NCASO Spring 2014

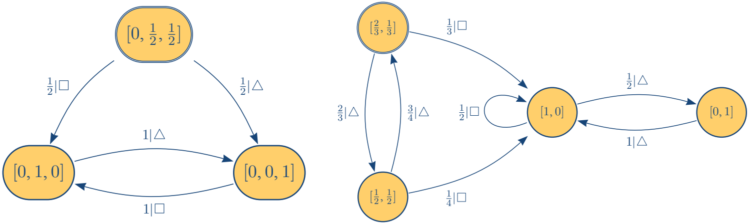
Complexity à la Mode: Spectral Methods for Complex Systems Part $2H_2(p)\big|_{p=\frac{1}{2}}$

Paul M. Riechers

Complexity Sciences Center
Department of Physics
University of California, Davis

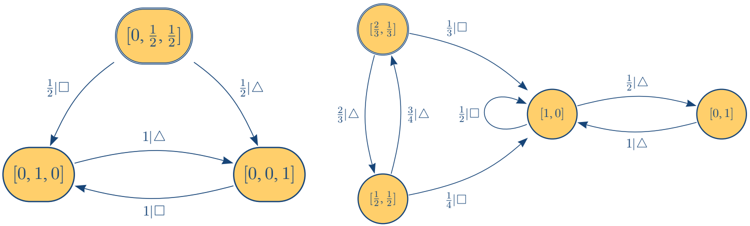
May 22nd 2014

Visualizing Modes



Implicitly, we already visualize modes.

Visualizing Modes



Implicitly, we already visualize modes.

Spectral methods formalize and empower our intuition.

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics
- Correlation functions

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics
- Correlation functions
- Power spectra (including diffraction spectra of disordered crystals)

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics
- Correlation functions
- Power spectra (including diffraction spectra of disordered crystals)
- Inter-Spike-Interval histograms and return maps
- ...

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics
- Correlation functions
- Power spectra (including diffraction spectra of disordered crystals)
- Inter-Spike-Interval histograms and return maps
- ...

Directly from the Mixed-State Presentation (MSP) of any HMM:

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics
- Correlation functions
- Power spectra (including diffraction spectra of disordered crystals)
- Inter-Spike-Interval histograms and return maps
- ...

Directly from the Mixed-State Presentation (MSP) of any HMM:

- Myopic entropy rates $h_\mu(L)$; asymptotic entropy rate h_μ

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics
- Correlation functions
- Power spectra (including diffraction spectra of disordered crystals)
- Inter-Spike-Interval histograms and return maps
- ...

Directly from the Mixed-State Presentation (MSP) of any HMM:

- Myopic entropy rates $h_\mu(L)$; asymptotic entropy rate h_μ
- Past–Future Mutual Information (i.e., Excess Entropy $\mathbf{E}$)

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics
- Correlation functions
- Power spectra (including diffraction spectra of disordered crystals)
- Inter-Spike-Interval histograms and return maps
- ...

Directly from the Mixed-State Presentation (MSP) of any HMM:

- Myopic entropy rates $h_\mu(L)$; asymptotic entropy rate h_μ
- Past–Future Mutual Information (i.e., Excess Entropy $\mathbf{E}$)
- Info shared btw. past and future, but *not* in the present (i.e., Elusive Information σ_μ)
- ...

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics
- Correlation functions
- Power spectra (including diffraction spectra of disordered crystals)
- Inter-Spike-Interval histograms and return maps
- ...

Directly from the Mixed-State Presentation (MSP) of any HMM:

- Myopic entropy rates $h_\mu(L)$; asymptotic entropy rate h_μ
- Past–Future Mutual Information (i.e., Excess Entropy $\mathbf{E}$)
- Info shared btw. past and future, but *not* in the present (i.e., Elusive Information σ_μ)
- ...

Directly from MSP of ϵ-machine:

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics
- Correlation functions
- Power spectra (including diffraction spectra of disordered crystals)
- Inter-Spike-Interval histograms and return maps
- ...

Directly from the Mixed-State Presentation (MSP) of any HMM:

- Myopic entropy rates $h_\mu(L)$; asymptotic entropy rate h_μ
- Past–Future Mutual Information (i.e., Excess Entropy $\mathbf{E}$)
- Info shared btw. past and future, but *not* in the present (i.e., Elusive Information σ_μ)
- ...

Directly from MSP of ϵ-machine:

- Average causal-state uncertainty $\mathcal{H}(L)$

Sample of Exact Results Obtained in Closed-Form:

Directly from *any* HMM presentation of a process:

- Dynamics
- Correlation functions
- Power spectra (including diffraction spectra of disordered crystals)
- Inter-Spike-Interval histograms and return maps
- ...

Directly from the Mixed-State Presentation (MSP) of any HMM:

- Myopic entropy rates $h_\mu(L)$; asymptotic entropy rate h_μ
- Past–Future Mutual Information (i.e., Excess Entropy $\mathbf{E}$)
- Info shared btw. past and future, but *not* in the present (i.e., Elusive Information σ_μ)
- ...

Directly from MSP of ϵ-machine:

- Average causal-state uncertainty $\mathcal{H}(L)$
- Synchronization Information $\mathbf{S}$

We covered:

We covered:

- Functions of operators; overview of spectral decomposition

We covered:

- Functions of operators; overview of spectral decomposition
- Dynamics, Correlation Functions, and Power Spectra from any HMM presentation of a process

Previously: Power Spectrum of Even Process

$$P_{\mathrm{c}}(\omega) = \langle\pi| T^{(1)} |\mathbf{1}\rangle + 2 \sum_{\lambda \in \Lambda_T} \mathrm{Re} \frac{\langle\pi| T^{(1)} T_\lambda T^{(1)} |\mathbf{1}\rangle}{e^{i\omega} - \lambda}$$

Previously: Power Spectrum of Even Process

$$P_c(\omega) = \langle \pi | T^{(1)} | \mathbf{1} \rangle + 2 \sum_{\lambda \in \Lambda_T} \mathrm{Re} \frac{\langle \pi | T^{(1)} T_\lambda T^{(1)} | \mathbf{1} \rangle}{e^{i\omega} - \lambda}$$

$$= \frac{2}{3} + 2 \left(\tfrac{4}{9} \mathrm{Re} \frac{1}{e^{i\omega} - 1} - \tfrac{1}{18} \mathrm{Re} \frac{1}{e^{i\omega} + \frac{1}{2}} \right)$$

