From Determinism to Stochasticity Measurement Theory II

Jim Crutchfield & Ryan James
Complexity Sciences Center
Physics Department
University of California at Davis

Complex Systems Summer School
Santa Fe Institute
St. Johns' College, Santa Fe, NM
9-24 June 2011

From Determinism to Stochasticity
Measurement Theory II

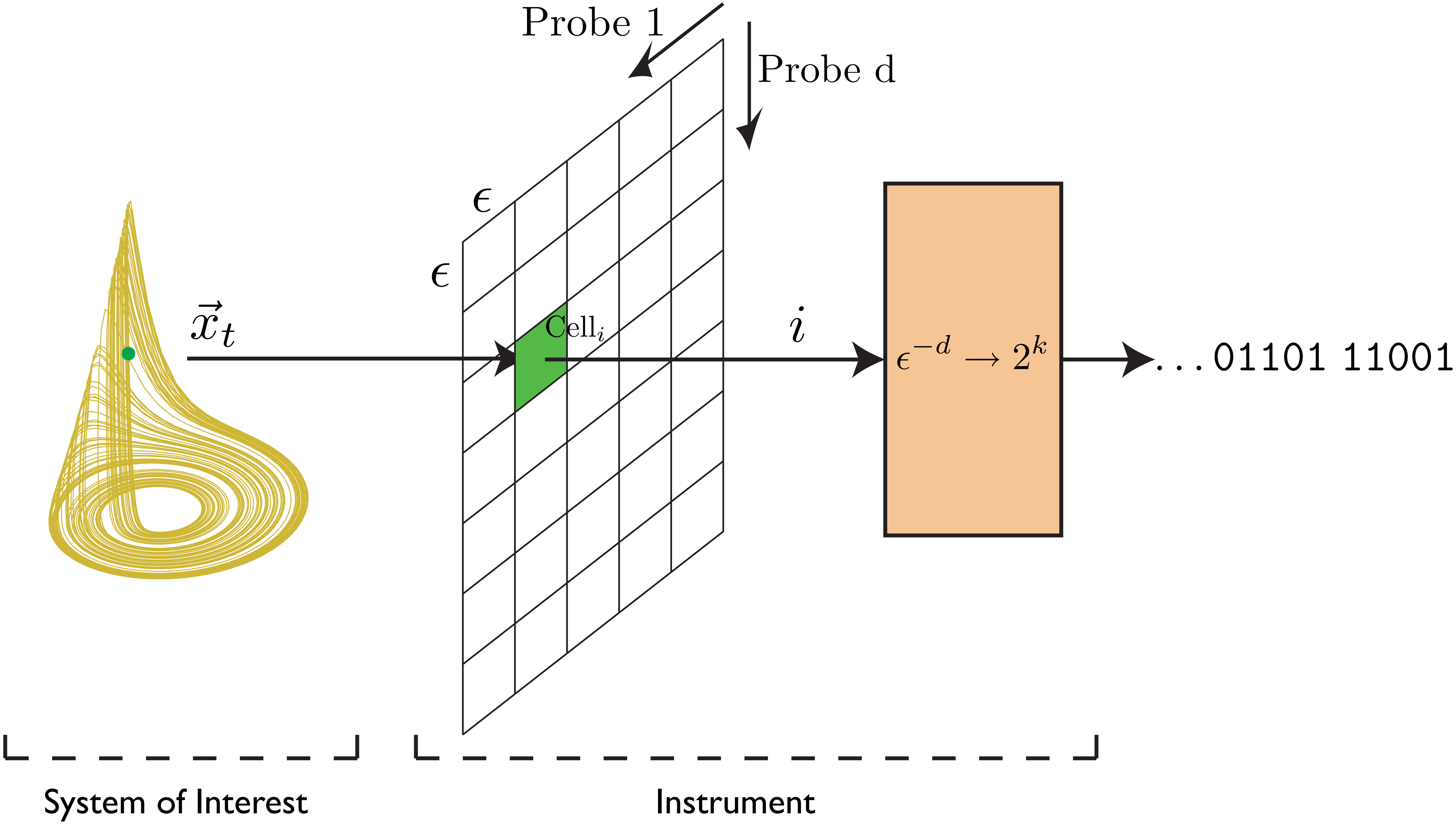
Previous Dynamics Lecture:

When measurements are faithful
study infinite discrete sequences
to learn about continuous-state dynamical system.

Now:

1. Most general faithful instrument
2. When measurements mislead
3. General stochastic processes

From Determinism to Stochasticity ...



Measurement Channel

From Determinism to Stochasticity ...

Measurement Theory ...

Markov partition very stringent.

Generating partitions:

Most general idea of "good" instrument.

Easier to find than Markov partitions, which may not exist.

Only requirement: Sequences track individual orbits

$$||\Delta(s^L)|| \rightarrow 0, \ \forall s^L \in \Sigma_f$$

Requires chaos:

Instability translates into reverse-time shrinking of cells.

From Determinism to Stochasticity ...

Measurement Theory ...

Some facts about partitions:

Markov partitions are generating.

Markov partition reduces Perron-Frobenius operator to
finite-dimensional stochastic transition matrix.

Resulting process is modeled by finite, probabilistic Markov chain.

Generating partitions may lead to
finite- or infinite-dimensional
hidden Markov models.

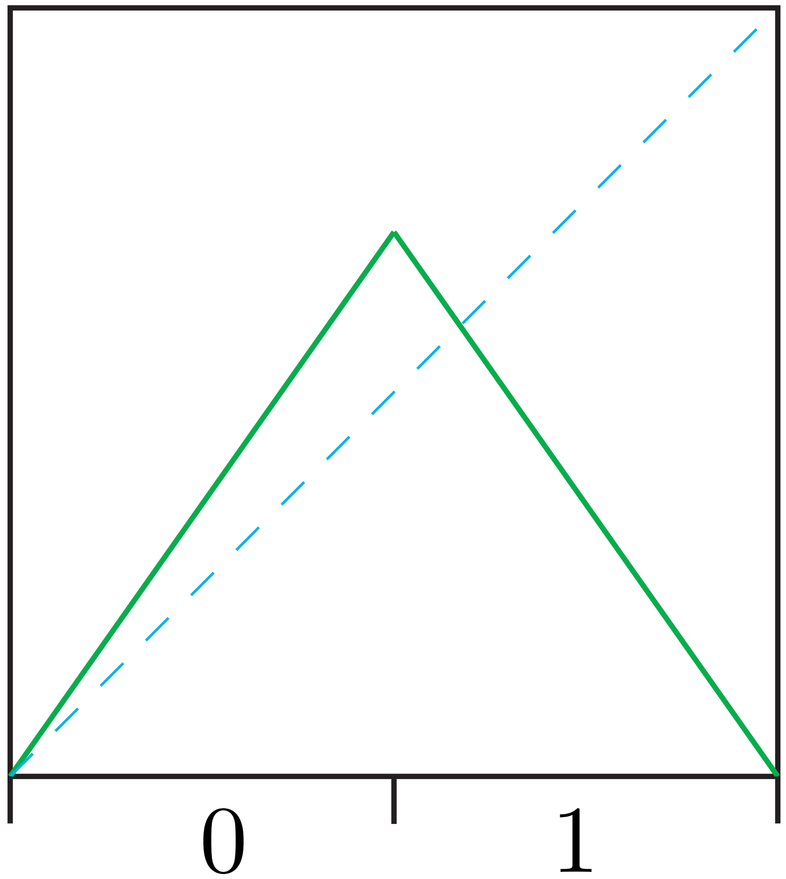
From Determinism to Stochasticity ...

Measurement Theory ...

Generating partition for Tent map:

At any parameter (slope > 1):

$$\mathcal{P} = \{0 \sim [0, \tfrac{1}{2}], 1 \sim [\tfrac{1}{2}, 1]\}$$



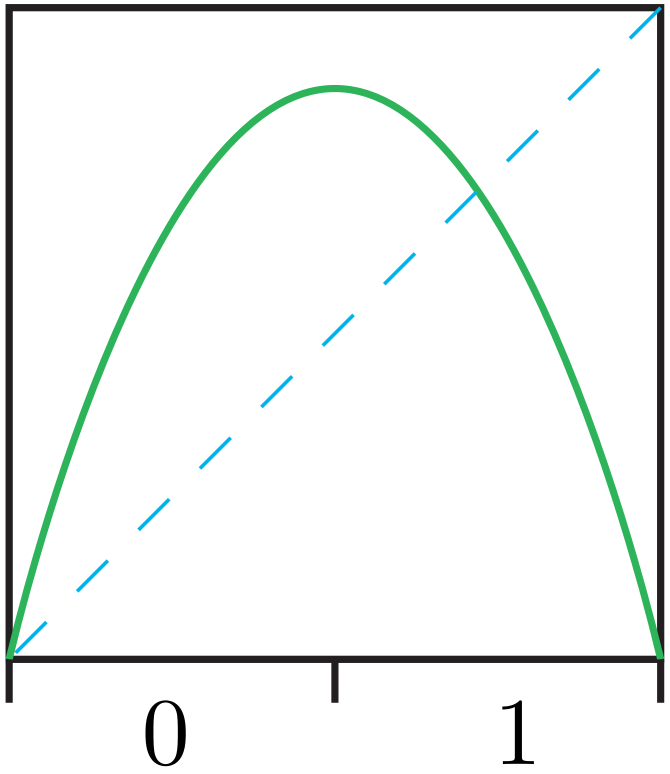
Not Markov!
except at slope = 2.

Measurement Theory ...

Generating partition for Logistic map:

At any parameter:

$$\mathcal{P} = \{0 \sim [0, \tfrac{1}{2}], 1 \sim [\tfrac{1}{2}, 1]\}$$



Not Markov!
except at r = 4.

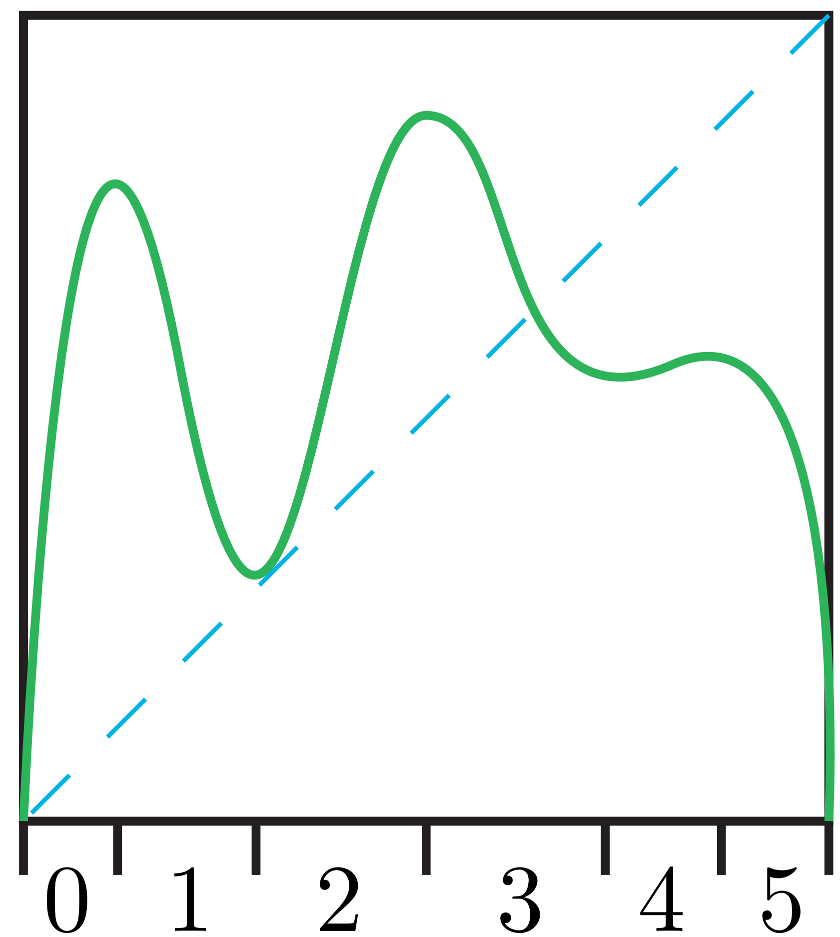
Measurement Theory ...

