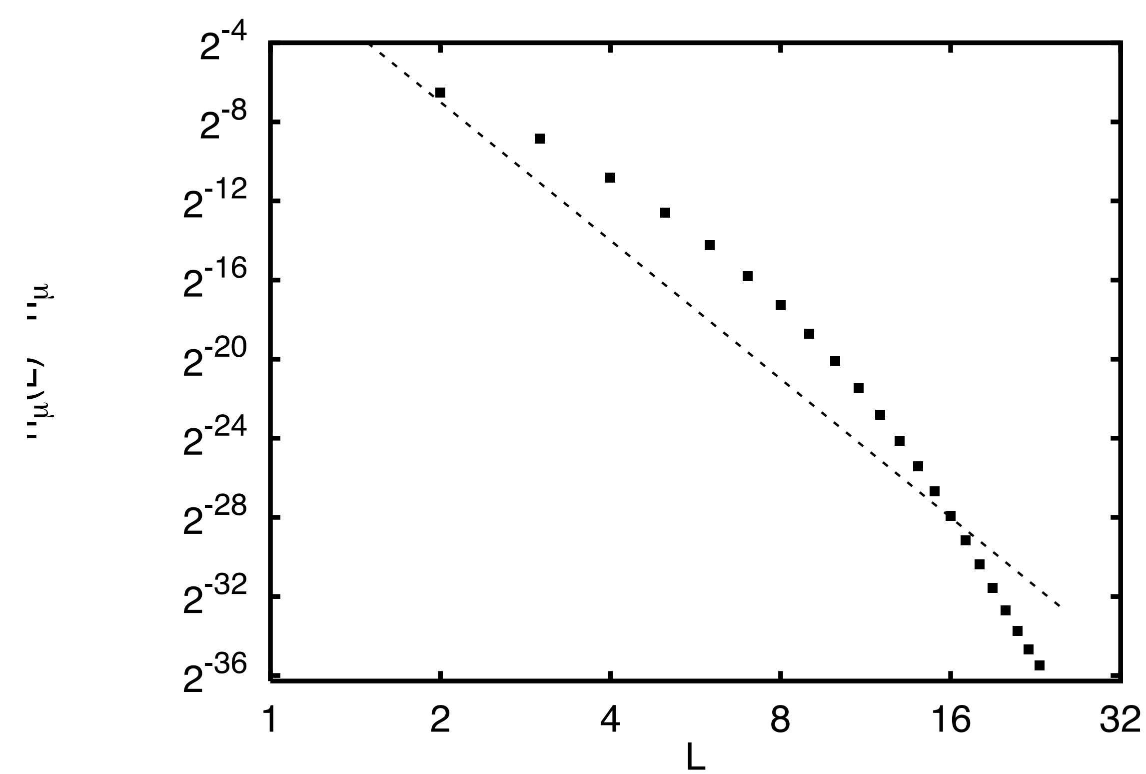
Neither exponential decay:

$$h_\mu(L) - h_\mu \propto 2^{-\gamma L}$$



Nor power-law decay:

$$h_\mu(L) - h_\mu \propto L^{\alpha}$$



Infinite state? "State"?

Memory in Processes ...

Synchronization:

Problem Statement:

You have a correct model of a process,
but you don't know it's current state.

Question: How much information
must you extract from measurements
to know which hidden state the process is in?

Memory in Processes ...

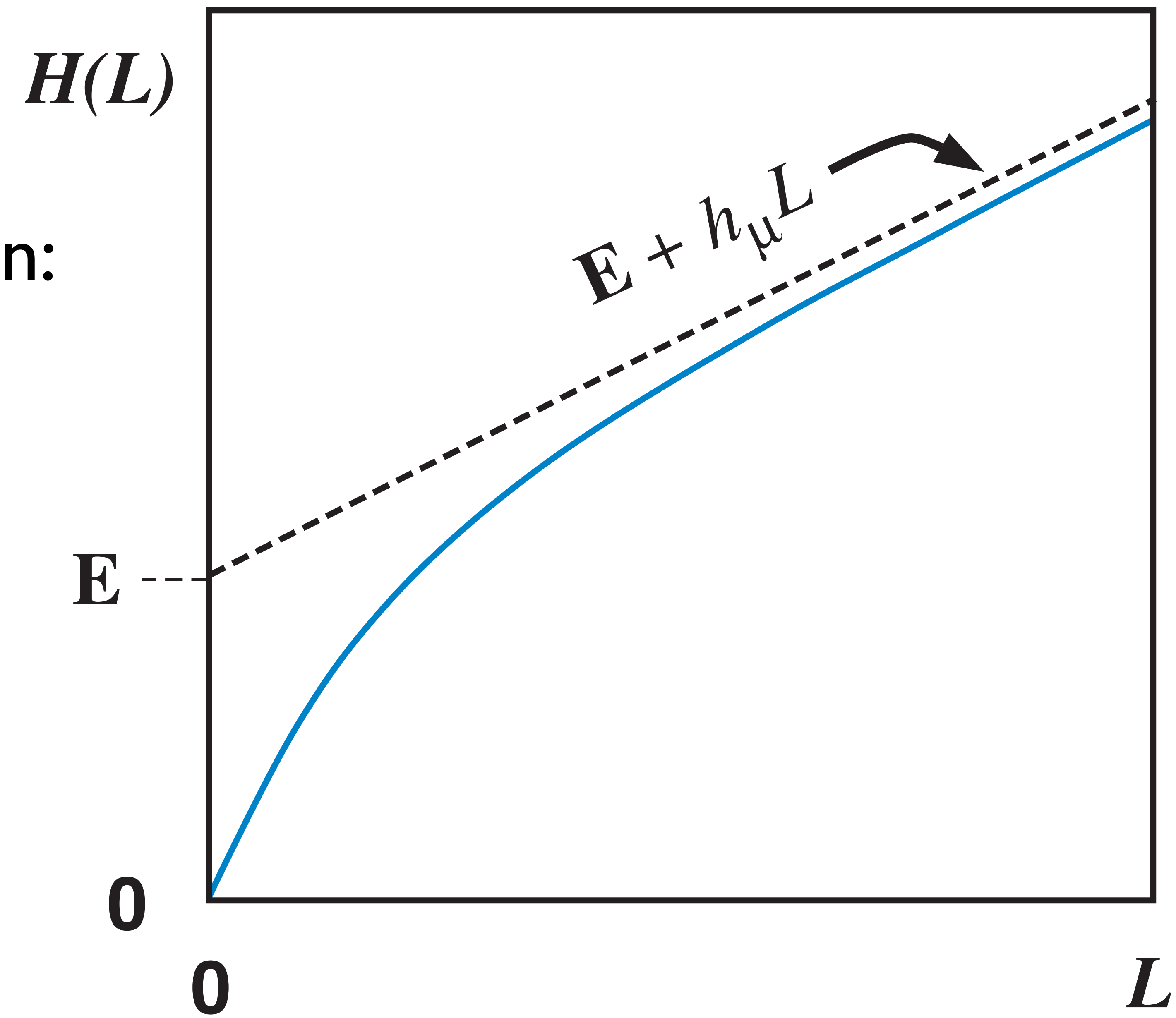
Transient Information:

Synchronized to source when:

$$L \geq L'$$

you have

$$H(L) \approx \mathbf{E} + h_\mu L$$



Synchronized:

At length L' at which you see true entropy rate.

Extracted sufficient information to do optimal prediction.

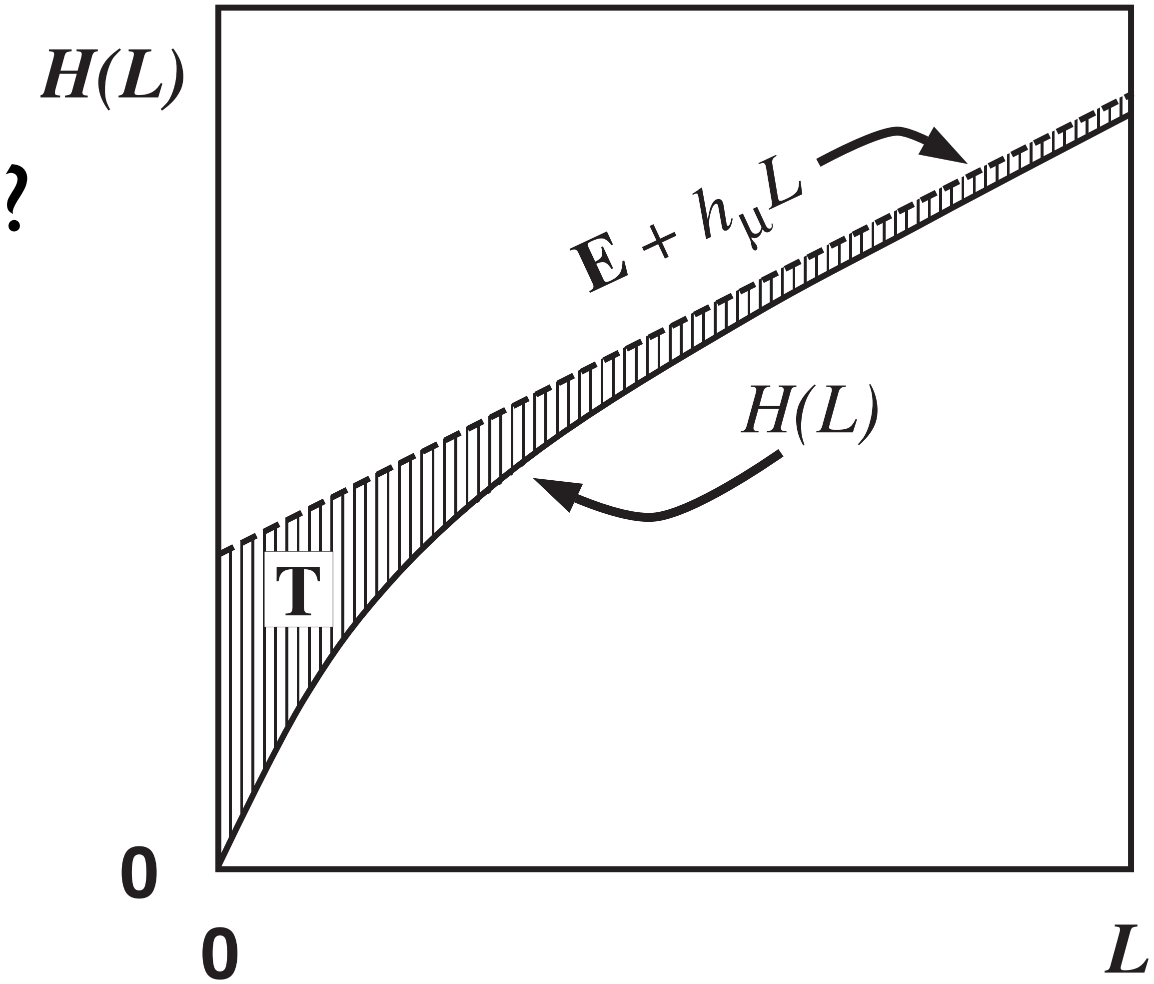
Memory in Processes ...

Transient Information ...

How much information to extract?

Transient Information:

$$\mathbf{T} = \sum_{L=0}^{\infty} [\mathbf{E} + h_\mu L - H(L)]$$



Controls convergence to synchronization.

Units: bits x symbols

Memory in Processes ...

Example of Transient Information:

Tahitian Vacation (3 days)!

- Weather has a 5 day cycle:
  - Two days of rain, followed by three of sun
- Weather is exactly predictable: $h_\mu = 0$ bits per day
- Weather has memory: $\mathbf{E} = \log_2 5$ bits

But,

- How to pack?
- What to pack?
- What to wear on trip?
- Dressed appropriately for arrival?

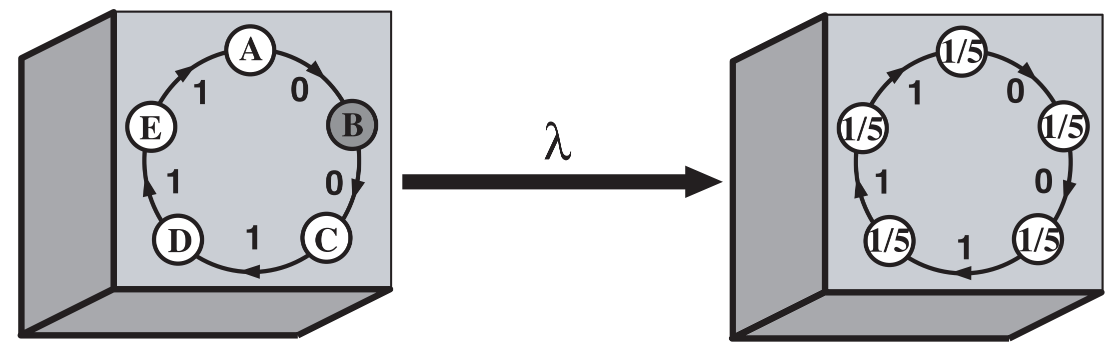
Memory in Processes ...

Example of Transient Information ...

Tahitian Vacation ... packing

0 = Rain
1 = Sun

No weather
reports yet.



