Lecture Summary: Information Theoretic Tools for Control

Vijay Gupta* and Richard M. Murray †

Relevance of information theory to networked control: Classical control is a one-block design process (design controller / estimator). Networked control is a two-block design question (design both encoder and decoder for every communication channel). Information theory provides tools to bound the performance for large classes of encoders and decoders.

Basic definitions: Consider random variables X and Y with p.d.f.s $f_X(x)$ and $f_Y(y)$, random process $X(k)$, and k-dimensional vector $X^k = (X(1), X(2), \cdots, X(k))$.

Entropy: $h(X) = -\int f_X(x) \log f_X(x) dx$.

Conditional Entropy: $h(X|Y) = -\int f_{XY}(x,y) \log f_{X|Y}(x|y) dxdy$. If X and Y are independent, $h(X|Y) = h(X)$. If random variable $Z = f(Y)$, $h(X|Y,Z) = h(X|Y)$.

Mutual Information: $I(X;Y) = h(X) - h(X|Y)$.

Entropy Rate: $h_\infty(X) = \limsup_{k\to\infty} \frac{h(X^k)}{k}$.

Basic Inequalities: The following relations are useful.

Capacity : For any channel, maximum rate at which transmission with as low a probability of error as desired is possible is called capacity C. For input X and output Y, $C \geq I(X;Y) \geq 0$.

Entropy for Gaussian variable: If X is an n-dimensional Gaussian vector with mean 0 and covariance K, $h(X) = \frac{1}{2} \ln \left((2\pi e)^n |\det(K)|\right)$. This is the maximum entropy for a given covariance K.

Information Processing Inequality: If $U = f(X)$, $Y = g(V)$, $I(X;Y) \leq I(U;V)$.

Chain Rule: $h(X^k) = \sum_{i=0}^{k} h(X(i)|X^{i-1})$.

Asymptotic Power Spectral Density: For an asymptotically stationary random process a, define the asymptotic power spectral density as $\hat{F}_a(\omega) = \sum_{k=-\infty}^{\infty} R_a(k) e^{-j\omega k}$, where $R_a(k) = \lim_{\gamma\to\infty} E[a(\gamma+k)a(k)]$. Then

$$h_\infty(a) \leq \frac{1}{4\pi} \int_{-\pi}^{\pi} \log(2\pi e \hat{F}_a(\omega)) d\omega,$$

with equality if a is an Gaussian auto-regressive process.

Results for the set-up in Figure 1 The following results hold.

- For any plant (possibly unstable) and arbitrary feedback, $h_\infty(e) \geq h_\infty(d) + \lim_{k\to\infty} \frac{I(x(0);e^k)}{k}$.
- Consider a plant described by

$$x(k+1) = \begin{bmatrix} x_s(k+1) \\ x_u(k+1) \end{bmatrix} = \begin{bmatrix} A_s & 0 \\ 0 & A_u \end{bmatrix} \begin{bmatrix} x_s(k) \\ x_u(k) \end{bmatrix} + Bu(k),$$

where $x_s(k)$ are the stable modes and $x_u(k)$ are the unstable modes. If the feedback is stabilizing, i.e., if $\sup_k E[x(k)x^T(k)] < \infty$ then

$$\liminf_{k\to\infty} \frac{I(e^{k-1}; x(0))}{k} \geq \log|\det(A_u)|.$$

Thus, for any plant and arbitrary but stabilizing feedback, $h_\infty(e) \geq h_\infty(d) + \sum_{\text{all unstable poles}} \log|\lambda_i(A)|$, where $\lambda_i(A)$ is a pole of matrix A defining the evolution of $x(k)$.