Generating partition for general 1D maps:

Lap: Monotone piece of $f(x)$

Partition:

$$\mathcal{P} = \{\text{domain of lap}(f)\}$$



Theorem: If map is chaotic, P is generating.

From Determinism to Stochasticity ...

Measurement Theory ...

What happens when there is

no Markov partition and

no generating partition?

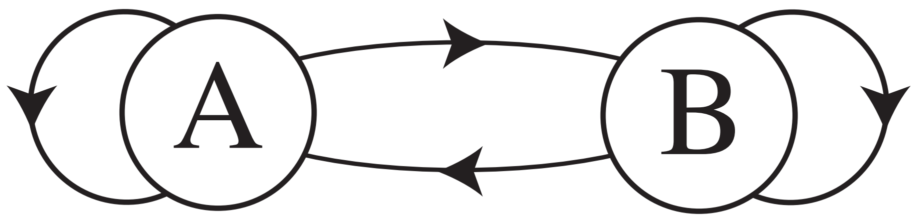
From Determinism to Stochasticity ...

Measurement Theory ...

Example of a nongenerating partition:

Logistic map at 2 onto 1

Internal Markov Model:



Measurement alphabet:

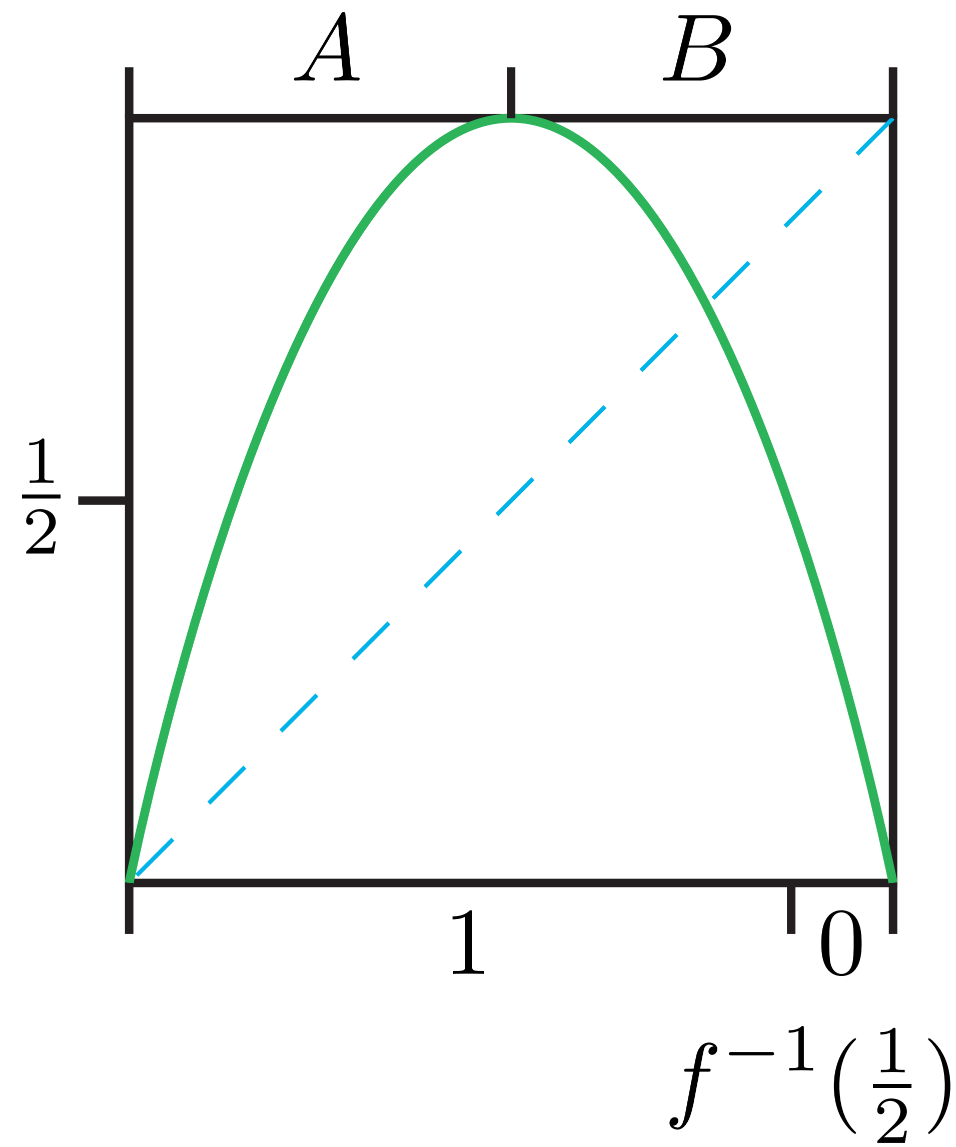
$\{0, 1\}$

Measurement partition:

$$\mathcal{P} = \{1 \sim [0, d], 0 \sim (d, 1]\}$$

Decision point:

$$d = \max \{f^{-1}(\tfrac{1}{2})\}$$



From Determinism to Stochasticity ...

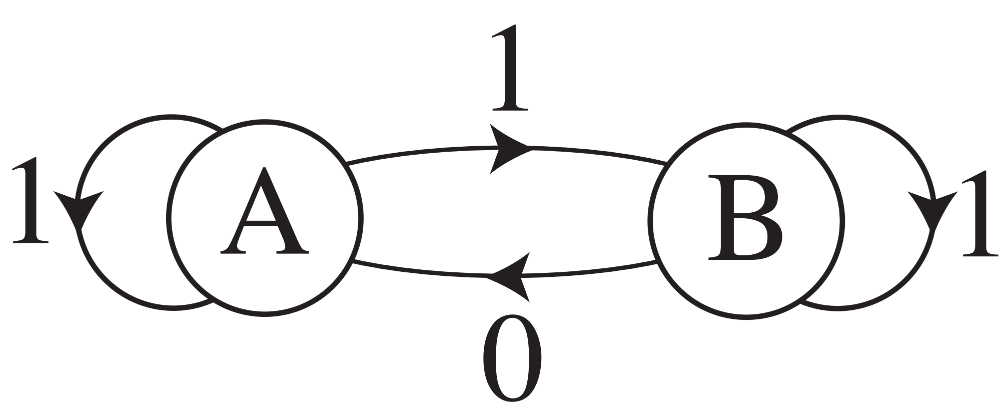
Measurement Theory ...

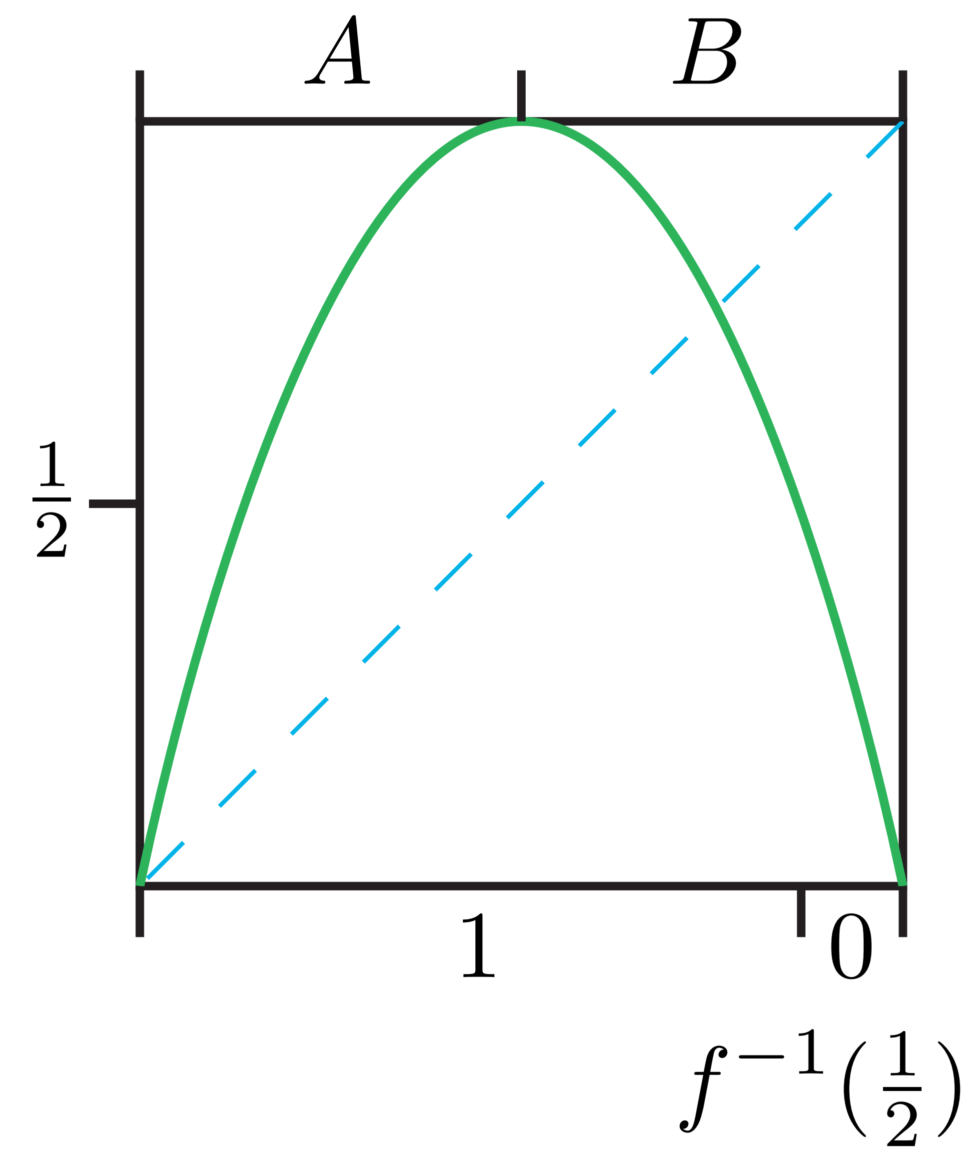
Example of a nongenerating partition ...

Decision point:

$$d = \max \{f^{-1}(\tfrac{1}{2})\}$$

Hidden Markov Model:





The Simple Nonunifilar Source.

Nonunifilar: Map from A-B sequences to observed 0-1 sequences throws away information!

From Determinism to Stochasticity ...

Measurement Theory ...

Synopsis:

How to model measurement process.

How chaos interacts with measurement partition.

End up with discrete-valued sequences.

What kind of stochastic process is the result?

Markov process? Sometimes.
Hidden Markov process? Sometimes.
Sometimes finite, sometimes infinite order.

From Determinism to Stochasticity ...

Measurement Theory ...

Main questions now:

How do we characterize the resulting process?

Measure degrees of unpredictability & randomness.

What correlational structure is there?

How do we build a model from the process itself?

How much can we reconstruct about the hidden internal dynamics?

Stochastic Processes:

Chain of random variables:

$$\overleftrightarrow{S} \equiv \ldots S_{-2} S_{-1} S_0 S_1 S_2 \ldots$$

Random variable: S_t

Alphabet: $\mathcal{A}$

Realization:

$$\cdots s_{-2} s_{-1} s_0 s_1 s_2 \cdots \; ; \; s_t \in \mathcal{A}$$

From Determinism to Stochasticity ...

Stochastic Processes:

Chain of random variables: $\overleftrightarrow{S} = \overleftarrow{S}_t \overrightarrow{S}_t$

Past: $\overleftarrow{S}_t = \ldots S_{t-3} S_{t-2} S_{t-1}$

Future: $\overrightarrow{S}_t = S_t S_{t+1} S_{t+2} \ldots$

L-Block: $S_t^L \equiv S_t S_{t+1} \ldots S_{t+L-1}$

Word: $s_t^L \equiv s_t s_{t+1} \ldots s_{t+L-1} \in \mathcal{A}^L$

From Determinism to Stochasticity ...

Stochastic Processes ...

