Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Epsilon-Entropy and H_∞ Entropy in Continuous Time Systems*

Hui ZHANG [a] Youxian SUN [b]

[a,b] *State Laboratory of Industrial Control, Institute of Industrial Process Control,*
Department of Control Science & Eigeneering, Zhejiang University, Hang Zhou, 310027, China
(Tel: 86-571-87951879; e-mail: [a] zhanghui_iipc@ zju.edu.cn, [b] yxsun@iipc.zju.edu.cn).

Abstract: Based on the analysis of $\varepsilon-$entropy, information in continuous time linear multivariable stochastic systems is discussed. To describe time average information variation after a process transmitted through a continuous time system, the concept of system variety defined by the difference between $\varepsilon-$entropy rates of system input and output is proposed. Furthermore, an equivalent relation between system variety and the H_∞ entropy, an auxiliary performance index in H_∞ control theory, is derived by using spectral factorization. This connection gives information theoretic interpretation to H_∞ entropy, and provides new potential method to analyze and design continuous time stochastic systems.

1. INTRODUCTION

Information theoretic approaches to control systems have attracted amounts of attention (Weidemann, 1969; Engell, 1984; Petersen, 2006). An important topic in this field is the description of information transmission and variation in dynamic control systems. For discrete time systems, information and uncertainty of variables can be measured by Shannon entropy (rate) (Ihara, 1993), with which the time average information variation in dynamic systems can be discussed (Papoulis, 1991; Zhang & Sun, 2005). Unfortunately, Shannon entropy is not suitable for measuring information of continuous time processes due to its ill-posedness for infinite dimensional objective. However, the concept of $\varepsilon-$entropy defined by Kolmogorov (Kolmogorov, 1956), which is also referred to as the rate-distortion function (Berger, 1971), was introduced for measuring information of continuous variables and is helpful in analyzing the case of continuous time. It will be used in the analysis of this paper to describe time average information variation after a signal transmitted through a continuous time system.

On the other hand, as a suboptimal robust control design method, the minimum entropy H_∞ control method has been developed in the past (Mustafa & Glover, 1990; Iglesias & Peters, 2000). It adopts an unintuitive function of system closed-loop transfer matrix, the H_∞ entropy, as its auxiliary performance index. Although the H_∞ entropy is different from Shannon entropy, it was proved to be connected with information theoretic measures firmly, for examples, in the system identification problem (Stoorvogel & Van Schuppen, 1996) and in the anisotropy-based theory of robust control (Vladimirov, Kurdyukov & Semyonov, 1995). Relationship between H_∞ entropy and the description of information variation in linear discrete time system was given by the authors (Zhang & Sun, 2005). This paper will give an analogous result for continuous time case.

In order to describe information variation in continuous time linear multivariable systems, the $\varepsilon-$entropy rate will be discussed, and the concept of system variety defined as the difference between $\varepsilon-$entropy rates of system input and output will be proposed, in section 2. In section 3, the relation between system variety and the H_∞ entropy will be given by using spectral factorization, and the equivalence between minimum entropy H_∞ control and minimum variety control will be derived. Section 4 is the conclusion.

2. EPSILON-ENTROPY AND SYSTEM VARIETY

In this section, the description of time average information variation in continuous time systems will be discussed by adopting the measure of $\varepsilon-$entropy rate.

In the case of discrete time, Shannon entropy rate (Ihara, 1993) measures the time average information of a stationary stochastic process. Let $\xi(k)$, $k\in\mathbb{Z}$, be a discrete time sequence of real-valued random variables where $\mathbb{Z}$ denotes the set of integers. Denote by $\xi_0^{N-1}=(\xi(0),\ \dots,\ \xi(N-1))$ a fragment of the sequence of N variables. The entropy rate $\bar{H}(\xi)\triangleq\lim_{N\to\infty}\frac{1}{N}H(\xi_0^{N-1})$, where $H(\xi_0^{N-1})$ is the differential entropy of the N-dimensional random vector ξ_0^{N-1} (with the N-variate Lebesgue measure), describes time average information or uncertainty in $\xi(k)$. For Gaussian sequence,

$$\bar{H}(\xi)=\frac{1}{2}\ln 2\pi e+\frac{1}{4\pi}\int_{-\pi}^{\pi}\ln\phi_\xi(\omega)\,\mathrm{d}\omega\,, \tag{1}$$

where ϕ_ξ is the spectral density of ξ. A description of information variation in discrete time linear multivariable

* This work is supported by the National Natural Science Foundation of China (60674028).

systems was given by the authors (Zhang & Sun, 2005). Let $F(z) \in \mathbb{H}_\infty^{m\times m}$ (where $\mathbb{H}_\infty$ denotes the set of all proper and stable transfer functions) be a full-rank square transfer function matrix of a discrete time linear multivariable system, i.e. $\det F(e^{j\omega}) \neq 0$ for almost all $\omega \in [-\pi, \pi]$. Suppose it is driven by a stationary stochastic input $\xi(k) \in \mathbb{R}^m$ (where $\mathbb{R}$ denotes the set of all real numbers). Then the entropy rate of system output $\zeta(k) \in \mathbb{R}^m$ is

$$\bar{H}(\zeta) = \bar{H}(\xi) + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln|\det F(e^{i\omega})|^2 \, \mathrm{d}\omega. \qquad (2)$$

The integral term in (2) is the difference between entropy rates of output and input. It describes the variation of information (or uncertainty) after a signal transmitted through a discrete time linear system, and was referred to as the 'variety' of system $F(z)$ (Zhang & Sun, 2003), which can be considered as the extending concept of the variety of a random variable (Ashby[1], 1957). Denote it as

$$V(F) \triangleq \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln|\det F(e^{i\omega})|^2 \, \mathrm{d}\omega. \qquad (3)$$

The variety is defined by system parameters, and independent of extraneous input.

The entropy rate and the variety discussed above are well defined for discrete time systems. However, no analogue of the expression (1) exists in continuous time. The obstacle in using directly the Shannon entropy to describe the information of a continuous time process is that the entropy of a path of continuous time process is always infinite, even if it is stationary. An alternative way to measure information of a process is to characterize the information contained in its observation (or reproduction). The value of a process for user can be defined by how precisely the user obtains its information when it is reproduced. Then the information of a stationary continuous time process $\xi(t)$ can be measured by means of a 'virtual reproduction' $\hat{\xi}(t)$ stationarily related to $\xi(t)$ (Kolmogorov, 1956). The concept of $\varepsilon-$entropy was introduced based on this interpretation. In this section, we will use $\hat{\xi}(t)$ to denote the reproduction of $\xi(t)$, etc.

The $\varepsilon-$ entropy is defined as the amount of mutual information between two processes. Let $\xi_0^T \triangleq \{\xi(t)\}$, $\zeta_0^T \triangleq \{\zeta(t)\}$ denote respectively paths of continuous time processes $\xi(t)$ and $\zeta(t)$, $t \in \mathbb{R}$, over $0 \le t \le T$. The mutual information rate between $\xi(t)$ and $\zeta(t)$ is defined as (Ihara, 1993)

$$\bar{I}(\xi;\zeta) \triangleq \lim_{T\to\infty}\frac{1}{T} I(\xi_0^T;\zeta_0^T) \qquad (4)$$

if the limit exists, where $I(\xi_0^T;\zeta_0^T)$ is the mutual information between the entire paths of $\xi(t)$ and $\zeta(t)$ over $0 \le t \le T$. If the processes are stationary, the limit exists.

It is natural to characterize the precision of reproduction of a continuous time stationary process $\xi(t)$ by using the mean square distortion measure $\sigma^2 = \mathbb{E}\{(\xi(t)-\hat{\xi}(t))^2\}$, where $\mathbb{E}$ denotes the mathematical expectation. Suppose a certain reproduction precision is required, i.e.,

$$\sigma^2 \le \varepsilon^2, \qquad (5)$$

then the $\varepsilon-$entropy rate of $\xi(t)$ is defined as (Kolmogorov, 1956)

$$\bar{H}_\varepsilon(\xi) \triangleq \inf_{\sigma^2\le\varepsilon^2} \bar{I}(\xi;\hat{\xi}). \qquad (6)$$

Function (6) can also be called as "the speed of creation of message in the process $\xi(t)$ for a certain reproduction exactness". In the Gaussian case, the $\varepsilon-$entropy rate can be calculated as (Kolmogorov, 1956; Berger, 1971)

$$\bar{H}_\varepsilon(\xi) = \frac{1}{4\pi}\int_{-\infty}^{\infty} \max\ [0,\ \ln\frac{\phi_\xi(\omega)}{\theta^2}]\, \mathrm{d}\omega, \qquad (7)$$

where ϕ_ξ is the spectral density of ξ, the parameter θ is uniquely determined from the equation

$$\varepsilon^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \min\ [\theta^2,\ \phi_\xi(\omega)]\, \mathrm{d}\omega. \qquad (8)$$

In follows, we will use (7) to get our conclusion. Firstly, the scalar case will be considered.

Suppose a stable system $f(s)$ is driven by the stationary Gaussian process $\xi(t)$, its output is $\zeta(t)$. It can be seen from (8) that if $\theta^2 \le \min_\omega \phi_\xi(\omega)$, then $\varepsilon^2 = \theta^2$, and

$$\bar{H}_\varepsilon(\xi) = \frac{1}{4\pi}\int_{-\infty}^{\infty} \ln\frac{\phi_\xi(\omega)}{\varepsilon^2}\, \mathrm{d}\omega. \qquad (9)$$

Conversely, if $\varepsilon^2 \le \min_\omega \phi_\xi(\omega)$, then we can also get that $\theta^2 \le \min_\omega \phi_\xi(\omega)$ and $\varepsilon^2 = \theta^2$. Hence, when the required reproduction distortion is small enough such that $\varepsilon^2 \le \min_\omega \phi_\xi(\omega)$, the $\varepsilon-$entropy rate of Gaussian process $\xi(t)$ can be calculated by (9). Suppose the mean square reproduction distortion of $\zeta(t)$ is bounded by ρ^2, i.e.

[1] Ashby, one precursor of cybernetics, defined the concept of 'variety' of a random variable, and proposed the 'law of requisite variety' of feedback control systems.

$\mathbb{E}\{(\zeta(t)-\hat{\zeta}(t))^2\}\le\rho^2$. If $\rho^2\le\min_\omega\phi_\zeta(\omega)$, where $\phi_\zeta(\omega)$ is the spectrum of $\zeta(t)$, then the $\varepsilon-$entropy rate of $\zeta(t)$ is

$$\bar{H}_\rho(\zeta)=\frac{1}{4\pi}\int_{-\infty}^{\infty}\ln\frac{\phi_\zeta(\omega)}{\rho^2}\,\mathrm{d}\omega. \tag{10}$$

Furthermore, if $\varepsilon=\rho$, i.e. the reproductions of system input and output are subjected to a common precision requirement, then

$$\bar{H}_\varepsilon(\zeta)-\bar{H}_\varepsilon(\xi)=\frac{1}{4\pi}\int_{-\infty}^{\infty}\ln|f(j\omega)|^2\,\mathrm{d}\omega. \tag{11}$$

The term in the right hand of equality in (11) is the scalar continuous time version of (3).

Now let us consider the case of multivariable systems under the following assumption.

Assumption 1: Suppose a multivariable full-rank continuous time linear system $F(s)\in\mathbb{H}_\infty^{m\times m}$ is driven by a Gaussian vector stationary processes $\xi(t)\in\mathbb{R}^m$, its output is $\zeta(t)\in\mathbb{R}^m$. The covariance matrices of $\xi(t)$ and $\zeta(t)$ are respectively $\varPi_\xi(t)$ and $\varPi_\zeta(t)$, and the corresponding spectrums are respectively $\varPhi_\xi(\omega)$, $\varPhi_\zeta(\omega)$.

Define the mean square distortion requirement of the virtual reproduction of $\xi(t)$ as

$$\mathbb{E}\{(\xi(t)-\hat{\xi}(t))^{\mathrm{T}}(\xi(t)-\hat{\xi}(t))\}\le\varepsilon^2. \tag{12}$$

For $\varPi_\xi(t)$ is positive, suppose $U(t)$ be the unitary matrix that diagonalizes $\varPi_\xi(t)$ for all time, i.e.,

$$\varOmega(t)\triangleq U^{\mathrm{T}}(t)\varPi_\xi(t)U(t)=\mathrm{diag}[\lambda_1(t),\lambda_2(t),...,\lambda_m(t)],$$

where $\lambda_i(t),\ i=1,2,\ldots,m$ are enginvalues of $\varPi_\xi(t)$. Denote $\vartheta(t)\triangleq U^{\mathrm{T}}(t)\xi(t)=[\vartheta(t)_1,\vartheta(t)_2,\cdots,\vartheta_m(t)]^{\mathrm{T}}$, where $\vartheta(t)_i$, $i=1,2,\cdots,m$, are mutual independent. Under the mean square distortion constraint

$$\begin{aligned}&\mathbb{E}\{(\vartheta(t)-\hat{\vartheta}(t))^{\mathrm{T}}(\vartheta(t)-\hat{\vartheta}(t))\}\\&=\mathbb{E}\{(\xi(t)-\hat{\xi}(t))^{\mathrm{T}}(\xi(t)-\hat{\xi}(t))\}\\&\le\varepsilon^2\end{aligned}$$

the $\varepsilon-$ entropy of the vector variable $\vartheta(t)$ at time t is (Berger, 1971)

$$H_\varepsilon(\vartheta(t))=\sum_{i=1}^m H_\varepsilon(\vartheta_i(t))=\frac{1}{2}\sum_{i=1}^m\ln\frac{\lambda_i(t)}{\delta_i(t)},$$

where

$$\delta_i(t)=\begin{cases}a(t), & \text{if } a(t)\le\lambda_i(t),\\ \lambda_i(t), & \text{if } a(t)>\lambda_i(t),\end{cases}$$

and $a(t)$ is chosen so that $\sum_{i=1}^m\delta_i(t)=\varepsilon^2$. This is the so-called water-filling solution. It is seen that if $a(t)\le\lambda_i(t)$ for all *i*, then $\delta_i(t)=a(t)$ and $a(t)=\frac{1}{m}\varepsilon^2$. Note that if

$$\frac{1}{m}\varepsilon^2\le\min_i\lambda_i(t), \tag{13}$$

then $a(t)\le\lambda_i(t)$, the $\varepsilon-$entropy of $\vartheta(t)$ at time *t* is

$$\begin{aligned}H_\varepsilon(\vartheta(t))&=\frac{1}{2}\sum_{i=1}^m\ln\frac{m\lambda_i(t)}{\varepsilon^2}\\&=\frac{1}{2}\ln(\frac{m}{\varepsilon^2}\det\varOmega(t)).\end{aligned} \tag{14}$$

Suppose a stricter requirement that

$$\frac{1}{m}\varepsilon^2\le\min_{i,t}\lambda_i(t), \tag{15}$$

i.e., the process reproduction is subjected to distortion requirement (13) for all time. For $\vartheta(t)_i$, $i=1,2,\cdots,m$, are mutual independent for all time, then under condition (15)

$$\begin{aligned}\bar{H}_\varepsilon(\vartheta)&=\sum_{i=1}^m\frac{1}{4\pi}\int_{-\infty}^{\infty}\ln\frac{m\lambda_i(\omega)}{\varepsilon^2}\,\mathrm{d}\omega\\&=\sum_{i=1}^m\bar{H}_\varepsilon(\vartheta_i),\end{aligned}$$

where the first equality is gotten from (9) and (14), $\lambda_i(\omega),\ i=1,\ldots,m$, are eigenvalues of $\varPhi_\xi(\omega)$.

For $\vartheta(t)$ is the unitary transformation of $\xi(t)$, then we get that under conditions (12) and (15), the $\varepsilon-$entropy of $\xi(t)$ at time *t* is

$$\begin{aligned}H_\varepsilon(\xi(t))&=H_\varepsilon(\vartheta(t))\\&=\frac{1}{2}\sum_{i=1}^m\ln\frac{m\lambda_i(t)}{\varepsilon^2}\\&=\frac{1}{2}\ln(\frac{m}{\varepsilon^2}\det\varPi_\xi(t)),\end{aligned}$$

and hence

$$\bar{H}_{\varepsilon}(\xi) = \sum_{i=1}^{m} \frac{1}{4\pi} \int_{-\infty}^{\infty} \ln \frac{m\lambda_i(\omega)}{\varepsilon^2} \, \mathrm{d}\omega \\ = \frac{1}{4\pi} \int_{-\infty}^{\infty} \ln(\frac{m}{\varepsilon^2} \det \Phi_{\xi}(\omega)) \, \mathrm{d}\omega, \tag{16}$$

if

$$\frac{1}{m}\varepsilon^2 \le \min_{i,\omega} \lambda_i(\omega) . \tag{17}$$

This implies that when the reproduction distortion is low enough (under condition (15) or (17)), the $\varepsilon-$entropy rate of multivariable Gaussian stationary process $\xi(t)$ is calculated by (16).

The above analysis for $\xi(t)$ is applicable to the output process $\zeta(t)$, for it is also Gaussian stationary. Suppose the distortion of the reproduction of $\zeta(t)$ is bounded by ρ^2, i.e., $\mathbb{E}\{(\zeta(t)-\hat{\zeta}(t))^{\mathrm{T}}(\zeta(t)-\hat{\zeta}(t))\} \le \rho^2$, and $\frac{1}{m}\rho^2 \le \min_{i,t} \kappa_i(t)$, where $\kappa_i(t)$, $i = 1,2,\dots,m$ are enginvalues of $\Pi_{\zeta}(t)$. Then the $\varepsilon-$entropy rate of $\zeta(t)$ is

$$\bar{H}_{\rho}(\zeta) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \ln(\frac{m}{\rho^2} \det \Phi_{\zeta}(\omega)) \, \mathrm{d}\omega. \tag{18}$$

Furthermore, when the reproductions of $\xi(t)$ and $\zeta(t)$ are subjected to a common low enough distortion, the following conclusion is derived.

Proposition 1: Consider the full-rank continuous time stable system $F(s) \in \mathbb{H}_{\infty}^{m\times m}$ under Assumption 1. Suppose the reproductions of input $\xi(t)$ and output $\zeta(t)$ are subjected to a common distortion constraint, i.e.,

$$\mathbb{E}\{(\xi(t)-\hat{\xi}(t))^{\mathrm{T}}(\xi(t)-\hat{\xi}(t))\} \le \varepsilon^2 ,$$
$$\mathbb{E}\{(\zeta(t)-\hat{\zeta}(t))^{\mathrm{T}}(\zeta(t)-\hat{\zeta}(t))\} \le \varepsilon^2 ,$$

and the distortion is low enough such that

$$\frac{1}{m}\varepsilon^2 \le \min_{i,t}\{\lambda_i(t),\kappa_i(t)\} \text{, or,}$$
$$\frac{1}{m}\varepsilon^2 \le \min_{i,\omega}\{\lambda_i(\omega),\kappa_i(\omega)\} ,$$

then

$$\bar{H}_{\varepsilon}(\zeta) - \bar{H}_{\varepsilon}(\xi) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \ln |\det F(j\omega)|^2 \, \mathrm{d}\omega . \tag{19}$$

■

Proof: Equation (19) is derived based on above analysis and the fact that $\Phi_{\zeta}(\omega) = F(j\omega)\Phi_{\xi}(\omega)F^{*}(j\omega)$.