Previously: Power Spectrum of Even Process

$$P_{\rm c}(\omega) = \langle \pi | \, T^{(1)} \, | \mathbf{1} \rangle + 2 \sum_{\lambda \in \Lambda_T} \mathrm{Re} \frac{\langle \pi | \, T^{(1)} T_\lambda T^{(1)} \, | \mathbf{1} \rangle}{e^{i\omega} - \lambda}$$

$$= \frac{2}{3} + 2 \left(\tfrac{4}{9} \mathrm{Re} \frac{1}{e^{i\omega} - 1} - \tfrac{1}{18} \mathrm{Re} \frac{1}{e^{i\omega} + \frac{1}{2}} \right)$$

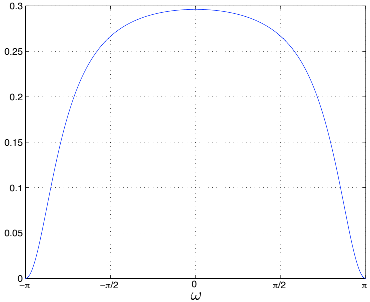
$$= \frac{1}{3} \left(1 - \frac{1}{5 + 4 \cos \omega} \right)$$

Previously: Power Spectrum of Even Process

$$P_c(\omega) = \langle \pi | T^{(1)} | \mathbf{1} \rangle + 2 \sum_{\lambda \in \Lambda_T} \mathrm{Re} \frac{\langle \pi | T^{(1)} T_\lambda T^{(1)} | \mathbf{1} \rangle}{e^{i\omega} - \lambda}$$

$$= \frac{2}{3} + 2 \left(\frac{4}{9} \mathrm{Re} \frac{1}{e^{i\omega} - 1} - \frac{1}{18} \mathrm{Re} \frac{1}{e^{i\omega} + \frac{1}{2}} \right)$$

$$= \frac{1}{3} \left(1 - \frac{1}{5 + 4\cos\omega} \right)$$



If $\mathcal{A} = \{0, 1\}$ and T is diagonalizable, then the power spectrum is simply:

$$P_{\text{c}}(\omega) = \langle \pi | \, T^{(1)} \, |\mathbf{1}\rangle + 2 \sum_{\lambda \in \Lambda_T} \operatorname{Re} \frac{\langle \pi | \, T^{(1)} T_\lambda T^{(1)} \, |\mathbf{1}\rangle}{e^{i\omega} - \lambda} .$$

If $\mathcal{A} = \{0, 1\}$ and T is diagonalizable, then the power spectrum is simply:

$$P_c(\omega) = \langle\pi| T^{(1)} |\mathbf{1}\rangle + 2 \sum_{\lambda \in \Lambda_T} \mathrm{Re} \frac{\langle\pi| T^{(1)} T_\lambda T^{(1)} |\mathbf{1}\rangle}{e^{i\omega} - \lambda} .$$

Consequences:

If $\mathcal{A} = \{0, 1\}$ and T is diagonalizable, then the power spectrum is simply:

$$P_{\mathrm{c}}(\omega) = \langle \pi | \, T^{(1)} \, |\mathbf{1}\rangle + 2 \sum_{\lambda \in \Lambda_T} \mathrm{Re} \frac{\langle \pi | \, T^{(1)} T_\lambda T^{(1)} \, |\mathbf{1}\rangle}{e^{i\omega} - \lambda}.$$

Consequences:

- # of states of any presentation puts upper bound on # of peaks in power spectrum

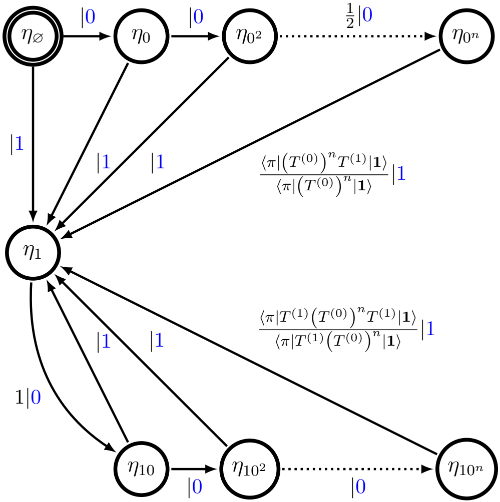
If $\mathcal{A} = \{0, 1\}$ and T is diagonalizable, then the power spectrum is simply:

$$P_c(\omega) = \langle\pi| T^{(1)} |\mathbf{1}\rangle + 2 \sum_{\lambda \in \Lambda_T} \mathrm{Re} \frac{\langle\pi| T^{(1)} T_\lambda T^{(1)} |\mathbf{1}\rangle}{e^{i\omega} - \lambda}.$$

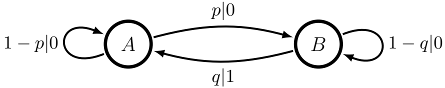
Consequences:

- # of states of any presentation puts upper bound on # of peaks in power spectrum
- # of peaks in power spectrum gives lower bound on # of states of any presentation

