# Dynamical Systems and Reliable Communication

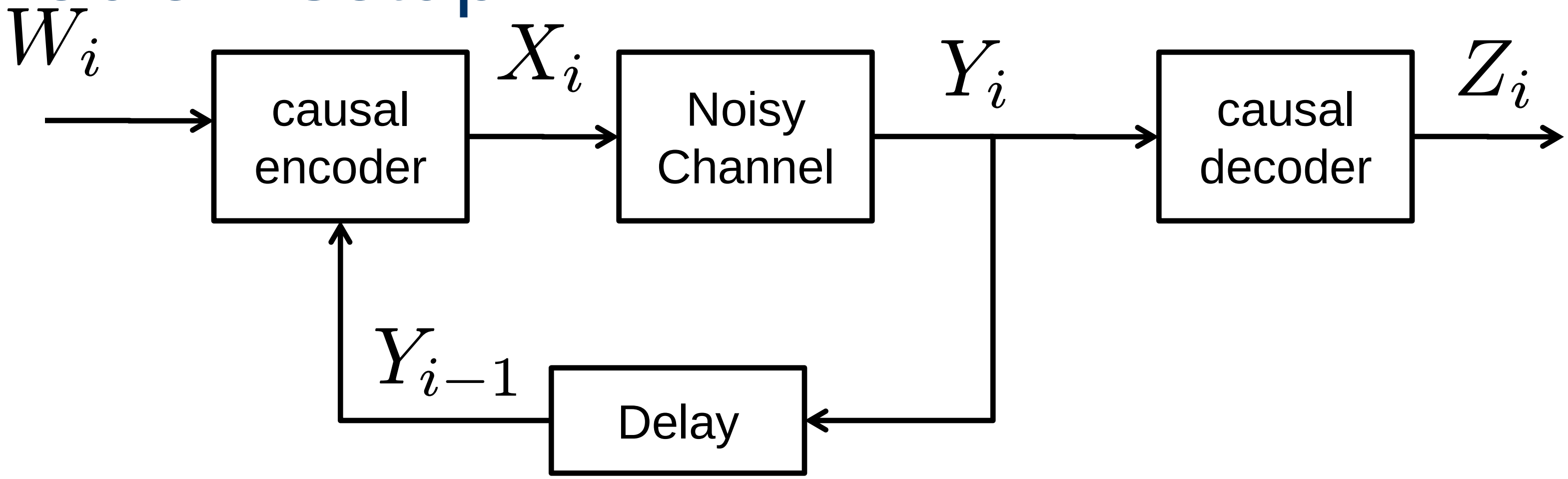
Todd P. Coleman, UIUC

January 26, 2011

- Typical optimal control problems have
  - a) time-horizon independent alphabets
  - b) sequential decision making
  - c) objective: minimize a sum of additive costs
- Typical information theory problems have
  - a) time-horizon dependent alphabets (e.g. one of 2^{nR} hypotheses)
  - b) decision-making is once at the end (decode msg at end of transmission)
  - c) objective: extremize mutual information.
- Observation:

$$
\begin{aligned}
I(W;Y^n) &= \sum_{i=1}^{n} \mathbb{E}\left[\log \frac{f_{W|Y^i}(W|Y^i)}{f_{W|Y^{i-1}}(W|Y^{i-1})}\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\log \frac{d\pi_i}{d\pi_{i-1}}(W)\right]
\end{aligned}
$$