Process:

$$\Pr(\overleftrightarrow{S}) = \Pr(\ldots S_{-2} S_{-1} S_0 S_1 S_2 \ldots)$$

Sequence (or word) distributions:

$$\{\Pr(S_t^L) = \Pr(S_t S_{t+1} \ldots S_{t+L-1}) : S_t \in \mathcal{A}\}$$

Process:

$$\{\Pr(S_t^L) : \forall t, L\}$$

Consistency conditions:

$$\Pr(S_t^{L-1}) = \sum_{S_{t+L-1}} \Pr(S_t^L) \qquad \Pr(S_{t+1}^{L-1}) = \sum_{S_t} \Pr(S_t^L)$$

From Determinism to Stochasticity ...

Types of Stochastic Process:

Stationary process:

$$\Pr(S_t S_{t+1} \ldots S_{t+L-1}) = \Pr(S_0 S_1 \ldots S_{L-1})$$

Assume stationarity, unless otherwise noted.

Notation: Drop time indices.

Types of Stochastic Process ...

Uniform Process:

Equal-length sequences occur with same probability

$$U^L : \ \Pr(s^L) \ = \ 1/|\mathcal{A}|^L$$

Example: Fair coin

$$\mathcal{A} = \{H, T\}$$

$$\Pr(H) = \Pr(T) = 1/2$$

$$\Pr(s^L) = 2^{-L}$$

Types of Stochastic Process ...

Independent, Identically Distributed (IID) Process:

$$\Pr(\overleftrightarrow{S}) = \ldots \Pr(S_t) \Pr(S_{t+1}) \Pr(S_{t+2}) \ldots$$

$$\Pr(S_t) = \Pr(S_\tau), \ \forall \ t, \tau$$

Example: Biased coin

$$\Pr(H) = p$$

$$\Pr(T) = 1 - p = q$$

$$\Pr(s^L) = p^n q^{L-n}$$

Number of heads in sequence: n

Types of Stochastic Process ...

Markov Process:

$$\Pr(\overleftrightarrow{S}) = \ldots \Pr(S_{t+1}|S_t)\Pr(S_{t+2}|S_{t+1})\Pr(S_{t+3}|S_{t+2})\ldots$$

Example: No Consecutive 0s (Golden Mean Process)

$$\mathcal{A} = \{0, 1\}$$

$$\Pr(0|0) = 0$$

$$\Pr(1|0) = 1$$

$$\Pr(0|1) = 1/2$$

$$\Pr(1|1) = 1/2$$

Not Noisy Period-2 Process: GMP @ L = 4 has 0110.

From Determinism to Stochasticity ...

Types of Stochastic Process ...

Hidden Markov Process:

Internal Order-R Markov Process: $\Pr(\overleftrightarrow{S})$

$$\Pr(S_t|\ldots S_{t-2}S_{t-1}) = \Pr(S_t|S_{t-R}\ldots S_{t-1})$$

$$s_t \in \mathcal{A}$$

Observed via a function of the internal sequences

$$\overleftrightarrow{Y} = f(\overleftrightarrow{S})$$

Measurement alphabet: $y_t \in \mathcal{B}$

Measurement random variables: $\overleftrightarrow{Y} = \ldots Y_{-2}Y_{-1}Y_0Y_1\ldots$

Observation process: $\Pr(\overleftrightarrow{Y} \mid \overleftrightarrow{S})$

Observed process: $\Pr(\overleftrightarrow{Y})$

Block Distribution: $\Pr(Y^L)$

From Determinism to Stochasticity ...

Types of Stochastic Process ...

Hidden Markov Process ...

Example: The Even Process

Internal Process: Golden Mean

$$s_t \in \{0, 1\}$$

Observation Process: $y_t \in \{a, b\}$

$$Y_t = f(S_{t-1} S_t)$$

$$y_t = \begin{cases} a, & s_{t-1} s_t = 11 \\ b, & s_{t-1} s_t = 01 \text{ or } 10 \end{cases}$$

$$\overleftrightarrow{s} = \texttt{1101110111101011111011} \ldots$$

$$\overleftrightarrow{y} = .\ \texttt{abbaabbaaabbbbaaaabba} \ldots$$

Models of Stochastic Processes:

Markov chain model of a Markov process:

States: $v \in \mathcal{A} = \{1, \ldots, k\}$

$$\overleftrightarrow{V} = \ldots V_{-2} V_{-1} V_0 V_1 \ldots$$

Transition matrix: $T_{ij} = \Pr(v_{t+1} | v_t) \equiv p_{vv'}$

$$T = \begin{pmatrix} p_{11} & \cdots & p_{1k} \\ \vdots & \ddots & \vdots \\ p_{k1} & \cdots & p_{kk} \end{pmatrix}$$

Stochastic matrix: $\sum_{i=1}^{k} T_{ij} = 1$

From Determinism to Stochasticity ...

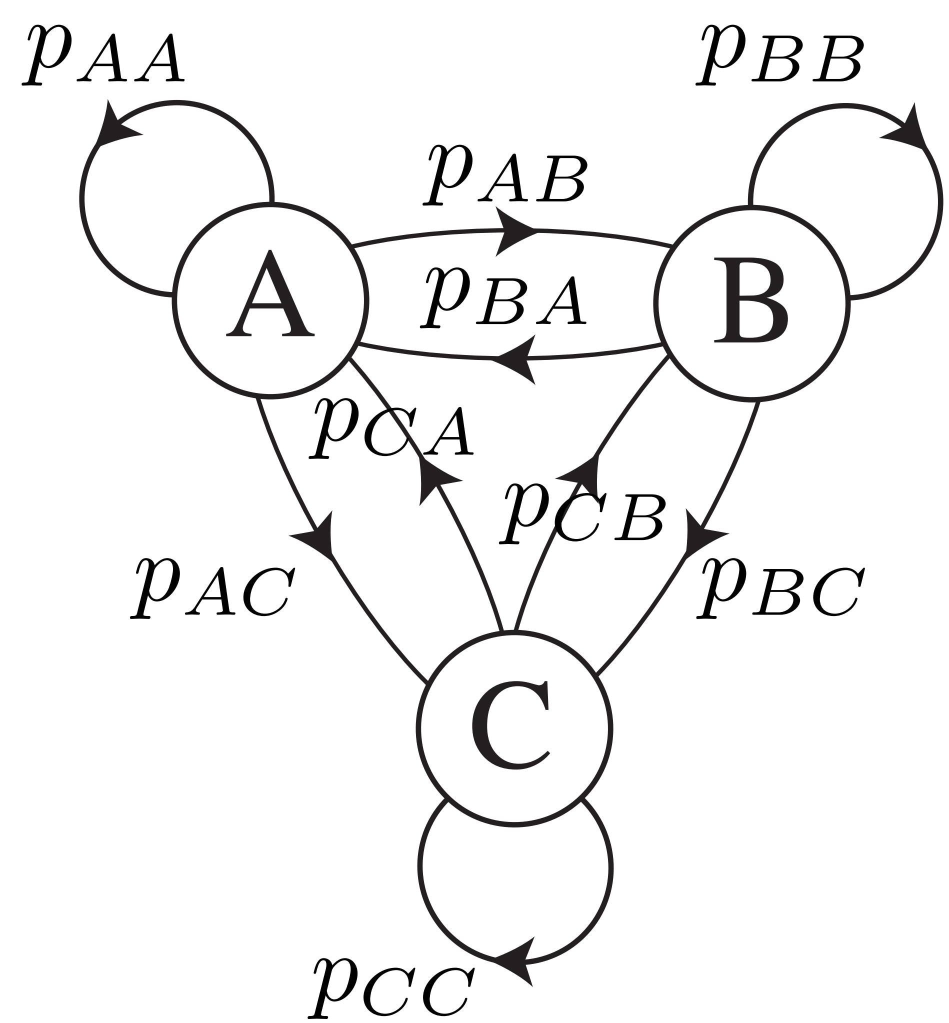
Models of Stochastic Processes ...

Markov chain ...

Example: $\mathcal{A} = \{A, B, C\}$

$$T = \begin{pmatrix} p_{AA} & p_{AB} & p_{AC} \\ p_{BA} & p_{BB} & p_{BC} \\ p_{CA} & p_{CB} & p_{CC} \end{pmatrix}$$

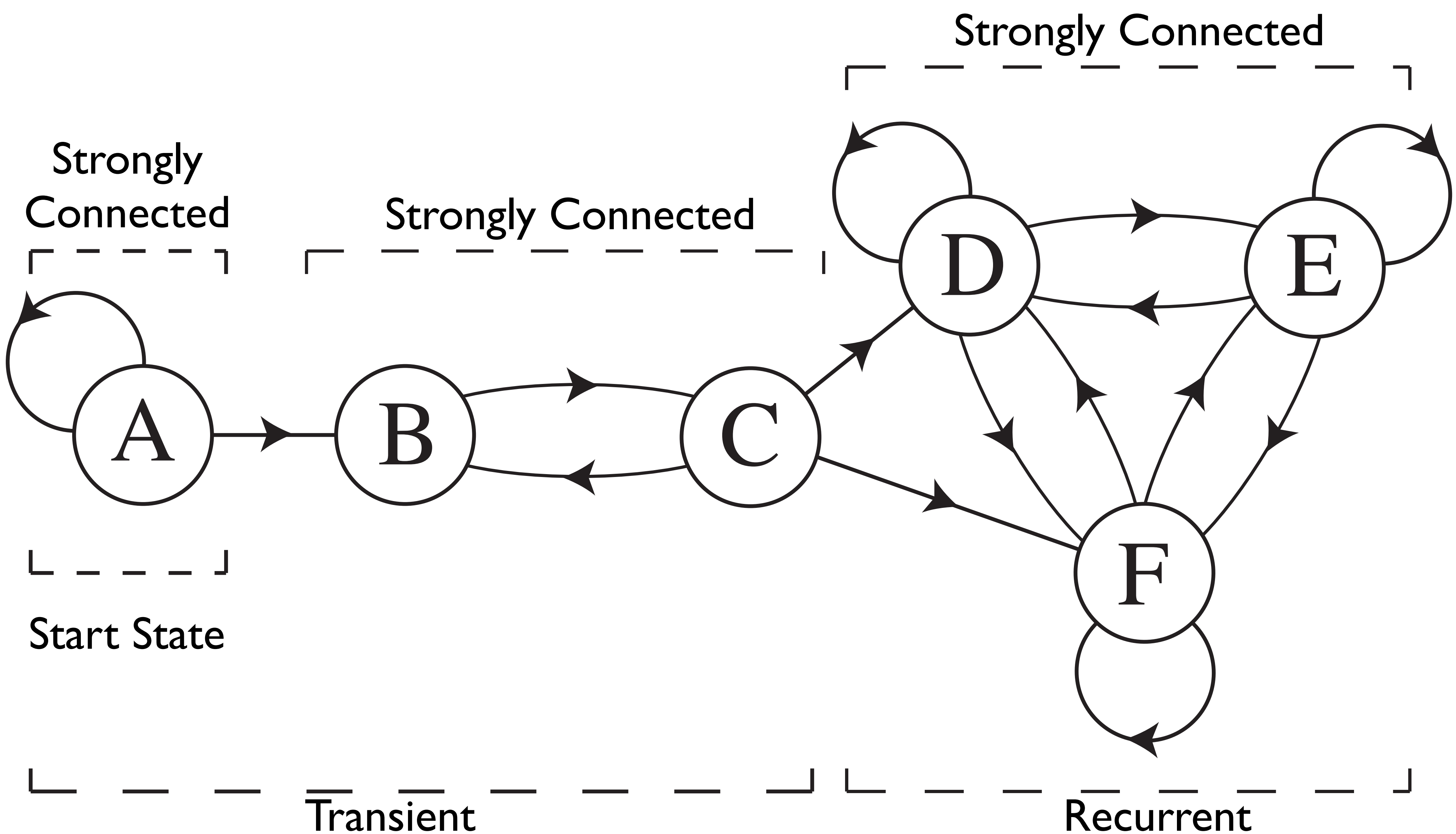
$$p_{AA} + p_{AB} + p_{AC} = 1$$
$$p_{BA} + p_{BB} + p_{BC} = 1$$
$$p_{CA} + p_{CB} + p_{CC} = 1$$



From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Kinds of state:



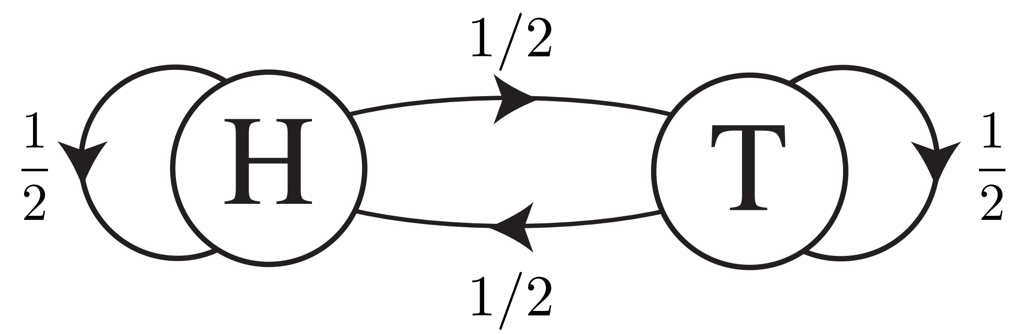
From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Example:

Fair Coin: $\mathcal{A} = \{H, T\}$

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$



$$\pi = (1/2, 1/2)$$

$$\Pr(H) = \Pr(T) = 1/2$$

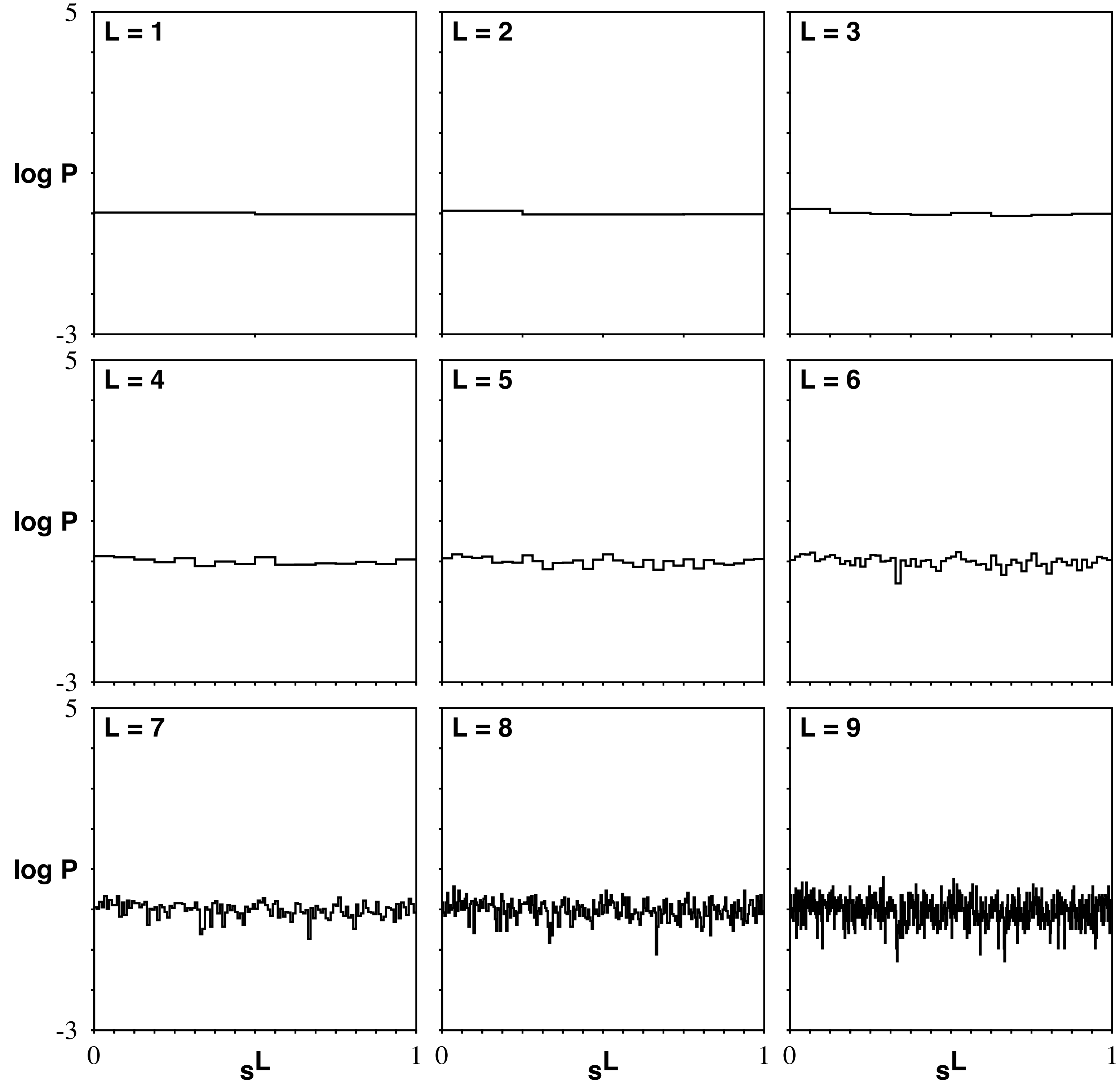
From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Example:

Fair Coin ...

Sequence Distribution: $\Pr(v^L) = 2^{-L}$



Word as binary fraction:

$$s^L = s_1 s_2 \ldots s_L$$

$$\text{“}s^L\text{”} = \sum_{i=1}^{L} \frac{s_i}{2^i}$$

$$s^L \in [0, 1]$$

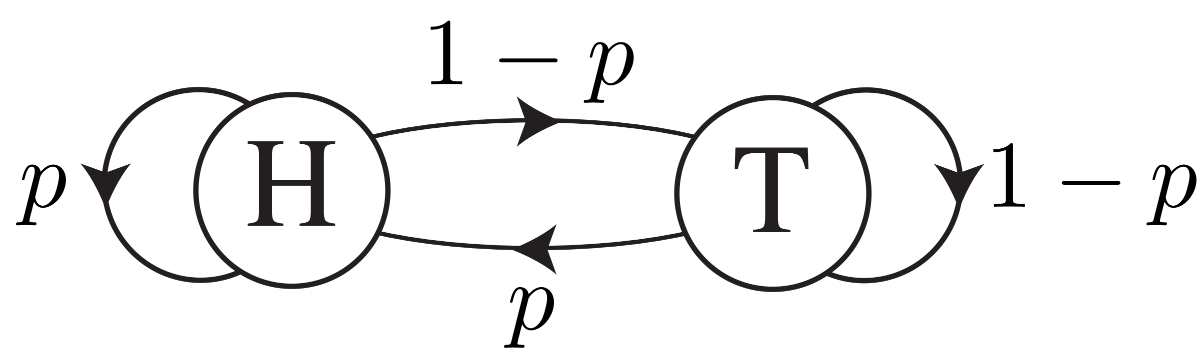
From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Example:

Biased Coin: $\mathcal{A} = \{H, T\}$

$$T = \begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix}$$



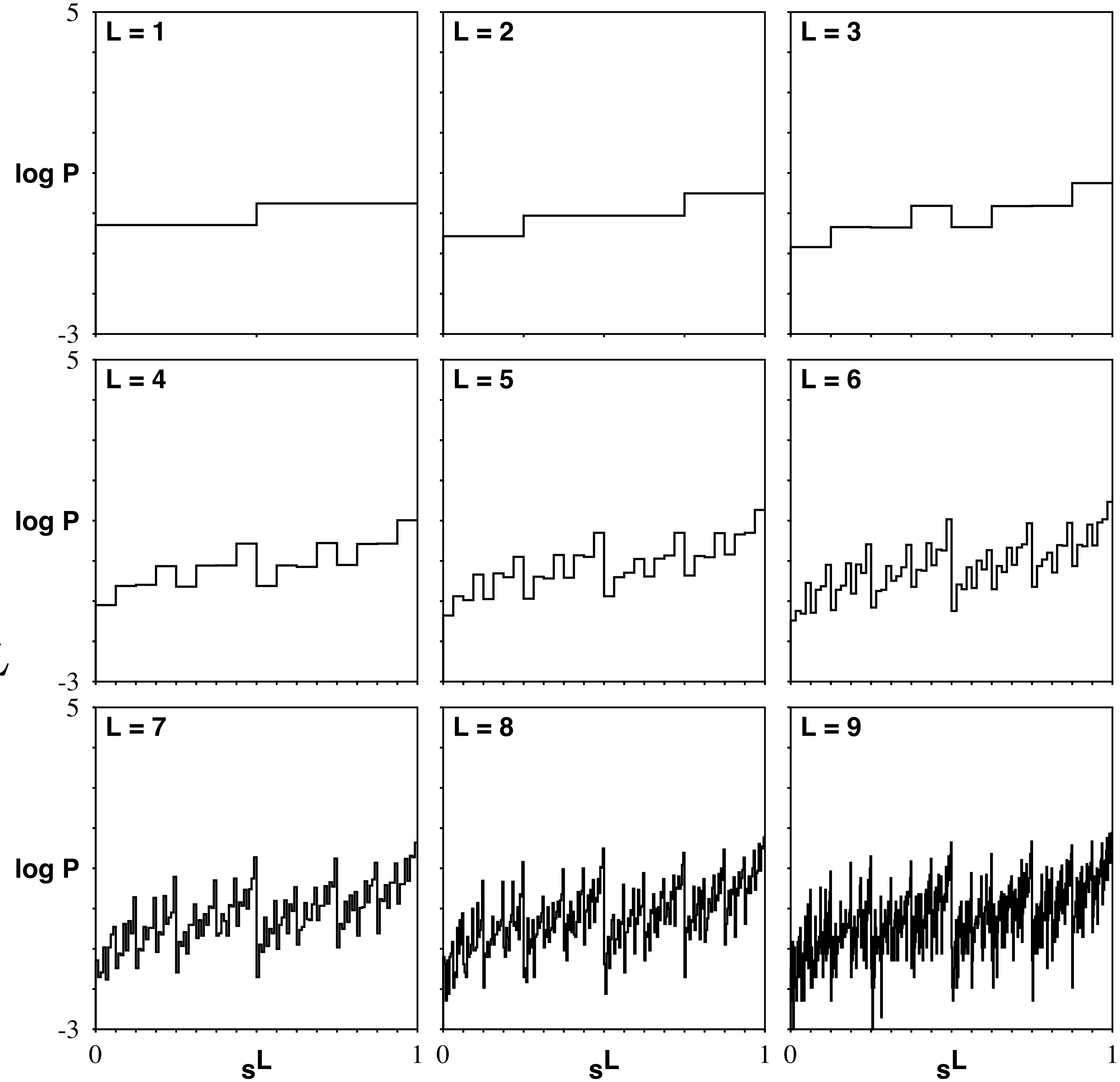
$$\pi = (p, 1-p)$$

From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Example:

Biased Coin ...



Sequence Distribution:

$$\Pr(s^L) = p^n (1-p)^{L-n},$$

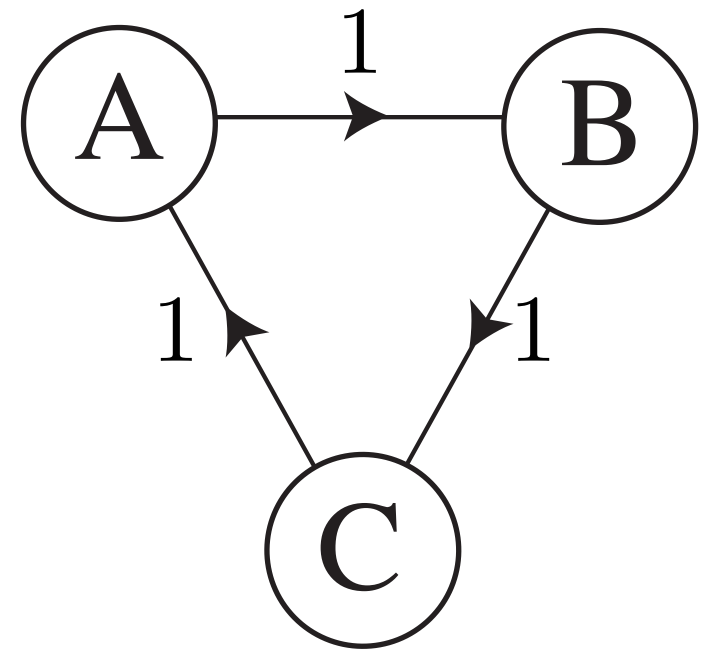
$$n = \text{Number } H\text{s in } s^L$$

From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Example:

Periodic: $\mathcal{A} = \{A, B, C\}$



$$T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$\pi = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ Careful!

