Robust Control over Uncertain Communication Channels

Charalambos D. Charalambous, Alireza Farhadi, Stojan Denic, and Farzad Rezaei

Abstract— In this paper, the mathematical framework for studying robust control over uncertain communication channels is introduced. The theory is developed by generalizing the classical information theoretic measures of information and the fundamental theorems of Shannon to the robust analogous, which are subject to uncertainty in the source and the communication channel. Then, by invoking these generalized measures of information, necessary conditions for robust control over communication channels subject to uncertainty are presented.

I. Introduction

One of the issues that has begun to emerge in a number of applications, such as sensor networking, large scale teleoperation, and etc., is how to control systems by communicating information reliably, through limited capacity channels, when the subsystems are subject to uncertainty. Typical examples are applications in which a single dynamical system sends feedback information to a distant controller via a communication link with finite capacity. In the absence of uncertainty in the control system and the communication channel, important results are derived in [1], [2], [3], [4], [5], [6], [7], [8], [9]. Specially, the aim of these articles is to find a necessary and sufficient condition for stabilizability, when there are channel capacity and power constraints. For finite-dimensional linear time invariant systems, it is shown that the transmission data rate (channel capacity) required to stabilize a controlled system must be at least equal to the sum of logarithms of the unstable open-loop eigenvalues.

The objective of this paper is to address similar question when there is uncertainty in the controlled system and communication link. In particular, to find necessary conditions on the channel capacity which ensure robust observability and stabilizability. The necessary steps in realizing such a study consists of the followings. 1. Give precise definitions of entropy, channel capacity, and rate distortion, when the communication blocks are subject to uncertainty. 2. Extend the fundamental theorems of Shannon to communication systems subject to uncertainty. 3. Derive necessary and sufficient conditions on the communication blocks which are subject to uncertainty in order to ensure robust observability and stabilizability of the controlled system. Clearly, the above questions are addressed by generalizing information theory to robust information theory which consists of the robust version of the classical source coding, channel coding, and rate distortion to their robust analog, which are subject to uncertainty. Then we show that the so called robust transmission rate of the channel must be at least equal to the robust entropy of the source in order to ensure reliable communication. Subsequently, we find necessary conditions for robust observability and stabilizability for uncertain systems over uncertain communication channels.

In the Section II, the precise notion of a robust communication system, and the corresponding information theoretic measures, which are necessary to analyze these systems are introduced. One of the fundamental results which is required to address issue 3 above is the derivation of a lower bound for robust rate distortion. In Section III, a robust version of information transmission theorem is introduced. This theorem provides an upper bound for robust rate distortion. In Section IV, a necessary condition for robust observability and stabilizability is derived for fully observed, finite dimensional, noiseless uncertain linear systems over uncertain channels.

II. Robust Communication systems

A. Communication System

Let $(\Omega, \mathcal{F}(\Omega))$ denote a measurable space in which $\mathcal{F}(\Omega)$ is the σ-field generated by Ω, and let $\mathcal{M}(\Omega)$ be the set of probability measure on $(\Omega, \mathcal{F}(\Omega))$.

Consider the communication diagram given in Figure $IV.1$. Here $(\mathcal{X}, \mathcal{F}(\mathcal{X}))$ is the source measurable space, and $(\widetilde{\mathcal{X}}, \mathcal{F}(\widetilde{\mathcal{X}}))$ is the source reproduction measurable space. The channel input and output measurable spaces are $(\mathcal{Z}, \mathcal{F}(\mathcal{Z}))$ and $(\widetilde{\mathcal{Z}}, \mathcal{F}(\widetilde{\mathcal{Z}}))$, respectively.

An information source is often specified by the probability measure $P_X : \mathcal{F}(\mathcal{X}) \rightarrow [0,1]$ induced by the source on $(\mathcal{X}, \mathcal{F}(\mathcal{X}))$(e.g., $P_X \in \mathcal{M}(\mathcal{X})$).

Communication Channel. A communication channel is a probabilistic mapping $P_{\widetilde{Z}|Z} : \mathcal{Z} \times \mathcal{F}(\widetilde{\mathcal{Z}}) \rightarrow [0,1]$,

$$P_{\widetilde{Z}|Z}(z, B) \stackrel{\triangle}{=} \Pr(\widetilde{Z} \in B | Z = z), \ z \in \mathcal{Z}, \ B \in \mathcal{F}(\widetilde{\mathcal{Z}}), \tag{II.1}$$

which satisfies the following conditions.

1) For every $z \in \mathcal{Z}$, the set function $P_{\widetilde{Z}|Z}(z, .)$ is a

This work is supported by the European Commission under the project ICCCSYSTEMS

C.D. Charalambous is with the Department of Electrical and Computer Engineering, University of Cyprus, 75 Kallipoleos Avenue, Nicosia, CYPRUS. Also with the School of Information Technology and Engineering, University of Ottawa, 161 Louis Pasteur, A519, Ottawa, Ontario, K1N 6N5, CANADA. E-mail: chadcha@ucy.ac.cy

A. Farhadi, S. Denic, and F. Rezaei are Ph.D. students with the School of Information Technology and Engineering, University of Ottawa, 161 Louis Pasteur, A519, Ottawa, Ontario, K1N 6N5, CANADA. E-mail: {afarhadi, sdenic, frezaei }@site.uottawa.ca

probability measure on $\mathcal{F}(\widetilde{\mathcal{Z}})$.
2)For every $B \in \mathcal{F}(\widetilde{\mathcal{Z}})$, the function $P_{\widetilde{Z}|Z}(.,B)$ is an $\mathcal{F}(\mathcal{Z})$-measurable function.
A mapping which satisfies 1), 2) is called a stochastic kernel, and clearly, $P_{\widetilde{Z}|Z}(z,.) \in \mathcal{M}(\widetilde{\mathcal{Z}}),\ \forall z \in \mathcal{Z}$.
The definition of the channel as a stochastic kernel implies that the probabilities of events $B \in \mathcal{F}(\widetilde{\mathcal{Z}})$ conditional on the input to the channel and the input of the encoder, namely $Z=z, X=x$, do not depend on the values of the input to the encoder $X=x$. That is,

$$\begin{aligned}&\Pr(\widetilde{Z} \in B|Z=z, X=x) = \Pr(\widetilde{Z} \in B|Z=z),\\ &\forall B \in \mathcal{F}(\widetilde{\mathcal{Z}}), z \in \mathcal{Z}, x \in \mathcal{X}.\end{aligned} \tag{II.2}$$

Hence, $\mathcal{F}(\widetilde{\mathcal{Z}})$ and $\mathcal{F}(\widetilde{X})$ are conditionally independent given $\mathcal{F}(\mathcal{Z})$. In the parlame of information theory, $(II.2)$ is denoted by $\mathcal{X} \rightarrow \mathcal{Z} \rightarrow \widetilde{\mathcal{Z}}$, and implies that the elements of $\mathcal{X}, \mathcal{Z}, \widetilde{\mathcal{Z}}$ form a Markov chain.
Encoder. An encoder is a stochastic kernel $P_{Z|X} : \mathcal{X} \times \mathcal{F}(\mathcal{Z}) \rightarrow [0,1]$,

$$\begin{aligned}&P_{Z|X}(x,A) = \Pr(Z \in A|X=x),\\ &x \in \mathcal{X},\ A \in \mathcal{F}(\mathcal{Z}).\end{aligned} \tag{II.3}$$

Decoder. A decoder is a stochastic kernel $P_{\widetilde{X}|\widetilde{Z}} : \widetilde{\mathcal{Z}} \times \mathcal{F}(\widetilde{\mathcal{X}}) \rightarrow [0,1]$,

$$\begin{aligned}&P_{\widetilde{X}|\widetilde{Z}}(\widetilde{z},C) = \Pr(\widetilde{X} \in C|\widetilde{Z}=\widetilde{z}),\\ &\widetilde{z} \in \widetilde{\mathcal{Z}},\ C \in \mathcal{F}(\widetilde{\mathcal{X}}).\end{aligned} \tag{II.4}$$

Deterministic encoders and decoders correspond to delta measures, and hence they follow from $(II.3)$ and $(II.4)$. Thus, the definition of the decoder as stochastic kernel implies that $\mathcal{Z} \rightarrow \widetilde{\mathcal{Z}} \rightarrow \widetilde{\mathcal{X}}$ forms a Markov chain, and therefore $\mathcal{X} \rightarrow \mathcal{Z} \rightarrow \widetilde{\mathcal{Z}} \rightarrow \widetilde{\mathcal{X}}$ forms a Markov chain as well.
Clearly, the above construction implies that the probability measure induced by the input of the channel on $(\mathcal{Z}, \mathcal{F}(\mathcal{Z}))$ can be defined through the Radon-Nikodym derivative

$$\begin{aligned}&P_Z(A) = \int_{\mathcal{X}} P_{Z|X}(x,A)dP_X(x),\\ &\forall A \in \mathcal{F}(\mathcal{Z}),\ x \in \mathcal{X}.\end{aligned} \tag{II.5}$$

However, often it is necessary to impose certain limitation on the input to the channel (such as average channel input power). These kind of limitations are introduced by assuming that the probability measure corresponding to the channel input measurable space $(\mathcal{Z}, \mathcal{F}(\mathcal{Z}))$ belongs to a smaller class $\mathcal{M}_{CI} \subset \mathcal{M}(\mathcal{Z})$.
Similarly, the probability measure induced by the output of the channel on $(\widetilde{\mathcal{Z}}, \mathcal{F}(\widetilde{\mathcal{Z}}))$, is defined through Radon-Nikodym derivative

$$\begin{aligned}&P_{\widetilde{Z}}(B) = \int_{\mathcal{Z}} P_{\widetilde{Z}|Z}(z,B)dP_Z(z),\\ &\forall B \in \mathcal{F}(\widetilde{\mathcal{Z}}),\ z \in \mathcal{Z}.\end{aligned} \tag{II.6}$$

Moreover, the probability measure induced by the output of the decoder on $(\widetilde{\mathcal{X}}, \mathcal{F}(\widetilde{\mathcal{X}}))$ is defined by

$$\begin{aligned}&P_{\widetilde{X}}(C) = \int_{\widetilde{\mathcal{Z}}} P_{\widetilde{X}|\widetilde{Z}}(\widetilde{z},C)dP_{\widetilde{Z}}(\widetilde{z}),\\ &\forall C \in \mathcal{F}(\widetilde{\mathcal{X}}),\ \widetilde{z} \in \widetilde{\mathcal{Z}}.\end{aligned} \tag{II.7}$$

Clearly, the above construction leads to the definition of the joint probability measure $P_Z \otimes P_{\widetilde{Z}|Z}$ induced on $\mathcal{F}(\mathcal{Z}) \times \mathcal{F}(\widetilde{\mathcal{Z}})$, via

$$\begin{aligned}&P_Z \otimes P_{\widetilde{Z}|Z}(G) = \int_{\mathcal{Z}} P_{\widetilde{Z}|z}(z,G_z)dP_Z(z),\\ &\forall G \in \mathcal{F}(\mathcal{Z}) \times \mathcal{F}(\widetilde{\mathcal{Z}}),\ z \in \mathcal{Z},\end{aligned} \tag{II.8}$$

where

$$G_z = \{\widetilde{z} \in \widetilde{\mathcal{Z}}; (z,\widetilde{z}) \in G\}. \tag{II.9}$$

Moreover, the product measure of P_Z and $P_{\widetilde{Z}}$ is denoted by $P_Z \otimes P_{\widetilde{Z}}$, and it is defined by

$$\begin{aligned}&P_Z \otimes P_{\widetilde{Z}}(G) = \int_{\mathcal{Z}} P_{\widetilde{Z}}(G_z)dP_Z(z),\\ &\forall G \in \mathcal{F}(\mathcal{Z}) \times \mathcal{F}(\widetilde{\mathcal{Z}}),\ z \in \mathcal{Z}.\end{aligned} \tag{II.10}$$

Consequently, using the conditional independence of the Markov chain $\mathcal{X} \rightarrow \mathcal{Z} \rightarrow \widetilde{\mathcal{Z}} \rightarrow \widetilde{\mathcal{X}}$, the joint probability measure induced on $\mathcal{F}(\mathcal{X}) \times \mathcal{F}(\mathcal{Z}) \times \mathcal{F}(\widetilde{\mathcal{Z}}) \times \mathcal{F}(\widetilde{\mathcal{X}})$ denoted by $P_{X,Z,\widetilde{Z},\widetilde{X}}$ is given by

$$\begin{aligned}&P_{X,Z,\widetilde{Z},\widetilde{X}}(dx,dz,d\widetilde{z},d\widetilde{x})\\ &= P_X(dx) \otimes P_{Z|X}(x,dz) \otimes P_{\widetilde{Z}|Z}(z,d\widetilde{z})\\ &\quad \otimes P_{\widetilde{X}|\widetilde{Z}}(\widetilde{z},d\widetilde{x}).\end{aligned} \tag{II.11}$$

Finally, the reconstruction of $\widetilde{X}$ from X is defined through stochastic kernel

$$\begin{aligned}&Q_{\widetilde{X}|X}(x,C) = \Pr(\widetilde{X} \in C|X=x),\\ &x \in \mathcal{X},\ C \in \mathcal{F}(\widetilde{\mathcal{X}}),\\ &Q_{\widetilde{X}|X} : \mathcal{X} \times \mathcal{F}(\widetilde{\mathcal{X}}) \rightarrow [0,1].\end{aligned} \tag{II.12}$$

Hence, $P_{X,\widetilde{X}}(dx,d\widetilde{x}) \stackrel{\triangle}{=} P_X(dx) \otimes Q_{\widetilde{X}|X}(x,d\widetilde{x})$ is the joint probability measure induced on $\mathcal{F}(\mathcal{X}) \times \mathcal{F}(\widetilde{\mathcal{X}})$ by P_X and $Q_{\widetilde{X}|X}$, and $P_X(dx) \otimes P_{\widetilde{X}}(d\widetilde{x})$ is the product measure of P_X and $P_{\widetilde{X}}$ on $\mathcal{F}(\mathcal{X}) \times \mathcal{F}(\widetilde{\mathcal{X}})$.

B. Robust Information Theoretic Measures

In this subsection, we first introduce robust definitions of the entropy of the source, channel capacity and the rate distortion. Second, we extend some of the fundamental theorems of Shannon to their robust analogs. Third, we consider specific source and channel uncertainty, and we present the forms of robust entropy and channel capacity.

1) Definitions of Information Theoretic Measures: First, we introduce robust definitions of information theoretic measures, for memoryless systems. These definitions are given using relative entropy between two measures which is defined below.

Definition 2.1: (Relative Entropy) The relative entropy of two probability measures π and ν on $(\Omega, \mathcal{F}(\Omega))$ is defined by

$$H(\pi|\nu) \stackrel{\triangle}{=} \tag{II.13}$$

$$\begin{cases} \int \log \frac{d\pi(x)}{d\nu(x)} d\pi(x) & if \pi << \nu, and \;\; log \frac{d\pi(x)}{d\nu(x)} \in L^1(\pi) \\ \infty & if \;\; otherwise \end{cases}$$

where " $<<$ " denotes absolute continuity of measures. Next, by invoking the relative entropy one can define the so called mutual information measure.

Definition 2.2: (Mutual Information) Let $(\mathcal{M}, \mathcal{F}(\mathcal{M}), P_M)$ and $(\mathcal{N}, \mathcal{F}(\mathcal{N}), P_N)$ be two probability spaces. Let $Q_{N|M} : \mathcal{M} \times \mathcal{F}(\mathcal{N}) \rightarrow [0, 1]$, be a stochastic kernel, and define the joint probability measure on $\mathcal{F}(\mathcal{M}) \times \mathcal{F}(\mathcal{N})$ by $P_{M,N}(dm, dn) = P_M(dm) \otimes Q_{N|M}(m, dn)$, and the product measure of the marginal measures on $\mathcal{F}(\mathcal{M}) \times \mathcal{F}(\mathcal{N})$ by $P_M(dm) \otimes P_N(dn)$.
The mutual information is defined by the relative entropy of $P_{M,N}$ and $P_M \otimes P_N$ via

$$\begin{aligned} I(M; N) &\stackrel{\triangle}{=} H(P_{M,N}|P_M \otimes P_N). \\ &= \int \log \frac{dP_{M,N}(n, m)}{dP_M(m) \otimes dP_N(n)} dP_{M,N}(m, n). \end{aligned} \tag{II.14}$$

Next we define the concept of robust entropy for a family of sources. This measure represents the amount of the information generated by the source symbols. This definition first appeared in [10].

Definition 2.3: (Robust Entropy of the Source) Consider the source probability measure $P_X \in \{P_X \in \mathcal{M}(\mathcal{X}); p_x = \frac{dP_X}{dx}\}$, where p_X is called source distribution. In many practical application, the true source distribution is unknown but it belongs to a set $\mathcal{M}^d_{SU}$, called the source distribution uncertainty set. For these sources when the source distribution belongs to $p_X \in \mathcal{M}^d_{SU}$, the robust entropy is defined by

$$H_{robust}(p^*_X) = \sup_{p_X \in \mathcal{M}^d_{SU}} H(p_X), \tag{II.15}$$

where

$$p^*_X = argsup_{\; p_X \in \mathcal{M}^d_{SU}} H(p_X). \tag{II.16}$$

The importance of the entropy of the source can be understood in terms of the so called Shannon first coding theorem [11], [12]. The robust Shannon first coding theorem derived in [10] states that source words of block length n produced by a discrete memoryless source (e.g. a finite alphabet source with i.i.d. outcomes) with unknown probability density p_X, which belongs to a uncertainty set $p_X \in \mathcal{M}^d_{SU} = \mathcal{M}^d_{SUR} \stackrel{\triangle}{=} \{p_X; H(p_X|\mu_X) \leq R\}$, ($\mu_X$ is nominal i.i.d. source distribution and $R \geq 0$ controls the size of uncertainty and it is known) can be encoded into codewords of block length r from a coding alphabet of size k with probability of block length decoding failure p_e arbitrary small for n-sufficiently large, regardless of the true source distribution, if and only if $\sup_{p_X \in \mathcal{M}_{SU}} H(p_X) \leq \frac{r}{n} logk$ [10].

Next, we define memoryless robust channel capacity. This measure provides the maximum achievable rate under which a reliable transmission of information is possible through memoryless channels (e.g, the output of the channel is only dependent on the input of the channel at that time and conditionally independent of previous channel inputs and outputs).

Definition 2.4: (Robust Memoryless Channel Capacity) In many practical applications the communication channel $P_{\widetilde{Z}|Z} : \mathcal{Z} \times \mathcal{F}(\widetilde{\mathcal{Z}}) \rightarrow [0, 1]$ belongs to the set $P_{\widetilde{Z}|Z} \in \mathcal{M}_{CU} \subset \mathcal{M}(\widetilde{\mathcal{Z}})$, called the channel uncertainty set. For these memoryless channels, the robust channel capacity is defined by

$$\begin{aligned} C_{robust} &= \sup_{P_Z \in \mathcal{M}_{CI}} \inf_{P_{\widetilde{Z}|Z} \in \mathcal{M}_{CU}} I(Z; \widetilde{Z}) \\ &= \sup_{P_Z \in \mathcal{M}_{CI}} \inf_{P_{\widetilde{Z}|Z} \in \mathcal{M}_{CU}} H(P_Z \otimes P_{\widetilde{Z}|Z} | P_Z \otimes P_{\widetilde{Z}}). \end{aligned} \tag{II.17}$$

The above definition for channel capacity is the so called information definition for channel capacity. The importance of this definition can be understood in terms of Shannon second coding theorem [11], [12] which relates the information channel capacity to the maximum transmission data rate for reliable communication, known as operational channel capacity. The robust Shannon second coding theorem states that for a memoryless uncertain additive white Gaussian channel, a transmission data rate is achievable (e.g. there exits a sequence of $(2^{nR}, n)$ code with maximum probability of error $\lambda^{(n)} \rightarrow 0$, uniformly over all uncertain channel models) if and only if $R \leq C_{robust}$ [13], [14].
Next we proceed by defining the memoryless robust rate distortion. This is a measure of the minimum rate under which an end to end transmission with distortion up to distortion level D is possible for memoryless sources. This definition first appeared in [15].

Definition 2.5: (Robust Memoryless Rate Distortion) Let $\mathcal{M}_{DC} = \{Q_{\widetilde{X}|X}; \int_{\mathcal{X} \times \widetilde{\mathcal{X}}} \rho(x, \widetilde{x}) dQ_{\widetilde{X}|X}(\widetilde{x}) P_X(dx) \leq D\}$ be the set of distortion constraints, in which $D \geq 0$ is the distortion level and $\rho : \mathcal{X} \times \widetilde{\mathcal{X}} \rightarrow [0, \infty)$ is the distortion measure. When the true source probability measure P_X belongs to the uncertainty set $P_X \in$

$\mathcal{M}_{SU} \subset \mathcal{M}(\mathcal{X})$, called the source uncertainty set, the robust rate distortion is defined by

$$\begin{aligned} R_{robust}(D) &= \inf_{Q_{\widetilde{X}|X} \in \mathcal{M}_{DC}} \sup_{P_X \in \mathcal{M}_{SU}} I(X;\widetilde{X}) \\ &= \inf_{Q_{\widetilde{X}|X} \in \mathcal{M}_{DC}} \sup_{P_X \in \mathcal{M}_{SU}} H(P_X \otimes Q_{\widetilde{X}|X} | P_X \otimes P_{\widetilde{X}}). \end{aligned} \quad \text{(II.18)}$$

The importance of this definition for the rate distortion can be understood in terms of the so called Shannon third coding theorem [11], [12]. The robust rate distortion theorem [15] considers an uncertain memoryless source $(\mathcal{X}, \mathcal{F}(\mathcal{X}), P_X)$ with single-letter fidelity criterion (e.g. in the case of sequence with length m of source symbols, the distortion measure is $\rho(x^{m-1}, \widetilde{x}^{m-1}) \triangleq \rho_m(x^{m-1}, \widetilde{x}^{m-1}) = \frac{1}{m}\sum_{t=0}^{m-1} \rho_1(x_t, \widetilde{x}_t)$) such that $P_X \in \mathcal{M}_{SU} = \mathcal{M}_{SUR} \triangleq \{P_X; H(P_X|M_X) \leq E\}$, where $M_X \in \mathcal{M}(\mathcal{X})$ is the nominal source probability measure, and $E \geq 0$ controls the uncertainty set which is known parameter. Let $R_{robust}(D)$ denote the rate distortion function of $(\mathcal{X}, \mathcal{F}(\mathcal{X}), P_X)$ with distortion level up to distortion level D. Then we can find a D-admissible code of block length n for n sufficiently large, regardless of the true source probability measure if and only if $R_{robust}(D) < R$. That is, $R_{robust}(D)$ is the minimum achievable rate for distortion up to the distortion level D. In information theoretic literature, this minimum achievable rate is known as operational rate distortion [15].

2) Robust Information Theoretic Measures Subject to Relative Entropy Uncertainty: In this part, we present calculations of the robust measures of information, when the uncertainty is described by a relative entropy constraint. Moreover, a lower bound for the robust rate distortion is derived.

Theorem 2.6: (Robust Entropy)[10], [16] Let $G(R) = \sup_{p_X \in \mathcal{M}^d_{SUR}} H(p_X)$, where $\mathcal{M}^d_{SUR} = \{p_X, H(p_X|\mu_X) \leq R\}$, where μ_X is fixed.
1. For memoryless discrete source with M possible outcomes, the supremum is attained at

$$p_i^* = \frac{\mu_i^{\frac{s}{1+s}}}{\sum_{j=1}^{M} \mu_j^{\frac{s}{1+s}}}, \quad i = 1, 2, ..., M, \quad \text{(II.19)}$$

where $0 \leq R \leq H(\eta|\mu_X)$,and η is uniform distributed (e.g. $\eta_j = \frac{1}{M},\ j = 1, 2, ..., M$).
Moreover, the robust entropy is given by

$$G(R) = min_{s>0}[sR + (1+s) log \sum_{j=1}^{M} \mu_j^{\frac{s}{1+s}}]. \quad \text{(II.20)}$$

2. For memoryless continuous sources, the supremum is attained at

$$p_X^*(x) = \frac{\mu_X(x)^{\frac{s}{1+s}}}{\int \mu_X(x)^{\frac{s}{1+s}} dx}, \ s > 0. \quad \text{(II.21)}$$

and the robust entropy is given by

$$G(R) = min_{s>0}[sR + (1+s) log \int_{\mathcal{X}} \mu_X(x)^{\frac{s}{1+s}} dx] \quad \text{(II.22)}$$

and $s > 0$ is such that $H(p_X^*|\mu_X) = R$.

Theorem 2.7: (Robust Rate Distortion) [15] Suppose $e^{s\rho} \in L_1(\widetilde{\mathcal{X}}, \mathcal{F}(\widetilde{\mathcal{X}}), P_{\widetilde{X}}),\ \forall s \in \Re$.
Then the solution to the problem $(II.18)$,with relative entropy constraint (e.g. $\mathcal{M}_{SU} = \mathcal{M}_{SUR} \triangleq \{P_X; H(P_X|M_X) \leq E\}$), when M_X is fixed is given by

$$\begin{aligned} &R(D) = sD + \lambda R \\ &+\lambda log \int_{\mathcal{X}} (\int_{\widetilde{\mathcal{X}}} e^{s\rho(x,\widetilde{x})} dP_{\widetilde{X}}(\widetilde{x}))^{-\frac{1}{\lambda}} dP_X(x), \end{aligned} \quad \text{(II.23)}$$

where $s \leq 0$ and $\lambda > 0$ are Lagrange multipliers.
Moreover the infimum is attained at

$$dP_X^*(x) = \frac{e^{\frac{l(x)}{\lambda}} dM_X(x)}{\int_{\mathcal{X}} e^{\frac{l(x)}{\lambda}} dM_X(x)} \quad \text{(II.24)}$$

$$l(x) = \int_{\widetilde{\mathcal{X}}} log(e^{-s\rho(x,\widetilde{x})} \frac{dQ^*_{\widetilde{X}|X}(x,\widetilde{x})}{dP_{\widetilde{X}}(\widetilde{x})}) dQ^*_{\widetilde{X}|X}(x,\widetilde{x}) \quad \text{(II.25)}$$

and the supremum is attained at

$$dQ^*_{\widetilde{X}|X}(x,\widetilde{x}) = \frac{e^{s\rho(x,\widetilde{x})} dP_{\widetilde{X}}(\widetilde{x})}{\int_{\widetilde{X}} e^{s\rho(x,\widetilde{x})} dP_{\widetilde{X}}(\widetilde{x})} \quad \text{(II.26)}$$

Since, the exact expression of robust rate distortion is difficult to obtain, it is desirable to have a lower bound which is easily computed. This lower bound is given below.

Lemma 2.8: (Lower Bound for Robust Rate Distortion) Let the true source distribution belongs to the set $\mathcal{M}^d_{SU}$ (not necessary relative entropy uncertainty set), and $\rho(x, \widetilde{x}) = \rho(x - \widetilde{x})$.
Then a lower bound for $R_{robust}(D)$ is given by

$$R_{robust}(D) \geq \sup_{p_X \in \mathcal{M}^d_{SU}} H(p_X) - \max_{g \in G_D} H(g), \quad \text{(II.27)}$$

where $G_D = \{g \in \mathcal{M}(\mathcal{X}); \int \rho(x) g(x) dx \leq D\}$.
Proof [16].

3) Communication Systems with Memory: In the beginning of this subsection, we defined channel capacity and rate distortion for memoryless channels and sources. In this part, we extend these definitions to channel and source with memory.
For channels with memory, the channel input and

output measurable spaces correspond to the sequences

$$(\mathcal{Z},\mathcal{F}(\mathcal{Z}))=(\mathcal{Z}_{0,n-1},\mathcal{F}^{\mathcal{Z}}_{0,n-1})$$
$$\triangleq \times_{k=0}^{n-1}(\mathcal{Z}_k,\mathcal{F}(\mathcal{Z}_k))\subset \times_{k=0}^{\infty}(\mathcal{Z}_k,\mathcal{F}(\mathcal{Z}_k)),$$
$$(\widetilde{\mathcal{Z}},\mathcal{F}(\widetilde{\mathcal{Z}}))=(\widetilde{\mathcal{Z}}_{0,n-1},\mathcal{F}^{\widetilde{\mathcal{Z}}}_{0,n-1})$$
$$\triangleq \times_{k=0}^{n-1}(\widetilde{\mathcal{Z}}_k,\mathcal{F}(\widetilde{\mathcal{Z}}_k))\subset \times_{k=0}^{\infty}(\widetilde{\mathcal{Z}}_k,\mathcal{F}(\widetilde{\mathcal{Z}}_k)). \tag{II.28}$$

An element in $\mathcal{Z}_{0,n-1}$ is defined by $Z^{n-1}=(Z_0,Z_1,...,Z_{n-1})$, and similarly, for an element in $\widetilde{Z}^{n-1}$.
In $(II.28)$, $(\mathcal{Z}_k,\mathcal{F}(\mathcal{Z}_k))$ and $(\widetilde{\mathcal{Z}}_k,\mathcal{F}(\widetilde{\mathcal{Z}}_k))$ are exemplars of measurable space $(\mathcal{Z}_I,\mathcal{F}(\mathcal{Z}_I))$ and $(\widetilde{\mathcal{Z}}_O,\mathcal{F}(\widetilde{\mathcal{Z}}_O))$,which are the channel input and output alphabet measurable space sets.

Definition 2.9: (Channel Capacity with Memory) When the channel is unknown but belongs to the uncertainty set $P_{\widetilde{Z}^{n-1}|Z^{n-1}}\in\mathcal{M}_{CU}\subset\mathcal{M}(\widetilde{\mathcal{Z}}_{0,n-1})$, the robust channel capacity is defined by

$$C^{cap}_{robust}=\lim_{n\to\infty}\frac{1}{n}C_{n,robust} \tag{II.29}$$

$$\triangleq \lim_{n\to\infty}\frac{1}{n}\sup_{P_{Z^{n-1}}\in\mathcal{M}_{CI}}\ \inf_{P_{\widetilde{Z}^{n-1}|Z^{n-1}}\in\mathcal{M}_{CU}} I(Z^{n-1};\widetilde{Z}^{n-1})$$