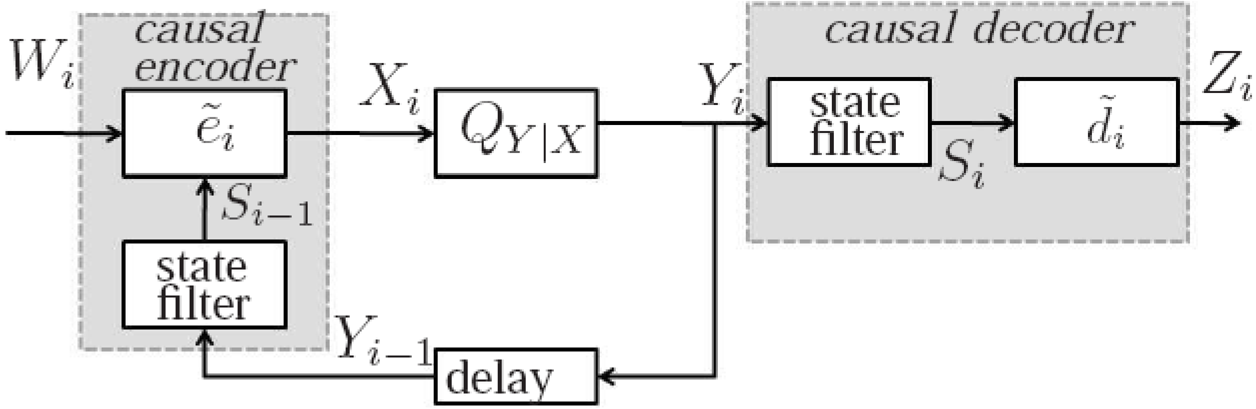
# Problem setup




- W is a Markov process
- Causal encoder $X_i = e_i(W^i, Y^{i-1})$
- Causal decoder $Z_i = d_i(Y^i)$
- Cost: $$J_n(e,d) = \mathbb{E}_{e,d}\left[\frac{1}{n}\sum_{i=1}^{n}\rho(W_i, Z_{i-1}, Z_i) + \alpha\eta(X_i)\right]$$
- Objective: find optimal causal encoder/decoders

$$J_n(e^*, d^*) \leq J_n(e,d) \text{ for all } e, d$$

# Result 1 (Gorantla, Coleman '11)

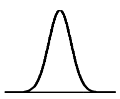



**Theorem (structural result): there always exist optimal $(\mu^*, d^*)$ of form**

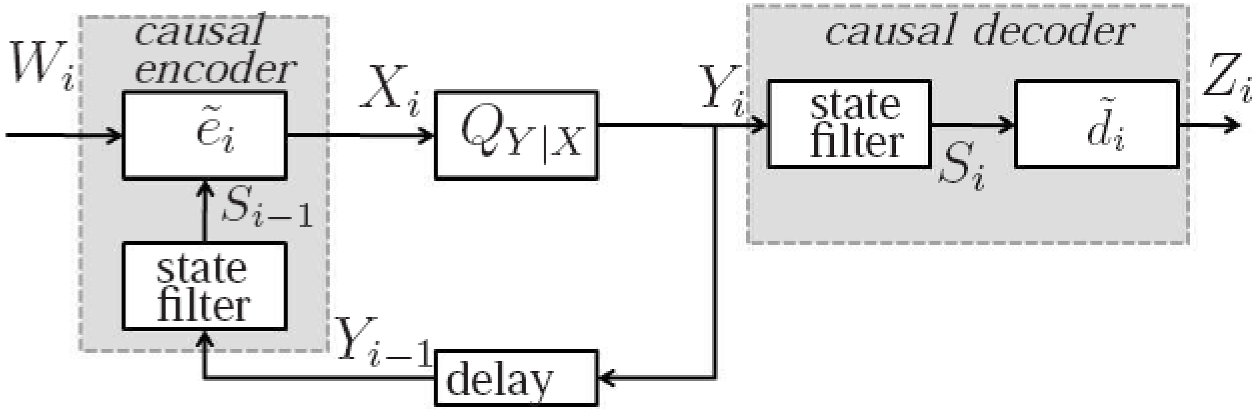
$$X_i = e_i^*(W^i, Y^{i-1}) \equiv e_i^*(W_i, \pi_{i-1}, Z_{i-1})$$

$$Z_i = d_i^*(Y^i) \equiv \tilde{d}_i^*\left(Z_{i-1}, \pi_i\right),$$

$$\pi_i(A) \triangleq \mathbb{P}\left(W_i \in A | Y^i\right)$$



# Result 1 (Gorantla, Coleman '11)

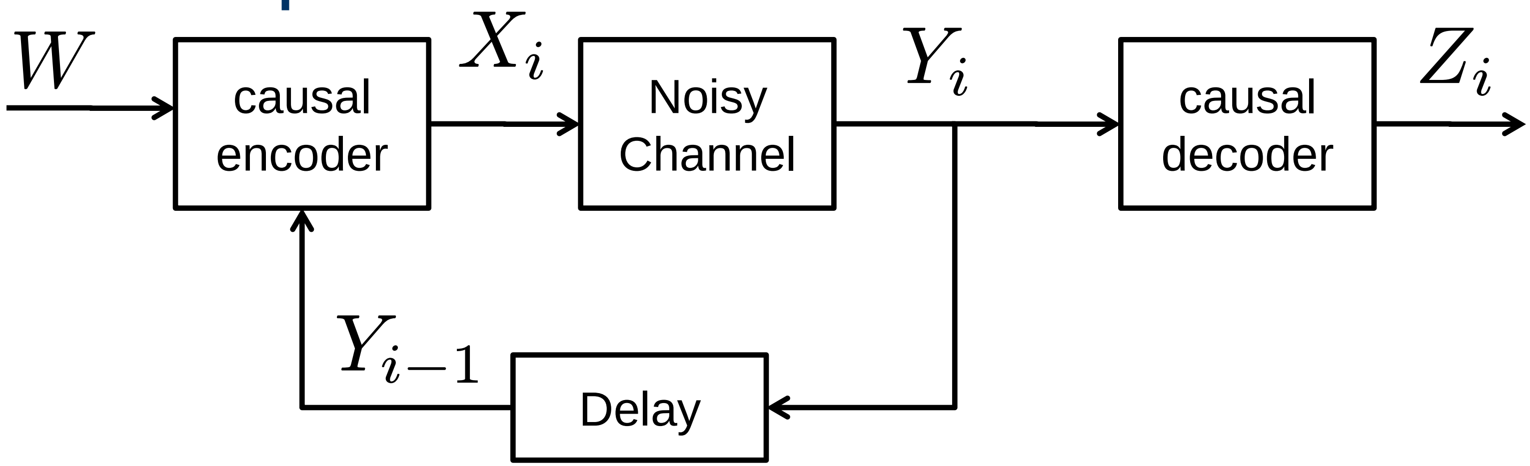



**Theorem (structural result): there always exist optimal $(\mu^*, d^*)$ of form**

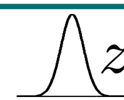
$$
\begin{aligned}
X_i = \mu_i^*(W_i, Y^{i-1}) \quad &\equiv \quad \tilde{\mu}_i^*(W_i, \pi_{i-1}, Z_{i-1}) \\
Z_i = d_i^*(Y^i) \quad &\equiv \quad \tilde{d}_i^*\left(Z_{i-1}, \pi_i\right), \\
\pi_i(A) \quad &\triangleq \quad \mathbb{P}\left(W_i \in A | Y^i\right)
\end{aligned}
$$

- A form of separation theorem (sufficient statistics)
- W need not be stationary nor ergodic (mobility)
- Proof uses dynamic programming argument

# An Example



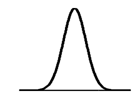

$$
\begin{aligned}
I(W;Y^n) &= \sum_{i=1}^{n} \mathbb{E}\left[\log \frac{f_{W|Y^i}(W|Y^i)}{f_{W|Y^{i-1}}(W|Y^{i-1})}\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\log \frac{d\pi_i}{d\pi_{i-1}}(W)\right]
\end{aligned}
$$

- Alphabets $\quad \mathcal{W} = [0,1], \quad \mathcal{Z} = \mathcal{M}([0,1])$ 
- W is Markov $\quad (W_i = W_{i-1} : i \geq 1), \quad W_0 \sim \mathrm{unif}[0,1]$