Sequence distribution:

$$\Pr(A) = \Pr(B) = \Pr(C) = \frac{1}{3}$$

$\Pr(AB) = \Pr(BC) = \Pr(CA) = \frac{1}{3}$ $\Pr(s^2) = 0$ otherwise

$\Pr(ABC) = \Pr(BCA) = \Pr(CAB) = \frac{1}{3}$ $\Pr(s^3) = 0$ otherwise

From Determinism to Stochasticity ...

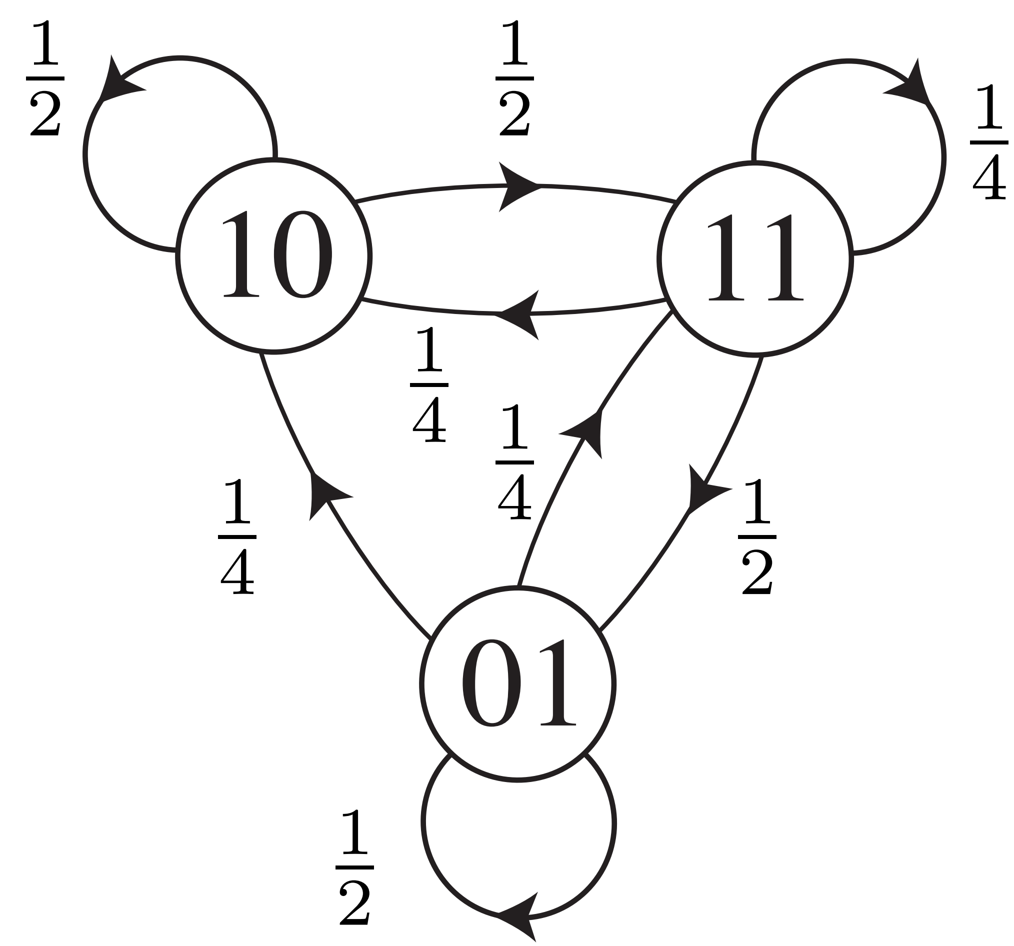
Models of Stochastic Processes ...

Example:

Golden Mean over 2-Blocks: $\mathcal{A} = \{10, 01, 11\}$

$$T = \begin{array}{c} \begin{matrix} \quad 10 & \quad 01 & \quad 11 \end{matrix} \\ \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix} \end{array}$$

$$\pi = \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right)$$



From Determinism to Stochasticity ...

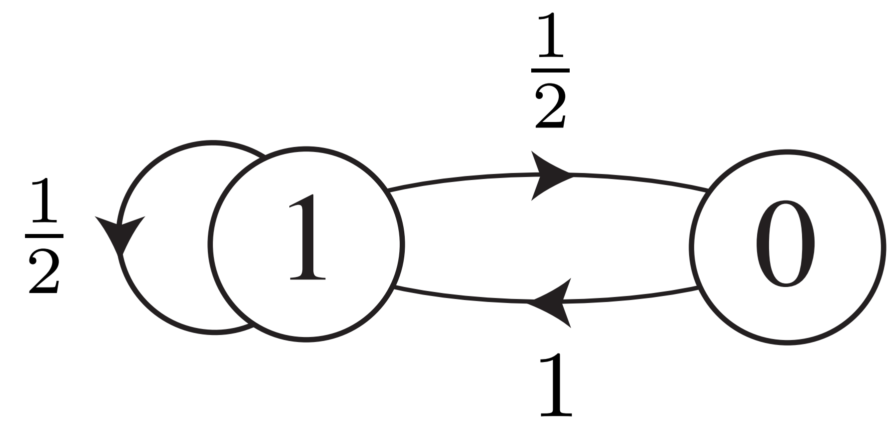
Models of Stochastic Processes ...

Example ...

Golden Mean over 1-Blocks: $\mathcal{A} = \{0, 1\}$

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix}$$

$$\pi = \left(\tfrac{2}{3}, \tfrac{1}{3}\right)$$



Also an order-1 Markov chain. Minimal order.

Previous model and this:

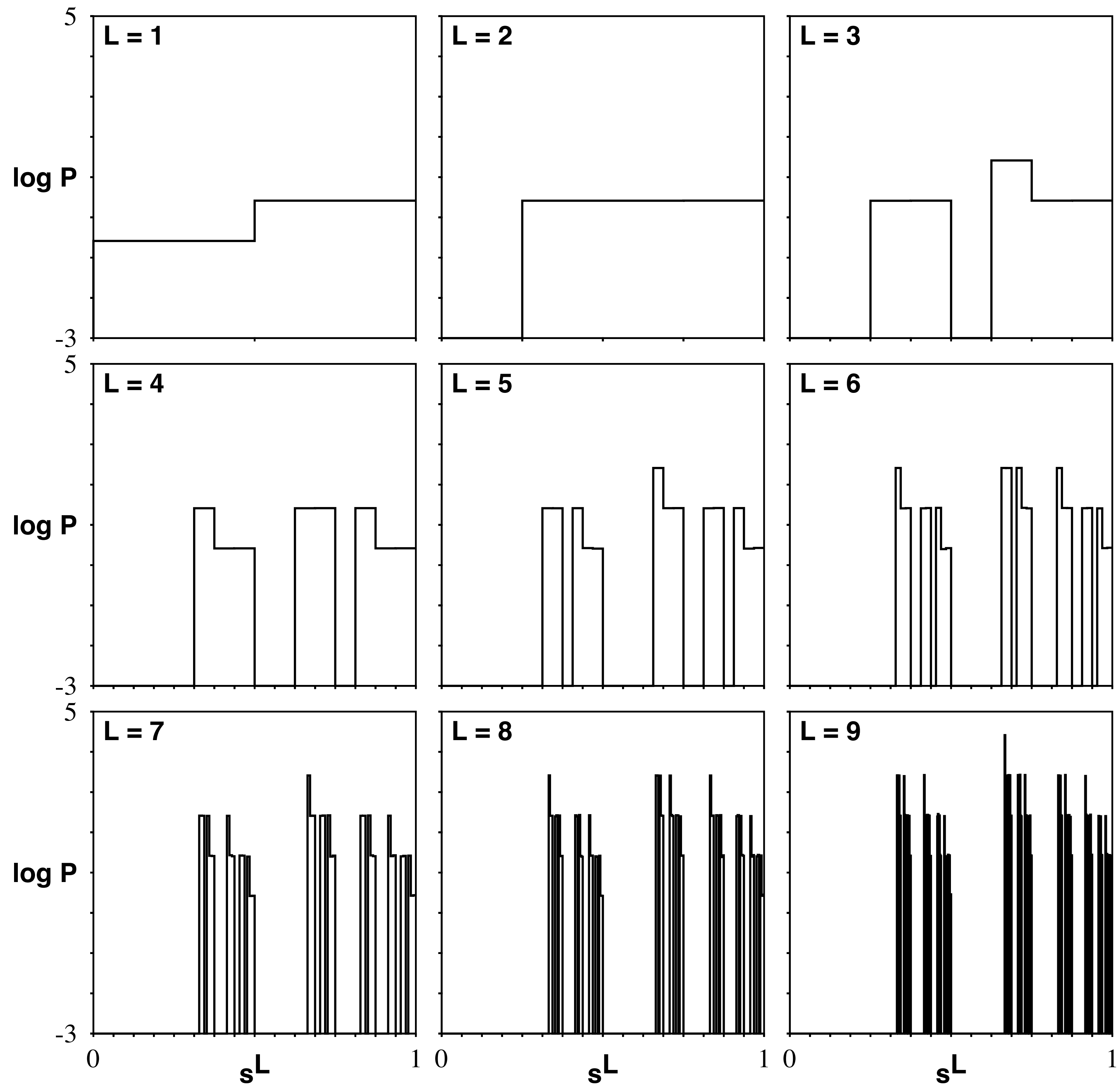
Different presentations of the Golden Mean Process

From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Example:

Golden mean:



From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Two Lessons:

Structure in the behavior: $\mathrm{supp}\ \Pr(s^L)$

Structure in the distribution of behaviors: $\Pr(s^L)$

Models of Stochastic Processes ...

Hidden Markov Models of Processes:

Internal states: $v \in \mathcal{A}$

Transition matrix: $T = \Pr(v'|v), \; v, v' \in \mathcal{A}$

Observation: Symbol-labeled transition matrices

$$T^{(s)} = \Pr(v', s|v), \; s \in \mathcal{B}$$

$$T = \sum_{s \in \mathcal{B}} T^{(s)}$$

Stochastic matrices:

$$\sum_j T_{ij} = \sum_j \sum_s T_{ij}^{(s)} = 1$$

From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Hidden Markov Models ...

Internal state distribution: $\vec{p}_V = (p_1, p_2, \ldots, p_k)$

Evolve internal distribution: $\vec{p}_n = \vec{p}_0 T^n$

State sequence distribution: $v^L = v_0 v_1 v_2 \ldots v_{L-1}$

$$\Pr(v^L) = \pi(v_0) p(v_1|v_0) p(v_2|v_1) \cdots p(v_{L-1}|v_{L-2})$$

Observed sequence distribution: $s^L = s_0 s_1 s_2 \ldots s_{L-1}$

$$\Pr(s^L) = \sum_{v^L \in \mathcal{A}^L} \pi(v_0) p(v_1, s_1|v_0) p(v_2, s_2|v_1) \cdots p(v_{L-1}, s_{L-1}|v_{L-2})$$

No longer 1-1 map between internal & observed sequences:

Multiple state sequences can produce *same* observed sequence.

From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Hidden Markov Models ...