*Department of Electrical Engineering, University of Notre Dame, vgupta2@nd.edu

†Control and Dynamical Systems, California Institute of Technology murray@caltech.edu

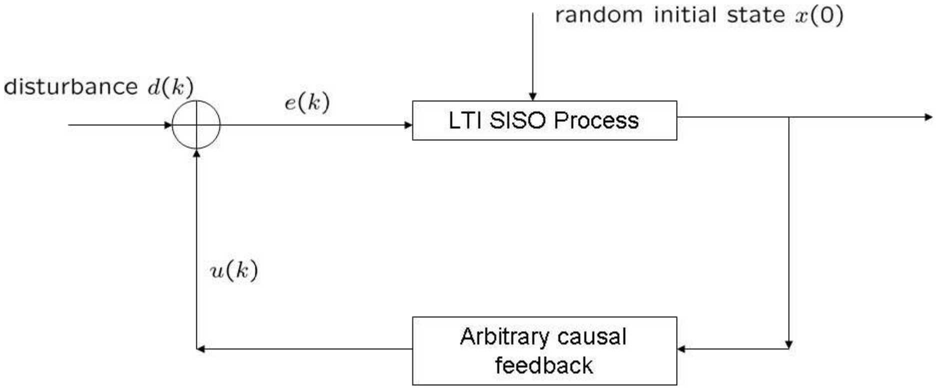

Figure 1:

- If the disturbance d is Gaussian auto-regressive,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\log(S_{e,d}(\omega))d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi}\log\left(\sqrt{\frac{\hat{F}_e(\omega)}{\hat{F}_d(\omega)}}\right)d\omega \geq \sum_{\text{all unstable poles}} \log|\lambda_i(A)|.$$

  For LTI systems $S_{e,d}(\omega)$ is the absolute value of the sensitivity function. Thus, Bode's formula holds for arbitrary causal feedback (not merely LTI control).

Results for the set-up in Figure 2 Let the feedback path have a finite capacity C_f. Then, for any LTI SISO process, any channels, any stabilizing controller, and any disturbance that is Gaussian auto-regressive, we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\min\{\log(S_{e,d}(\omega)),0\}d\omega \geq \sum_{\text{all unstable poles}} \log|\lambda_i(A)| - C_f.$$

Thus, performance degrades due to finite capacity. Further, a data rate of $\sum_{\text{all unstable poles}} \log|\lambda_i(A)|$ is necessary for stability even with a digital noiseless channel. The remaining rate is used for improving the performance. Finally, the bound is tight for Gaussian channels with linear encoding and decoding.

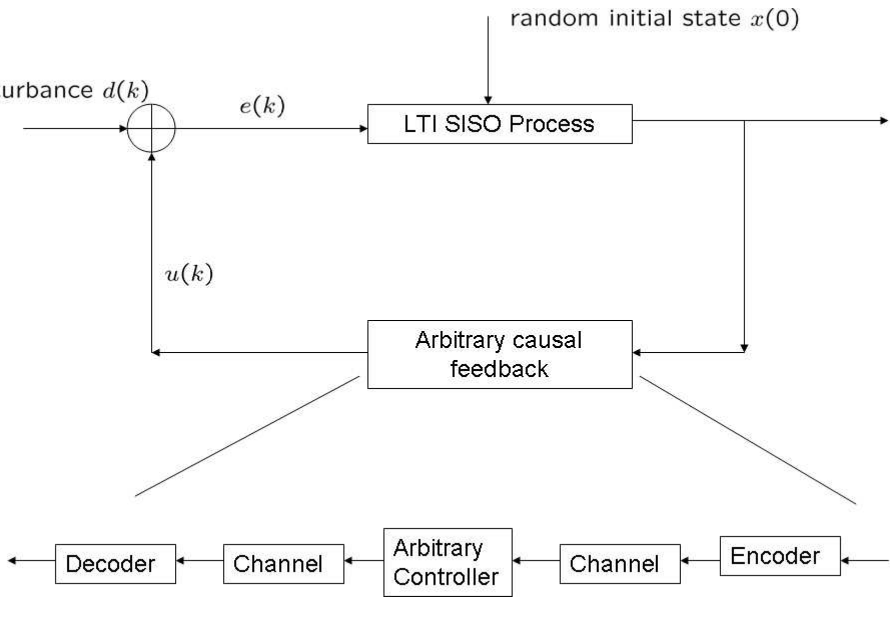

Figure 2: