Abstract
This paper deals with disturbance rejection, when the external disturbance signal is a chaotic process. We measure the amount of chaos by the Kolmogorov-Sinai entropy. The natural question is the extent to which the Kolmogorov-Sinai entropy is reduced by means of a feedback. This answer is that it is not possible to reduce the Kolmogorov-Sinai entropy by linear, stationary feedback if the open loop system is stable. In general, the Shannon entropy rate can't be decreased because of the Bode limitation.

1 Introduction
The Kolmogorov-Sinai entropy (KS entropy) appears to be a reliable measure of the amount of chaos in a dynamical system. More importantly, it draws sharp lines between classical dynamical systems, chaotic systems, and stochastic systems. More precisely, the KS entropy classifies all dynamical systems taking value in the continuum as follows [8, p. 3]:

- KS entropy < 0 \Rightarrow \text{"classical" system}
- KS entropy = 0 \Rightarrow \text{chaotic system}
- KS entropy = \infty \Rightarrow \text{stochastic system}

The restriction that the process takes value in the continuum is motivated by Markov chain problems. Indeed, these processes are genuinely stochastic. Nevertheless, they have finite Shannon entropy, because there are finitely many states. Under the conditions of finitely many possible outcomes (or countably infinitely many outcomes) the KS entropy reduces to the usual Shannon entropy and is therefore finite, even though the process is stochastic. For that reason, we shall assume from now on that the dynamical systems under consideration, whether they be classical, chaotic, or stochastic, evolve in a continuum.

1.1 Abstract Dynamical Systems
In order to be able to classify all dynamical systems according to their KS entropy, we must first develop a concept of dynamical system, general enough to encompass deterministic and stochastic processes. This is essentially the notion of "abstract dynamical system" [1, p. 2] introduced by Kolmogorov.

Definition: An abstract dynamical system is a quadruple \((\Omega, \Sigma, \mu, T)\) where

- \(\Omega\) is the sample space which could be the usual sample space of a stochastic process; it could also be the phase space of a Hamiltonian system, or the usual state space, or a manifold on which the dynamical system evolves.
- \(\Sigma\) is a \(\sigma\)-algebra of subsets of \(\Omega\).
- \(\mu\) is a finite measure, that is, a nonnegative, \(\sigma\)-additive set function defined on \(\Sigma\), i.e., \(\mu : \Sigma \rightarrow \mathbb{R}\), such that \(\mu(\Omega) = 1\).
- \(T : \Omega \rightarrow \Omega\) is the measure preserving dynamical system, that is, \(\mu(A) = \mu(T^{-1}A)\) for any measurable set \(A \in \Sigma\).

It is important to understand the general generality of the concept of "abstract dynamical system". The most intuitive interpretation of this concept is the case where \(\Omega\) is the phase space \((p, q)\) of a Hamiltonian system:

\[
\dot{p} = \frac{\partial H}{\partial q}, \quad \dot{q} = -\frac{\partial H}{\partial p}
\]

in which case \(\Sigma\) is the collection of Borel sets of \(\mathbb{R}^2\) and \(\mu\) is the usual Lebesgue measure, i.e., \(\mu( dp dq) = dp dq\). The shift \(T\) is the Poincaré map [7, p. 81] (or the so-called Birkhoff associated recursion):

\[
T \left( \begin{array}{c} p(t) \\ q(t) \end{array} \right) = \left( \begin{array}{c} p(t + \Delta t) \\ q(t + \Delta t) \end{array} \right)
\]

The fact that \(T\) is measure preserving is the celebrated Liouville theorem [15, p. 165], stating that along a Hamiltonian flow the area is preserved.

If we consider the Logistic Map \(z_{t+1} = \alpha z_t(1 - z_t)\), in this case \(\Omega = [0,1]\), \(\Sigma\) is the collection of Borel sets of \([0,1]\), the measure is \(\mu( dz) = \frac{dz}{\sqrt{\alpha(1 - \alpha)}}\), and it is readily verified that the measure is preserved, that is,

\[
\mu(A) = \int_A \frac{dz}{\sqrt{\alpha(1 - \alpha)}} = \int_A \mu( dz) = \mu(T^{-1}A)
\]

where \(A \subseteq [0,1]\) and \(T^{-1}A = [a_1, b_1] \cup [a_2, b_2]\).

It takes more work to recast a stationary stochastic process \(\{y_t\}\) in the abstract dynamical system context. Clearly, \(\Omega\) is the sample space, in the usual sense of the word, of the stochastic process, \(\Sigma\) is the \(\sigma\)-algebra of "events", and \(\mu\) is the usual probability measure on the sample space. It is, however, not so clear, at this stage, what the shift is, nor is it clear that the shift is measure preserving. To define the shift, consider on the real line an event of the form \(y_0 \in B_0, y_1 \in B_1, \ldots\) where \(B_0, B_1, \ldots\) are Borel sets of \(\mathbb{R}\). We trace this event back in the sample space, i.e., we define the measurable set \(A = \{y \in \Omega : y_0 \in B_0, y_1 \in B_1, \ldots\}\). Likewise, consider the shifted event \(\{y_t \in B_0, y_{t+1} \in B_1, \ldots\}\) and we trace it back into the sample space to get \(A' = \{y \in \Omega : y_t \in B_t, y_{t+1} \in B_{t+1}, \ldots\}\). This defines a set mapping \(T : A' \rightarrow A\). Whether it is possible, by restricting the events, to "squeeze" the set \(A'\) to a singleton \(\omega\) and, therefore, obtain a point mapping \(T : \omega \rightarrow \omega\) is a difficult question that has only been tackled in full generality by Halmos [5]. Here, we will take the standpoint of Doob [3, p. 460] and agree that it is possible to extend the set mapping \(T : \Sigma \rightarrow \Sigma\) to a point mapping \(T : \Omega \rightarrow \Omega\), provided the sample space is correctly defined. Assuming that \(T\) has been extended to a point mapping on the sample space, the fundamental set equality \(T(\omega') = \{y_0 \in B_0, y_1 \in B_1, \ldots\}\) together with \(\omega' = \omega\) yields \(y_{t+1} = y_t(T\omega) = y_t(\omega')\). Therefore, we get a "state space" model

\[
\omega_{t+1} = T\omega_t, \quad \omega_0 = \omega(\omega_0)
\]

If we assume that the process is strictly stationary, we have

\[
\Pr(\omega \in B) = \Pr(y_0 \in B_0, y_1 \in B_1, \ldots) = \Pr(q(\omega) \in B_0, q(\omega') \in B_1, \ldots)
\]

Tracing back in the sample space yields

\[
\mu(A) = \mu(T^{-1}A)
\]

that is, the measure preserving property in the sense of an abstract dynamical system. If the process is weakly stationary, we follow the argument of Doob [3, p. 461] by changing the probability measure to a Gaussian one, in which sense the process becomes strictly stationary and the stochastic shift is still measure preserving.

The above provides an example that in many problem set up, the shift, and sometimes even the sample space, are not to be taken for granted. If we just have a sequence \(\{y_n\}\) of data, the most natural choice for the sample space \(\Omega\) is the space where every single point \(\omega\) is a whole history

\[
\omega = \left( \begin{array}{c} \omega_0 \\ \vdots \\ \omega_{\infty} \end{array} \right)
\]

in other words, \(\Omega = \mathbb{R}^{\infty}\) [17, p. 144,160] and the shift is

\[
T \left( \begin{array}{c} \omega_0 \\ \vdots \\ \omega_{\infty} \end{array} \right) = \left( \begin{array}{c} \omega_1 \\ \vdots \\ \omega_{\infty} \end{array} \right)
\]
If \( P_1 \) denotes the projection on the 1st component, we obtain the realization
\[
\omega_{k+1} = T \omega_k, \quad \omega_k = P_1(\omega_k)
\]
However, quite often, this sample space is much too big and has to be compressed. Let \( M \) be the smallest manifold that contains all \((\omega_0, \omega_1, \ldots, \omega_T)\). Let \( x \) be a system of Riemann coordinates [10]. The shift on \( \Omega \) induces a shift on \( M \), viz, \( x_{k+1} = T^k x_k \). Therefore, we obtain the "state space realization"
\[
x_{k+1} = T^k x_k, \quad x_k = P_1(x_k)
\]
1.2 KS entropy
Let us assume now that the "abstract dynamical system" set up is available either directly from a state-space like model or indirectly through some realisation arguments. The KS entropy [10, 18] is a measure of the amount of chaos of the dynamical shift in the sample space. Consider a sequence of words, the following diagram commutes
\[
\begin{array}{c}
\text{words} \\
\text{random variables}
\end{array}
\]
related to the usual Shannon entropy. Consider a sequence of words, the following diagram commutes
\[
\begin{array}{c}
\text{words} \\
\text{random variables}
\end{array}
\]
where the vertical arrows represents homeomorphisms. (The dynamical systems are also said to be topologically conjugate.) Then we have
\[
KS(\Omega, \Sigma, \mu, T) = KS(S(\Omega, \Sigma, \mu, T))
\]
This so-called metric invariance property, that guarantees invariance of the amount of chaos under nonlinear change of variables, will appear to be of paramount importance. The Lyapunov exponents, yet another way to measure amount of chaos, do not always enjoy the required metric invariance property.
We finally examine the extent to which the KS entropy can be related to the usual Shannon entropy. Consider a sequence of random variables \( y_k \) on the probability space \((\Omega, \Sigma, \mu)\); in other words, \( y_k : \Omega \rightarrow R \). \( \Sigma \) is chosen to be the \( \sigma \)-algebra generated by \( y_k \) for all \( k \)'s and all Borel sets \( B \) of \( R \). We further assume that the dynamical shift exists, in which case \( y_k(T^j \omega) = y(T^j \omega) \). Let \( \mathcal{B}_0 = \mathcal{B} \) be a partitioning of the real line into Borel sets. Observe that \( y_k(T^j \omega) = \{ \omega : y_k(\omega) \in B \} = y(T^j \omega) \in \mathcal{B} \).
\[
T^j y_k^{-1}(B_k) = \{ \omega : y_k(\omega) \in B \} = y(T^j \omega) \in \mathcal{B}
\]
Therefore, we have
\[
\Pr(\omega_0 = B_0, \ldots, \omega_T = B_T) = \mu(\cap_{i=0}^T y_k^{-1}(B_{i+1})) = \mu(\cap_{i=0}^T (T^{-i+1} A_{i+1}) = h(A, T)
\]
Therefore, if the random variables are observed within the resolution afforded by the partitioning \( \mathcal{B}_0 \) of the real line, which amounts to finitely many outcomes, the Shannon entropy rate amounts to \( h(A, T) \). Now, taking finer and finer partitionings of the real line, which amounts to a process \( y_k \), taking value in the continuum, and assuming that \( \Pr \) is absolutely continuous with respect to the Lebesgue measure [6, p. 312], viz, \( \Pr(y_0 \in B_0, \ldots, y_T \in B_T) \rightarrow p(y_0, \ldots, y_T)dy \), we get
\[
\sup_A h(A, T) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln p(y_0, \ldots, y_{i+1})dy
\]
The left hand side of the above equality is by definition the KS entropy, \( KS(Y) \). In the right hand side of the equality, we recognize the usual Shannon entropy rate
\[
SER(Y) = \lim_{n \to \infty} \frac{1}{n} \int p(y_0, \ldots, y_{i+1})dy
\]
that is finite under mild conditions. Therefore, we get
\[
KS(Y) = SER(Y) = \lim_{n \to \infty} \frac{1}{n} \int p(y_0, \ldots, y_{i+1})dy
\]
The above reveals a subtle difference between the KS entropy and the Shannon entropy rate. If \( Y \) is a genuinely stochastic process, we have \( KS(Y) = 0 = \lim_{n \to \infty} \int p(y_0, \ldots, y_{i+1})dy \) because \( \int p(y_0, \ldots, y_{i+1})dy \) diverges. The singularity of the measure \( \Pr \) in the chaotic case can easily be understood from the case of the Logistic Map. Indeed, the measure \( \Pr(y_1, y_2) \) in Fig 1 vanishes for any set \( A \) of nonvanishing Lebesgue measure included in \( A \) does not intersect the graph \( y_1 = 4y_2(1 - y_2) \). So, \( \Pr(y_1, y_2) \) is not absolutely continuous with respect to Lebesgue measure [6, p. 312]. Therefore, Eq.(1) can't be splitted into two parts as in Eq.(2), that is, Eq.(3), which links the KS entropy and Shannon entropy rate, is not true for chaotic processes. Consequently, we conclude that, for a chaotic process, there is no intuitive relationship between the KS entropy and the Shannon entropy rate when the sample space is the continuum.

Figure 1: The Logistic Map
2 KS entropy in Linear Systems

2.1 SISO Linear Systems

The following two lemmas from Rudin [14] are useful in the proof of the relationship between the KS entropies, $K(S(X))$ and $K(S(Y))$, of the input and output, respectively, of a linear system.

Lemma 1: If $f$ is continuous on $[a, b]$ then $f$ is integrable on $[a, b]$.

Lemma 2: Suppose $f$ is integrable on $[a, b]$ and $m \leq f \leq M$. Let $\phi$ be continuous on $[m, M]$, and define $\phi(f)$. Then $\phi$ is integrable on $[a, b]$.

Theorem 2: Discrete-time systems, let $\{x_n\} \in \mathbb{R}^n$ [17, p. 144,160] be the input to the linear time-invariant system with transfer function $H(z)$. $H(z)$ is assumed to be stable and minimum-phase. In the frequency domain, the simpler relationship is

$$H(z). H(z)$$

The following two lemmas from Rudin [14] are useful in the proof of the relationship between the KS entropies.

Lemma 1: If $f$ is continuous on $[a, b]$, then $f$ is integrable on $[a, b]$.

Lemma 2: For continuous-time systems, let $\{x(t)\} \in \mathbb{R}^n$ [17, p. 168] be the input to the linear time-invariant system with transfer function $H(z)$ which is assumed to be stable and minimum-phase. In the frequency domain, we compute the KS entropy in the frequency domain because of the simpler relationship. Let $z(t)$ be the output of $H(z)$, and $\check{y}, \check{x}$ be the $L$ transforms of $y(t), x(t)$, respectively. We have

$$K(S(Y)) = K(S(X))$$

If $X$ is stochastic, then

$$\text{SER}(Y) - \text{SER}(X) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln |H(e^{j\omega})| d\omega \quad (4)$$

(b). For continuous-time systems, let $\{x(t)\} \in \mathbb{R}^n$ [17, p. 168] be the input to the linear time-invariant system with transfer function $H(z)$ which is assumed to be stable and minimum-phase. In the frequency domain, we compute the KS entropy in the frequency domain because of the simpler relationship. Let $z(t)$ be the output of $H(z)$, and $\check{y}, \check{x}$ be the $L$ transforms of $y(t), x(t)$, respectively. We have

$$K(S(Y)) = K(S(X))$$

If $X$ is stochastic, then

$$\text{SER}(Y) - \text{SER}(X) = \lim_{n \to \infty} \frac{1}{2\pi} \sum_{k=0}^{\infty} \int_{-\infty}^{\infty} \ln |H(e^{j\omega})| d\omega \quad (5)$$

Proof:

(a). For discrete-time systems, we have $\Omega = \mathbb{R}^n$. For $X = \{x_n\} \in \Omega$, let $T_1$ be the dynamical shift in the time domain. $T_1 : \Omega \to \Omega$. Since $x$ is a memory transform of $x$, it is not easy to compute $K(S(Y))$ in time domain. Hence, we need to consider its frequency domain. The Fourier Transform of chaotic or stochastic processes will not converge in the usual $L^1$ or $L^2$ sense. However, the $Z$-Transform $Z$ carries the time domain sample space element $X$ to its frequency domain counterpart $X(z) \triangleq ZX$ will converge in some region $|z| = r > 1$. It is observed that the dynamical shift $T_1$ in the frequency domain sample space $\Omega'$ is just the multiplication by $H(z)$. Since $Z$ is linear and bounded, $Z$ is continuous [11, p. 340] and is $Z^{-1}$. Therefore, $Z$ is a metric isomorphism [10] between the time domain and the frequency domain. We compute the KS entropy in the frequency domain because of the simpler relationship, $\check{Y}(z) = H(z)\check{X}(z)$ where $|z| = r > 1$, between the two signals. Since $H(z)$ is stable and minimum-phase, $\check{Y}(z)$ and $\check{X}(z)$ are related by the $L^{-1}$ correspondence $H(z)$. Therefore, we have the following commutative diagram:

$$\begin{array}{ccc}
Z & X & \check{X} \quad T_1(x_0, x_1, \ldots) \to (x_1, x_2, \ldots) \\
H(z) & \check{X} & \check{Y} \quad T_2(\check{X}) \to \check{Y}(z) \\
Z^{-1} & \check{Y} & T(\check{Y}_0, \check{y}_1, \ldots) \to (\check{y}_1, \check{y}_2, \ldots)
\end{array}$$

Since the KS entropy is a metric invariant [10], the KS entropy in the frequency domain is the same as the KS entropy in the time domain, thus we have

$$K(S(X)) = K(S(\check{X})), \quad K(S(\check{X})) = K(S(\check{Y})), \quad K(S(\check{Y})) = K(S(\check{X}))$$

(6)

It follows that $K(S(Y)) = K(S(X))$.

If $X$ is stochastic, then $\check{Y}$ is also stochastic. For each sample frequency $\omega_k$, $k = 0, \ldots, n-1$, define

$$Y_k \triangleq \check{Y}(z_k), \quad X_k \triangleq \check{X}(z_k), \quad H_k \triangleq H(z_k)$$

$$Y_k = H_k X_k$$

This is a coordinate transformation. The invariant measure $\mu'$ on the frequency domain sample space $\Omega'$ induces a measure $f \, d\check{X}$ on the space $\mathbb{R}^n$ of $n$ sampled data values of $X(re^{j\omega})$ as follows:

$$f = \mu' \{ X(re^{j\omega}) \in \Omega' : |X_r| \leq X(re^{j\omega}) < |X_r + dX_r|$$

$$= \frac{2\pi}{n} \sum_{k=0}^{n-1} k \leq \omega_k < \frac{2\pi}{n} (k+1); \quad k = 0, \ldots, n-1 \}$$

(7)

Similarly, we can define $f \, d\check{Y}$. Thus, the density of invariant measure of $\check{Y}$, $f \, d\check{Y}$, is related to the density of invariant measure of $X$, $f \, d\check{X}$, via the Jacobian factor by the Perron-Frobenius equation [7, p. 85]. Therefore, we obtain

$$f \{ Y_0, \ldots, Y_n-1 \} = f \{ X_0, \ldots, X_n-1 \}$$

Then, the difference of Shannon entropy rate in frequency domain is

$$\text{SER}(Y) - \text{SER}(X) = \lim_{n \to \infty} \frac{1}{2\pi} \sum_{k=0}^{n-1} \ln |H(e^{j\omega})| d\omega \quad (9)$$

Let $z = re^{j\omega}$. Since $H(re^{j\omega})$ is a bounded rational function of $\omega$ on $[\omega, \pi]$, $|H(re^{j\omega})|$ is continuous, i.e., integrable by Lemma 1. Moreover, $|H(re^{j\omega})|$ is bounded and $\omega$ is continuous, so, by Lemma 2, $\ln |H(re^{j\omega})|$ is integrable on $[\omega, \pi]$. Thus, the limit of the summation in Eq.(9) exists and can be written as an integral

$$\lim_{n \to \infty} \frac{1}{2\pi} \sum_{k=0}^{n-1} \ln |H(e^{j\omega})| d\omega = \int_{\omega}^{\pi} \ln |H(re^{j\omega})| d\omega$$

