The Input Design Analysis for Non Causal Finite Impulse Response Actuator: An Entropy-Rate 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: Information theory has been adapted to control in various aspects including optimal metric, bound constraint and system explanation. 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. However, the mutual information is usually a computational heavy work when used as optimal metric in control system. It is a result of both the increasing time steps and the posterior pdf. In this paper, we will show how to model the FIR system with entropy-rate framwork and give a practical algorithm foucsing on the posterior entropy to calculate the mutual information in the case of pdf with regenerative property.

Key Words: Gaussian, FIR, Actuator, Quantization noise, Bayesian, Information Theory, Entropy

1 INTRODUCTION

Most traditional control system analysis assumes that observed signal is available to controller in its entire decision time period [1]; however, the real situation is that controller receives a quantized version of original observed signal, which is accompanied by transmission delays and distortion subject to bit rate constraints [2] [3].

There exists numerous work using information theory to explain and resolve problems in optimal control. Touchette and Lloyd [4] showed that controllability and observability can be defined through using the concepts of information theory. They proved that the one-step reduction in entropy of the final state is upper bounded by mutual information between the control variables and the current state of the system.

Minimum of the one-step reduction in entropy can be found by dynamic programming (DP), but the computational complexity of DP grows exponentially with the number of time steps and control variables. Some methods is used to overcome this problem but not profound. For example, in 2020 [6], an approximation method is deveoped to avoid this problem. In [5], the mutual information is used as a utility function to be maximized under the signal energy constraint. We focus on the main computaional complexity: posterior probability, in mutual information calculation within the control systems. To be in detail, a random sampling method is used to get the posterior pdf.

One more problem, which is better addressed in communication, is the bound of specific system parameter. If the bound can be derived, then more quantitative comparisons can be carried out basing on the theoretical limit derived. Recently [7] [8] [9] has done some fundamental limit research in some aspects in control field using information theory.

We investigate the information theory modeling of FIR actuator, which is a commonly used convolutional structure. The main novel contribution is an algorithm to calculate the mutual information between system state and output basing on a kind of pdf with reproductivity.

This contribution is twofold: On the one hand, we demonstrate how to derive the entropy-rate function about a given actuator with noise, which is based on traditional rate-distortion theory. On the other hand, we define the conditional entropy as a distortion metric in an actuator system, then we derive and simulate the mutual information between system state and output.

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 [10]. 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

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

in feedback control [11]. There is a trade-off between the fineness of the control and the time delay required to send the data over the link. In [12] J. Baillieul has shown that reaching a good data quantization requires a larger channel capacity [13] 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 [14], 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 [8] [9], 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 [13], 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 Rate-Distortion

Quantization errors exist in a large body of work in control theory, one class of which [15] 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 [11] 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 [8] [16] [17] [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.Current feasible approach is related to the theory of optimal quantization.This inspires us to use entropy-rate theory to analyse the actuator model.
This paper is organized as follows.
Section 2 introduces the two preliminaries of our analysis process: entropy-rate function and Sampling A Posterior method.
In Section 3, we present modeling of FIR actuator, and then derive the posterior estimation.
In Section 4, we derive and calculate the mutual information between system state and output. The algorithm we used is given in matlab. Simulation of a non causal track system iss given in section 5.

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.
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 basic system with state-space model given by

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

, where $\mathbf{A}$ is the FIR actuator matrix.
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 [20].

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

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$N_S \times N$

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.

2.1 Entropy-Rate Function

We want to explore the information theoretic bound in a FIR system. Our framework is the entropy-rate function, which connects the distortion metric to mutual information. In this section, a brief derivation is given and we can observe that the main computation workload is on posterior entropy, particularly, the posterior probability density function sampling.