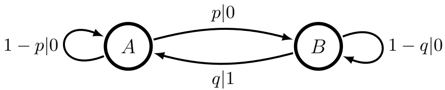
Power Spectrum of SNS

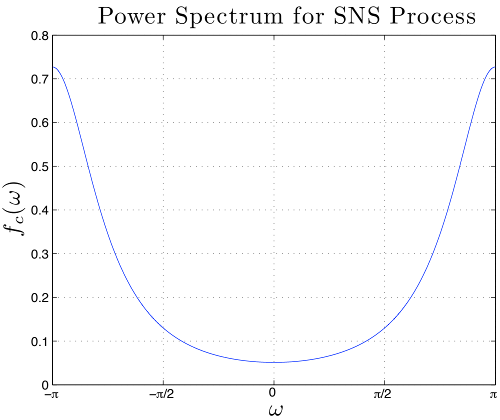


Power Spectrum of SNS

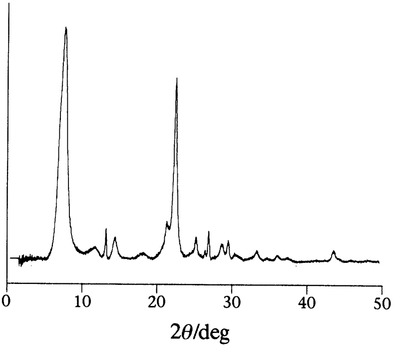


Power Spectrum of SNS





Diffraction Pattern (Power Spectrum of Structure Factors) of Chaotic Crystal



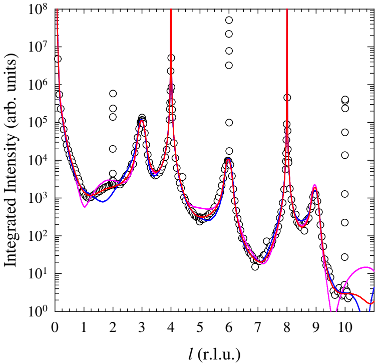
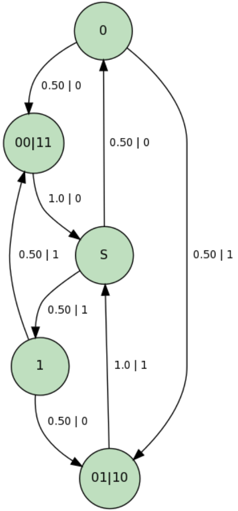


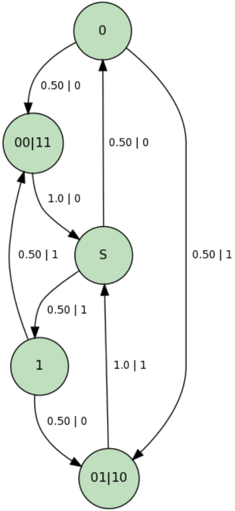
Figure of XRD from typical zeolite beta sample from Treacy et al.

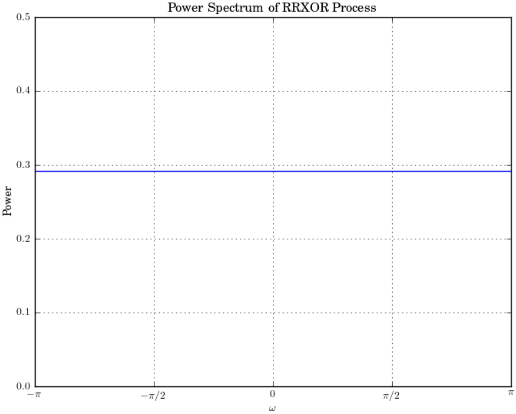
Figure of XRD of multi-layer grapheme on SiC from Hass et al.

RRXOR Process

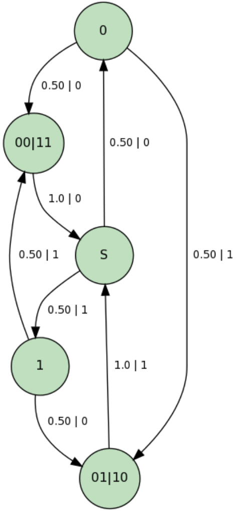


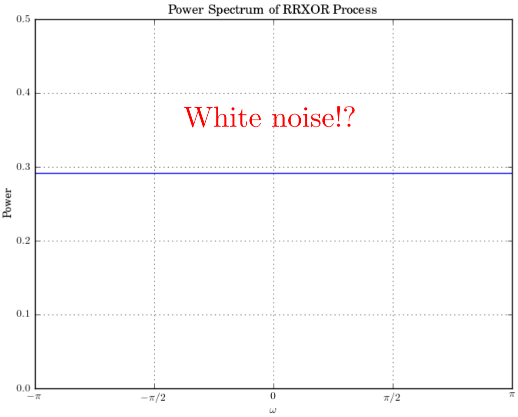
RRXOR Process



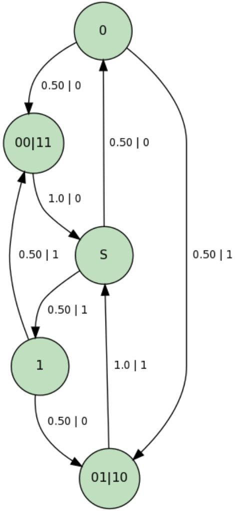


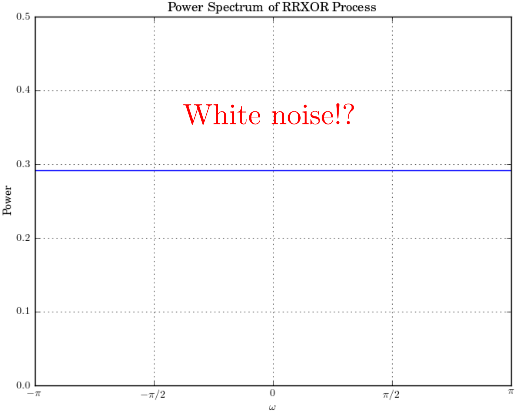
RRXOR Process





RRXOR Process





Regularities unseen; randomness observed.

RRXOR Process

Regularities unseen; randomness observed.
$\rightarrow$ A need for more sophisticated probes of structure: measures of information transduction.

RRXOR Process

Regularities unseen; randomness observed.
$\rightarrow$ A need for more sophisticated probes of structure: measures of information transduction.

- Especially as processes become more complex, the interesting behavior gets processed primarily within very high-order correlations.

Definitions

- The set of *eigenvalues* of a square matrix A is:

$$\Lambda_A \equiv \{\lambda \in \mathbb{C} : \det(\lambda I - A) = 0\}$$

- The set of *eigenvalues* of a square matrix A is:

$$\Lambda_A \equiv \{\lambda \in \mathbb{C} : \det(\lambda I - A) = 0\}$$

- The *projection operator* associated with the eigenvalue λ is:

$$A_\lambda = \text{Res}\left[(zI - A)^{-1}, z \to \lambda\right]$$

$\to A_\lambda = \prod_{\zeta \in \Lambda_A \setminus \{\lambda\}} \frac{A - \zeta I}{\lambda - \zeta}$ if A is diagonalizable.