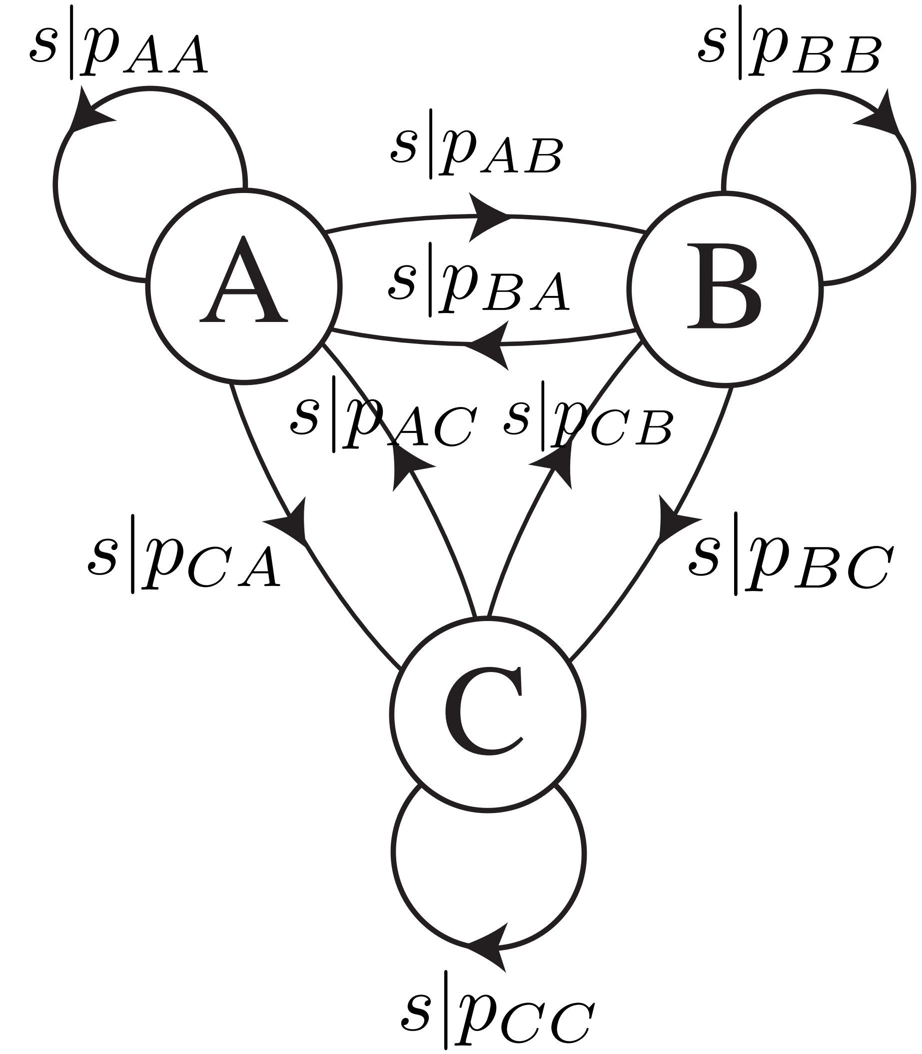
Internal: $\mathcal{A} = \{A, B, C\}$

$$T = \begin{pmatrix} p_{AA} & p_{AB} & p_{AC} \\ p_{BA} & p_{BB} & p_{BC} \\ p_{CA} & p_{CB} & p_{CC} \end{pmatrix}$$

Observed: $\mathcal{B} = \{0, 1\}$

$$T^{(s)} = \begin{pmatrix} p_{AA;s} & p_{AB;s} & p_{AC;s} \\ p_{BA;s} & p_{BB;s} & p_{BC;s} \\ p_{CA;s} & p_{CB;s} & p_{CC;s} \end{pmatrix}$$

$$p_{AA} = \sum_{s \in \mathcal{B}} p_{AA;s}$$



symbol | transition probability

Models of Stochastic Processes ...

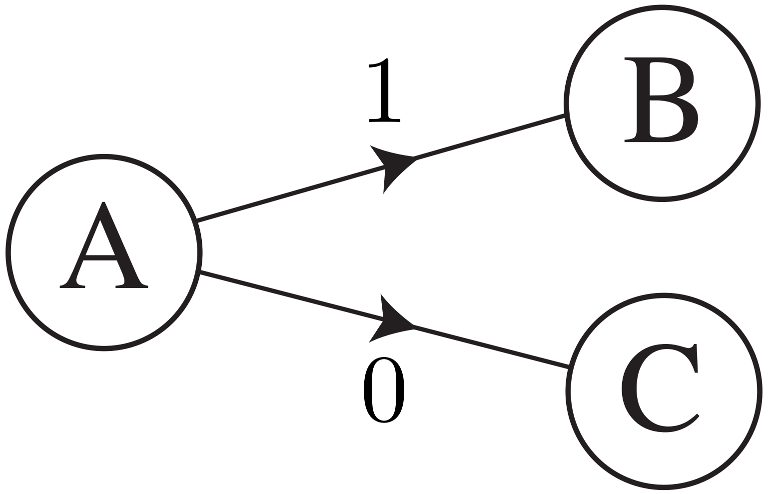
Types of Hidden Markov Model:

"Unifilar": current state + symbol "determine" next state

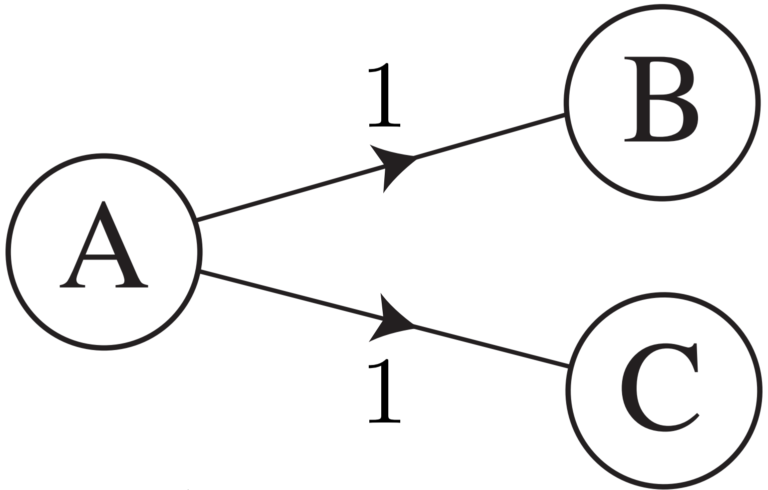
$$\Pr(v'|v,s) = \begin{cases} 1 \\ 0 \end{cases}$$

$$\Pr(v',s|v) = p(s|v)$$

$$\Pr(v'|v) = \sum_{s \in \mathcal{A}} p(s|v)$$



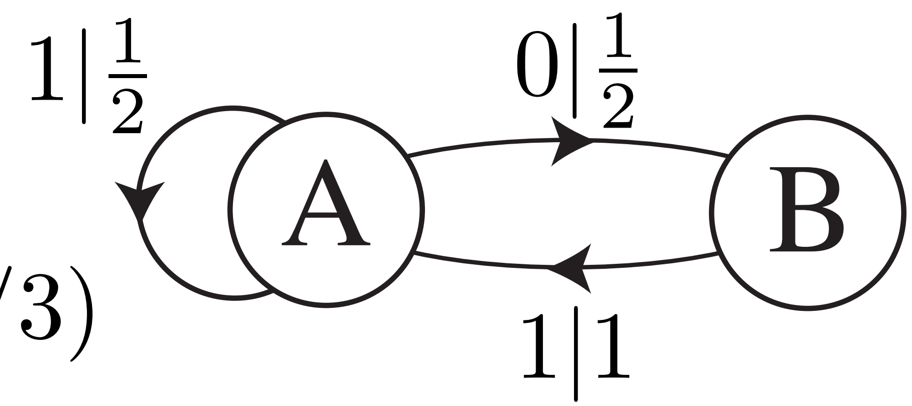
"Nonunifilar": no restriction



Multiple internal edge paths can generate same observed sequence.

Example:

Golden Mean Process as a unifilar HMM:



Internal: $\mathcal{A} = \{A, B\}$

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix} \quad \pi_V = (2/3, 1/3)$$

Observed: $\mathcal{B} = \{0, 1\}$

$$T^{(0)} = \begin{pmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{pmatrix} \quad T^{(1)} = \begin{pmatrix} \frac{1}{2} & 0 \\ 1 & 0 \end{pmatrix}$$

Initial ambiguity only: At most 2-to-1 mapping

$$BA^n = 1^n$$

$$AA^n = 1^n$$

Sync'd: $s = 0 \Rightarrow v = B$
$s = 1 \Rightarrow v = A$

Irreducible forbidden words: $\mathcal{F} = \{00\}$

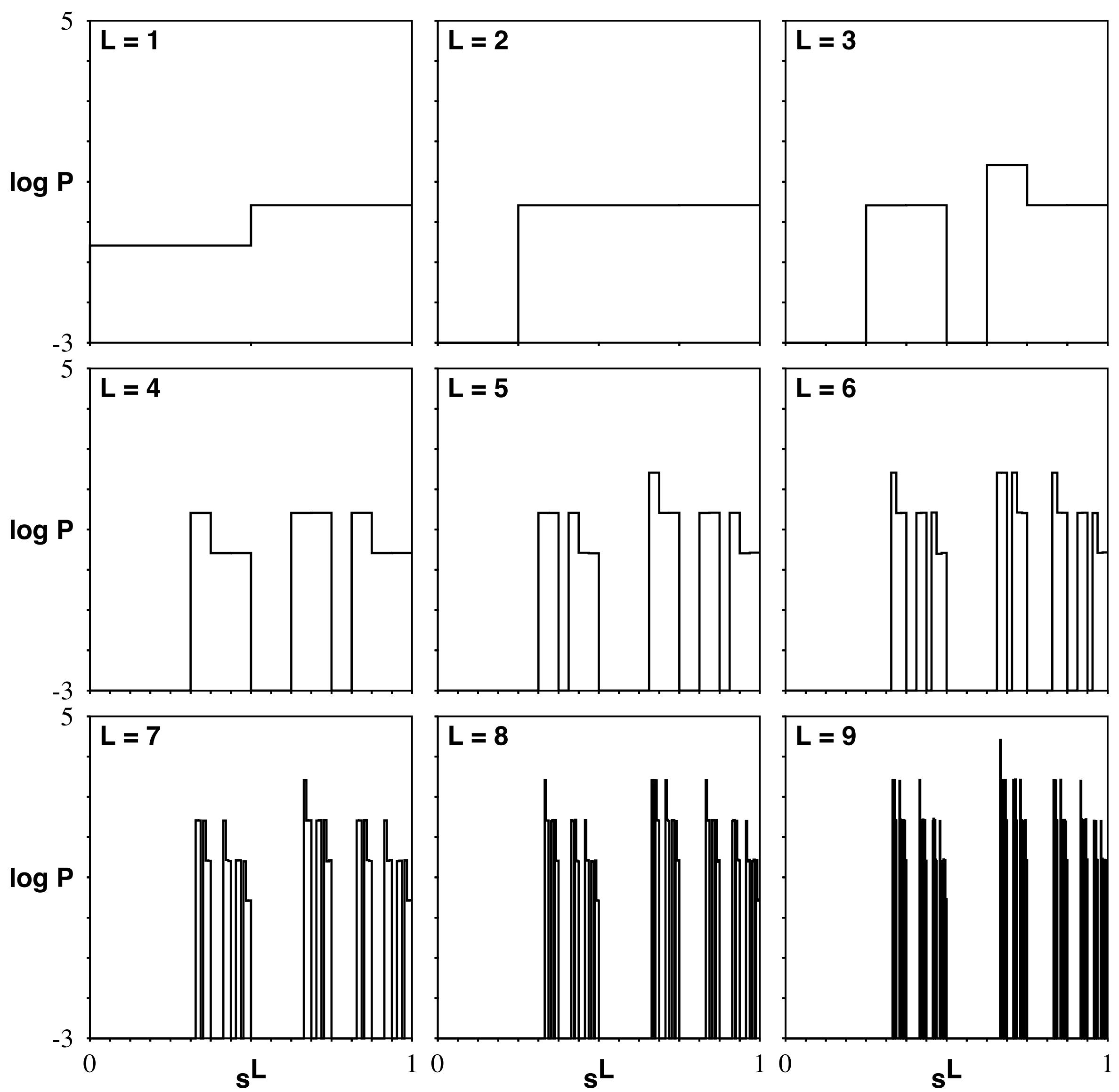
From Determinism to Stochasticity ...

Models of Stochastic Processes ...

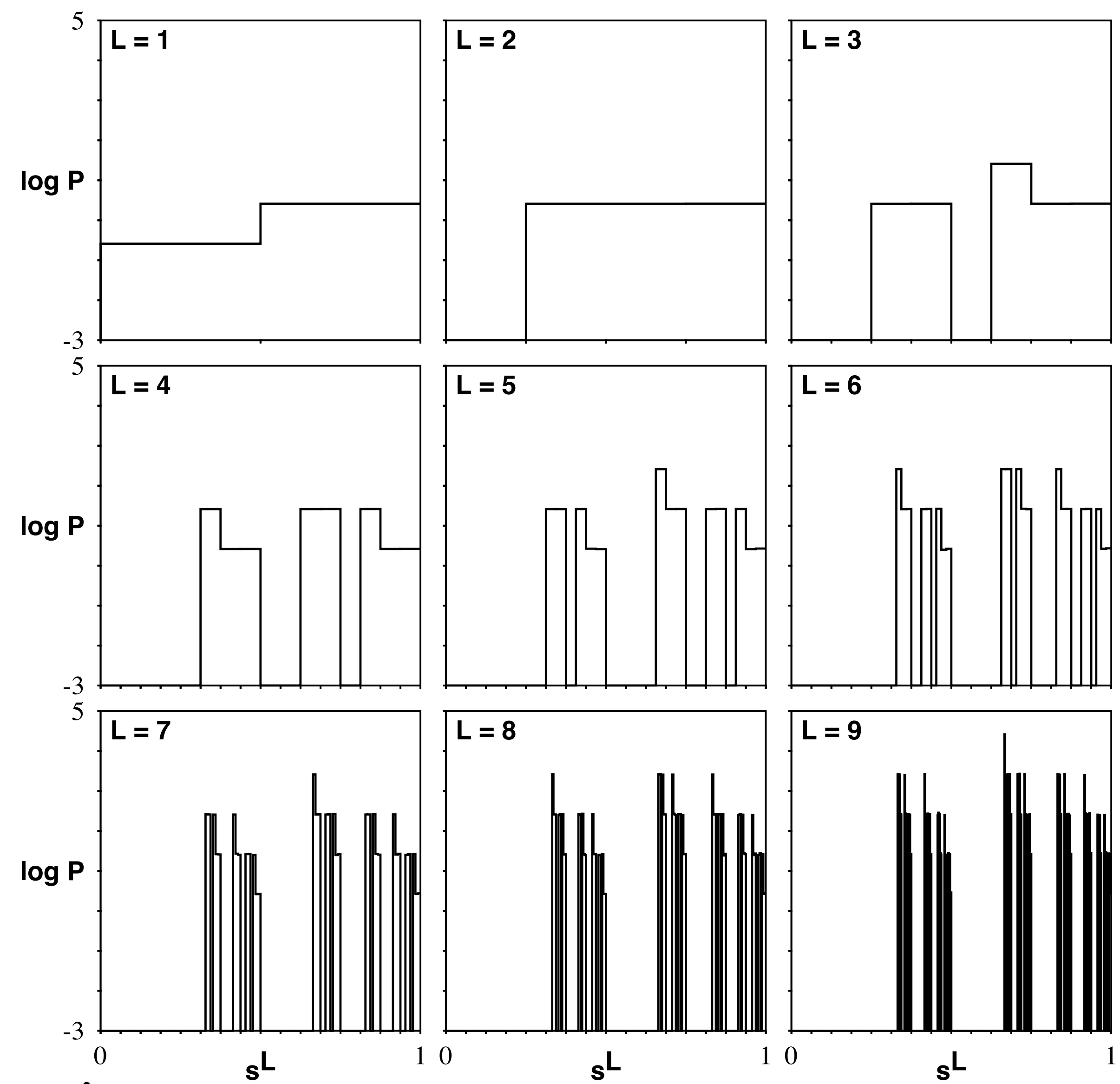
Example:

Golden Mean Process ... Sequence distributions:

Internal state sequences
(A = 1; B = 0)



Observed sequences



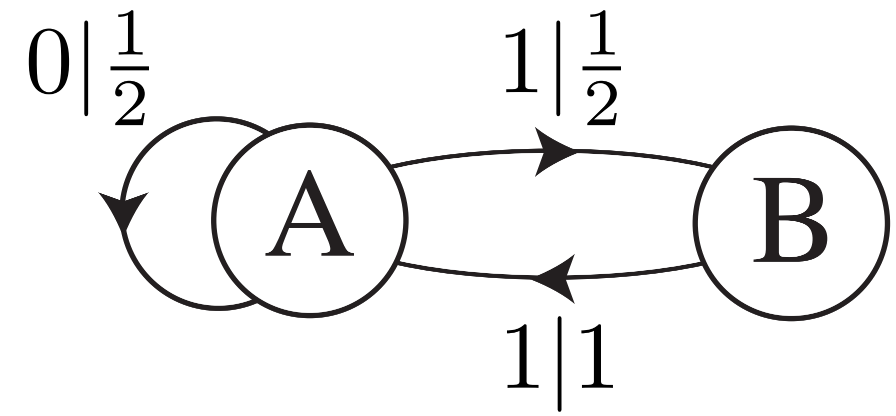
Same!

From Determinism to Stochasticity ...

Models of Stochastic Processes ...

Example:

Even Process as a unifilar HMM:



Internal (= GMP): $\mathcal{A} = \{A, B\}$

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix} \quad \pi_V = (2/3, 1/3)$$

Observed: $\mathcal{B} = \{0, 1\}$

$$T^{(0)} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix} \quad T^{(1)} = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & 0 \end{pmatrix}$$

$$v^L = \ldots AABAABABAA \ldots$$

$$s^L = \ldots .\ \ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 1\ 0 \ldots \qquad s^L = \{\ldots 01^{2n}0 \ldots\}$$

Irreducible forbidden words: $\mathcal{F} = \{010, 01110, 0111110, \ldots\}$

No finite-order Markov process can model the Even process!

Lesson: Finite Markov Chains are a subset of HMMs.

From Determinism to Stochasticity ...

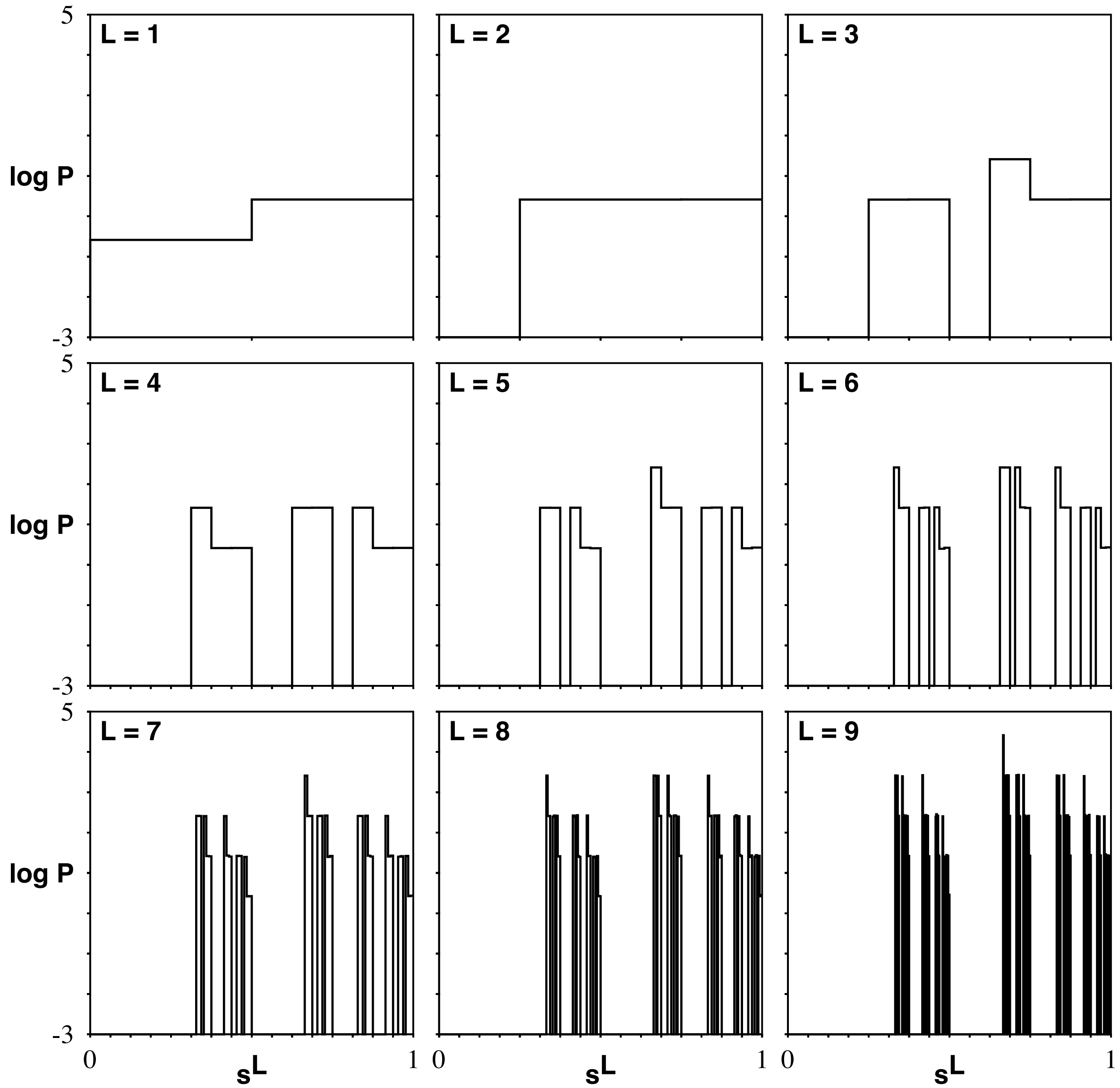
Models of Stochastic Processes ...

Example:

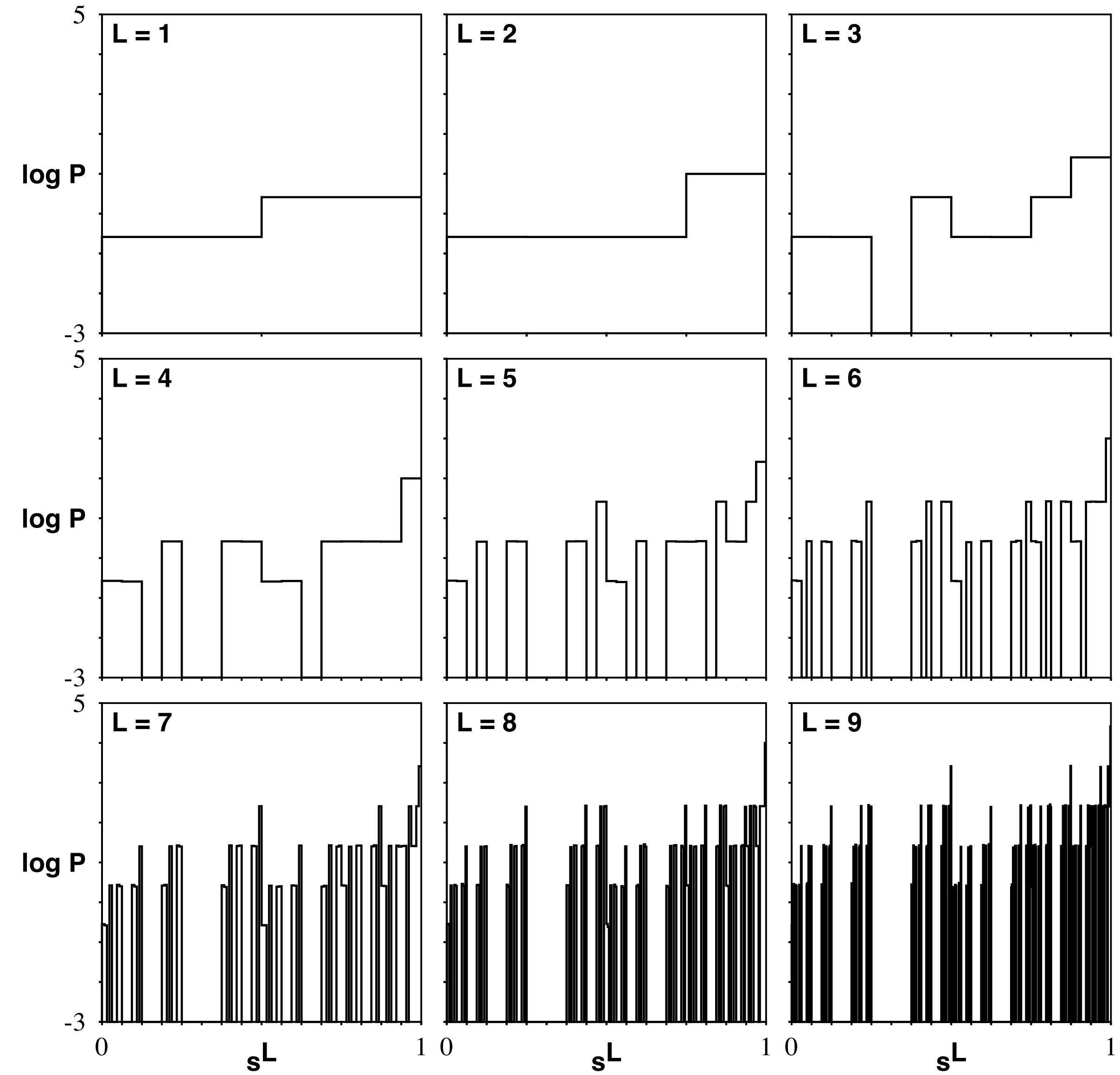
Even Process ... Sequence distributions:

Internal states (= GMP)

(A = 1; B = 0)



Observed sequences



Rather different!

From Determinism to Stochasticity ...

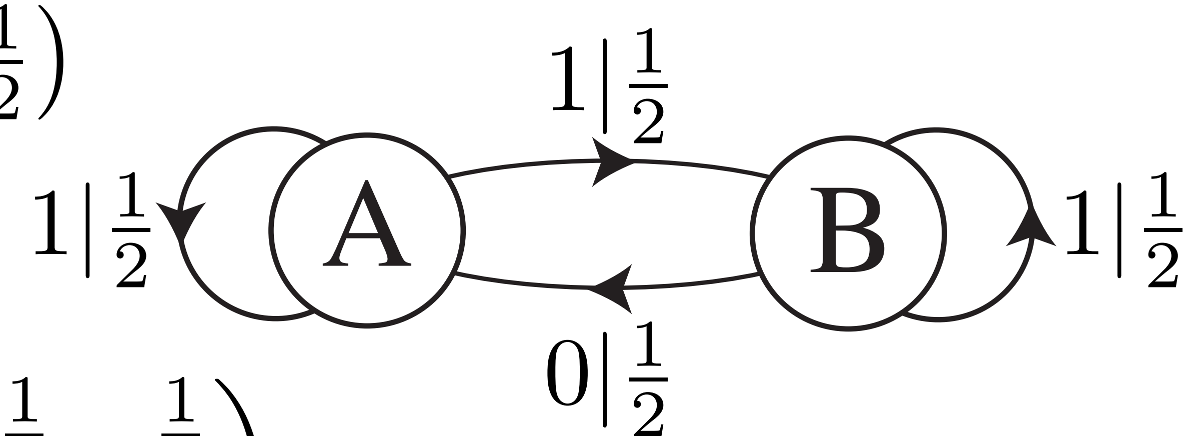
Models of Stochastic Processes ...

Example:

Simple Nonunifilar Source:

Internal (= Fair Coin): $\mathcal{A} = \{A, B\}$

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad \pi_V = \left(\frac{1}{2}, \frac{1}{2}\right)$$



Observed: $\mathcal{B} = \{0, 1\}$

$$T^{(0)} = \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{pmatrix} \quad T^{(1)} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}$$

Many to one: $1111111 \Leftarrow \begin{cases} AAAAAAAA\ldots \\ ABBBBBBB\ldots \\ AABBBBBB\ldots \\ AAABBBBB\ldots \\ \ldots \\ BBBBBBBB\ldots \end{cases}$

Is there a unifilar HMM presentation of the observed process?

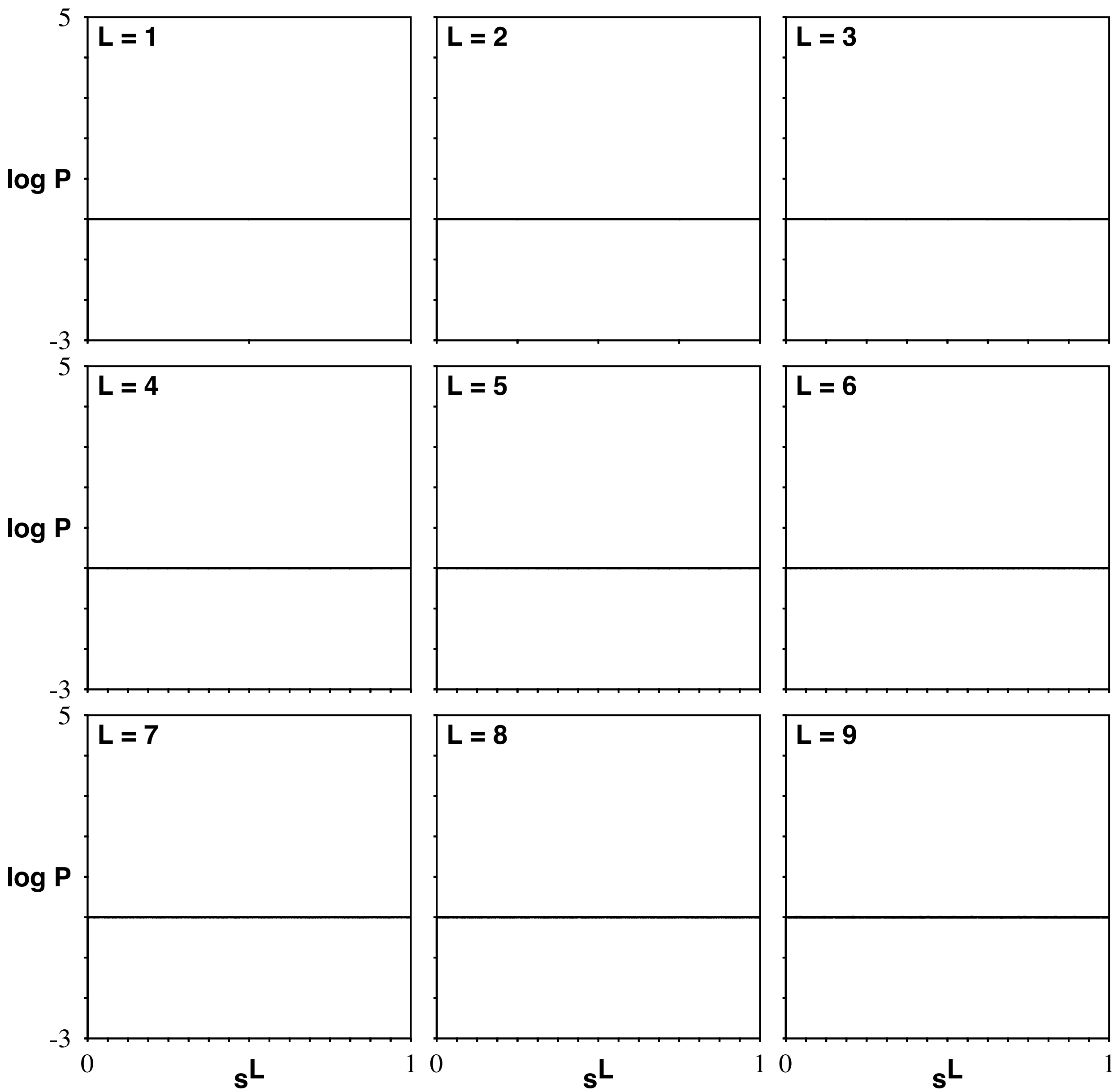
From Determinism to Stochasticity ...

Models of Stochastic Processes ...

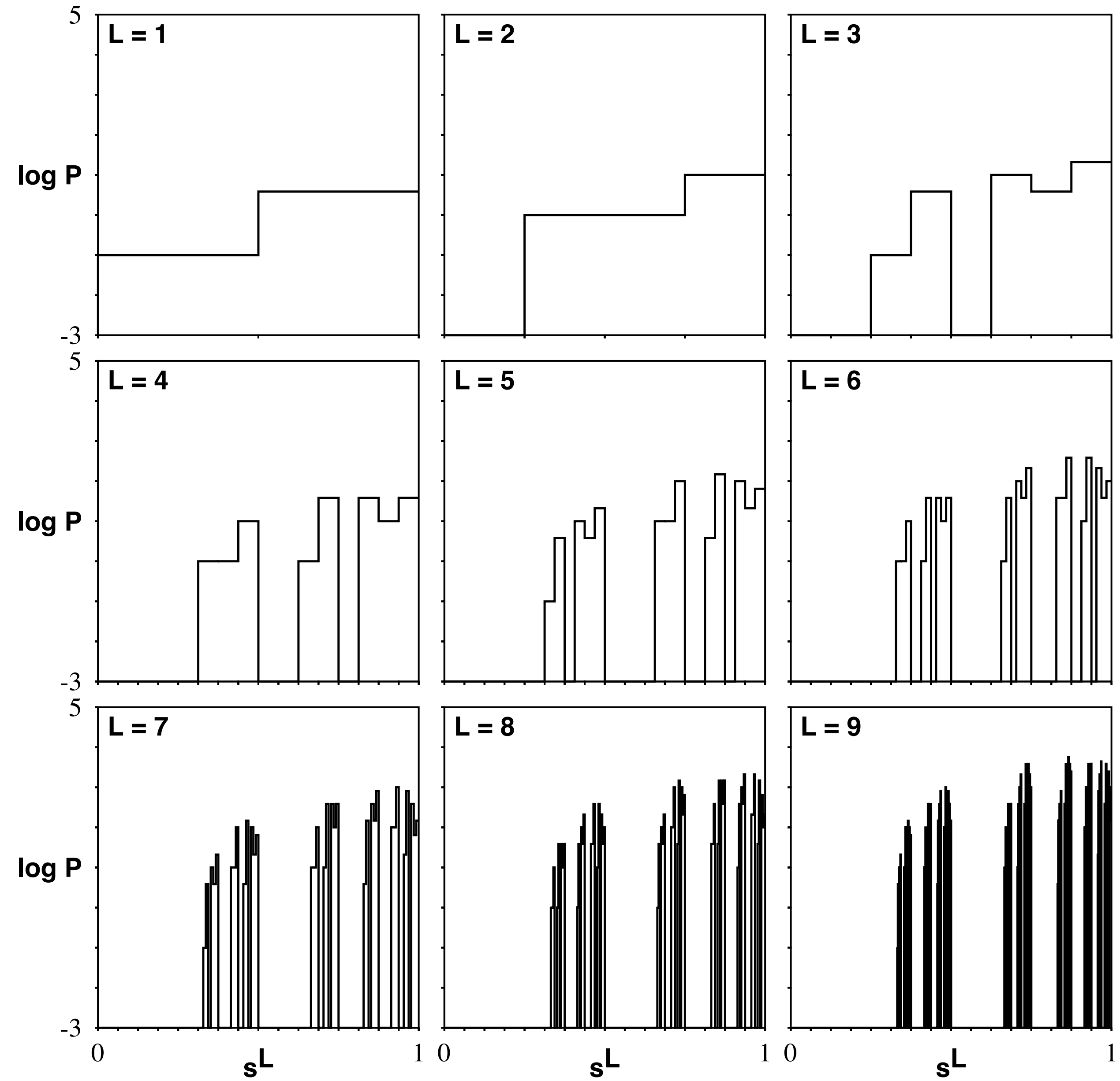
Example:

Simple Nonunifilar Process ...

Internal states (= Fair coin)
(A = 1; B = 0)



Observed sequences



From Determinism to Stochasticity ...

Next lectures: “Complexity Module”

1. Stochastic processes and models of them

2. Information theory for general stochastic processes

3. Measures of complexity

4. Optimal model and how to build them