The distortion between y and $\hat{y}$ is a non-negative real number $d(y,\hat{y})$, and the mean distortion is $\overline{d}=E[d(y,\hat{y})]$. All test channels which satisfies $I(X;\hat{X})\leq R$ is defined as the set $P_R=\{f(\hat{x}\mid x):I(X;\hat{X})\leq R\}$, then the distortion-rate function is defined as:

$$D(R)=\inf_{f(\hat{x}|x)\in P_R} d(X,\hat{X}) \tag{2}$$

The concrete formulation of distortion function is decided upon real application scene. The common used distortion functions are square distortion $d(y,\hat{y})=(y-\hat{y})^2$ or absolute distortion $d(y,\hat{y})=\mid y-\hat{y}\mid$ which are corresponded to mean square distortion criterion and absolute distortion criterion.

The traditional solving method to rate-distortion function is recognized as the variational problem of mutual information [**?**] or directed information [2], which is hard to obtain the lower bound. Besides, currently only closed-form solutions exist for the rate distortion functions of a few types of sources, such as Gaussian distributions.

There are three distinct distributions which we are interested in: conditional distribution $f(\hat{y}\mid y)$ denotes the actuator, inverse conditional distribution $g(y\mid\hat{y})$, and the actuator output distribution $q(\hat{y})$.

Consider a posterior probability distributions of the following form: $g(y\mid\hat{y})=g(y-\hat{y})$, which yields that the shape of posterior distribution function has nothing to do with and only affects the location of the posterior distribution. Basing on the fact that joint Distribution $p(y,\hat{y})=g(y\mid\hat{y})q(\hat{y})$, we have:

$$p(y,\hat{y})=g(y-\hat{y})q(\hat{y}) \tag{3}$$

Integrate over $\hat{y}$:

$$p(y)=\int g(y-\hat{y})q(\hat{y})d\hat{y} \tag{4}$$

It is easy to notice that the pdf of deisgned input y is the convolution of output pdf and posterior pdf. Let $\zeta=y-\hat{y}$, then $y=\zeta+\hat{y}$. Under the case that ζ is independent of $\hat{y}$, such convolution formula established:

$$p(y)=\int g(\zeta)*q(\hat{y}) \tag{5}$$

Since pdf has convolutional relation, the corresponding characteristic functions have the multiplication relation:

$$P_Y(jw)=G_Z(jw)Q_{\hat{Y}}(jw) \tag{6}$$

Next,we consider a kind of distribution family possessing the characteristic of "regenerative", which guarantees the input and output share the same pdf form. There are many probability distributions with regenerative properties, such as Gaussian, Gamma, Chi-square and Cauchy distribution, etc. For discrete pdf, we know that Posisson, Exponential and Bernouli are pdf with regenerative property.

In this paper, we call the sources with regenerative properties of probability distribution as regenerative sources. There exist simple methods to solve the rate entropy function and rate distortion function of regenerative sources.

$$P(jw,\theta_y)=P(jw,\theta_\zeta+\theta_{\hat{y}}) \tag{7}$$

Then, we get the posterior distribution from its characteristic function:

$$G_Z(jw)=P_Y(jw)/Q_{\hat{Y}}(jw) \tag{8}$$

Given , the differential entropy of posterior distribution is:

$$h(Y\mid\hat{y})=-\int g(y\mid\hat{y})\log g(y\mid\hat{y})dy \tag{9}$$

which denotes the uncertainty at $\hat{y}$.

$$\begin{aligned} h(Y\mid\hat{Y})&=\int q(\hat{y})h(Y\mid\hat{y})d\hat{y}\\ &=-\int q(\hat{y})g(y\mid\hat{y})\log g(y\mid\hat{y})dy \end{aligned} \tag{10}$$

denotes the mean uncertainty of all $\hat{y}$.

The conditional distribution $f(\hat{y}\mid y)$ and inverse conditional distribution $g(y\mid\hat{y})$ satisfies the Bayesian formulation $p(y)f(\hat{y}\mid y)=q(\hat{y})g(y\mid\hat{y})$, we have the marginal pdf formulation

$$\int q(\hat{y})g(y\mid y)d\hat{y}=p(y) \tag{11}$$

Define the pdf set $(g(y\mid\hat{y}),q(\hat{y}))$, let

$$\begin{aligned} G_R=\{&g(y|\hat{y}),q(\hat{y})|I(Y;\hat{Y})\leq R,\\ &\int q(\hat{y})g(y\mid\hat{y})d\hat{y}=p(y)\} \end{aligned} \tag{12}$$

denotes the set $(g(y\mid\hat{y}),q(\hat{y}))$ which satisfied the rate constraint and marginal pdf formulation

$$H(R)=\inf_{(g(y|\hat{y}),q(\hat{y}))\in P_R} d(X,\hat{X}) \tag{13}$$

The entropy=rate function has two distinctively difference with distortion-rate function:

(1) distortion-rate function is defined from $f(\hat{y}\mid y)$ side, while entropy-rate function is defined from $g(y\mid\hat{y})$ side.

(2) if $d(y,\hat{y})=-\log g(y\mid\hat{y})$, then

$$\begin{aligned} \overline{d}&=E[d(y,\hat{y})]\\ &=-\iint q(\hat{y})g(y\mid\hat{y})\log g(y\mid\hat{y})dyd\hat{y}\\ &=h(Y\mid\hat{Y}) \end{aligned} \tag{14}$$

In other words, the differential entropy is also a distortion measure,but it may not satisfy non-negative. To guarantee its non-negativity, we define D(H) as entropy exponential distortion. The D(H) varies for diverse input pdf, e.g. $\frac{1}{2\pi e}e^{2h(Y|\hat{Y})}$ for Gaussian pdf, $e^{h((Y|\hat{Y})}$ for exponential pdf. In fact, entropy exponential function is the inverse of input's entropy function. To solve the entropy-rate function is traditionally a variational problem and do not seem to has a common solution.Therefore, our idea is to construct a specified pdf group satisfied (10) and achieve the lower bound.

Theorem 1: If there exist a special solution $(g(y \mid \hat{y}), q(\hat{y}))$ which make $I(Y;\hat{Y}) = R$ and satisfy (10), then we have

$$H(R) = h(Y) - R \tag{15}$$

(1) Non-negativity $H(R) \geq 0$, the equaility holds iff $\hat{X}$ and X are independent.

(2) H(R) is monotonically decreasing function of R.

For convenience, we just give the gaussian example here.

Example 1:

Assume the gaussian pdf $p(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}}\exp\left(-\frac{x^2}{2\sigma_X^2}\right)$, it is known that the differential entropy is
$h(X) = \frac{1}{2}\log 2\pi e\sigma_X^2$

$$\begin{aligned} H(R) &= \frac{1}{2}\log(2\pi eD) \\ &= \frac{1}{2}\log\left(2\pi e\sigma_X^2\right) - R \\ \Leftrightarrow D &\geq \frac{e^{2R}}{\sigma_X^2} \end{aligned} \tag{16}$$

2.2 Sampling A Posterior Method

Stochastic estimation is an estimation method analogous to the concept of random coding in Shannon's information theory, which is different from deterministic estimation. The difference between stochastic estimation and deterministic estimation is that stochastic estimation may yield different estimates for the same set of samples, which is not uniquely deterministic. The posterior probability estimation of sampling proposed in this paper is a stochastic estimation method. It is based on a random sampling of the posterior probability generated by each received signal. The statistical distribution of this estimate will naturally approximate the theoretical posterior probability distribution of the parameter to be estimated. The sampling posterior probability estimation method is described below:

1.Generating received signal vector.

2.Matching conditional PDF of received signal vector.

3.Calculate posterior PDF of the parameter to be estimated.

4.Generating an estimate by random sampling of the posterior PDF.

5.Multiple estimations, calculate the empirical entropy, and obtain the average estimation performance.

3 FIR Actuator Model and Posteriori Estimation

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 is:

$$\begin{aligned} y(n) &= \hat{x}(n) * h(n) \\ &= \sum_{l=n-d+1}^{n} \hat{x}(l)h(n-l) \\ &= \hat{x}(n)h(0) + \sum_{l=n-d+1}^{n-1} \hat{x}(l)h(n-l) \end{aligned} \tag{17}$$

The system state equations thus can be written as:

$$\begin{cases} \mathbf{x}_k = A\mathbf{x}_{k-1} + B\mathbf{e}_{k-1} + \mathbf{v}_k \\ \mathbf{y}_k = C\mathbf{x}_k + D\mathbf{e}_k + \mathbf{w}_k \end{cases} \tag{18}$$

Combined with section 2.2, we give the whole system model:

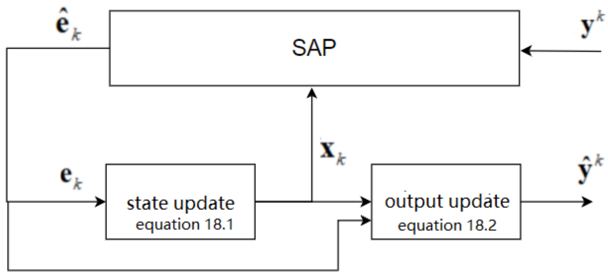

Figure 1: System Input Design with SAP method

Basing on the state equations in (18)

$$\begin{cases} p\left(\mathbf{y}_k \mid \mathbf{e}_k, \mathbf{x}_k\right) = f_w\left(\mathbf{y}_k - C\mathbf{x}_k - D\mathbf{e}_k\right) \\ p\left(\mathbf{x}_k \mid \mathbf{x}_{k-1}, \mathbf{e}_{k-1}\right) = f_v\left(\mathbf{x}_k - A\mathbf{x}_{k-1} - B\mathbf{e}_{k-1}\right) \end{cases} \tag{19}$$

Gaussian noise v_k, w_k are independent.

$$\begin{aligned} & p\left(e_k \mid y_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right) \\ &= \frac{p\left(e_k, y_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right)}{p\left(y_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right)} \end{aligned} \tag{20}$$

the molecular of (20) have the following set of relations

$$\begin{aligned} p\left(e_k, y_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right) = p\left(y_k \mid e_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right) \cdot \\ p\left(e_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right) \end{aligned} \tag{21}$$

The first item in(21)

$$\begin{aligned} & p\left(y_k \mid e_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right) \\ &= p\left(y_k \mid e_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right) p\left(y^{k-1} \mid e_k, e^{k-1}, x^{k-1}, x_k\right) \\ &= p\left(y_k \mid e_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right) p\left(y^{k-1} \mid e^{k-1}, x^{k-1}\right) \end{aligned} \tag{22}$$

The second item in(21)

$$\begin{aligned} & p\left(e_k, e^{k-1}, x^{k-1}, x_k\right) \\ &= p\left(x_k \mid e_k, e^{k-1}, x^{k-1}\right) p\left(e_k, e^{k-1}, x^{k-1}\right) \\ &= p\left(x_k \mid e_k, e^{k-1}, x^{k-1}\right) p\left(x^{k-1} \mid e_k, e^{k-1}\right) p\left(e_k, e^{k-1}\right) \\ &= p\left(x_k \mid e_k, e^{k-1}, x^{k-1}\right) p\left(x^{k-1} \mid e^{k-1}\right) p\left(e_k\right) p\left(e^{k-1}\right) \\ &= p\left(x_k \mid e_k, e^{k-1}, x^{k-1}\right) p\left(x^{k-1}, e^{k-1}\right) p\left(e_k\right) \end{aligned} \tag{23}$$

Combine (22) and (23), we have

$$p\left(e_k, y_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right) \\ = f_w f_v p\left(y^{k-1}, e^{k-1}, x^{k-1}\right) p\left(e_k\right) \tag{24}$$

Return to (20), we get the equation

$$p\left(e_k \mid y_k, y^{k-1}, e^{k-1}, x^{k-1}, x_k\right) \\ = \frac{f_w f_v p\left(y^{k-1}, e^{k-1}, x^{k-1}\right) p\left(e_k\right)}{E_{p(e_k)}\left[f_w f_v p\left(y^{k-1}, e^{k-1}, x^{k-1}\right)\right]} \tag{25}$$

For example, consder the gaussian case and start from the first step, using Bayesian equation we have:

$$p\left(e_1 \mid y_1\right) = \frac{p_w\left(y_1 \mid e_1\right) p\left(e_1\right)}{p\left(y_1\right)}$$

$$p_w\left(y_1 \mid e_1\right) = \frac{1}{\sqrt{2\pi\sigma_w^2}} e^{-\frac{(y-C_1-D_\varphi)^2}{2\sigma_w^2}}$$

$$p\left(e_1\right) = \frac{1}{\sqrt{2\pi\sigma_{e_1}^2}} e^{-\frac{(\varphi_1)^2}{2\sigma_{\theta_1}^2}}$$

$$p\left(e_1 \mid y_1\right) = \frac{1}{2\pi\sqrt{\sigma_w^2\sigma_{e_1}^2}} e^{-\frac{(y_1-C_1-D_1)^2}{2\sigma_w^2} - \frac{(\varphi_1)^2}{2\sigma_{\ell_1}^2}}$$

let $\frac{\partial \ln p(e_1|y_1)}{\partial e_1} = 0$, we can get the ML estimation of input e in the current stage:

$$e_1 = \frac{D\left(y_1 - Cx_1\right)}{\left(\frac{\sigma_w^2}{\sigma_{e_1}^2} + D^2\right)} \tag{26}$$

the next stages are derived in the same way.

4 Mutual Information of FIR Actuator

In this section, we examine the fundamental limits of Finite Response Actuator for linear Gaussian systems. The main problem is on how to handle mutual information in actuator model. We give an algorithm based on sampling method to quantitatively figure out the mutual information between the system output and state.

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

$$R(d) = E\left[\hat{\mathbf{x}}h(0)[\hat{\mathbf{x}}h(0)]^T\right] \\ = \left[\begin{array}{cccc} R_{xx}(0) & R_{xx}(-1) & \cdots & R_{xx}(-d+1) \\ R_{xx}(1) & R_{xx}(0) & \cdots & R_{xx}(-d+2) \\ \vdots & \vdots & \ddots & \vdots \\ R_{xx}(d-1) & R_{xx}(d-2) & \cdots & R_{xx}(0) \end{array}\right] \tag{27}$$

where

$$\hat{\mathbf{x}}^T h(0) = \left[\begin{array}{c} \left.y(d) - \sum_{l=+1}^{d-1} \hat{x}(l)h(d-l)\right) \\ \left.y(d+1) - \sum_{l=2}^{d} \hat{x}(l)h(d+1-l)\right) \\ \vdots \\ \left.y(2d-1) - \sum_{l=d}^{2d-2} \hat{x}(l)h(2d-1-l)\right) \end{array}\right]$$

The designed output $y \sim N(0, \Sigma_Y)$, distortion $(y-\hat{y}) \sim N(0, \mathbf{D})$. 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:

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

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

The joint distribution of state and output is written as:

$$p(\hat{\mathbf{x}}, \mathbf{y}) = p(\mathbf{y}) p(\hat{\mathbf{x}} \mid \mathbf{y}) = p(\mathbf{y}) p_w(\hat{\mathbf{x}} - \mathbf{y}) \\ = \frac{1}{(2\pi)^N |\mathbf{\Sigma_Y}||\mathbf{D}|} \exp\left(-\frac{\mathbf{y}^{\mathrm{T}} \mathbf{\Sigma_Y}^{-1} \mathbf{y}}{2} \right. \\ \left. - \frac{(\hat{\mathbf{x}} - \mathbf{y})^T \mathbf{D}^{-1} (\hat{\mathbf{x}} - \mathbf{y})}{2}\right) \tag{28}$$

Integrate about y,

$$p(\hat{\mathbf{x}}) = \int p(\mathbf{y}) p(\hat{\mathbf{x}} \mid \mathbf{y}) dy \\ = \frac{1}{\sqrt{(2\pi)^N (|\mathbf{\Sigma_Y}| + |\mathbf{D}|)}} \exp\left(-\frac{\hat{\mathbf{x}}^T \left(\mathbf{\Sigma_Y} + \mathbf{D}\right)^{-1} \hat{\mathbf{x}}}{2}\right) \tag{29}$$

Substitute into Bayesian equation, we have the posterior pdf:

$$p(\mathbf{y} \mid \hat{\mathbf{x}}) = \frac{p(\mathbf{y}) p(\hat{\mathbf{x}} \mid \mathbf{y})}{p(\hat{\mathbf{x}})} \\ = \frac{\sqrt{(2\pi)^N (|\mathbf{\Sigma_Y}| + |\mathbf{D}|)}}{(2\pi)^N |\mathbf{\Sigma_Y}| |\mathbf{D}|} \exp\left(-\frac{\mathbf{y}^T \left(\mathbf{\Sigma_Y}\right) \mathbf{y}}{2} \right. \\ \left. - \frac{(\hat{\mathbf{x}} - \mathbf{y})^T \mathbf{D}^{-1} (\hat{\mathbf{x}} - \mathbf{y})}{2} + \frac{\hat{\mathbf{x}}^T \left(\mathbf{\Sigma_Y} + \mathbf{D}\right)^{-1} \hat{\mathbf{x}}}{2}\right. \tag{30}$$

Use SAP algorithm described below to calculate the posterior entropy:

$$h(\mathbf{y} \mid \mathbf{x}) = -\iint p(\mathbf{x}, \mathbf{y}) \ln p(\mathbf{y} \mid \mathbf{x}) dx dy \\ = \frac{\sqrt{(2\pi)^N (|\mathbf{\Sigma_Y}| + |\mathbf{D}|)}}{|\mathbf{D}|} \left[\ln \frac{\sqrt{(2\pi)^N (|\mathbf{\Sigma_Y}| + |\mathbf{D}|)}}{(2\pi)^N |\mathbf{\Sigma_Y}| |\mathbf{D}|} - 1\right] \tag{31}$$

Then, we can get the mutual information from posterior pdf(31).

$$I(\mathbf{y}; \mathbf{x}) = \iint p(\mathbf{x}, \mathbf{y}) \ln \frac{p(\mathbf{y} \mid \mathbf{x})}{p(\mathbf{y})} d\mathbf{x} d\mathbf{y} = h(\mathbf{y}) - h(\mathbf{y} \mid \mathbf{x}) \\ = \frac{1}{2} \ln(2\pi)^N |\mathbf{\Sigma_Y}| + \frac{\sqrt{(2\pi)^N (|\mathbf{\Sigma_Y}| + |\mathbf{D}|)}}{|\mathbf{D}|} \\ \left[\ln \frac{\sqrt{(2\pi)^N (|\mathbf{\Sigma_Y}| + |\mathbf{D}|)}}{(2\pi)^N |\mathbf{\Sigma_Y}| |\mathbf{D}|} - 1\right] \tag{32}$$

Algorithm 1: Sampling from A Posterior Method

Data: received signal vector, posterior pdf, prior entropy H_prior
Result: posterior entropy H, mutual information

initialization;
for $j = 1; j \leq sampling\_iteration\_num$ **do**
  Use random sampling method to get sample points from posterior pdf, eg. randsample() in matlab ;
  while *sampling_num not end* **do**
    H = H - $\log(sampling\_point)$;
  end
  H = H / sampling_num;
  I = H_prior - H;
end
return I;

5 Simulation of a Non Causal Track System

In figure 2, we calculate the mutual information using algorithm1. We use signal to noise ratio(SNR) here to describe magnitude relation between designed output y and gaussian noise. Besides, we also do the sampling using Metropolis–Hastings algorithm in figure3 for comparison on this posterior entropy calculation task. From figure 2 3, it is obvious that the posterior pdf sampling effect of Algorithm 1 can be adaptive to both high snr and low snr.

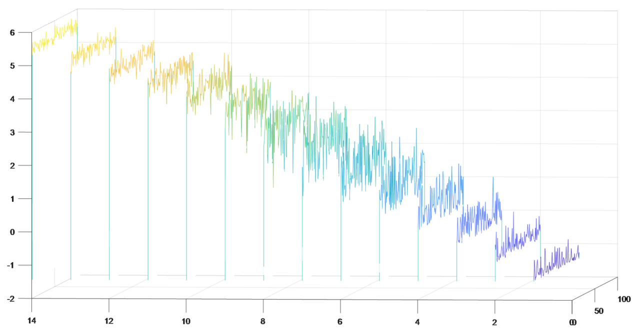

Figure 2: posterior entropy vs. different SNR using Algorithm 1

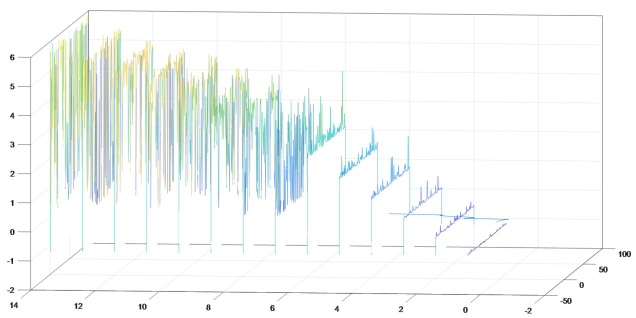

Figure 3: posterior entropy vs. different SNR using M-H sampling

Next, we carry out the simulation based on equation (17). The FIR system has a bandwidth of 0 1hz, and it functions as a track system with reference input y. As can be shown in figure 4, the left column is the FFT of time series and the right column is the time series. The first row is the designed output y, while the second row is the actual output y. The third row is input e basing on the estimation equation(26). The forth row is the noise w in system equations. The track system is a lowpass filter with 1Hz bandwidth. We can observe that as the input frequency getting closer to the system cut-off frequency, the tracking process performs worse. This result is reasonable in entropy-rate framework, because the control system has a control capacity quantified by mutual information.

In figure 5, MSE is calculated between y and $\hat{y}$. As the input frequency exceeds the allowed entropy-rate relation, the signal tracking process will not keep up, which is the MSE arises with the trackced signal frequency arising.

Mutual inforamtion result in figure 6 reflects the capacity of this control system. Control system possess its limit to track signal, $I(y|\hat{y})$ tells us the information from y to $\hat{y}$. We can see the same phenomenon in MSE: as input frequency getting closer to the FIR system‘s cutoff frequency , mutual inforamtion decreases at the same time.

6 CONCLUSION

We show that entropy-rate framework works for the input design of a FIR actuator model. The main method is a SAP algorithm for the mutual information calculation. For this reason, we examine a kind of regenerative pdf, then it is handy to obtain the actuator pdf function. Finally, we give the posterior pdf recursive deduction, which shows the probability density relation between input and system state. We believe that the entropy-rate relation can be a feasible framework to explain control system problems.

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] Touchette, H. and Lloyd, S. (2004). Information-theoretic approach to the study of control systems, Physica A 331(1): 140–172.

[5] Bania, Piotr. ”An Information Based Approach to Stochastic Control Problems” International Journal of Applied Mathematics and Computer Science, vol.30, no.1, 2020, pp.23-34.

[6] Bania, Piotr. 2019. ”Bayesian Input Design for Linear Dynamical Model Discrimination” Entropy 21, no. 4: 351

[7] 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.

[8] 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.

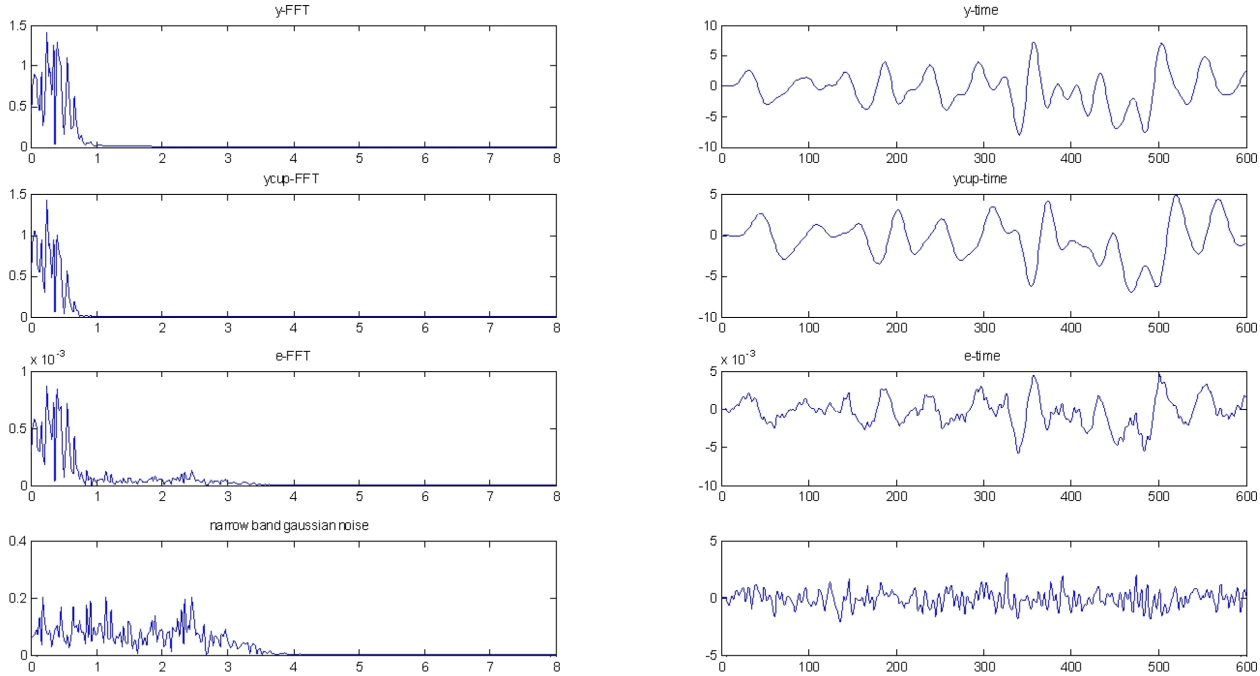

Figure 4: Simulation of the track system

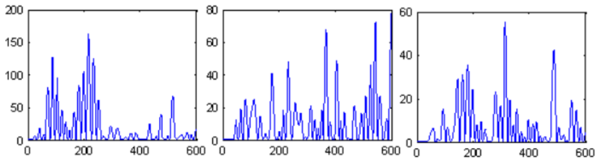

Figure 5: left - middle - right: y with frequency 0.95 - 0.8 - 0.5 hz

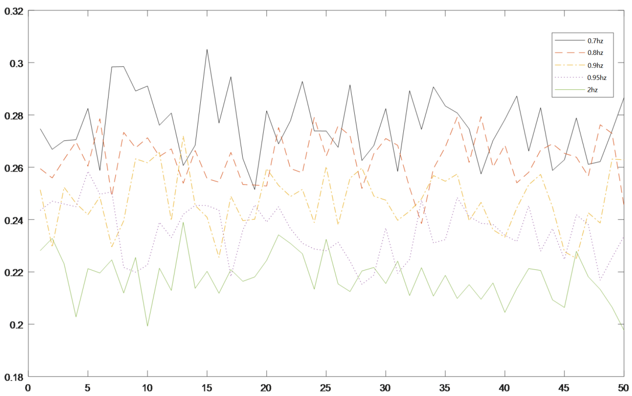

Figure 6: Mutual Information $I(y|\hat{y})$ as input frequency y increase from 0.7hz to 2hz

[9] 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.

[10] 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

[11] 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.

[12] 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

[13] 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.

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

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

[16] 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.

[17] 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.

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

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

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