# The Sequential Information Gain Cost

$$I(W;Y^n) = \sum_{i=1}^{n} \mathbb{E}\left[\log \frac{f_{W|Y^i}(W|Y^i)}{f_{W|Y^{i-1}}(W|Y^{i-1})}\right]$$

$$= \sum_{i=1}^{n} \mathbb{E}\left[\log \frac{d\pi_i}{d\pi_{i-1}}(W)\right]$$

- Z lies in the space of beliefs on W $\mathcal{Z} = \mathcal{P}(\mathcal{W})$ 
- Sequential information gain cost

$$\rho(w_i, z_i, z_{i-1}) = -\log \frac{dz_i}{dz_{i-1}}(w_i)$$

- Theorem:
  - For any encoder, optimal decoder is given by

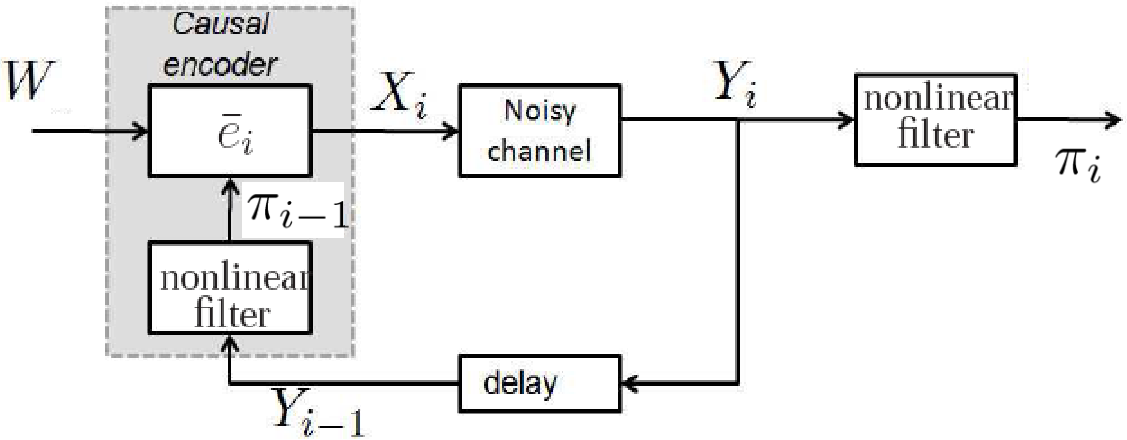
$$Z_i = \pi_i$$

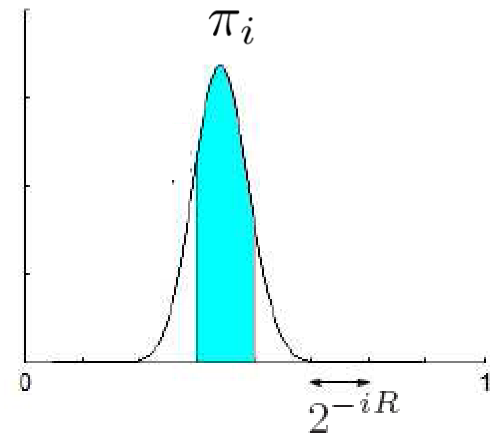
  - Total optimal cost is a constrained mutual information