Tahiti

Weather
Reports

Update
Traveler's
Model

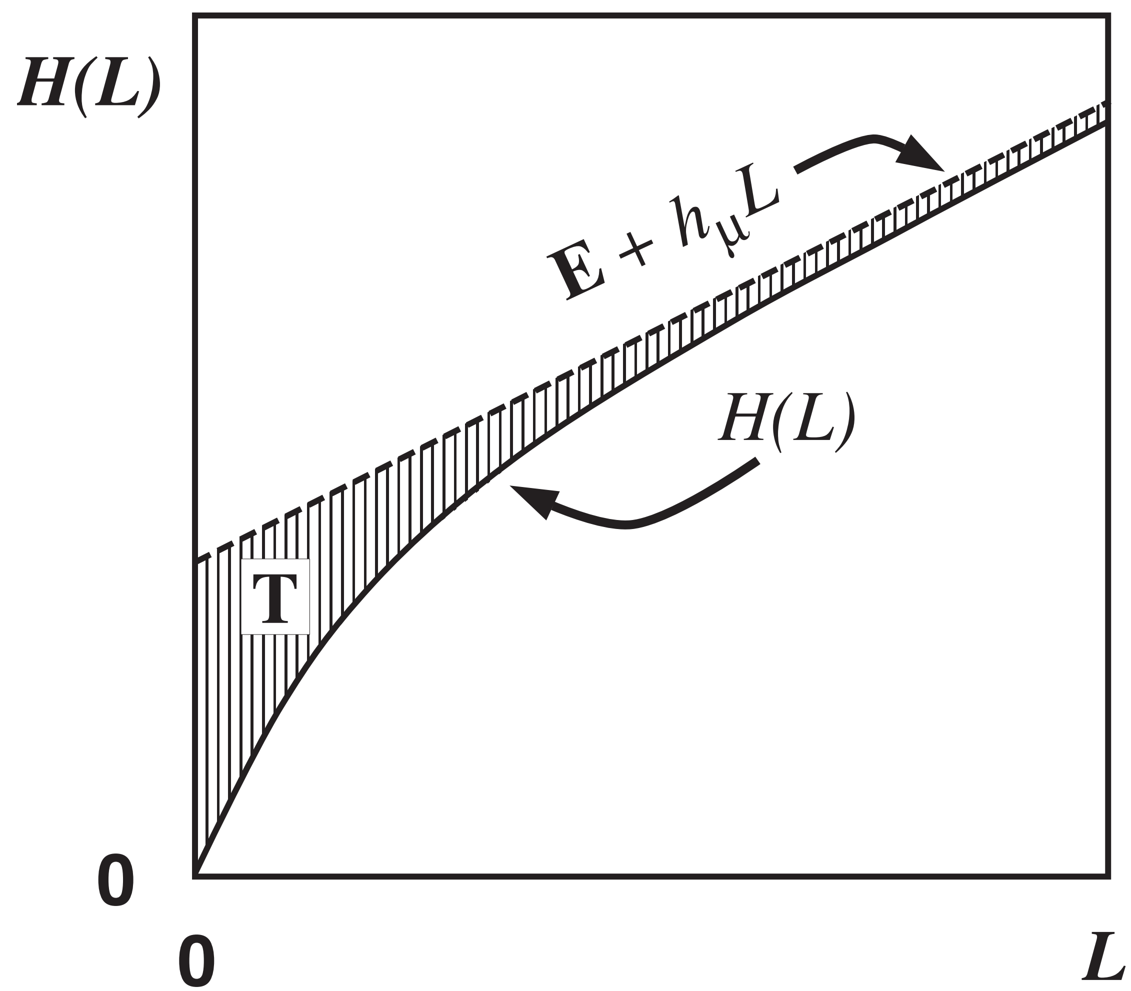
Pack umbrella,
wear shorts on plane

$$\mathbf{T} \approx 4.073 \text{ bit} \times \text{symbols}$$

Memory in Processes ...

Transient Information ...

How to interpret?



Memory in Processes ...

Transient Information ...

Synchronization information:

Observer has correct model of a Markov chain: $\mathcal{M} = \{V, T\}$

Observer Synchronized to Process:

$$\mathbf{T}(L) \equiv \mathbf{E} + h_\mu L - H(L) = 0$$

Observer knows with certainty in which state the process is:

$$\Pr(v_0, v_1, \ldots, v_k) = (0, \ldots, 1, \ldots, 0)$$

Average per-symbol uncertainty is exactly h_μ.

Memory in Processes ...

Transient Information ...

Synchronization information ...

Average state-uncertainty:

$$\mathcal{H}(L) \equiv - \sum_{s^L \in \mathcal{A}^L} \Pr(s^L) \sum_{v \in \mathcal{V}} \Pr(v|s^L) \log_2 \Pr(v|s^L)$$

Synchronization information:

$$\mathbf{S} \equiv \sum_{L=0}^{\infty} \mathcal{H}(L)$$

Memory in Processes ...

Transient Information ...

Synchronization information ...

Theorem: For a R-block (spin-block) process, the synchronization information is given by:

$$\mathbf{S} = \mathbf{T} + \frac{1}{2} R(R+1) h_\mu$$

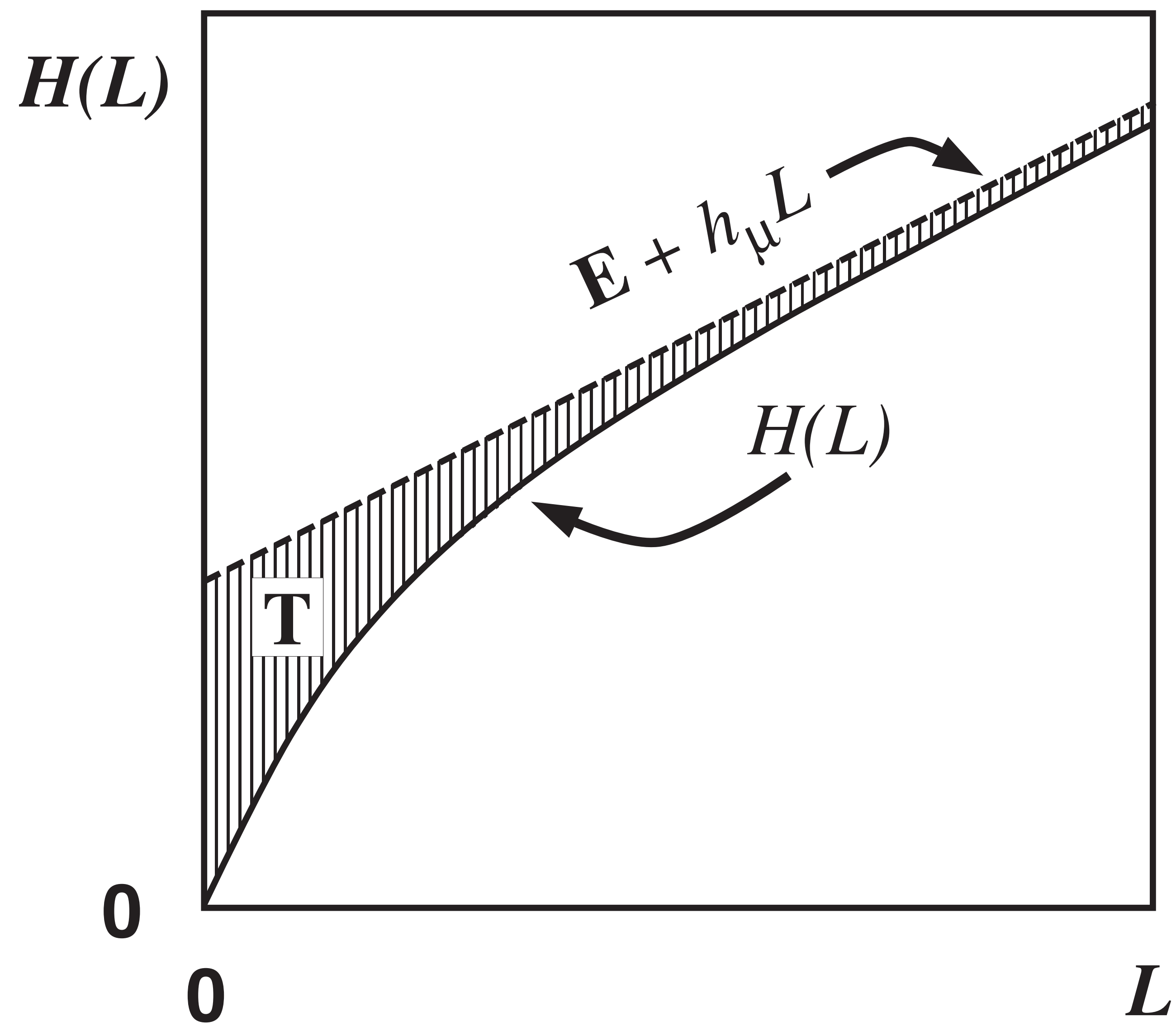
Corollary: For periodic process:

$$\mathbf{S} = \mathbf{T}$$

Memory in Processes ...

Transient Information ...

How to interpret?



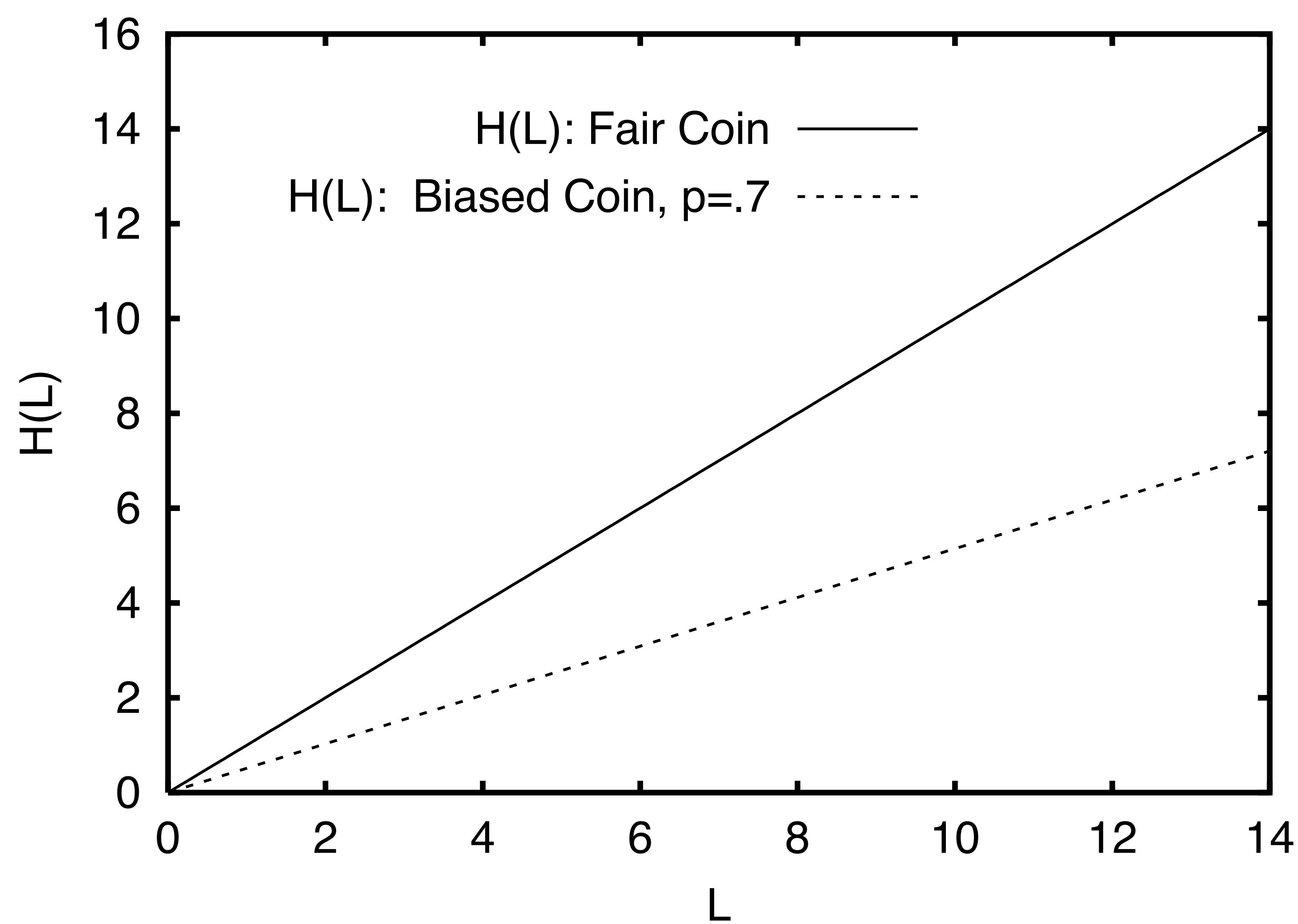
1. Total uncertainty observed while synchronizing.

2. Information to extract to be synchronized.

Memory in Processes ...

Examples of Transient Information:

Fair & Biased Coins
& IID Processes: $\mathbf{T} = 0$



Memory in Processes ...
Examples of Transient Information ...

Period-5 Processes:

There are three distinct:

$$(11000)^\infty$$
$$(10101)^\infty$$
$$(10000)^\infty$$

All:

Predictable: $h_\mu = 0 \text{ bits}$

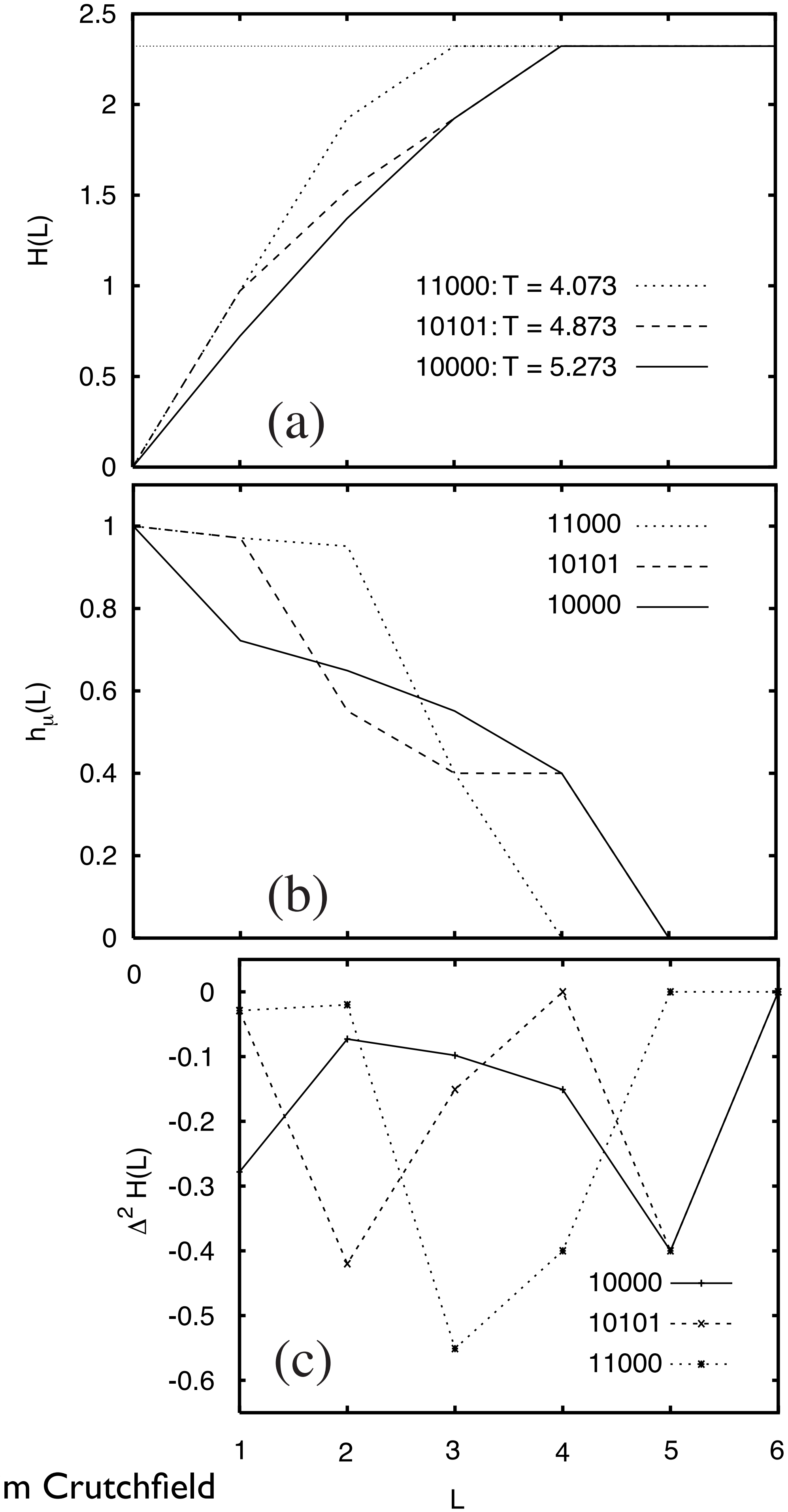
Memory: $\mathbf{E} = \log_2 5 \text{ bits}$

Memory in Processes ...

Examples of Transient Information ...

Period-5 Processes ...

But different ways to sync:



Memory in Processes ...

Examples of Transient Information ...

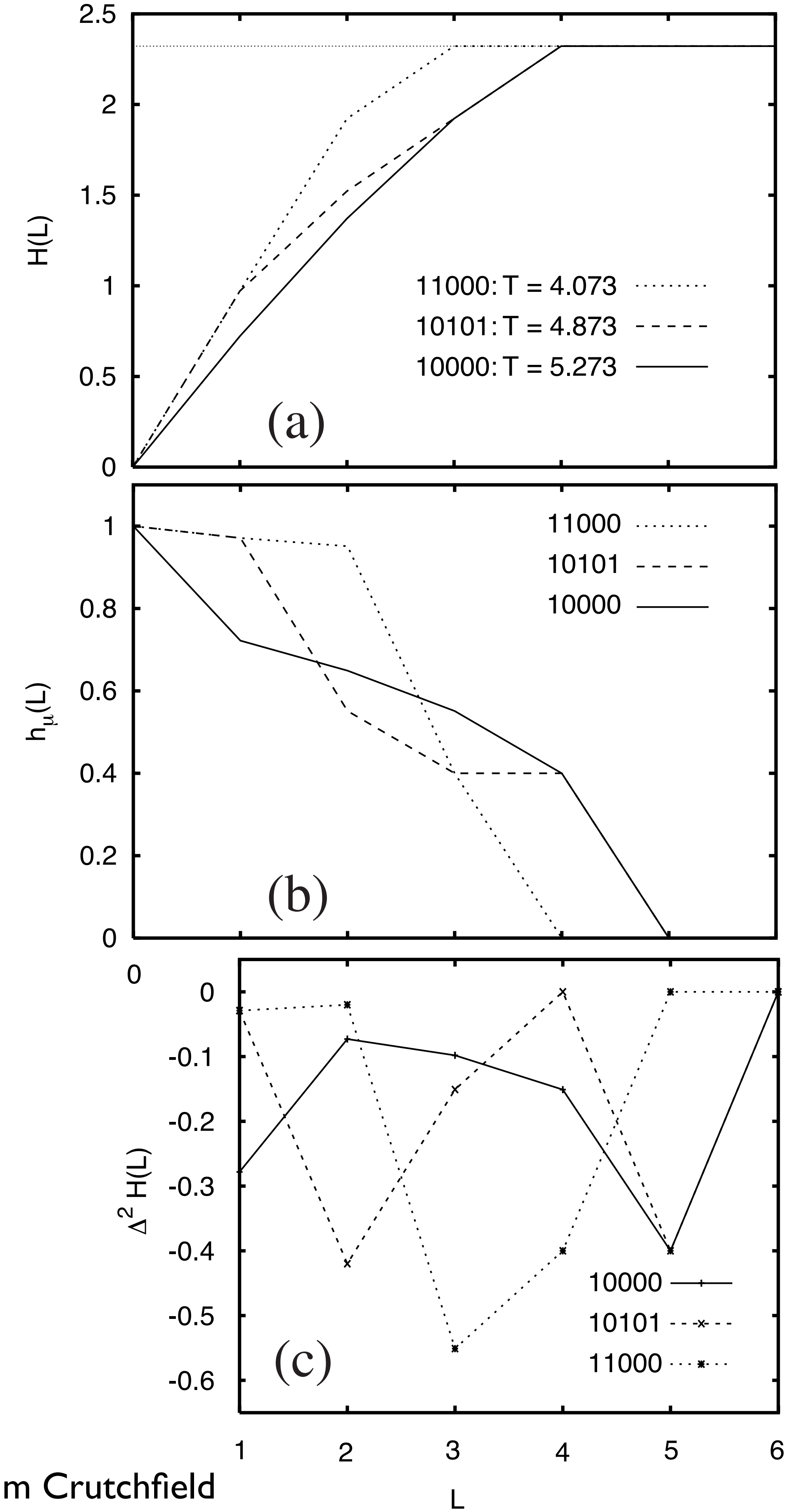
Period-5 Processes ...

But different ways to sync:

$(11000)^\infty$ $\mathbf{T} \approx 4.073$ bit $\times$ symbols

$(10101)^\infty$ $\mathbf{T} \approx 4.873$ bit $\times$ symbols

$(10000)^\infty$ $\mathbf{T} \approx 5.273$ bit $\times$ symbols



Memory in Processes ...
Examples of Transient Information ...

Period-P Processes:

Entropy rate vanishes.

Excess entropy same for all.

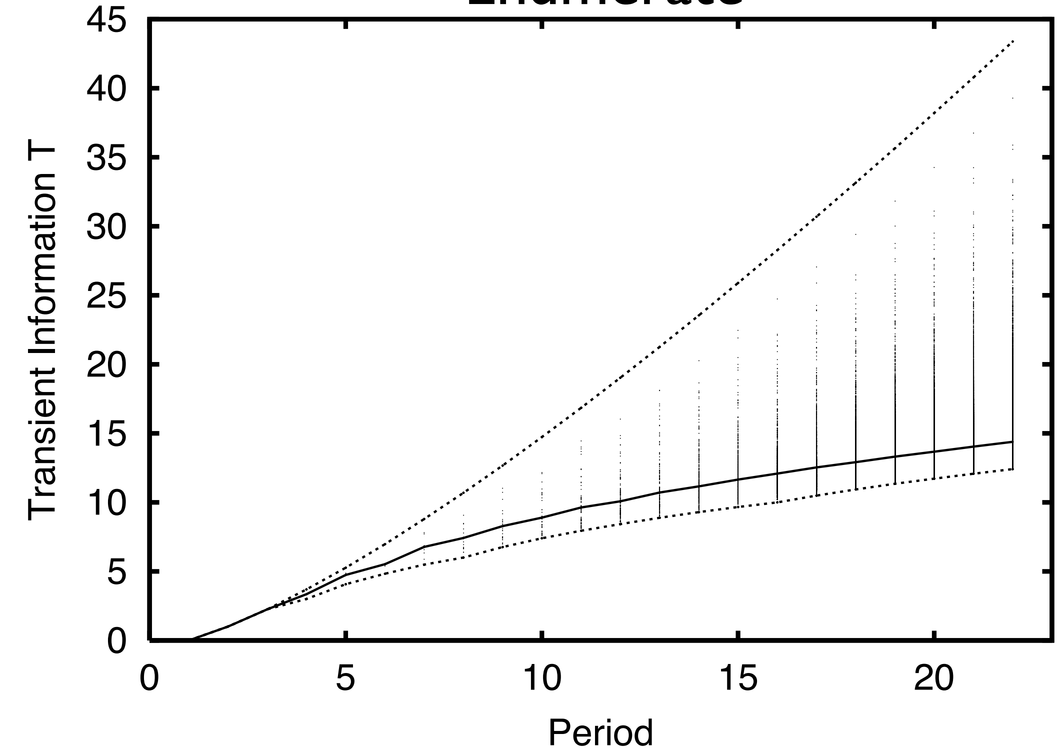
But T distinguishes periodic processes.

Memory in Processes ...

Transient Information For Periodic Processes

Period: P

Enumerate



Max T:

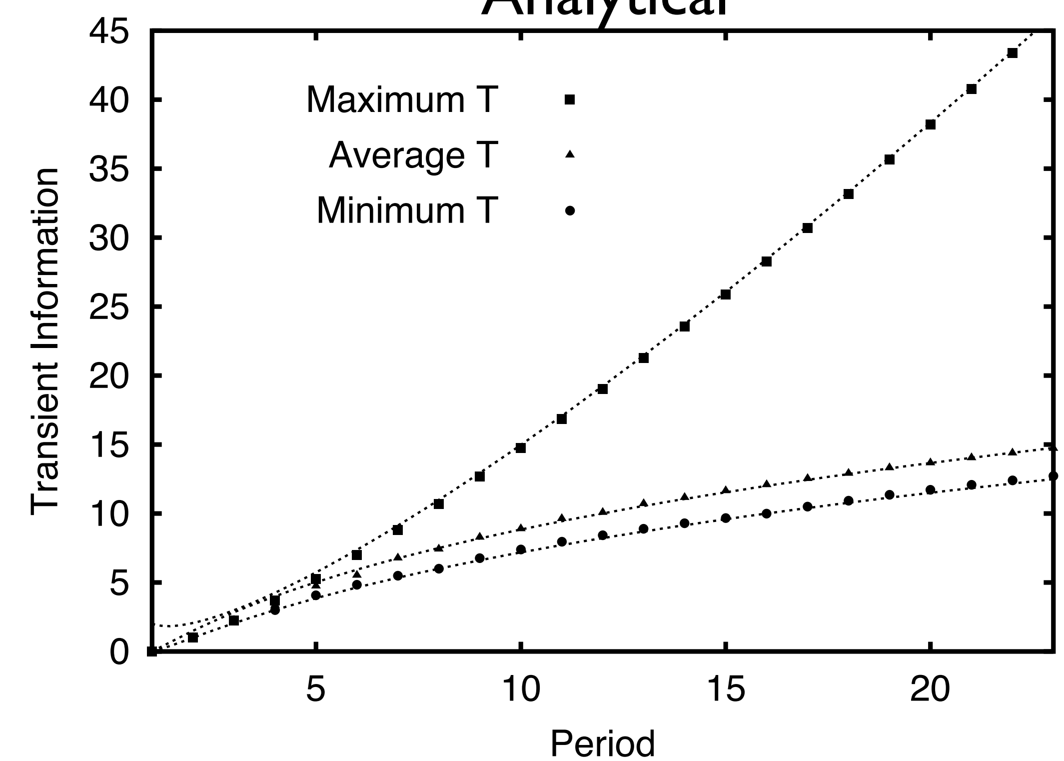
Slow convergence

Most nonuniform word dist.

P−1 0's followed by isolated 1

$$\mathbf{T}_{\max} \approx \frac{1}{2} P \log_2 P$$

Analytical



Min T:

Fast convergence

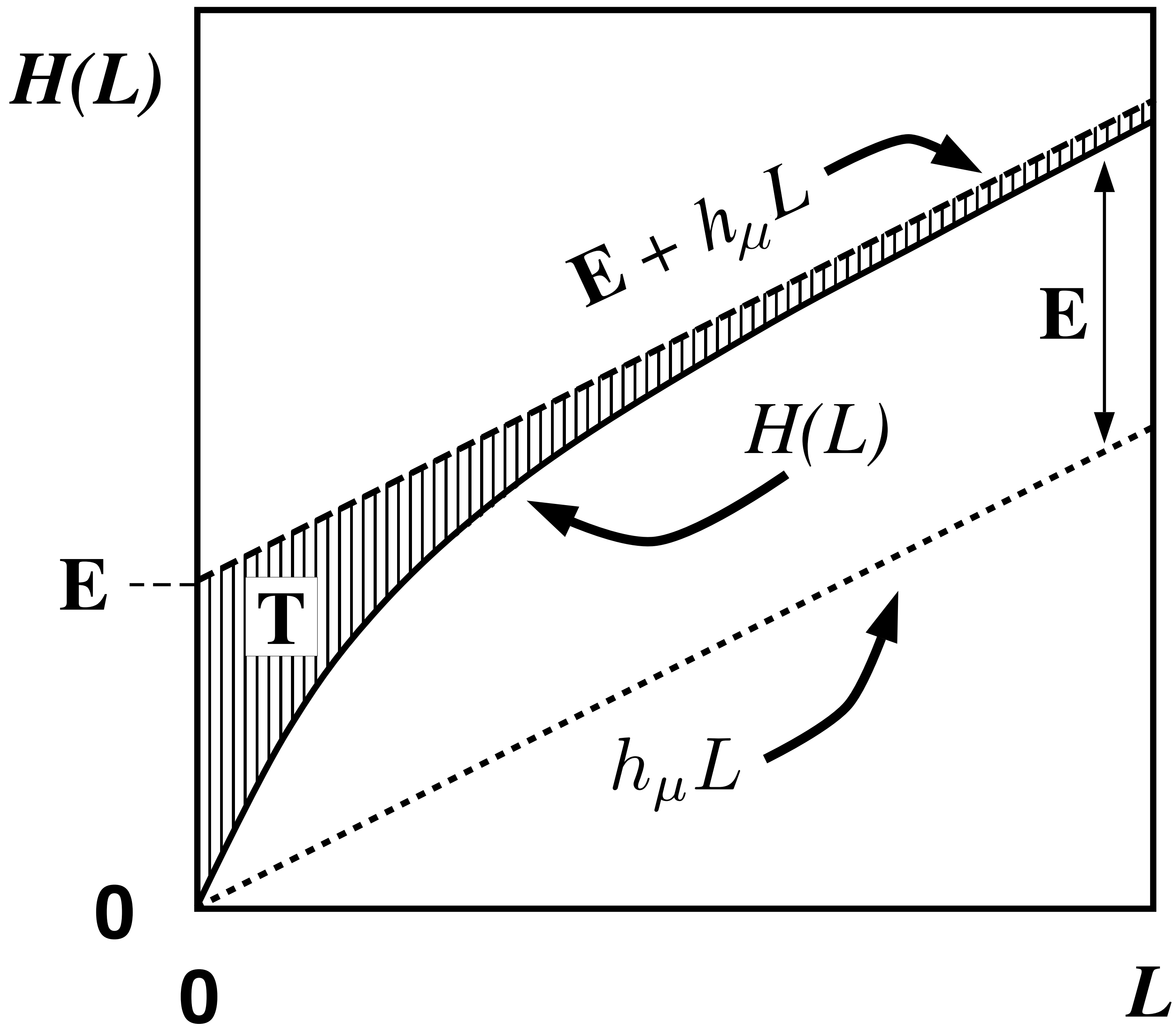
Flattest word distribution

P = 2^L: deBruin sequences

$$\mathbf{T}_{\min} \approx \frac{1}{2} \left(\log_2^2 P + \log_2 P \right)$$

Memory in Processes ...

Information-Entropy Roadmap for a Stochastic Process:



Memory in Processes ...

Regularities Unseen, Randomness Observed:

- Untangle distinct sources of apparent randomness?
- Estimates of entropy rate if ignore a process's structure?

Consequences:

- When an observer ignores entropy-rate convergence?
- When the process's apparent memory is ignored?
- If the observer ignores synchronization?
- If the observer assumes it is synchronized?

Memory in Processes ...

Disorder is the Price of Ignorance:

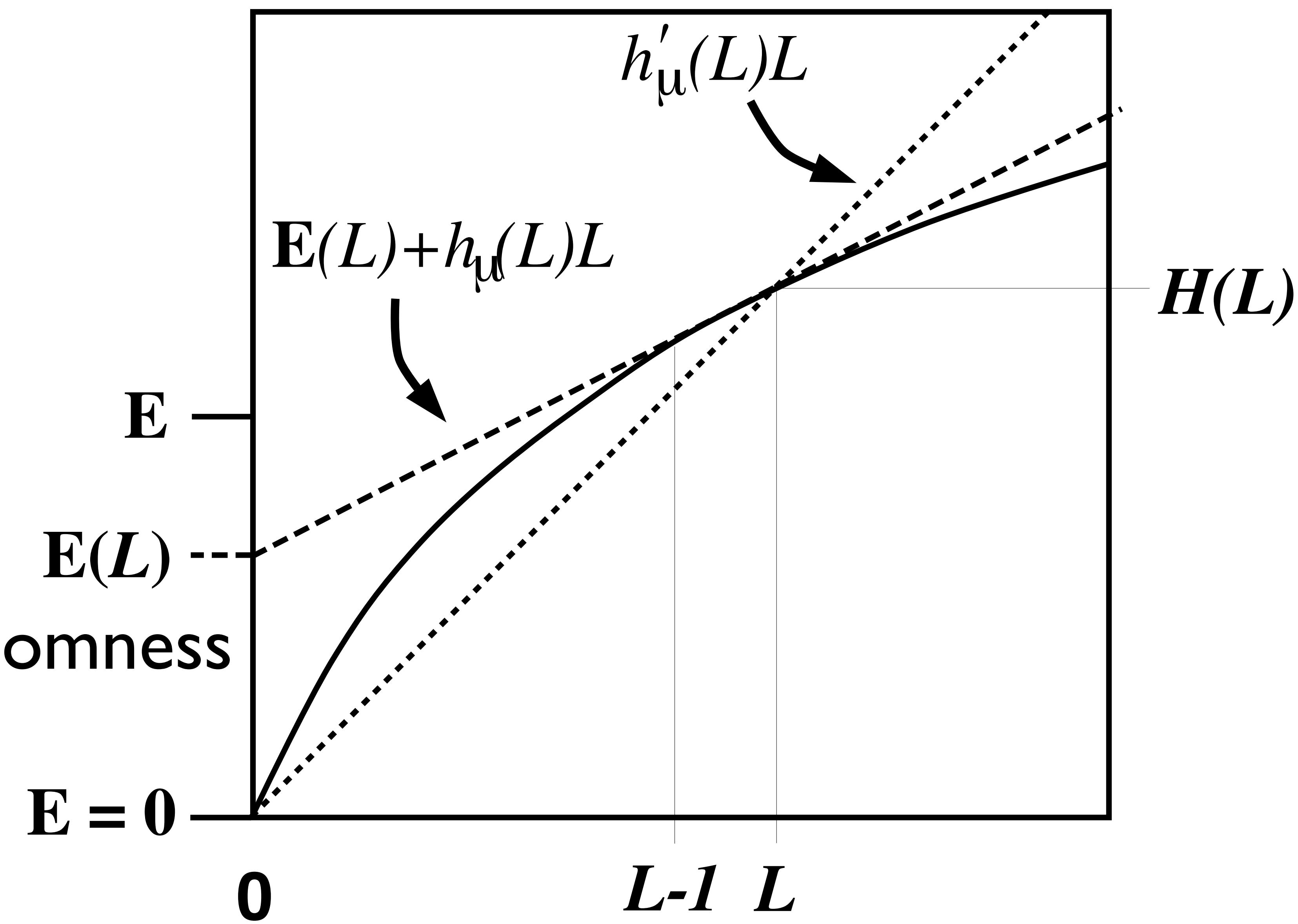
Ignore process's memory

By assuming

$$\mathbf{E} = 0$$

Over-estimate true randomness

$$h_\mu{}' > h_\mu$$



Lesson:

Structure $(\mathbf{E}\ \&\ \mathbf{T})$ converted to apparent randomness (h_μ).

Memory in Processes ...

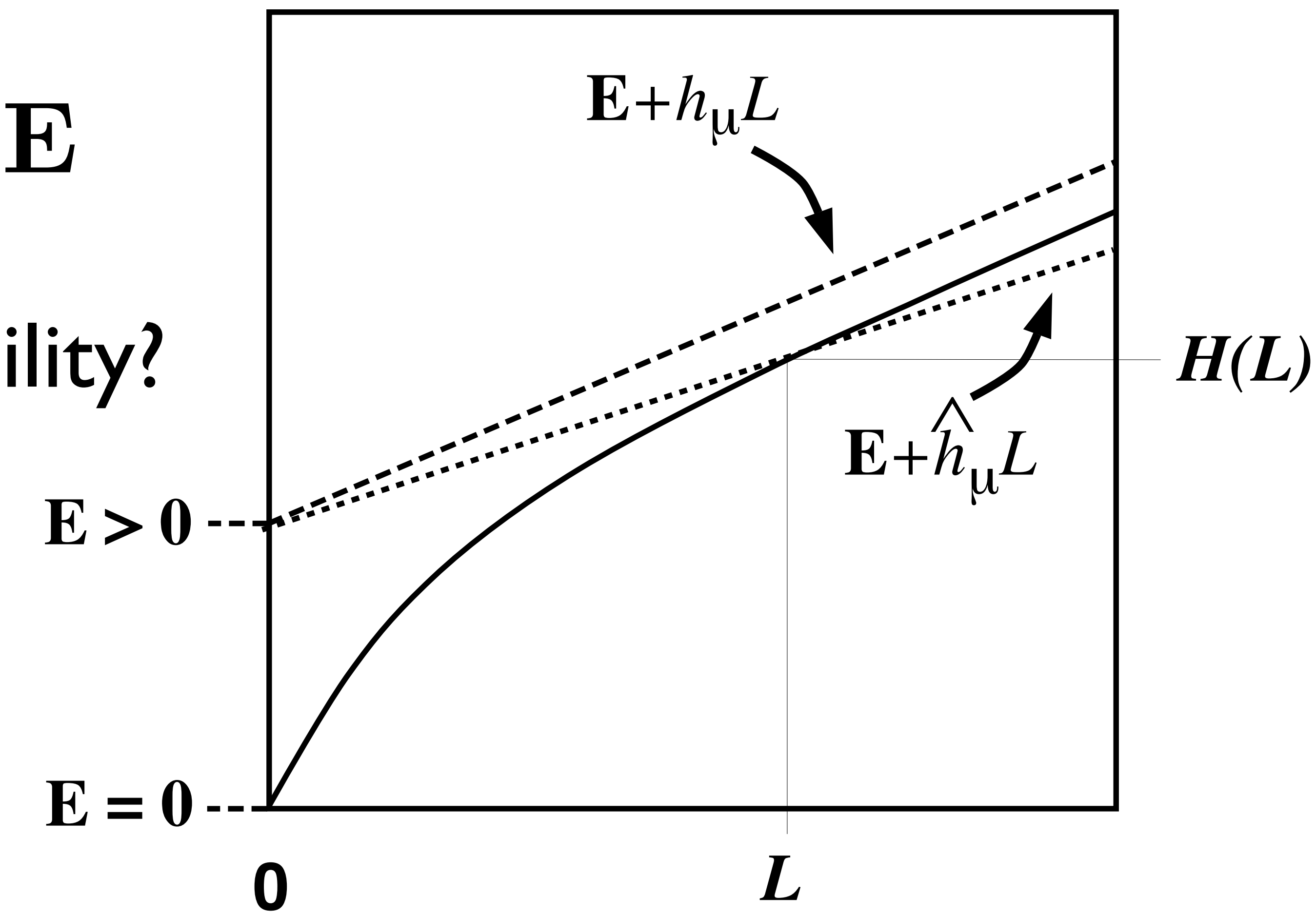
Predictability and Instantaneous Synchronization:

Instant Sync:

Assume you know memory $\mathbf{E}$

Your estimate $\widehat{h_\mu}$ of unpredictability?

$$\widehat{h_\mu} < h_\mu$$



Lesson:

Assumed synchronization converted to false predictability.

Memory in Processes ...

Assumed Synchronization Implies Reduced Apparent Memory:

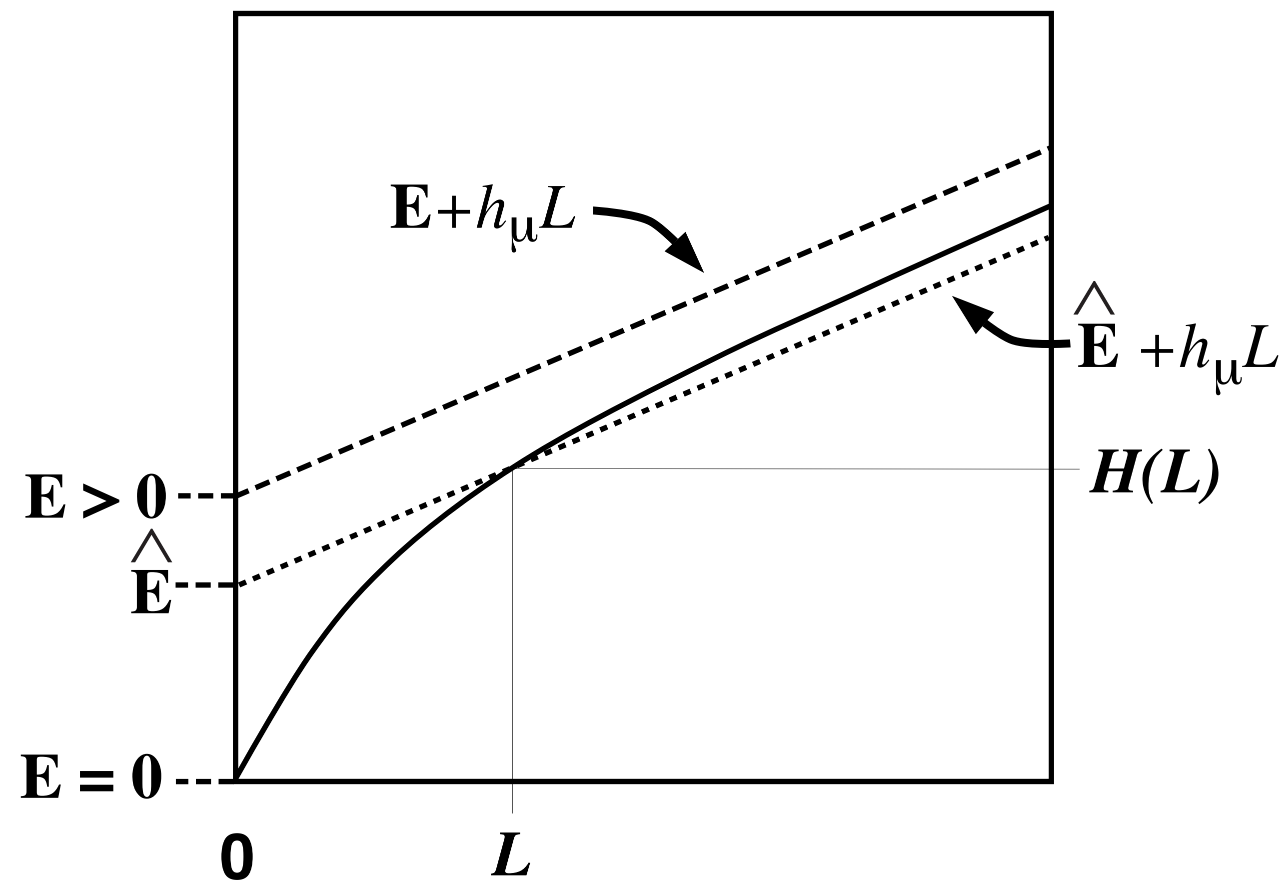
Assume you're sync'd:

$$H(L) = \mathbf{E} + h_\mu L$$

$$h_\mu(L) = h_\mu$$

Estimate of memory?

$$\widehat{\mathbf{E}} < \mathbf{E}$$



Lesson:

The world appears less structured.

Memory in Processes ...

Calculus of the Entropy Hierarchy:

Via Discrete-Time Derivatives and Integrals

Level	Gain (Derivative)	Information (Integral)
0	Block Entropy $H(L)$	Transient Information $\mathbf{T} = \sum_{L=0}^{\infty} [\mathbf{E} + h_\mu L - H(L)]$
1	Entropy Rate Loss $h_\mu(L) = \Delta H(L)$	Excess Entropy $\mathbf{E} = \sum_{L=1}^{\infty} [h_\mu(L) - h_\mu]$
2	Predictability Gain $\Delta^2 H(L)$	Total Predictability (Redundancy) $\mathbf{G} = -\mathcal{R}$
...	...	...

Memory in Processes ...

Reading for next lecture:

Yeung and *Anatomy* in CMech Reader.