- The set of *eigenvalues* of a square matrix A is:

$$\Lambda_A \equiv \{\lambda \in \mathbb{C} : \det(\lambda I - A) = 0\}$$

- The *projection operator* associated with the eigenvalue λ is:

$$A_\lambda = \text{Res}\left[(zI - A)^{-1}, z \to \lambda\right]$$

$\to A_\lambda = \prod_{\zeta \in \Lambda_A \setminus \{\lambda\}} \frac{A - \zeta I}{\lambda - \zeta}$ if A is diagonalizable.

- The projection operators are mutually orthogonal:

$$A_\zeta A_\lambda = \delta_{\zeta,\lambda} A_\lambda$$

- The set of *eigenvalues* of a square matrix A is:

$$\Lambda_A \equiv \{\lambda \in \mathbb{C} : \det(\lambda I - A) = 0\}$$

- The *projection operator* associated with the eigenvalue λ is:

$$A_\lambda = \text{Res}\left[(zI - A)^{-1}, z \to \lambda\right]$$

$\to A_\lambda = \prod_{\zeta \in \Lambda_A \setminus \{\lambda\}} \frac{A - \zeta I}{\lambda - \zeta}$ if A is diagonalizable.

- The projection operators are mutually orthogonal:

$$A_\zeta A_\lambda = \delta_{\zeta,\lambda} A_\lambda$$

- For a *stochastic matrix* (i.e., all rows sum to unity), all eigenvalues lie on or within the unit circle; the eigenvalue of unity is guaranteed with $a_1 = g_1 = 1$; single attractor $\to$ $T_1 = |\mathbf{1}\rangle \langle \pi|$.

Reasons for Using Projection Operators

- They play an important role in determining the amplitudes of constituent exponential decays. E.g.:

$$\gamma[\tau] = \sum_{\lambda \in \Lambda_T} \frac{1}{\lambda} \langle \pi | \Big(\sum_{x \in \mathcal{A}} \overline{x} T^{(x)} \Big) T_\lambda \Big(\sum_{x' \in \mathcal{A}} x' T^{(x')} \Big) | \mathbf{1} \rangle \, \lambda^\tau$$

for $\tau > 0$ and diagonalizable T.

- They play an important role in determining the amplitudes of constituent exponential decays. E.g.:

$$\gamma[\tau] = \sum_{\lambda \in \Lambda_T} \frac{1}{\lambda} \langle \pi | \Big(\sum_{x \in \mathcal{A}} \overline{x} T^{(x)} \Big) T_\lambda \Big(\sum_{x' \in \mathcal{A}} x' T^{(x')} \Big) | \mathbf{1} \rangle \, \lambda^\tau$$

  for $\tau > 0$ and diagonalizable T.
- Isolating spectral contributions allows for unusual ease in extracting poorly behaved modes. E.g.:

$$\mathbf{E} = \sum_{\lambda \in \Lambda_W \setminus \{1\}} (\dots)$$

  where $(\dots)$ will soon be revealed.

- They play an important role in determining the amplitudes of constituent exponential decays. E.g.:

$$\gamma[\tau] = \sum_{\lambda \in \Lambda_T} \frac{1}{\lambda} \langle \pi | \Big(\sum_{x \in \mathcal{A}} \overline{x} T^{(x)} \Big) T_\lambda \Big(\sum_{x' \in \mathcal{A}} x' T^{(x')} \Big) | \mathbf{1} \rangle \, \lambda^\tau$$

  for $\tau > 0$ and diagonalizable T.
- Isolating spectral contributions allows for unusual ease in extracting poorly behaved modes. E.g.:

$$\mathbf{E} = \sum_{\lambda \in \Lambda_W \setminus \{1\}} (\dots)$$

  where $(\dots)$ will soon be revealed.
- Degeneracy of eigenvalues unproblematic so long as $a_\lambda = g_\lambda$ (compare to fuss over 'degenerate perturbation theory' although H is always diagonalizable!)

- They play an important role in determining the amplitudes of constituent exponential decays. E.g.:

$$\gamma[\tau] = \sum_{\lambda \in \Lambda_T} \frac{1}{\lambda} \langle \pi | \Big(\sum_{x \in \mathcal{A}} \overline{x} T^{(x)} \Big) T_\lambda \Big(\sum_{x' \in \mathcal{A}} x' T^{(x')} \Big) | \mathbf{1} \rangle \, \lambda^\tau$$

  for $\tau > 0$ and diagonalizable T.
- Isolating spectral contributions allows for unusual ease in extracting poorly behaved modes. E.g.:

$$\mathbf{E} = \sum_{\lambda \in \Lambda_W \setminus \{1\}} (\dots)$$

  where $(\dots)$ will soon be revealed.
- Degeneracy of eigenvalues unproblematic so long as $a_\lambda = g_\lambda$ (compare to fuss over 'degenerate perturbation theory' although H is always diagonalizable!)
- They play nicely with non-diagonalizability and (optimistically) also ∞-state extensions.

The length-L entropy rate

$$h_\mu(L) \equiv H(L) - H(L-1)$$

The length-L entropy rate

$$
\begin{aligned}
h_\mu(L) &\equiv H(L) - H(L-1) \\
&= H[X_{0:L} \,|\, \mu_0 = \pi] - H[X_{0:L-1} \,|\, \mu_0 = \pi]
\end{aligned}
$$

The length-L entropy rate

$$
\begin{aligned}
h_\mu(L) &\equiv H(L) - H(L-1) \\
&= H[X_{0:L} | \mu_0 = \pi] - H[X_{0:L-1} | \mu_0 = \pi] \\
&= H[X_{L-1}, X_{0:L-1} | \mu_0 = \pi] - H[X_{0:L-1} | \mu_0 = \pi]
\end{aligned}
$$

The length-L entropy rate

$$\begin{aligned}
h_\mu(L) &\equiv H(L) - H(L-1) \\
&= H[X_{0:L} \,|\, \mu_0 = \pi] - H[X_{0:L-1} \,|\, \mu_0 = \pi] \\
&= H[X_{L-1}, X_{0:L-1} \,|\, \mu_0 = \pi] - H[X_{0:L-1} \,|\, \mu_0 = \pi] \\
&= H[X_{L-1} \,|\, X_{0:L-1}, \mu_0 = \pi]
\end{aligned}$$

The length-L entropy rate

$$
\begin{aligned}
h_\mu(L) &\equiv H(L) - H(L-1) \\
&= H[X_{0:L} | \mu_0 = \pi] - H[X_{0:L-1} | \mu_0 = \pi] \\
&= H[X_{L-1}, X_{0:L-1} | \mu_0 = \pi] - H[X_{0:L-1} | \mu_0 = \pi] \\
&= H[X_{L-1} | X_{0:L-1}, \mu_0 = \pi] \\
&= H[X_{L-1} | \mathcal{R}_{L-1}, \mathcal{R}_0 = \pi]
\end{aligned}
$$

The length-L entropy rate

$$
\begin{aligned}
h_\mu(L) &\equiv H(L) - H(L-1) \\
&= H[X_{0:L} \,|\, \mu_0 = \pi] - H[X_{0:L-1} \,|\, \mu_0 = \pi] \\
&= H[X_{L-1},\, X_{0:L-1} \,|\, \mu_0 = \pi] - H[X_{0:L-1} \,|\, \mu_0 = \pi] \\
&= H[X_{L-1} \,|\, X_{0:L-1},\, \mu_0 = \pi] \\
&= H[X_{L-1} \,|\, \mathcal{R}_{L-1},\, \mathcal{R}_0 = \pi] \\
&= H\left[X_{L-1} | \left(\mathcal{R}_{L-1} | \mathcal{R}_0 = \pi\right)\right].
\end{aligned}
$$

The states $\eta \in \boldsymbol{\mathcal{R}}$ of the Mixed State Presentation (MSP) are *distributions* over the states $\boldsymbol{\mathcal{S}}$ of another presentation. Starting from the MSP's start state—the stationary distribution π—the MSP gives the observation-induced evolution of probability density over distributions-over-$\boldsymbol{\mathcal{S}}$.

The initial distribution over MSP—assuming no prior observations of the known process—is denoted δ_π with $\langle \delta_\pi | = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \end{bmatrix}$.

The transition matrix W of the MSP is a stochastic matrix with $W_1 = |\mathbf{1}\rangle \langle \pi_W|$.

The length-L entropy rate

$$h_\mu(L) = H\left[X_{L-1} \middle| \left(\mathcal{R}_{L-1} \middle| \mathcal{R}_0 = \pi\right)\right]$$

The length-L entropy rate

$$
\begin{aligned}
h_\mu(L) &= H\left[X_{L-1} \middle| \left(\mathcal{R}_{L-1} | \mathcal{R}_0 = \pi\right)\right] \\
&= \sum_{\eta \in \boldsymbol{\mathcal{R}}} \Pr(\mathcal{R}_{L-1} = \eta | \mathcal{R}_0 = \pi)\, H[X_{L-1} |\, \mathcal{R}_{L-1} = \eta]
\end{aligned}
$$

The length-L entropy rate

$$
\begin{aligned}
h_\mu(L) &= H\Big[X_{L-1} \big| \big(\mathcal{R}_{L-1} | \mathcal{R}_0 = \pi\big)\Big] \\
&= \sum_{\eta \in \boldsymbol{\mathcal{R}}} \Pr(\mathcal{R}_{L-1} = \eta | \mathcal{R}_0 = \pi)\, H[X_{L-1} | \, \mathcal{R}_{L-1} = \eta] \\
&= \sum_{\eta \in \boldsymbol{\mathcal{R}}} \langle \delta_\pi | \, W^{L-1} \, | \eta \rangle \times - \sum_{x \in \mathcal{A}} \langle \delta_\eta | \, W^{(x)} \, | \mathbf{1} \rangle \, \log_2\Big(\langle \delta_\eta | \, W^{(x)} \, | \mathbf{1} \rangle \Big)
\end{aligned}
$$

The length-L entropy rate

$$\begin{aligned}
h_\mu(L) &= H\Big[X_{L-1} \big| \big(\mathcal{R}_{L-1} | \mathcal{R}_0 = \pi\big)\Big] \\
&= \sum_{\eta \in \boldsymbol{\mathcal{R}}} \Pr(\mathcal{R}_{L-1} = \eta | \mathcal{R}_0 = \pi)\, H[X_{L-1} |\, \mathcal{R}_{L-1} = \eta] \\
&= \sum_{\eta \in \boldsymbol{\mathcal{R}}} \langle \delta_\pi |\, W^{L-1} \,| \eta \rangle \times - \sum_{x \in \mathcal{A}} \langle \delta_\eta |\, W^{(x)} \,| \mathbf{1} \rangle \, \log_2 \Big(\langle \delta_\eta |\, W^{(x)} \,| \mathbf{1} \rangle \Big) \\
&= \langle \delta_\pi |\, W^{L-1} \left[- \sum_{\eta \in \boldsymbol{\mathcal{R}}} | \delta_\eta \rangle \sum_{x \in \mathcal{A}} \langle \delta_\eta |\, W^{(x)} \,| \mathbf{1} \rangle \, \log_2 \Big(\langle \delta_\eta |\, W^{(x)} \,| \mathbf{1} \rangle \Big) \right]
\end{aligned}$$