$$I(W^n; Y^n)$$

- Proof: uses DP and 2nd law of thermodynamics

# Feedback communication: posterior matching



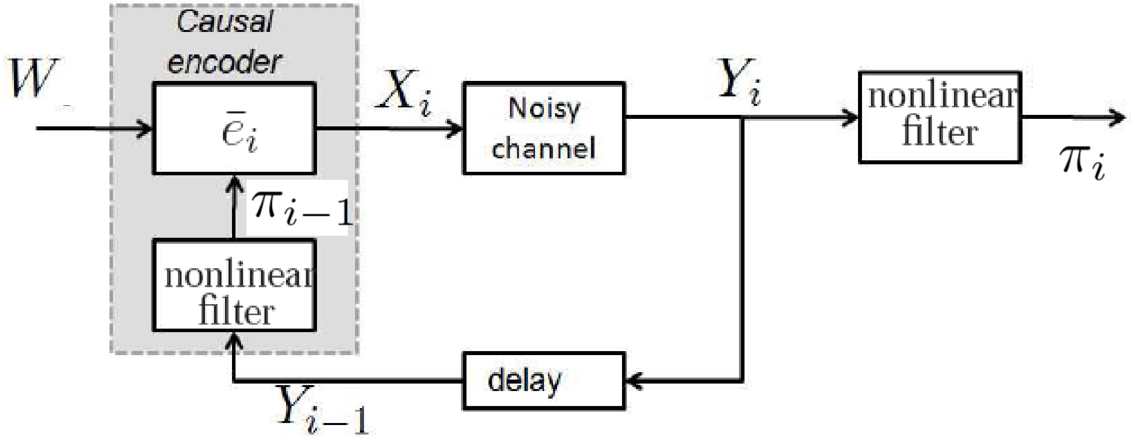



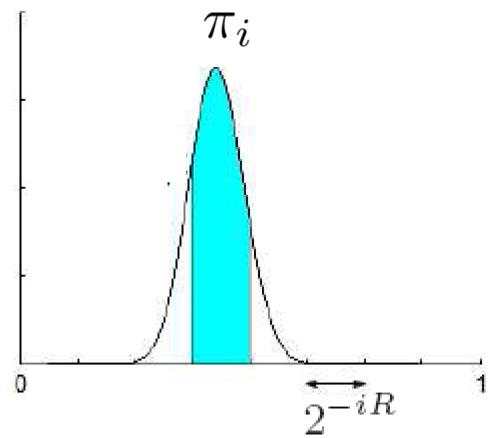

$$\pi_i(\cdot) = \mathbb{P}(W \in \cdot | Y^i)$$

$$X_i = F_X^{-1}(\underbrace{\pi_{i-1}([0, W])}_{\text{what is missing}})$$

- Non-standard notion of communication
  - Message point W on [0,1] line
  - No block length, no FEC, no pre-specified rate
  - Achieves capacity on arbitrary memoryless channels
  - Optimal solution to our problem: $\frac{1}{n} I(W; Y^n) = C$ for **all** $n$!

# Achievability: HMMs and the stability of the nonlinear filter




$$\pi_n(dw) = \frac{dP_{Y|X}(\cdot|\bar{e}(w,\pi_{i-1}))}{dP_Y(\cdot)}(y_i)\pi_{n-1}(dw)$$

Subject to initial condition $\pi_0 = \nu$

- $\nu$: uniform over blue region
- $\bar{\nu}$: uniform over $[0,1]$

Achievability occurs if in $\bar{P}$, $D\left(\pi_n \| \bar{\pi}_n\right) \to 0$

*stability* of the nonlinear filter

## Nec & Suf Conditions on Achievability based on Stability of NLF

- Optimal solution to sequential information gain problem: (X,Y) form a Markov chain
- As a consequence:

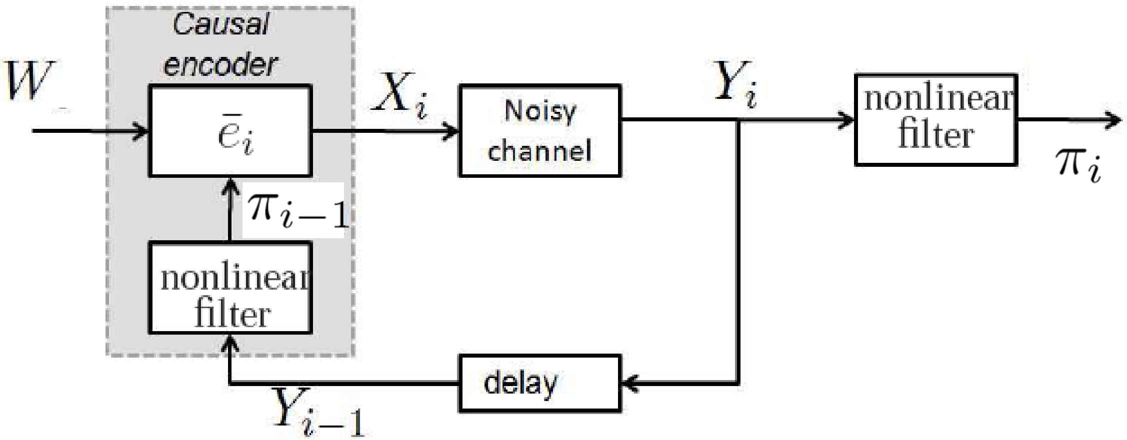
$$
\begin{aligned}
D(\pi_n \| \bar{\pi}_n) &= \mathbb{E}\left[\log \overline{\mathbb{E}}\left[\frac{d\nu}{d\bar{\nu}}(W) | \mathcal{F}^Y_{1,\infty} \vee \sigma(W)\right] \Big| \mathcal{F}^Y_{1,n}\right] \\
&- \log \overline{\mathbb{E}}\left[\frac{d\nu}{d\bar{\nu}}(W) | \mathcal{F}^Y_{1,n}\right] \\
&= -\log(\epsilon) - \log \overline{\mathbb{E}}\left[\frac{d\nu}{d\bar{\nu}}(W) | \mathcal{F}^Y_{1,n}\right]
\end{aligned}
$$

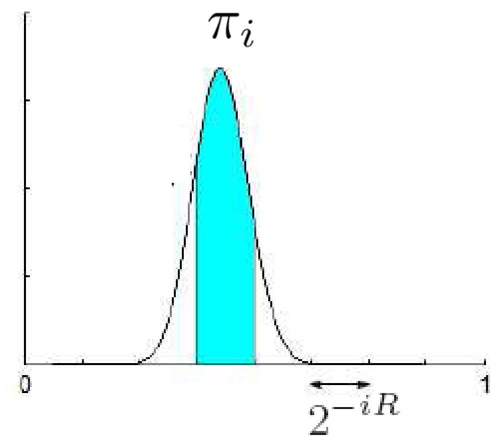
- Theorem: achievability occurs if and only if

$$
\overline{\mathbb{E}}\left[\frac{d\nu}{d\bar{\nu}}(W) | \mathcal{F}^Y_{1,\infty}\right] = \overline{\mathbb{E}}\left[\frac{d\nu}{d\bar{\nu}}(W) | \mathcal{F}^Y_{1,\infty} \vee \sigma(W)\right]
$$

- Generalizes work of Shayevitz & Feder.
    - Sufficient conditions: X is ergodic and channel not degenerate.
- Relates message point communication to stability to nonlinear filter

# Converses Based on Dynamical Systems



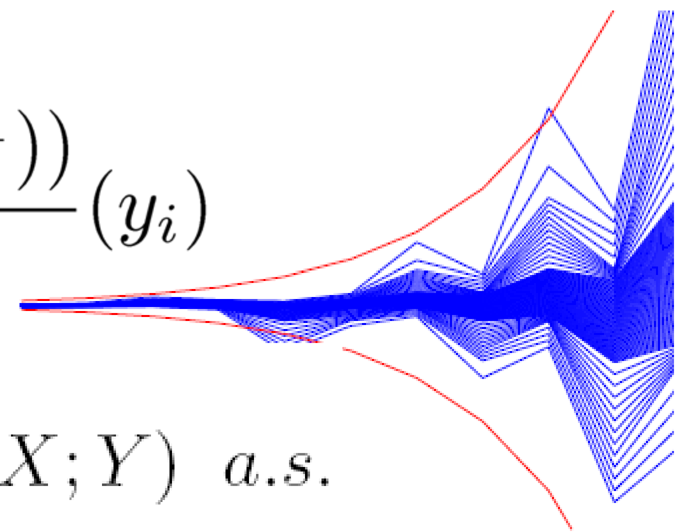




– Motivation: under the PM scheme:

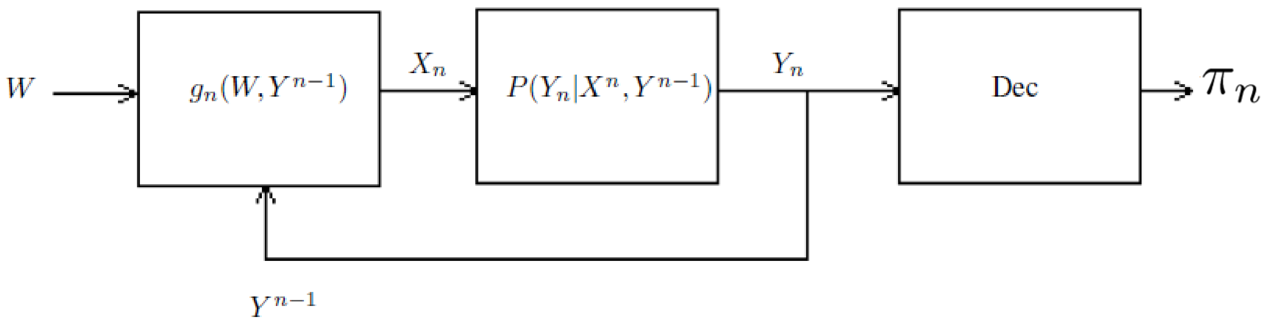
$$X_i = g_i(W, Y^{i-1})$$

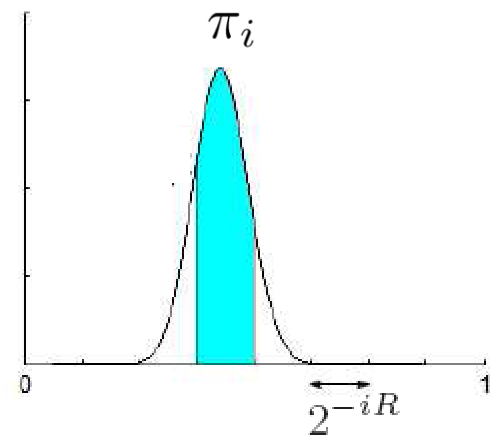
$$\pi_n(dw) = \prod_{i=1}^{n} \frac{dP_{Y|X}(\cdot|g_i(w, y^{i-1}))}{dP_Y(\cdot)}(y_i)$$

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{\partial g_n(w, Y^{n-1})}{\partial w}\Big|_{w=W} = I(X;Y) \quad a.s.$$



# Converses Based on Dynamical Systems



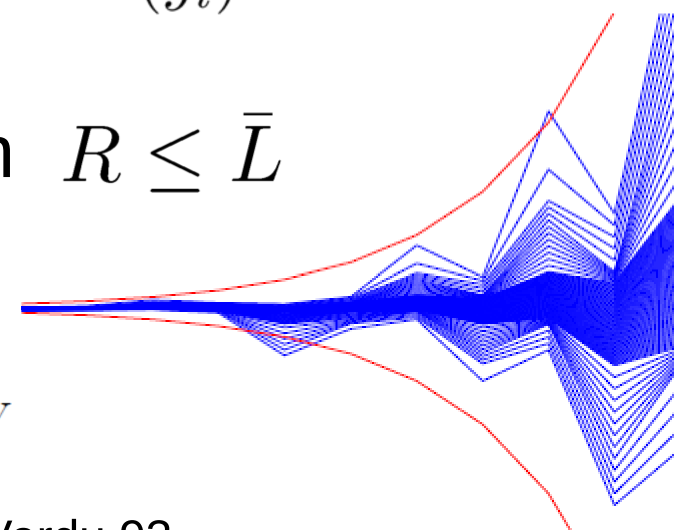




$$X_i = g_i(W, Y^{i-1})$$

$$\pi_n(dw) = \prod_{i=1}^{n} \frac{dP_{Y_i|Y^{i-1},X^i}(\cdot|y^{i-1}, g^i(w, y^{i-1}))}{dP_{Y_i|Y^{i-1}}(\cdot|y^{i-1})}(y_i)$$

Theorem: if R is achievable then $R \leq \bar{L}$

$$\bar{L}(W) \triangleq \limsup_{\text{in p.}} \frac{1}{n} \log \left| \frac{\partial}{\partial u} g_n(u, Y^{n-1}) \right|_{u=W}$$



Where the limsup in probability is defined in Han & Verdu 93

# Conclusion

Lots of cool relationships between systems theory, control theory, and information theory **if we formulate problems carefully**.