The Performance Analysis on Finite Impulse Response Actuator: An Information Theoretic view

Zhang desen[1] Xu dazhuan[2]
1. School of Information Science and Engineering,NUAA, Nanjing, 211106
E-mail: dszhang@nuaa.edu.cn
2. School of Information Science and Engineering,NUAA, Nanjing, 211106
E-mail: xudazhuan@nuaa.edu.cn

Abstract: It is well known that information theory and cybernetics share in a sense a common theoretical and engineering background, both of which are concerned with signals and dynamic systems. Communication theory is primarily concerned with the reliable transmission of information from one point to another and is relatively unconcerned with the purpose of the transmitted information or whether it is eventually fed back to the source. In contrast, control theory uses the information in the feedback loop to achieve performance goals and usually assumes limitations in the communication link do not significantly affect performance. In this paper, we analyse the actuator part of a control system basing on information theory, particularly, rate-distortion theory. The rate-distortion formulation is derived in this finite impulse response actuator with input noise to analyse the difference between expected output and designed output. Our mutual information bound is shown to be reduced to the bound given by Oohama in the one-terminal case. Spectrum analysis is also done afterwards, which is shown to be reduced to classic Kolmogorov–Szegö formula and Wiener–Masani formula in prediction theory.

Key Words: Actuator, Gaussian, FIR, Quatization noise, Information theory, Rate-Distortion

1 INTRODUCTION

Most traditional control system analysis assumes that the observed signal is available to the controller in its entire decision time period [1]; however, the real situation is that the controller receives a quantized version of the original observed signal, which is transmitted through the system processing, and is accompanied by transmission delays and distortion subject to bit rate constraints [2] [3]. One more problem, which is better addressed in communication, is the limit problem. There are limits given by information theory in communication, and then various coding methods converge to the limits, therefore, more quantitative comparisons can be carried out basing on the theoretical limit derived. Recently [4] [5] [6] has done some fundamental limit research in some aspects in control field using information theory. Our team has also done some information limit research in detection and radar field [7].

In this article, we investigate the information-theoretic modeling of FIR actuator, which is a commonly used convolutional structure. The main novel contribution is an approach to exploiting detailed knowledge about a given FIR actuator in order to obtain the rate-distortion relation. This contribution is twofold: On the one hand, we demonstrate how to derive the rate-distortion function about a given actuator with input noise, and we show that the result is coincide with the data rate requirement in [8]. On the other hand, we define the conditional entropy as a distortion metric.

This work is supported in part by National Nature Science Foundation of China under Grant 61471192 and Grant 61371169.

1.1 Bandwidth-limited control

There are many types of limitation that an actuator may impose, including quantization, delay, noise, data loss, and bandwidth constraints. A unified treatment of these issues is yet unavailable. Thus, researchers focus on simplified models that highlight certain aspects of the overall problem. In fact, the bandwidth limitation in feedback control poses a quantization problem, where information is exchanged between the controller and the controlled dynamic object through an actuator channel that can only transmit a finite number of states at any given time, where the observations and responses made by the controller are quantified with finite accuracy in both time and space [9]. We adopt the idea of finite accuracy and use finite accuracy input before the actuator model. The bandwidth-limited control problem is further complicated by the inherent delay in feedback control [10]. There is a trade-off between the fineness of the control and the time delay required to send the data over the link. In [11] J. Baillieul has shown that reaching a good data quantization requires a larger channel capacity [12] or longer time allocated for data transmission. Therefore, the FIR actuator model with input noise in our paper also possesses the problem about the trade off between information transferring and distortion. By exploiting the properties of differential entropy power, a universal lower bound is obtained on the time-asymptotic mean square state norm.In particular [13], this bound implies that as the data rate approaches the intrinsic entropy rate H of the plant, the mean square state becomes arbitrarily large, regardless of the coding and control scheme. This point direct us to find the information limit in actuator model.

1.2 Generic Information Bounds on control distortion

The generic bound on control system has many works using diverse methods. Recently, in [5] [6], information theory is adopted as the main mathematical tool to obtain generic bounds on the variance of estimation and prediction errors in time series analysis by studying the basic entropy relationships of the data points that make up the time series. In the special case of Gaussian processes and 1-step (ahead) prediction, his bounds can be simplified to the well-known Kolmogorov-Szegö formula [14] and Wiener-Masani formula [15]. However, such formulas are lacking for m-step predictions and non-Gaussian processes.This bound also has a specific background in information leakage. In [12], directed information quantity is adopted as a measure of information transmission in feedback control system. An explicit lower bound on the rate-cost function is derived from the set of conditional probabilities obeying the LQR quadratic cost function by selecting a suitable system distribution to optimize the directed information quantity. And through this way, the optimal transmission rate is obtained. However, the specific expressions for distortion measures other than quadratic distortion and non-Gaussian pdf cases are not studied.

1.3 Quatization error and Rate-distortion

Quantization errors exist in a large body of work in control theory, one class of which [16] continues to apply stochastic control theory solution ideas by modeling quantization factors as additive white noise. While this approach is reasonable when the communication bandwidth is not constrained, it will not work if the communication bandwidth is constrained, and it fails to consider [10] the existence of a critical data rate below which no quantization and control scheme exists that will stabilize the system. This effectively shows that the constrained communication capacity has a significant negative impact on the control performance [1]. However, optimal control problems under data rate constraints are generally difficult to formulate and solve except for special cases [5] [17] [8] [18]. In 1988,Saridis [19] gave the formulation of the optimal control problem by choosing a controller pdf that minimizes a given performance metric, and the differential entropy minimization of this pdf satisfying the maximum entropy principle is equivalent to the minimization of the given performance metric. The current feasible approach is related to the theory of optimal quantization. This inspires us to use rate-distortion theory to analyse the actuator model.