□

Remark 1: High reproduction precision is a natural requirement when we try to characterize the information contained in a process, i.e., we tend to make the reproduction distortion as lower as possible. Moreover, when we investigate the information difference between two processes, it is 'justice' to ask a common precision requirement to reproductions. Hence, the conditions in Proposition 1 is rational. In following analysis, the $\varepsilon-$ entropy rates will always be defined by means of these conditions. The term in the right hand of (19), which is the difference of $\varepsilon-$entropy rates of system input and output, can be considered as the information (or uncertainty) variation after the input process transmitted through system $F(s)$. Denote it as

$$V(F) \triangleq \frac{1}{4\pi} \int_{-\infty}^{\infty} \ln |\det F(j\omega)|^2 \, \mathrm{d}\omega . \tag{20}$$

It is the continuous time version of (3), and can be referred to as the 'variety' of the continuous time system. It is also defined by system parameters, and independent of extraneous input.

3. VARIETY AND H_{∞} ENTROPY

In this section, we will discuss the relation between variety and H_{∞} entropy in continuous time linear systems. We will use $G(s) = \{A, B, C, D\}$ to denote a system realization, which means $G(s) = C(sI - A)^{-1}B + D$.

Assumption 2: Consider a continuous time linear multivariable system $G(s) \in \mathbb{RH}_{\infty}^{m\times m}$ (where $\mathbb{RH}_{\infty}$ denotes the set of all proper, rational and stable transfer functions). Its input and output are respectively $w(t) \in \mathbb{R}^m$ and $z(t) \in \mathbb{R}^m$, and

$$\| G(s)\|_{\infty} < \eta . \tag{21}$$

It has a stablizable and detectable realization as

$$G(s) = \{A,B,C\} \tag{22}$$

where A, B and C are constant matrices with corresponding dimensions.

The minimum entropy H_{∞} control of $G(s)$ is to minimize the H_{∞} entropy (Mustafa & Glover, 1990) under the given constraint (21):

$$E(G, \eta) \triangleq \frac{-\eta^2}{2\pi} \int_{-\infty}^{\infty} \ln \det[I - \eta^{-2}G^{*}(j\omega)G(j\omega)]\mathrm{d}\omega. \tag{23}$$

Regardless the fact that the H_∞ control and information theoretic method have different philosophies, the connection between these two methodologies will be analyzed in follows by using spectral factorization.

From the spectral factorization theorem (Zhou, 1998), Assumption 2 allows one has

$$\eta^2 I - G^*(s)G(s) = U^*(s)U(s)\,, \tag{25}$$

where $U(s),\ U^{-1}(s) \in \mathbb{RH}_\infty^{m\times m}$,

$$\| U \|_\infty < \eta\,, \tag{26}$$

and $U(s)$ has the state space realization

$$U(s) = \{A, B, -R^{\frac{1}{2}}F, R^{\frac{1}{2}}\}\,, \tag{27}$$

where

$$F = R^{-1}B^{\mathrm{T}}X,$$
$$R = \eta^2 I,$$

and $X = X^{\mathrm{T}} \geq 0$ is the stabilizing solution of an algebraic Riccati equation

$$X = \mathrm{Ric}\begin{bmatrix} A & BR^{-1}B^{\mathrm{T}} \\ -C^{\mathrm{T}}C & -A^{\mathrm{T}} \end{bmatrix}. \tag{28}$$

Let $\Lambda(s) = U^{-1}(s)$. $\Lambda(s)$ has the realization

$$\Lambda(s) = \{A + BF, -BR^{-\frac{1}{2}}, -F, R^{-\frac{1}{2}}\}\,. \tag{29}$$

Then,

$$[\eta^2 I - G^*(s)G(s)]^{-1} = \Lambda^*(s)\Lambda(s)\,. \tag{30}$$

Suppose $\Lambda(s)$ is the closed-loop transfer function of a system driven by a stationary process. Let $\bar{H}_\varepsilon(\alpha)$ and $\bar{H}_\varepsilon(\beta)$ be $\varepsilon-$entropy rates of its input $\alpha(t)$ and output $\beta(t)$, respectively. Then the variety of $\Lambda(s)$ is

$$\begin{aligned} V(\Lambda) &\triangleq \frac{1}{4\pi}\int_{-\infty}^{\infty} \ln|\det\Lambda(j\omega)|^2 \, \mathrm{d}\omega \\ &= \bar{H}_\varepsilon(\beta) - \bar{H}_\varepsilon(\alpha). \end{aligned} \tag{31}$$

From (23), (30) and (31), the following conclusion can be gotten directly.

Proposition 2: For system $G(s) \in \mathbb{RH}_\infty^{m\times m}$ satisfies the conditions stated in Assumption 2, there exists a system $\Lambda(s)$ with state space realization (29) satisfying the spectral factorization (30), so that

$$E(G, \eta) = 2\eta^2 V(\Lambda) - 2m\eta^2 \ln \eta^{-2}\,. \tag{32}$$

■

Remark 2: Proposition 2 gives an information theoretic interpretation for the H_∞ entropy of system $G(s)$: Minimizing the H_∞ entropy of system $G(s)$ is equivalent to minimizing the variety of system $\Lambda(s)$.

4. CONCLUSIONS

Based on the analysis of $\varepsilon-$entropy, the concept of variety was introduced to describe information variation in continuous time linear multivariable systems, and an equivalent relation between system variety and the H_∞ entropy was derived. This connection provides new potential methods to analyze and design continuous time stochastic systems under the framework of information theory.

ACKNOWLEDGMENTS

The authors thank the anonymous reviewers, especially reviewer 1 (Review ID: 3990), whose comments and suggestions were very helpful in improving our work.

REFERENCES

Ashby W. R. (1957). *An Introduction to Cybernetics*. Chapman & Hall Ltd., London.

Berger T. (1971). *Rate Distortion Theory*. Prentice-Hall, New Jersey.

Engell, S. (1984). Variety, information and feedback. *Kybernetes*. 13, 73-77.

Iglesias, P. A., & Peters, M. A. (2000). An entropy formula for nonlinear systems. *J. of the Franklin Institute*, 337: 859-874.

Ihara, S. (1993). *Information Theory for Continuous Systems*. World Scientific Publishing Co. Pte. Ltd., Singapore.

Kolmogorov, A. N. (1956). On the Shannon theory of information transmission in the case of continuous signals, *IRE Transactions on Information Theory*, 102-108.

Mustafa, D., & Glover, K. (1990), *Minimum Entropy H_∞ Control*. Springer-Verlay, Berlin.

Papoulis, A. (1991). *Probability, Random Variables, and Stochastic Processes*, 3rd edition, McGraw-Hill, Inc., New York.

Petersen, I. R. (2006). Minimax LQG control. *Int. J. Appl. Math. Comput.*, 16(3): 309-323.

Stoorvogel, A. A. & Van Schuppen, J. H. (1996). System identification with information theoretic criteria. In: *Identification, Adaptation, Learning*. (S.Bittanti, G.Picc, (Eds)), 289-338. Springer, Berlin.

Vladimirov, I. G., Kurdyukov, A. P., & Semyonov, A. V. (1995) Anisotropy of signals and entropy of linear time-invariant systems. *Dokl. Akad. Nauk*, 342(5): 583-585 (in Russian) (English translation in 51(3): 388-390).

Weidemann, H. L. (1969). Entropy analysis of feedback control systems. In: *Advances in Control Systems*, 7: 225-255. Academic Press.

Zhang, H., & Sun, Y. X. (2003). Bode integrals and laws of variety in linear control systems. In: *Proceedings of the 2003 American Control Conference*, 1: 66-70, Denver, CO, United States, Jun 4-6 2003.

Zhang, H., & Sun, Y. X. (2005). Information theoretic interpretations for H_∞ entropy. *Proceedings of 16th IFAC World Congress*, Prague, July 3-8, 2005.

Zhou, K. (1998). *Essential of robust control*. Prentice-Hall, New Jersey.