The length-L entropy rate

$$
\begin{aligned}
h_\mu(L) &= H\left[X_{L-1}\middle|\left(\mathcal{R}_{L-1}\middle|\mathcal{R}_0 = \pi\right)\right] \\
&= \sum_{\eta \in \boldsymbol{\mathcal{R}}} \Pr(\mathcal{R}_{L-1} = \eta | \mathcal{R}_0 = \pi)\, H[X_{L-1} | \mathcal{R}_{L-1} = \eta] \\
&= \sum_{\eta \in \boldsymbol{\mathcal{R}}} \langle \delta_\pi | W^{L-1} | \eta \rangle \times - \sum_{x \in \mathcal{A}} \langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle \log_2 \left(\langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle \right) \\
&= \langle \delta_\pi | W^{L-1} \left[- \sum_{\eta \in \boldsymbol{\mathcal{R}}} | \delta_\eta \rangle \sum_{x \in \mathcal{A}} \langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle \log_2 \left(\langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle \right) \right] \\
&= \langle \delta_\pi | W^{L-1} | H(W^{\mathcal{A}}) \rangle ,
\end{aligned}
$$

where

$$
| H(W^{\mathcal{A}}) \rangle \equiv - \sum_{\eta \in \boldsymbol{\mathcal{R}}} | \delta_\eta \rangle \sum_{x \in \mathcal{A}} \langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle \log_2 \left(\langle \delta_\eta | W^{(x)} | \mathbf{1} \rangle \right)
$$

The length-L entropy rate

If W is diagonalizable, then

$$\begin{aligned} \langle \delta_\pi | W^{L-1} &= \langle \delta_\pi | \sum_{\lambda \in \Lambda_W} \lambda^{L-1} W_\lambda \\ &= \sum_{\lambda \in \Lambda_W} \lambda^{L-1} \langle \delta_\pi | W_\lambda \end{aligned}$$

where $\{W_\lambda\}$ can by obtained via

$$W_\lambda = \prod_{\substack{\zeta \in \Lambda_W \\ \zeta \neq \lambda}} \frac{W - \zeta I}{\lambda - \zeta}.$$

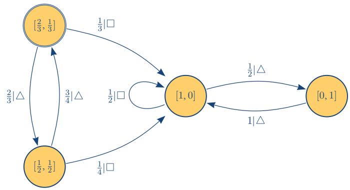
The length-L entropy rate

Hence, for diagonalizable W, the length-L entropy rate can be rewritten as

$$\begin{aligned} h_\mu(L) &= \langle \delta_\pi | \, W^{L-1} \, | H(W^{\mathcal{A}}) \rangle \\ &= \sum_{\lambda \in \Lambda_W} \lambda^{L-1} \, \langle \delta_\pi | \, W_\lambda \, | H(W^{\mathcal{A}}) \rangle \, , \end{aligned}$$

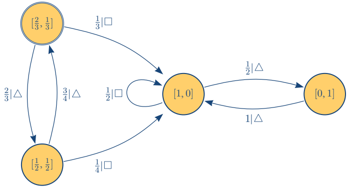
which is a closed-form expression for $h_\mu(L)$ with scalar exponentiation as the only L-dependence.

Example Strictly Sofic Processes



$\mathcal{M}_{\text{msp}}$

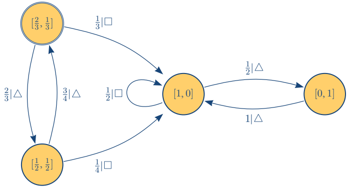
Example Strictly Sofic Processes



$\mathcal{M}_{\text{msp}}$

$$W = \begin{bmatrix} 0 & 2/3 & 1/3 & 0 \\ 3/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Example Strictly Sofic Processes



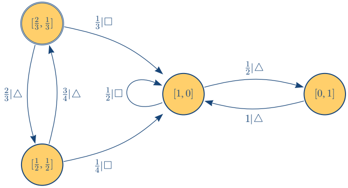
$$W = \begin{bmatrix} 0 & 2/3 & 1/3 & 0 \\ 3/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

$\mathcal{M}_{\text{msp}}$

$\det(\lambda I - W) =$

$(\lambda - 1)(\lambda + \frac{1}{2})(\lambda^2 - \frac{1}{2}) = 0$

Example Strictly Sofic Processes



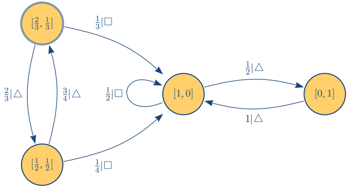
$\mathcal{M}_{\text{msp}}$

$$W = \begin{bmatrix} 0 & 2/3 & 1/3 & 0 \\ 3/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

$$\det(\lambda I - W) = (\lambda - 1)(\lambda + \tfrac{1}{2})(\lambda^2 - \tfrac{1}{2}) = 0$$

$$\Lambda_W = \left\{1,\ \frac{\sqrt{2}}{2},\ -\frac{\sqrt{2}}{2},\ -\frac{1}{2}\right\}$$

Example Strictly Sofic Processes



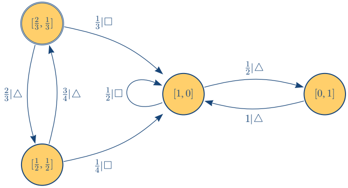
$\mathcal{M}_{\text{msp}}$

$$W = \begin{bmatrix} 0 & 2/3 & 1/3 & 0 \\ 3/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

$$\det(\lambda I - W) = (\lambda - 1)(\lambda + \tfrac{1}{2})(\lambda^2 - \tfrac{1}{2}) = 0$$

$$\Lambda_W = \left\{1,\ \frac{\sqrt{2}}{2},\ -\frac{\sqrt{2}}{2},\ -\frac{1}{2}\right\} \approx \{1,\ 0.707,\ -0.707,\ -0.5\}$$

Example Strictly Sofic Processes



$\mathcal{M}_{\text{msp}}$

$$W = \begin{bmatrix} 0 & ^2/_3 & ^1/_3 & 0 \\ ^3/_4 & 0 & ^1/_4 & 0 \\ 0 & 0 & ^1/_2 & ^1/_2 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

$$\det(\lambda I - W) = (\lambda - 1)(\lambda + \tfrac{1}{2})(\lambda^2 - \tfrac{1}{2}) = 0$$

$$\Lambda_W = \left\{1,\ \frac{\sqrt{2}}{2},\ -\frac{\sqrt{2}}{2},\ -\frac{1}{2}\right\} \approx \{1,\ 0.707,\ -0.707,\ -0.5\}$$

$$\langle \delta_\pi | = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}.$$

Example Strictly Sofic Processes

λ	$\langle \delta_\pi \vert W_\lambda$			
1	$[0$	0	$\frac{2}{3}$	$\frac{1}{3}]$
$\sqrt{2}/2$	$[\frac{1}{2}$	$\frac{\sqrt{2}}{3}$	$\frac{-2-\sqrt{2}}{6}$	$\frac{-\sqrt{2}-1}{6}]$
$-\sqrt{2}/2$	$[\frac{1}{2}$	$\frac{-\sqrt{2}}{3}$	$\frac{-2+\sqrt{2}}{6}$	$\frac{\sqrt{2}-1}{6}]$
$-1/2$	$[0$	0	0	$0]$

$$|H(W^{\mathcal{A}})\rangle = \begin{bmatrix} \log_2(3) - 2/3 \\ 2 - \frac{3}{4}\log_2(3) \\ 1 \\ 0 \end{bmatrix}$$

Example Strictly Sofic Processes

$$h_\mu(L) = \sum_{\lambda \in \Lambda_W} \lambda^{L-1} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \rangle$$

Example Strictly Sofic Processes

$$
\begin{aligned}
h_\mu(L) &= \sum_{\lambda \in \Lambda_W} \lambda^{L-1} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \\
&= \frac{2}{3} + \left(\frac{\sqrt{2}}{2}\right)^{L-1} \left(\frac{1}{2}\log_2(3) - \frac{\sqrt{2}}{4}\log_2(3) - \frac{2}{3} + \frac{\sqrt{2}}{2}\right) \\
&\quad + \left(-\frac{\sqrt{2}}{2}\right)^{L-1} \left(\frac{1}{2}\log_2(3) + \frac{\sqrt{2}}{4}\log_2(3) - \frac{2}{3} - \frac{\sqrt{2}}{2}\right)
\end{aligned}
$$

Example Strictly Sofic Processes

$$h_\mu(L) = \sum_{\lambda \in \Lambda_W} \lambda^{L-1} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \rangle$$

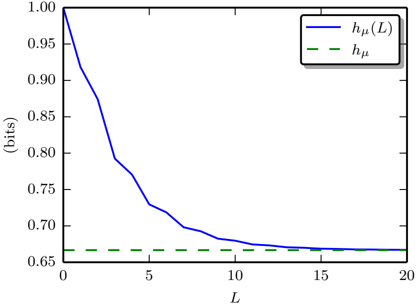
$$= \frac{2}{3} + \left(\frac{\sqrt{2}}{2}\right)^{L-1} \left(\frac{1}{2}\log_2(3) - \frac{\sqrt{2}}{4}\log_2(3) - \frac{2}{3} + \frac{\sqrt{2}}{2}\right)$$

$$+ \left(-\frac{\sqrt{2}}{2}\right)^{L-1} \left(\frac{1}{2}\log_2(3) + \frac{\sqrt{2}}{4}\log_2(3) - \frac{2}{3} - \frac{\sqrt{2}}{2}\right)$$

$$= \begin{cases} \frac{2}{3} + \left(\frac{\sqrt{2}}{2}\right)^{L-1} \left(-\frac{\sqrt{2}}{2}\log_2(3) + \sqrt{2}\right) & \text{for even } L \\ \frac{2}{3} + \left(\frac{\sqrt{2}}{2}\right)^{L-1} \left(\log_2(3) - \frac{4}{3}\right) & \text{for odd } L \end{cases}$$

Example Strictly Sofic Processes

$$h_\mu(L) = \begin{cases} \frac{2}{3} + \left(\frac{\sqrt{2}}{2}\right)^{L-1} \left(-\frac{\sqrt{2}}{2}\log_2(3) + \sqrt{2}\right) & \text{for even } L \\ \frac{2}{3} + \left(\frac{\sqrt{2}}{2}\right)^{L-1} \left(\log_2(3) - \frac{4}{3}\right) & \text{for odd } L \end{cases}$$



Excess Entropy

For diagonalizable W,

$$\mathbf{E} \equiv \sum_{L=1}^{\infty} [h_\mu(L) - h_\mu]$$

Excess Entropy

For diagonalizable W,

$$\begin{aligned}\mathbf{E} &\equiv \sum_{L=1}^{\infty} [h_\mu(L) - h_\mu] \\ &= \sum_{L=1}^{\infty} \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda|<1}} \lambda^{L-1} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle\end{aligned}$$

Excess Entropy

For diagonalizable W,

$$\begin{aligned}
\mathbf{E} &\equiv \sum_{L=1}^{\infty} \left[h_\mu(L) - h_\mu \right] \\
&= \sum_{L=1}^{\infty} \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \lambda^{L-1} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \\
&= \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \underbrace{\sum_{L=1}^{\infty} \lambda^{L-1}}_{= \sum_{L=0}^{\infty} \lambda^L = \frac{1}{1-\lambda}}
\end{aligned}$$

Excess Entropy

For diagonalizable W,

$$\begin{aligned}
\mathbf{E} &\equiv \sum_{L=1}^{\infty} \left[h_\mu(L) - h_\mu \right] \\
&= \sum_{L=1}^{\infty} \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \lambda^{L-1} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \\
&= \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle \underbrace{\sum_{L=1}^{\infty} \lambda^{L-1}}_{=\sum_{L=0}^{\infty} \lambda^L = \frac{1}{1-\lambda}} \\
&= \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{1-\lambda} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle
\end{aligned}$$