The objectives of control under data rate constraints have striking similarities to the objectives of source coding and rate distortion in information theory [20] [21]. It is well known that for rate distortion theory, a noncausal classical rate distortion function is known under Gaussian source mean square distortion. Rate distortion theory describes the fundamental restriction between the desired bit rate of a particular source and corresponding achievable distortion, the converse also holds [21]. The application of distortion rate functions to quantization control problems was described in the seminal work of [22]: when a sequence is encoded by an arbitrary finite-state encoder at a given fixed rate, the distortion rate function is defined as the asymptotically achievable minimum distortion in a single sequence. The model allows for arbitrary delays but at a fixed rate [23].

Therefore, we will not consider the factor of delay in our model. For such a communication bandwidth-constrained actuator model, our goal is to derive a distortion rate relation, thus describing the relationship between control accuracy, system bandwidth, and transmission rate. We further give the simuilation of the model used in deduction, and the cases in non-gaussian pdf.

The remainder of the paper is organized as follows. Section 2 introduces the technical preliminaries. Section 3 presents the fundamental rate-distortion trade-off for FIR actuator, spectrum analysis is also given in Section 4. Section 5 presents several examples on the FIR actuator model used. Conclusions are given in Section 6.

2 PRELIMINARIES

Throughout the paper, we consider real-valued continuous random variables and random vectors, as well as discrete time stochastic processes. All random variables, random vectors, and stochastic processes are assumed to be zero mean. Given a stochastic process $\{x_k\}$, we assume the excitation of the control system be a set of finite-length K-dimensional i.d.d. Gaussian vectors, the system is linear and has memory. The output signal is an N-dimensional Gaussian vector Y.

2.1 FIR Actuator Model and Assumptions

While classical linear system theory uses frequency response theory to characterize noise and system transfer functions, modern control theory uses linear algebra and state space theory to establish dynamic equations, such as the Kalman filter algorithm assumes that all noise is white noise with zero mean, by affecting the value of the Kalman gain K weights with Q and R representing the confidence in the predicted and measured values respectively, which is corresponding to input noise and quatization noise in stochastic control. Specifically, we consider the dynamical system with state-space model given by

$$\begin{cases} \hat{\mathbf{x}}_k = \mathbf{y}_k + \mathbf{w}_k \\ \hat{\mathbf{y}}_k = A\hat{\mathbf{x}}_k + \mathbf{n}_k \end{cases}$$

We define $\mathbf{y}_k$ as a sequence of designed inputs over some number of timesteps k, $\hat{\mathbf{y}}_k$ is the real actuator output. We assume that direct control over designed input $\mathbf{y}_k$ exists, but not over $\hat{\mathbf{y}}_k$. This is a reasonable noise assumption for many robotic systems in which command input must pass through a lower-level controller. A prototypical example is the steering and throttle inputs for a car which are then used as set-point targets for low level servomotor controllers.

The formulation of Actuator Model: Assuming that the channel is constant (slow fading) during one symbol period , its finite impulse response sequence in the i_{th} symbol period sequence without noise is:

$$
\begin{aligned}
y(n) &= \sum_{l=n-d+1}^{n} \hat{x}(l)h(n-l) \\
&= \hat{x}(n)h(0) + \sum_{l=n-d+1}^{n-1} \hat{x}(l)h(n-l)
\end{aligned} \tag{1}
$$

The impluse response sequence h(d) has the length of d. The sequence is then filled with zero to the length of $N_S = N_{cp} + d$, where N_{cp} is the length of circulant prefix. The finite implues response $\mathbf{A}$ during the i^{th} symbol, which is NxN toeplitz [24].

$$
\begin{aligned}
&\mathbf{A} = \mathbf{R}_{cp}\mathbf{H}\mathbf{T}_{cp} \\
&= \begin{bmatrix}
h(0) & 0 & \cdots & 0 & h(d-1) & \cdots & h(1) \\
h(1) & h(0) & 0 & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \ddots & \cdots & \cdots & \cdots & h(d-1) \\
h(d-1) & \cdots & \cdots & \ddots & \cdots & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & \ddots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \ddots & 0 \\
0 & \cdots & 0 & h(d-1) & \cdots & \cdots & h(0)
\end{bmatrix}
\end{aligned} \tag{2}
$$

where $\mathbf{R}_{cp} = \begin{bmatrix} 0_{N\times(N_S-N)} & I_N \end{bmatrix}$ is $N \times N_S$

$\mathbf{T}_{cp} = \begin{bmatrix} 0_{N-Ncp} & I_{Ncp} \\ & I_N \end{bmatrix}$ is$N_S \times N$

Therefore, the matrix equation of the actuator model is:

$$
\mathbf{Y} = \mathbf{A}\hat{\mathbf{X}} + \mathbf{N}
$$

Calculate the covariance matrix

$$
\begin{aligned}
\mathbf{R}_{\mathbf{y}} &= E_{\hat{\mathbf{x}},\mathbf{N}}\left[\mathbf{y}\mathbf{y}^H\right] \\
&= E_{\hat{\mathbf{x}},\mathbf{N}}\left[(\mathbf{A}\hat{\mathbf{x}} + \mathbf{N})(\mathbf{A}\hat{\mathbf{x}} + \mathbf{N})^H\right] \\
&= \mathbf{A}E\left[\hat{\mathbf{x}}\hat{\mathbf{x}}^H\right]\mathbf{A}^H + E\left[\mathbf{N}\mathbf{N}^H\right] \\
&= \mathbf{A}\mathbf{R}_{\hat{\mathbf{x}}}\mathbf{A}^H + N_0\mathbf{I}
\end{aligned} \tag{3}
$$

Since A is toeplitz, it can be diagonalized by Fourier matrix [25].

$$
\begin{aligned}
\mid \mathbf{R}_{\mathbf{y}} \mid &= \mid \mathbf{R}_{\hat{\mathbf{x}}}\mathbf{A}\mathbf{A}^H + N_0\mathbf{I} \mid \\
&= \left| \mathbf{Q}^H E\left[\Sigma_{\hat{X}}\right]\mathbf{Q}\mathbf{A}\mathbf{A}^H + N_0\mathbf{I} \right| \\
&= \left| E\left[\Sigma_{\hat{X}}\right]\mathbf{Q}^H\mathbf{Q}\Lambda_{\mathrm{A}} + N_0\mathbf{I} \right|
\end{aligned} \tag{4}
$$

where Λ_{A} is the diagonalization matrix of $\mathbf{A}\mathbf{A}^H$

$$
= \prod_{i=1}^{N}\left(\lambda_{\hat{X}}\lambda_{\mathrm{A}} + N_0\right) = N_0^N \prod_{i=1}^{N}\left(\frac{\lambda_{\hat{X}}\lambda_{\mathrm{A}}}{N_0} + 1\right) \tag{5}
$$

Therefore, we get the mutual information in the Gaussian case:

$$
\begin{aligned}
I(\mathbf{Y} \mid \hat{\mathbf{X}} = \hat{x}) &= h(\mathbf{Y}) - h(\mathbf{Y} \mid \hat{\mathbf{X}} = \hat{x}) \\
&= \sum_{i=1}^{N} \log\left(\frac{\lambda_{\hat{X}}\lambda_{\mathrm{A}}}{N_0} + 1\right)
\end{aligned} \tag{6}
$$

3 Performance Limit of Finite Response Actuator

In this section, we examine the fundamental limits of Finite Response Actuator for linear Gaussian dynamical systems. The main problem is on how to handle the noise in actuator model.

The assumption on stochastic process: Basing on the actuator model shown in (1),it is reasonable to assume that the stochastic process in both sides of (1) share the same statistical characteristic, thus the auto correlation of a stochastic process $\{\hat{x}_k\}$ is

$$
\begin{aligned}
R(d) &= E\left[\hat{\mathbf{x}}h(0)[\hat{\mathbf{x}}h(0)]^T\right] \\
&= \begin{bmatrix}
R_{xx}(0) & R_{xx}(-1) & \cdots & R_{xx}(-d+1) \\
R_{xx}(1) & R_{xx}(0) & \cdots & R_{xx}(-d+2) \\
\vdots & \vdots & \ddots & \vdots \\
R_{xx}(d-1) & R_{xx}(d-2) & \cdots & R_{xx}(0)
\end{bmatrix}
\end{aligned} \tag{7}
$$

where

$$
\hat{\mathbf{x}}^T h(0) = \begin{bmatrix}
y(d) - \sum_{l=+1}^{d-1} \hat{x}(l)h(d-l)\big) \\
y(d+1) - \sum_{l=2}^{d} \hat{x}(l)h(d+1-l)\big) \\
\vdots \\
y(2d-1) - \sum_{l=d}^{2d-2} \hat{x}(l)h(2d-1-l)\big)
\end{bmatrix}
$$

The power spectrum of a stationary process $\{\hat{x}_k\}$, $x_k \in R^m$, is defined as

$$
\phi(w) = \sum_{d=-\infty}^{\infty} \frac{R(d)}{h^2(0)} e^{-jwd} \tag{8}
$$

The variance of $\hat{x}(n)$ is given by

$$
R_{\hat{X}} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(w)dw
$$

hence, the entropy of $\{\hat{x}_k\}$ is

$$
h(\hat{\mathbf{X}}) = \frac{1}{2}\ln(2\pi e)^N \mid R_{\hat{X}} \mid \tag{9}
$$

Moreover, it is known that circulation matrix $\mathbf{A}$ can be diagonalized by Fourier matrices.Hence we have $\hat{\mathbf{Y}} = \Sigma\hat{\mathbf{X}}$, To the same effect [if necessary there is a demonstration in my appendix], the conditional probability distribution that represents the relationship between the finite precision output and the ideal output is:

$$
p(\hat{\boldsymbol{y}} \mid \boldsymbol{y}) = \frac{1}{\sqrt{(2\pi)^N |R_\delta|}} \exp\left\{-\frac{1}{2}(\hat{y} - \boldsymbol{y})R_\delta^{-1}(\hat{\boldsymbol{y}} - \boldsymbol{y})^T\right\}
$$

and basing on the rate-distortion, write the conditional entropy as the distortion metric constraint for infimum: $\frac{1}{2}\log(2\pi e)^N D = H(\mathbf{Y} \mid \hat{\mathbf{Y}})$

Step1:
The mutual information between target output and real output is

$$\begin{aligned} I(\mathbf{Y};\hat{\mathbf{Y}}) &= h(\mathbf{Y}) - h(\mathbf{Y} \mid \hat{\mathbf{Y}}) \\ &= \frac{1}{2}\ln\left[(2\pi e)^N |R_{\mathbf{Y}}|\right] - \frac{1}{2}\ln\left[(2\pi e)^N |R_{\dot{o}}|\right] \\ &= \frac{1}{2}\ln\frac{|R_{\mathrm{Y}}|}{|R_{\dot{o}}|} \\ h(\mathbf{Y} \mid \hat{\mathbf{Y}}) &= \frac{1}{2}\ln\left[(2\pi e)^N |R_{\mathbf{Y}}|\right] - I(\mathbf{Y};\hat{\mathbf{Y}}) \end{aligned} \tag{10}$$

A classical method in information theory to describe the transform model is Jaccobian matrix, we have
$J\left(\frac{\hat{\mathbf{Y}}}{\hat{\mathbf{X}}}\right) = J\left(\frac{\hat{Y}_1\hat{Y}_2\ldots\hat{Y}_N}{\hat{X}_1\hat{X}_2\ldots\hat{X}_N}\right) = |\mathbf{A}|$
Using the Jaccobian relation to descirbe the actuator model, the entropy of finite precision output :

$$h(\hat{\mathbf{Y}}) = h(\hat{\mathbf{X}}) + \int p(\hat{\mathbf{x}})\log\left|J\left(\frac{\hat{\mathbf{Y}}}{\hat{\mathbf{X}}}\right)\right| d\mathbf{x} \tag{11}$$

$$= h(\hat{\mathbf{X}}) + E\left[\log\left|J\left(\frac{\hat{\mathbf{Y}}}{\hat{\mathbf{X}}}\right)\right|\right] \tag{12}$$

$$\tag{13}$$

From the derivation in (9),

$$= \frac{1}{2}\ln(2\pi e)^N \mid R_{\hat{X}} \mid + E\left[\log\left|J\left(\frac{\hat{\mathbf{Y}}}{\hat{\mathbf{X}}}\right)\right|\right] \tag{14}$$

To the same effect, we have the results in covariance matrix form:

$$\begin{aligned} I(\hat{\mathbf{Y}};\hat{\mathbf{X}}) &= h(\hat{\mathbf{X}}) - h(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) \\ &= \frac{1}{2}\ln(2\pi e)^N \mid R_{\hat{X}} \mid - h(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) \end{aligned} \tag{15}$$

Recalling formulation(12), it informs us the place where FIR actuator matrix take.

$$\log|\mathbf{A}| = h(\hat{\mathbf{Y}} \mid \hat{\mathbf{X}}) - h(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) \tag{16}$$

$$I(\hat{\mathbf{Y}};\hat{\mathbf{X}}) + h(\hat{\mathbf{Y}} \mid \hat{\mathbf{X}}) = I(\mathbf{Y};\hat{\mathbf{Y}}) + h(\hat{\mathbf{Y}} \mid \mathbf{Y}) \tag{17}$$

Step2:
Next we will derive the information rate-distortion function. Note first that

$$H(\hat{\mathbf{X}}) = H(\hat{\mathbf{X}} \mid \mathbf{Y}) + I(\mathbf{Y};\hat{\mathbf{X}})$$

$$H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) = H(\hat{\mathbf{X}} \mid \mathbf{Y}\hat{\mathbf{Y}}) + I(\mathbf{Y};\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) \tag{18}$$

$$\begin{aligned} I(\mathbf{Y};\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) &= I(\mathbf{Y};\hat{\mathbf{X}}) - I(\mathbf{Y};\hat{\mathbf{X}};\hat{\mathbf{Y}}) \\ &\geq \frac{1}{2}\log\left(\frac{|R_{\mathbf{Y}}|}{|R_{\mathbf{W}}|}\right) - I(\hat{\mathbf{Y}};\hat{\mathbf{X}}) \end{aligned} \tag{19}$$

where, $I(\mathbf{Y};\hat{\mathbf{X}}) \geq H(\mathbf{Y}) - H(\mathbf{Y} \mid \hat{\mathbf{X}}) = \frac{1}{2}\log\frac{|R_{\mathbf{Y}}|}{|R_W|}$

$$\begin{aligned} H(\hat{\mathbf{X}} \mid \mathbf{Y}\hat{\mathbf{Y}}) &= \frac{1}{2}\log\left(2\pi e\,|R_W|\right) - H(\hat{\mathbf{Y}} \mid \mathbf{Y}) + H(\hat{\mathbf{Y}} \mid \hat{\mathbf{X}}) \\ &\geq \frac{1}{2}\log\left(\frac{|R_{\mathbf{Y}}|}{|R_{\mathbf{W}}|}\right) + I(\mathbf{Y};\hat{\mathbf{Y}}) + H(\hat{\mathbf{Y}} \mid \hat{\mathbf{X}}) \end{aligned} \tag{20}$$

where,

$$\begin{aligned} &H(\hat{\mathbf{Y}} \mid \mathbf{Y}) = I(\hat{\mathbf{X}};\hat{\mathbf{Y}} \mid \mathbf{Y}) + H(\hat{\mathbf{Y}} \mid \hat{\mathbf{X}}\mathbf{Y}) \\ &= H(\hat{\mathbf{X}} \mid \mathbf{Y}) - H(\hat{\mathbf{X}} \mid \mathbf{Y}\hat{\mathbf{Y}}) + H(\hat{\mathbf{Y}} \mid \hat{\mathbf{X}}\mathbf{Y}) \\ &= \frac{1}{2}\log\left((2\pi e)^N |R_W|\right) - H(\hat{\mathbf{X}} \mid \mathbf{Y}\hat{\mathbf{Y}}) + H(\hat{\mathbf{Y}} \mid \hat{\mathbf{X}}\mathbf{Y}) \\ &= \frac{1}{2}\log\left((2\pi e)^N |R_W|\right) - H(\hat{\mathbf{X}} \mid \mathbf{Y}\hat{\mathbf{Y}}) + H(\hat{\mathbf{Y}} \mid \hat{\mathbf{X}}) \end{aligned}$$

Substitute (19)(20) into (18)

$$\begin{aligned} H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) &= I(\mathbf{Y};\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) + H(\hat{\mathbf{X}} \mid \mathbf{Y}\hat{\mathbf{Y}}) \\ &\geq -I(\hat{\mathbf{Y}};\hat{\mathbf{X}}) + I(\mathbf{Y};\hat{\mathbf{Y}}) + H(\hat{\mathbf{Y}} \mid \hat{\mathbf{X}}) \end{aligned} \tag{21}$$

Substitute (15) into (21), we have

$$\begin{aligned} &\geq I(\mathbf{Y};\hat{\mathbf{Y}}) - \frac{1}{2}\ln(2\pi e)^N \mid R_{\hat{X}} \mid \\ &\quad - \log|\mathbf{A}| + 2h(\hat{\mathbf{Y}} \mid \hat{\mathbf{X}}) \end{aligned} \tag{22}$$

From (16)(17), we have

$$\begin{aligned} &\frac{1}{2}\ln(2\pi e)^N \mid R_{\hat{X}} \mid \\ &\geq I(\mathbf{Y};\hat{\mathbf{Y}}) + \log|\mathbf{A}| + h(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) \end{aligned} \tag{23}$$

Step3:
Basing on the well-known entropy power inequality, we have

$$\exp\left\{\frac{2}{n}H(\mathbf{Y} \mid \hat{\mathbf{Y}})\right\} - \exp(H(\mathbf{W})) \geq \exp\left\{\frac{2}{n}H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}})\right\} \tag{24}$$

substitute the transmission rate - mutual information relation $R - I(\mathbf{Y};\hat{\mathbf{Y}}) \geq H(\hat{\mathbf{Y}} \mid \mathbf{Y})$

$$R \geq \log\left[\exp\{H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}})\} + (2\pi e)^N |\Sigma_W|\right] + I(\mathbf{Y};\hat{\mathbf{Y}}) \tag{25}$$

It is reasonable to assume that $\exp\{H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}})\} > (2\pi e)^N |\Sigma_W|$, which indicates that the posterior entropy measures both the uncertainty of actuator and noise. Use Taylor expansion on the log item

$$\begin{aligned} &\geq \log\left[\exp\{H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}})\} + \frac{(2\pi e)^N |\Sigma_W|}{\exp\{H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}})\}}\right] + I(\mathbf{Y};\hat{\mathbf{Y}}) \\ &\geq H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) + \frac{(2\pi e)^N |\Sigma_W|}{\exp\{H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}})\}} + I(\mathbf{Y};\hat{\mathbf{Y}}) \\ &= H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}}) + \frac{(2\pi e)^N |\Sigma_W|}{\exp\{H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}})\}} + H(\mathbf{Y}) - H(\mathbf{Y} \mid \hat{\mathbf{Y}}) \end{aligned} \tag{26}$$

write the posterior entropy as distortion metric constraint: $H(\mathbf{Y} \mid \hat{\mathbf{Y}}) \leq \frac{1}{2}\log(2\pi e)^N D$

$$\geq \frac{(2\pi e)^N |\Sigma_W|}{\exp\{H(\hat{\mathbf{X}} \mid \hat{\mathbf{Y}})\}} + \frac{1}{2}\log(2\pi e)^N |\Sigma_Y| - \frac{1}{2}\log(2\pi e)^N D \tag{27}$$

from (23), we have the following

$$\geq \frac{(2\pi e)^N \left|\Sigma_W\right|}{\exp\left\{\frac{1}{2}\log(2\pi e)^N \mid R_{\hat{X}} \mid -I(\mathbf{Y};\hat{\mathbf{Y}})+\log\left|\mathbf{A}\right|\right\}} + \frac{1}{2}\log\frac{\left|\Sigma_Y\right|}{D} \tag{28}$$

Hence, we have the final result on information rate of the FIR actuator model with noise W and designed input Y with conditional entropy constraint.

$$R \geq \frac{1}{2}\log\frac{\left|\Sigma_Y\right|}{D} + \frac{(2\pi e)^N \left|\Sigma_W\right|}{\sqrt{\frac{(2\pi e)^N \left|R_{\hat{X}}\Sigma_Y\right|\left|A\right|^2}{D}}} \tag{29}$$

When the actuator matrix A is adjusted, the formulation (29) can be reduced to the classic result given by Oohama [26] on multi-terminal rate -distortion theory. The lower bound also coincides with the theorem given in [13] that any coder-controller which stabilizes the plant in the mean square sense $\sup_{k\in \mathrm{W}} \mathrm{E}\left\|\mathbf{X}_k\right\|^2 < \infty$ must have a data rate R strictly satisfying

$$R > \sum_{\left|\eta_j\right|\geq 1} \log_2\left|\eta_j\right| \tag{30}$$

where $\eta_1,\ldots,\eta_n$ are the open-loop eigenvalues.
The inequality is also tight i.e. for any number $H(\mathbf{Y}\mid\hat{\mathbf{Y}}) \leq \hat{D}$, a conditional entropy stabilizing actuator at data rate (29) can be constructed, which furthermore is finite-dimensional with periodic alphabet.

4 Perspective of Power Spectrum Density Analysis

This section consider the output noise $\{v_k\}$ to derive the Rate-Distortion function under constraint in the FIR control system the same as section 3, where the FIR matrix C share the same characteristic as A.

$$\begin{cases} \mathbf{z}_k = C\mathbf{x}_k, \\ \mathbf{y}_k = \mathbf{z}_k + \mathbf{v}_k, \end{cases}$$

The power spectrum of $\{z_k\}$ in this actuator model can be found in details in [15].

$$\begin{aligned} W &= \left[e^{jw}, e^{2jw}, \ldots, e^{djw}\right] I \\ W_I &= \left[e^{-jw}, e^{-2jw}, \ldots, e^{-djw}\right] I \\ Z(w) &= C\left(W\right)^{-1} X \left(W_I\right)^{-T} C^T \end{aligned}$$

If v is Gaussian, then y is also Gaussian, it is recognized as

$$\begin{aligned} &I(x;y) = h(y) - h(y\mid x) = h(y) - h(v\mid x) = h(y) - h(v) \\ &= \frac{1}{2}\ln\left[(2\pi e)^m \left|\Sigma_y\right|\right] - \frac{1}{2}\ln\left[(2\pi e)^m\left|\Sigma_v\right|\right] \\ &= \frac{1}{2}\ln\left[\frac{\left|\Sigma_y\right|}{\left|\Sigma_v\right|}\right] = \frac{1}{2}\ln\left[\frac{\left|\Sigma_z+\Sigma_v\right|}{\left|\Sigma_v\right|}\right] \\ &= \frac{1}{2}\ln\left[\frac{\left|U_z\Lambda_z U_z^T + \Sigma_v\right|}{\left|\Sigma_v\right|}\right] \\ &= \frac{1}{2}\ln\left[\frac{\left|\Lambda_z + U_z^T\Sigma_v U_z\right|}{\left|U_z^T\Sigma_v U_z\right|}\right] \end{aligned} \tag{31}$$

where, the eigen-decomposition

$$\begin{aligned} &U_z^T\Sigma_v U_z = U_z^T\sigma_v^2 U_z = \sigma_v^2 U_z^T U_z = \sigma_v^2 I \\ &\Lambda_z = \operatorname{diag}\left(\lambda_1,\ldots,\lambda_{im}\right) \end{aligned}$$

Then, substitute the decomposition of matrix into the basic mutual information above.

$$\begin{aligned} &\frac{1}{2}\ln\left[\frac{\left|\Lambda_z + U_z^T\sum_v U_z\right|}{\left|U_z^T\sum_v U_z\right|}\right] = \frac{1}{2}\ln\left[\frac{\left|\Lambda_z+\sigma_v^2 I\right|}{\left|\sigma_v^2 I\right|}\right] \\ &= \frac{1}{2}\ln\left[\frac{\left[\prod_{i=1}^m\left[\lambda_i+\sigma_v^2\right]\right.}{\left(\sigma_v^2\right)^m}\right] \\ &= \frac{1}{2}\ln\left[\prod_{i=1}^m\left[\lambda_i+\sigma_v^2\right]\right] - \frac{1}{2}\ln\left(\sigma_v^2\right)^m \\ &= \frac{1}{2}\sum_{i=1}^m\ln\left[\lambda_i+\sigma_v^2\right] - \frac{1}{2}\ln\left(\sigma_v^2\right)^m \end{aligned} \tag{32}$$

Since the processes are stationary, the covariance matrices are Toeplitz [24], and their eigenvalues approach their limits as $k\to\infty$. Moreover, the densities of eigenvalues on the real line tend to the power spectra of the processes [27]. Accordingly,

$$\begin{aligned} I_\infty(x;y) &= \lim_{k\to\infty}\frac{I\left(x_{0,\ldots,k};y_{0,\ldots,k}\right)}{k+1} \\ &= \lim_{k\to\infty}\frac{1}{(k+1)}\left\{\frac{1}{2}\sum_{i=1}^k\ln\left[\lambda_i+\sigma_v^2\right]-\frac{1}{2}\ln\left(\sigma_v^2\right)^k\right\} \\ &= \lim_{k\to\infty}\frac{1}{(k+1)}\left\{\frac{1}{2}\sum_{i=1}^k\ln\left[\frac{\lambda_i+\sigma_v^2}{\sigma_v^2}\right]\right\} \end{aligned} \tag{33}$$

where $\sigma_v^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} V(w)dw$

$$\begin{aligned} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\ln\left[\frac{V(w)+Z(w)}{V(w)}\right]dw \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\ln\left[\frac{V(w)+C\left(W\right)^{-1}X\left(W_I\right)^{-T}C^T}{V(w)}\right]dw \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\ln\left[1+\frac{CXC^T}{V(w)}\right]dw \end{aligned} \tag{34}$$

Next, consider the target constraint$E\left[(y-\hat{y})^2\right]\leq D$

$$\begin{aligned} &\inf_{E[(y-\hat{y})^2]\leq D} I_\infty(x,\hat{y}) \\ &= \lim_{k\to\infty\infty}\min_{E[\mathbf{n}^2]\leq(k+1)D}\frac{I\left(x_{0,\ldots,k};y_{0,\ldots,k}\right)}{k+1} \\ &= \lim_{k\to\infty}\frac{1}{(k+1)}\left\{\frac{1}{2}\sum_{i=1}^k\ln\left[\frac{\lambda_i+\sigma_v^2+N_i}{\sigma_v^2+N}\right]\right\} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\ln\left[1+\frac{Z(w)}{\sigma_v^2+N(w)}\right]dw \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\ln\left[1+\frac{CXC^T}{\sigma_v^2+N(w)}\right]dw \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2}\ln\left[1+\frac{CXC^T}{\sigma_v^2+D}\right]dw \end{aligned} \tag{35}$$

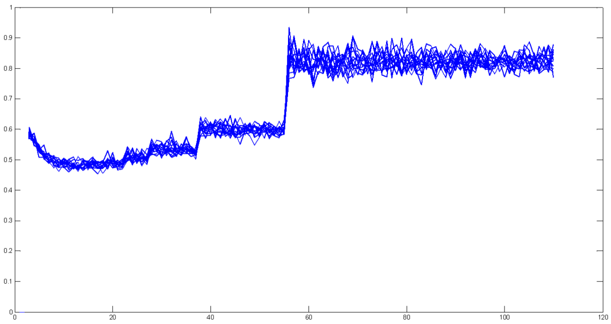

Figure 1: MSE vs. Length of FIR matrix

The 1-step prediction Wiener–Masani [28] and Kolmogorov–Szegö [29] [5]formula

$$\begin{aligned}&\liminf_{k\to\infty} \det \mathrm{E}\left[\left(\mathbf{y}_k-\overline{\mathbf{y}}_k\right)\left(\mathbf{y}_k-\overline{\mathbf{y}}_k\right)^T\right]\\&\geq 2^{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\det\ Y(\omega)\mathrm{d}\omega}\end{aligned} \tag{36}$$

$$\liminf_{k\to\infty} \mathrm{E}\left[\left(\mathbf{y}_k-\overline{\mathbf{y}}_k\right)\left(\mathbf{y}_k-\overline{\mathbf{y}}_k\right)^T\right] \geq 2^{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log Y(\omega)\mathrm{d}\omega} \tag{37}$$

are coincide with result given in (35).

5 Simulation of the FIR Actuator

Take the impulse response of length d, for the simplicity of simulation we assume it to be constant number.

Example 1: Given FIR matrix's length d as 3,the actuator model is

$$\begin{bmatrix}\hat{x}(3)\\\hat{x}(4)\\\hat{x}(5)\end{bmatrix} h(0) = \begin{bmatrix}y(3)\\y(4)\\y(5)\end{bmatrix} - \begin{bmatrix}\hat{x}(1)h(2)+\hat{x}(2)h(1)\\\hat{x}(2)h(2)+\hat{x}(3)h(1)\\\hat{x}(3)h(2)+\hat{x}(4)h(1)\end{bmatrix}$$
$$= \begin{bmatrix}y(3)\\y(4)\\y(5)\end{bmatrix} - \begin{bmatrix}h(2) & h(1) & 0 & 0\\0 & h(2) & h(1) & 0\\0 & 0 & h(2) & h(1)\end{bmatrix}\cdot$$

$$\begin{bmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\-\frac{h(2)}{h(0)} & -\frac{h(1)}{h(0)} & \frac{1}{h(0)} & 0\\\frac{h(1)h(2)}{h^2(0)} & \frac{h^2(1)-h(1)h(2)}{h^2(0)} & -\frac{h(1)}{h^2(0)} & \frac{1}{h(0)}\end{bmatrix}\begin{bmatrix}\hat{x}(1)\\\hat{x}(2)\\y(3)\\y(4)\end{bmatrix} \tag{38}$$

Given the max number of input array, which is d, as 111, use LBG algorithm [30] to quantify the input data, we can get the MSE of this FIR actuator model, the simulation results is shown in Figure 1, the x-axis means the growing number d of input array(influencing the LBG quantization order),and the y-axis is the MSE. It can be observed that the distortion between expected $\mathbf{y}$ and real output $\hat{\mathbf{y}}$ decreases as d is less than 20, then increases stairly to the max number.

Example 2: Consider the feedback matrix F in the system, given the model as

$$\begin{cases} s_{k+1} = Fs_k + x_k \\ y_k = Cs_k + v_k \end{cases} \quad F = \begin{bmatrix}0 & 0\\-4 & 4\end{bmatrix}, C = \begin{bmatrix}-1 & 1\end{bmatrix}$$

The spectrum can be figured out by

$$Z(w) = \begin{bmatrix}\frac{-e^{jw}}{e^{2jw}-4e^{jw}} & \frac{e^{jw}}{e^{2jw}-4e^{jw}}\end{bmatrix} X \begin{bmatrix}\frac{-e^{jw}}{e^{2jw}-4e^{jw}}\\\frac{e^{jw}}{e^{2jw}-4e^{jw}}\end{bmatrix}$$
$$= \begin{bmatrix}\frac{-e^{jw}}{e^{2jw}-4e^{jw}} & \frac{e^{jw}}{e^{2jw}-4e^{jw}}\end{bmatrix}\begin{bmatrix}\sigma_{x1}^2 & \rho\sigma_{x1}\sigma_{x2}\\\rho\sigma_{x1}\sigma_{x2} & \sigma_{x2}^2\end{bmatrix}\begin{bmatrix}\frac{-e^{jw}}{e^{2jw}-4e^{jw}}\\\frac{e^{jw}}{e^{2jw}-4e^{jw}}\end{bmatrix}$$
$$= \frac{\left(e^{jw}\sigma_{x1}-e^{jw}\sigma_{x2}\right)^2}{\left(e^{2jw}-4e^{jw}\right)^2}$$

Then, determine the noise matrix and substitute the result into (31) or (25).

6 CONCLUSION

Recent informational analysis on control system can also be found in [12], their work use direct information to give a causal case analysis in feedback control system. Our bound results in FIR actuator can be recognized as a more feasible case compared to theirs. We consider the actuator system from two perspective: multivariable information theory and power spectrum analysis to figure out the rate-distortion lower bound in this FIR actuator model. Further, our group can examine the cases of non-gaussian basing on the mathematical foundation of reproductivity of pdf. For the reason that the power density function possess the characteristic of reproductivity, then it is handy to obtain the actuator function in the model.

REFERENCES

[1] G. N. Nair, F. Fagnani, S. Zampieri and R. J. Evans, "Feedback Control Under Data Rate Constraints: An Overview," in Proceedings of the IEEE, vol. 95, no. 1, pp. 108-137, Jan. 2007

[2] S.Tatikonda and S. Mitter, "Control under communication constraints," in IEEE Transactions on Automatic Control, vol. 49, no. 7, pp. 1056-1068, July 2004.

[3] Borkar and S. Mitter, "LQG control with communication constraints," in Communications, Computation, Control, and Signal Processing: A Tribute to Thomas Kailath. Norwell, MA: Kluwer, 1997.

[4] Ehsan Nekouei, Takashi Tanaka, Mikael Skoglund, Karl H. Johansson, Information-theoretic approaches to privacy in estimation and control, Annual Reviews in Control,Volume 47,2019, pp.412-422.

[5] S. Fang, M. Skoglund, K. H. Johansson, H. Ishii and Q. Zhu, "Generic Variance Bounds on Estimation and Prediction Errors in Time Series Analysis: An Entropy Perspective," 2019 IEEE Information Theory Workshop (ITW), 2019, pp. 1-5.

[6] S. Fang and Q. Zhu, "Fundamental Limits of Obfuscation for Linear Gaussian Dynamical Systems: An Information-Theoretic Approach," 2021 American Control Conference (ACC), 2021, pp. 4574-4579.

[7] D. Xu, F. ZHANG. Spatial Information Theory, Beijing: Science Press, 2021.

[8] G. N. Nair, M. Huang, and R. J. Evans,Optimal infinite horizon control under a low data rate, 14th IFAC Symp. Sys. Ident.,Newcastle, Australia, Mar. 2006.

[9] Keyong Li and J. Baillieul, ”Robust quantization for digital finite communication bandwidth (DFCB) control,” in IEEE Transactions on Automatic Control, vol. 49, no. 9, pp. 1573-1584, Sept. 2004

[10] Wing Shing Wong and R. W. Brockett, ”Systems with finite communication bandwidth constraints. II. Stabilization with limited information feedback,” in IEEE Transactions on Automatic Control, vol. 44, no. 5, pp. 1049-1053, May 1999.

[11] J. Baillieul, ”Matching conditions and geometric invariants for second-order control systems,” Proceedings of the 38th IEEE Conference on Decision and Control, 1999, pp. 1664-1670 vol.2

[12] C. D. Charalambous, C. K. Kourtellaris and I. Tzortzis, ”Information Transfer of Control Strategies: Dualities of Stochastic Optimal Control Theory and Feedback Capacity of Information Theory,” in IEEE Transactions on Automatic Control, vol. 62, no. 10, pp. 5010-5025, Oct. 2017.

[13] Nair, Girish Evans, Rob. Stabilizability of Stochastic Linear Systems with Finite Feedback Data Rates. SIAM J. Control and Optimization. 43. 413-436.2004.

[14] M. Pourahmadi, Foundations of Time Series Analysis and Prediction Theory. John Wiley Sons, 2001.

[15] A. Papoulis and S. U. Pillai, Probability, Random Variables and Stochastic Processes. New York: McGraw-Hill, 2002

[16] W. R. Bennett, ”Spectra of quantized signals,” in The Bell System Technical Journal, vol. 27, no. 3, pp. 446-472, July 1948.

[17] A. S. Matveev and A. V. Savkin, The problem of LQG optimal control via a limited capacity communication channel,[Sys. Control Lett., vol. 53, no. 1, pp. 51–64, 2004.

[18] S. Yuksel, O. C. Imer, and T. Basar, Constrained state estimation and control over communication networks, Proc. 38th Annu. Conf. Inf. Sci. Syst., Princeton Univ.,Princeton, NJ, 2004.

[19] G. N. Saridis, ”Entropy formulation of optimal and adaptive control”, IEEE Transactions on Automatic Control, vol. 33, no. 8, pp. 713-721, Aug. 1988.

[20] C. E. Shannon, A mathematical theory of communication,[Bell Syst. Tech. J., 1948, Reprinted in FClaude Elwood Shannon

[21] T. Berger, Rate Distortion Theory:A Mathematical Basis for Data Compression. Englewood Cliffs, NJ: Prentice-Hall, 1971

[22] J. Ziv, ”Distortion-rate theory for individual sequences,” in IEEE Transactions on Information Theory, vol. 26, no. 2, pp. 137-143, March 1980.

[23] T.Linder and G. Lagosi, ”A zero-delay sequential scheme for lossy coding of individual sequences,” in IEEE Transactions on Information Theory, vol. 47, no. 6, pp. 2533-2538, Sept. 2001.

[24] U. Grenander and G. Szego, Toeplitz Forms and Their Applications. University of California Press, 1958.

[25] Robert M. Gray, Toeplitz and Circulant Matrices: A Review , now, 2006.

[26] Y. Oohama, ”Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder,” in IEEE Transactions on Information Theory, vol. 51, no. 7, pp. 2577-2593, July 2005.

[27] J. Gutierrez-Gutierrez and P. M. Crespo, “Asymptotically equivalent sequences of matrices and Hermitian block Toeplitz matrices with continuous symbols: Applications to MIMO systems,” IEEE Transactions on Information Theory, vol. 54, no. 12, pp. 5671–5680, 2008

[28] P. P. Vaidyanathan, The Theory of Linear Prediction. Morgan Claypool Publishers, 2007.

[29] Y. Chen, J. Karlsson and T. T. Georgiou, ”The Role of the Time-Arrow in Mean-Square Estimation of Stochastic Processes,” in IEEE Control Systems Letters, vol. 2, no. 1, pp. 85-90, Jan. 2018.

[30] Y.Linde, A. Buzo and R. Gray, ”An Algorithm for Vector Quantizer Design,” in IEEE Transactions on Communications, vol. 28, no. 1, pp. 84-95, January 1980.