$$= \ln H(z) dz$$

In the above, $H(z)$ is the complex argument $H$ is defined as the unique branch of the In function that has its Riemann sheet characterized by a " slit" along the negative real axis $ReH < 0$. $C$ is the trigonometrically oriented, closed contour along a circle of radius $r$, see Fig. 2, except that the contour is not allowed to cross the " slits" and therefore goes via the edges of the slits around the branch points. $S$ denotes the path along the edges of the Slits and $B$ denotes the small circles around the Branching points after removal of the small area that cross the slits. Since $H(z)$ is stable and minimum-phase, $\ln H(z)$ has no singularities in the annulus $1 \leq |z| \leq r$ and in $|z| < 1$ all singularities of $\ln H(z)$ are branching points. Therefore, by elementary residue calculation

$$\int_{C} \ln H(z) dz = 2\pi i \text{Res} \left( \frac{\ln H(z)}{z}, z = 0 \right)$$

(assuming that $H(z)$ is analytic at $z = 0$; otherwise $z = 0$ becomes a branching point that is encircled and hence excluded from the interior of the contour $C$). The relevant conclusion is that $f \, d\check{Y}$ does not depend on $r$. Likewise $f \, d\check{X}$ does not depend on $r$ either. Finally $f \, d\check{Y}$ does depend on $r$, but is purely imaginary, so that $\text{Re} f \, d\check{Y} = 0$. Therefore, $f \, d\check{X} \ln |H(re^{j\omega})| d\omega$ does not depend on $r$, as large as $r \geq 1$, so that we can let $r = 1$. It follows that

$$\text{SER}(Y) - \text{SER}(X) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |H(e^{j\omega})| d\omega$$

2229
This result agrees with a related result by Papoulis [13]. The Z transform is a homeomorphism between the time domain and the frequency domain. However, Shannon entropy rate is not metric invariant. The SER in the frequency domain is related with the SER in the time domain by a Jacobian matrix $J$, that is:

$$SER(X) = SER(Y) + |J|, \quad SER(Y) = SER(Y) + |J|$$

The $|J|$ factor can be cancelled if we consider the difference of SER. Thus, the difference of Shannon entropy rate in the frequency domain is equal to the difference of Shannon entropy rate in the time domain.

(b) For continuous-time systems, we have $\Omega = \mathbb{R}^{0,\infty}$. The approach is similar to that of part (a). For the same reason as part (a), we use the Laplace Transform $L$ instead of the Fourier Transform. The Laplace transform carries the time domain sample space element $X$ to its frequency domain counterpart $\hat{X}(s) \triangleq \mathcal{L}X$ where $s = \sigma + j\omega$ and $\sigma > 0$. Then $\mathcal{L}$ is a metric isomorphism between the time domain and the frequency domain. Then, similar to Eq.(9), we have

$$\lim_{n \to -\infty} \frac{1}{n} \sum_{n=0}^{n-1} \log |H(sa)| = \frac{1}{2W} \int_{-W}^{W} \log |H(s = \sigma + j\omega)|ds$$

Using the same complex integration argument, we conclude that $\int_{-\infty}^{\infty} \log |H(s + j\omega)|ds$ does not depend on $\sigma$. Therefore, setting $\sigma = 0$, and letting $W \to \infty$, we get

$$SER(Y) - SER(X) = \lim_{W \to \infty} \int_{-W}^{W} \log |H(j\omega)|ds$$

and the result is proved. This result is in agreement with a related result of Shannon [16].

### 2.2 MIMO Linear Systems

We now extend our result of Theorem 2 to MIMO linear time-invariant systems.

Theorem 3: Let $X$ be the stochastic input and $Y$ be the stochastic output of the MIMO linear time-invariant system with transfer function matrix $H$. If $\dim(X) = \dim(Y)$ and $H$ is stable and minimum phase, then we have the following two equalities:

(a) For discrete-time MIMO systems, we have

$$SER(Y) - SER(X) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |H(e^{j\omega})|d\omega$$

(b) For continuous-time MIMO systems, we have

$$SER(Y) - SER(X) = \lim_{W \to \infty} \frac{1}{2W} \int_{-W}^{W} \log |H(j\omega)|d\omega$$

**Proof:** For simplicity, we assume that the MIMO system has two inputs $X_1, X_2$ and two outputs $Y_1, Y_2$. Let's define $X \triangleq [X_1, X_2]^T$ and $Y \triangleq [Y_1, Y_2]^T$. The transfer function is a $2 \times 2$ matrix $H$. Consider case (a) first.

$$Y_1 = H_{11}(z)X_1 + H_{12}(z)X_2$$

$$Y_2 = H_{21}(z)X_1 + H_{22}(z)X_2$$

For each sample frequency $\omega_k$, $k = 1, \ldots, n$, we define

$$Y_{ia} \triangleq \hat{Y}(z_k), \quad X_{ia} \triangleq \hat{X}(z_k), \quad H_{imi} \triangleq \hat{H}(z_k)$$

where $i = 1, 2; m = 1, 2; k = 1, \ldots, n$ and $z_k \triangleq re^{j\omega_k}, r > 1$. Therefore, we have the following equalities:

$$Y_{1a} = H_{11a}X_{1a} + H_{12a}X_{2a}$$

$$Y_{2a} = H_{21a}X_{1a} + H_{22a}X_{2a}$$

Then,

$$f_X(Y_{1a}, Y_{2a}, \ldots, Y_{1n}, Y_{2n}) = f_X(X_{1a}, X_{2a}, \ldots, X_{1n}, X_{2n})$$

where

$$J = \prod_{k=1}^{n} \det H(z_k)$$

Following the same argument as in Section 2.1, Eq.(10) holds for the special case where the dimension of the input is 2. It is obvious that we can extend the result to any number of inputs and outputs provided they have the same dimension. Since the difference of SER in the frequency domain is equal to the difference of SER in the time domain, Eq.(11) is also easily derived from the same argument as in this proof.

### 3 Applications

We now apply Theorem 2 to control systems — chaotic disturbance rejection and communication systems — difference of KS entropies of original signals and modulated signals.

#### 3.1 Reduction of Entropy in Linear Systems

Consider the SISO linear system of Fig. 3 where $P$ is a linear time-invariant plant, $C$ is a linear time-invariant controller to be designed, and $d$ is the external disturbance or noise. Define the open loop transfer function $L \triangleq \frac{P}{C}$. We have $y = d$ where $S \triangleq \frac{1}{1-L}$ is the sensitivity function. $S$ is always stable.

(i) Discrete-Time Linear Systems: $S = S(s)$

At this stage, we need the results related to the Bode integral

$$\int_{-\infty}^{\infty} \log |S(j\omega)|d\omega = 2\pi \sum_{k} \log (p_k - 1)$$

where the $p_k$'s are the (i.e. outside the closed unit disk) open-loop poles and $\gamma \triangleq \lim_{s \to \infty} L(s)$.

Let $L(s) \triangleq \frac{N(s)}{D(s)}$ and $\nu \triangleq \deg D(s) - \deg N(s)$. By the Discrete Generalized Bode Theorem, if $\nu \geq 1$, we have $\gamma = 0 \Rightarrow SER(y) = SER(d)$. Therefore, we can't reduce the uncertainty as measured by the SER entropy of a chaotic disturbance $d$ by means of a linear time-invariant feedback. In the case where $\nu = 0$, whether $SER(y) \leq SER(d)$ depends on the gain $\gamma$. For KS entropy, $KS(Y) = KS(X)$. It is no way to reduce the KS entropy in Linear time-invariant feedback too.

(ii) Continuous-Time Linear Systems: $S = S(s)$

We need to assess the value of the Bode integral in Eq.(5). The **Revised Generalized Bode Theorem** (see Wu and Jonckheere [20]) states the following:

Let $K \triangleq \lim_{s \to \infty} L(s), L(s) \triangleq \frac{N(s)}{D(s)}$ and $\nu \triangleq \deg D(s) - \deg N(s)$. Then

$$\int_{-\infty}^{\infty} \log |S(j\omega)|d\omega$$

**Figure 3:** Linear System with Disturbance

In the time domain, Eq.(11) is also easily derived from the same argument as in this proof.
Consider the original signal \( z(t) \) and the carrier signal \( z(t) \). Thus, similar to Eq.(8), we have

\[
\text{SER}(z) - \text{SER}(d) = \lim_{w \to \infty} \frac{1}{2W} \int_{-W}^{W} |S(j\omega)| d\omega = 0
\]

So, \( \text{SER}(y) = \text{SER}(d) \) i.e. we can't reduce the uncertainty caused by a chaotic disturbance \( d \) by means of a linear time-invariant feedback.

In the case where \( \nu = 0 \), by [30], we have

\[
\text{SER}(y) - \text{SER}(d) = \lim_{w \to \infty} \frac{1}{2W} \int_{-W}^{W} |S(j\omega)| d\omega = -\ln|K| + 1
\]

Whether \( \text{SER}(y) < \text{SER}(d) \) depends on the gain \( K \). For KS entropy, since \( K_S(Y) = K_S(X) \), it is no way to reduce the KS entropy by linear time-invariant feedback.

In the proof of Theorem 2, we considered the difference of KS entropy when a chaotic signal \( x \) is multiplied by a deterministic signal \( y \) in the time domain?

(1) : Continuous-time causal signals: \( x = x(t) \), chaotic and \( y = y(t) \), deterministic and bounded.

Let \( z(t) = x(t)y(t) \). Since the difference of Shannon entropy rate in time domain is the same as in frequency domain, we obtain

\[
\text{SER}(x) - \text{SER}(y) = \int_{0}^{T} \text{ln} |L(z(t))| - \text{SER}(X) dt
\]

where \([0, T]\) is the support of \( L^{-1} \{ \tilde{Y}(s) \} \).

\[
\text{SER}(x) - \text{SER}(y) = \frac{1}{T} \int_{0}^{T} \text{ln} |Y(t)| dt
\]

If \( \text{SER}(y) \) is periodic, then Eq.(12) can be rewritten as

\[
\text{SER}(x) - \text{SER}(y) = \frac{1}{T} \int_{0}^{T} \text{ln} |y(t)| dt
\]

where \( T \) is the period of \( |y(t)| \).

(II) : Discrete-time causal signals: \( x = x(k) \) and \( y = y(k) \).

Let \( z_k = x_ky_k \) and \( z = (z_k) \). The joint probability density function of \( z_0, \ldots, z_{n-1} \) is related to the joint probability density function of \( x_0, \ldots, x_{n-1} \) via the Jacobian factor.

\[
f_x(z_0, \ldots, z_{n-1}) = f_x(x_0, \ldots, x_{n-1}) |\text{det} J|
\]

Thus, similar to Eq.(8), we have

\[
\text{SER}(x) - \text{SER}(y) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \text{ln} |y_k|
\]

Consider the original signal \( x(t) \) and the carrier signal \( g(t) = A \cos(\omega_0 t + \phi) \) where \( A, \omega_0, \phi \) are constant. Let \( z(t) = x(t)y(t) \) denote the modulated signal. \( |y(t)| \) is periodic with \( T = \frac{2\pi}{\omega_0} \). From Eq.(13) and integral formula in [4, p. 526], we have

\[
\text{SER}(x) - \text{SER}(y) = \ln |A| + \frac{1}{2} \int_{0}^{\frac{T}{2}} \text{ln} |A| \text{cos}(\omega_0 t + \phi)| dt
\]

Thus, \( \text{SER}(x) \leq \text{SER}(z) \) if \( |\omega| \leq 2 \). Whether \( \text{SER}(z) \) is less than \( \text{SER}(x) \) is independent of the choice of \( \omega_0 \) and \( \phi \).

4 Conclusions

Because of the Bode integral, it is not, in general, possible to reduce the amount of chaos by means of a linear time-invariant feedback in a SISO system. Recent results [9, 12, 20], however, suggest that chaos can be reduced using a time-varying and/or nonlinear controller. To be more precise, [2, 12] report that a chaotic attractor (with KS entropy > 0) can be controlled into an attracting periodic motion (with KS entropy = 0) by time varying perturbation of some parameters, the perturbation depending on output measurements. This latter is a nonlinear time varying controller. This combined with our results suggest that a time varying nonlinear controller may be the only way to reduce the KS entropy.

References