Alternatively, for any non-diagonalizable W,

$$\begin{aligned}
\mathbf{E} &\equiv \sum_{L=1}^{\infty} \left[h_\mu(L) - h_\mu \right] \\
&= \sum_{L=1}^{\infty} \left[\langle \delta_\pi | \, W^{L-1} \, | H(W^{\mathcal{A}}) \rangle - \langle \delta_\pi | \, W_1 \, | H(W^{\mathcal{A}}) \rangle \right] \\
&= \sum_{L=0}^{\infty} \left[\langle \delta_\pi | \, W^{L} \, | H(W^{\mathcal{A}}) \rangle - \langle \delta_\pi | \, W_1 \, | H(W^{\mathcal{A}}) \rangle \right] \\
&= \sum_{L=0}^{\infty} \langle \delta_\pi | \left[(\underbrace{W - W_1}_{\equiv Q})^L - \delta_{L,0} W_1 \right] | H(W^{\mathcal{A}}) \rangle \\
&= - \underbrace{\langle \delta_\pi | \, W_1 \, | H(W^{\mathcal{A}}) \rangle}_{= \langle \pi_W | H(W^{\mathcal{A}}) \rangle = h_\mu} + \sum_{L=0}^{\infty} \underbrace{\langle \delta_\pi | \, Q^L}_{= \langle \delta_\pi | \mathcal{Q}^L} \, | H(W^{\mathcal{A}}) \rangle \\
&= \langle \delta_\pi | \left(\sum_{L=0}^{\infty} \mathcal{Q}^L \right) | H(W^{\mathcal{A}}) \rangle - h_\mu \\
&= \langle \delta_\pi | \, (I - \mathcal{Q})^{-1} \, | H(W^{\mathcal{A}}) \rangle - h_\mu \\
&= \langle \delta_\pi | \, (I - Q)^{-1} \, | H(W^{\mathcal{A}}) \rangle - h_\mu \; .
\end{aligned}$$

Similar statements result for all of the familiar complexity measures; **T**, **S**, and so on.

Interestingly, they all appear more alike than ever, yielding new insight for how they are all related. E.g.:

$$\mathbf{E} = \langle \delta_\pi | \left(I - Q\right)^{-1} | H(W^{\mathcal{A}}) \rangle - h_\mu ,$$

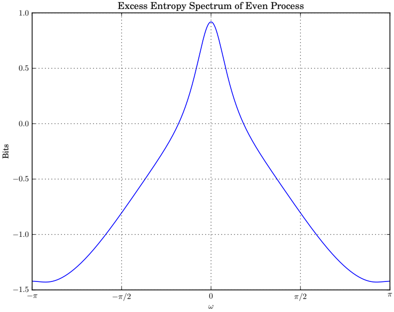
$$\mathbf{S} = \langle \delta_\pi | \left(I - Q\right)^{-1} | H[\eta] \rangle - \mathcal{H} ,$$

and

$$\mathbf{T} = \langle \delta_\pi | \left(I - Q\right)^{-2} | H(W^{\mathcal{A}}) \rangle - h_\mu .$$

Complexity Measures Anew: Complexity Spectra

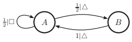
$$\mathbf{E}(\omega) \equiv \mathrm{Re} \langle \delta_\pi | \left(e^{i\omega} I - Q\right)^{-1} | H(W^{\mathcal{A}}) \rangle - h_\mu$$



Thank you!

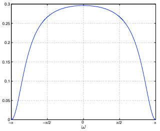
Example: Complexity Measures for a Prototypic Sofic System: The Even Process

$\mathcal{M} = (\mathcal{A} = \{\square, \triangle\}, \boldsymbol{\mathcal{S}} = \{A, B\}, T^{\mathcal{A}})$.

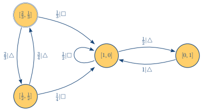


For $\square = 0$ and $\triangle = 1$, we find the continuous part of the power spectrum to be:

$$P_c(\omega) = \frac{1}{3}\left(1 - \frac{1}{5 + 4\cos\omega}\right).$$



$\mathcal{M}_{\text{MSP}} = (\mathcal{A} = \{\square, \triangle\}, \mathcal{R}, W^{\mathcal{A}})$, with $\mathcal{R} = \{(\frac{2}{3}, \frac{1}{3}), (\frac{1}{2}, \frac{1}{2}), (1, 0), (0, 1)\}$.



$$W = \sum_{x \in \mathcal{A}} W^{(x)} = \begin{bmatrix} 0 & 2/3 & 1/3 & 0 \\ 3/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

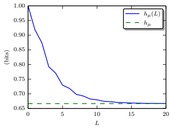
Solving $\det(\lambda I - W) = 0$ gives W's eigenvalues:

$$\Lambda_W = \left\{1, \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, -\frac{1}{2}\right\}.$$

$$\mathcal{H}(L) = \begin{cases} (\log_2(3) - 2/3)\left(\frac{\sqrt{2}}{2}\right)^L & \text{for even } L \\ \frac{2\sqrt{2}}{3}\left(\frac{\sqrt{2}}{2}\right)^L & \text{for odd } L \end{cases}.$$

$$h_\mu(L) = \begin{cases} \frac{2}{3} + \left(\frac{\sqrt{2}}{2}\right)^{L-1}\left(-\frac{\sqrt{2}}{2}\log_2(3) + \sqrt{2}\right) & \text{for even } L \\ \frac{2}{3} + \left(\frac{\sqrt{2}}{2}\right)^{L-1}\left(\log_2(3) - \frac{4}{3}\right) & \text{for odd } L \end{cases}.$$

$h_\mu = 2/3$ bits per step,
$C_\mu = \log_2(3) - \frac{2}{3}$ bits,
$\mathbf{E} = \log_2(3) - \frac{2}{3}$ bits,
$\mathbf{T} = 2\log_2(3)$ bits-symbols,
$\mathbf{S} = 2\log_2(3)$ bits.



In short:

$$\begin{aligned} h_\mu(L) &= H(L) - H(L-1) \\ &= H[X_{L-1} | X_{0:L-1}] \\ &= H\big[X_{L-1} | \big(\mathcal{R}_{L-1} | \mathcal{R}_0 = \pi\big)\big] \\ &= \langle \delta_\pi | \, W^{L-1} \, | H(W^{\mathcal{A}}) \rangle \, , \end{aligned}$$

where
$|H(W^{\mathcal{A}})\rangle = -\sum_{\eta \in \boldsymbol{\mathcal{R}}} |\delta_\eta\rangle \sum_{x \in \mathcal{A}} \langle \delta_\eta | \, W^{(x)} \, |\mathbf{1}\rangle \log_2 \langle \delta_\eta | \, W^{(x)} \, |\mathbf{1}\rangle$.

$$h_\mu(L) = \langle \delta_\pi | \, W_1 \, | H(W^{\mathcal{A}}) \rangle + \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \lambda^{L-1} \langle \delta_\pi | \, W_\lambda \, | H(W^{\mathcal{A}}) \rangle$$

for diagonalizable W.