By Shannon second theorem, under certain assumptions it is shown that C^{cap}_{robust} is equal to the operational capacity [13], [14].
For sources with memory, the source and reproduction measurable spaces correspond to the sequences

$$(\mathcal{X},\mathcal{F}(\mathcal{X}))=(\mathcal{X}_{0,m-1},\mathcal{F}^{\mathcal{X}}_{0,m-1})$$
$$\triangleq \times_{k=0}^{m-1}(\mathcal{X}_k,\mathcal{F}(\mathcal{X}_k))\subset \times_{k=0}^{\infty}(\mathcal{X}_k,\mathcal{F}(\mathcal{X}_k)),$$
$$(\widetilde{\mathcal{X}},\mathcal{F}(\widetilde{\mathcal{X}}))=(\widetilde{\mathcal{X}}_{0,m-1},\mathcal{F}^{\widetilde{\mathcal{X}}}_{0,m-1})$$
$$\triangleq \times_{k=0}^{m-1}(\widetilde{\mathcal{X}}_k,\mathcal{F}(\widetilde{\mathcal{X}}_k))\subset \times_{k=0}^{\infty}(\widetilde{\mathcal{X}}_k,\mathcal{F}(\widetilde{\mathcal{X}}_k)) \tag{II.30}$$

An element in $\mathcal{X}_{0,m-1}$ is denoted by $X^{m-1}=(X_0,X_1,...,X_{m-1})$, and similarly for an element in $\mathcal{X}_{0,m-1}$. In $(II.30)$, $(\mathcal{X}_k,\mathcal{F}(\mathcal{X}_k))$ and $(\widetilde{\mathcal{X}}_k,\mathcal{F}(\widetilde{\mathcal{X}}_k))$ are exemplars of measurable space $(\mathcal{X}_S,\mathcal{F}(\mathcal{X}_S))$ and $(\widetilde{\mathcal{X}}_R,\mathcal{F}(\widetilde{\mathcal{X}}_R))$ which are the source and reproduction alphabet set measurable spaces.

Definition 2.10: When the true probability of the source belongs to the uncertainty set $P_{X^{m-1}}\in\mathcal{M}_{SU}\subset\mathcal{M}(\mathcal{X}_{0,m-1})$, the robust rate distortion is defined by

$$R^{ra}_{robust}(D)=\lim_{m\to\infty}\frac{1}{m}R_{m,robust}(D)\triangleq \tag{II.31}$$

$$\lim_{m\to\infty}\frac{1}{m}\inf_{Q_{\widetilde{X}^{m-1}|X^{m-1}}\in\mathcal{M}_{DC}}\ \sup_{P_{X^{m-1}}\in\mathcal{M}_{SU}} I(X^{m-1};\widetilde{X}^{m-1})$$

4) Examples: In this part, we provide specific examples for robust entropy of the source, and channel capacity, for which these quantities can be computed explicitly.

Example 2.11: (Entropy of the Source)[16] Let the source distribution is unknown but belongs to the uncertainty set $p_X\in\mathcal{M}^d_{SUR}=\{p_X;H(p_X|\mu_X)\leq R\}$, where μ_X is d-dimensional Gaussian distribution, that is $\mu_X\sim G(m,\Gamma_X)$. Then, $\forall R\geq 0$

$$\sup_{p_X\in\mathcal{M}^d_{SU}} H(p_X)=\frac{1}{2}ln(2\pi e)^d|det\frac{1+s}{s}\Gamma_X|,\ nats, \tag{II.32}$$

where $s\geq 0$ is the solution of the following equation $e^{-2R}=(\frac{s+1}{s})e^{-\frac{d}{s}}$.

Example 2.12: (Robust Memoryless Channel Capacity) [17]. Consider a Binary symmetric channel (BSC) with crossover probability p given in Figure $IV.2$. When the crossover probability p is partially known and belongs to the uncertainty set Θ, the robust channel capacity is given by

$$C_{robust}=\inf_{p\in\Theta}(1-H_b(p)). \tag{II.33}$$

where

$$H_b(x)=-x\log_b x-(1-x)\log_b(1-x), x\in[0,1]. \tag{II.34}$$

Example 2.13: (Robust Channel Capacity with Memory)[14].Consider additive white Gaussian channel given in Figure $IV.3$. Here we assume that the power spectral density of the noise is unknown but it belongs to the uncertainty set $\{S_V(f);\ \int_{-\infty}^{+\infty}S_V(f)df\leq P_n\}$. Assume $\frac{S_X(f)|H(f)|^2}{S_V(f)},\forall f\in(-\infty,+\infty)$ is bounded and integrable.
Then

$$C^{cap}_{robust}=\frac{1}{2}\int_{-\infty}^{+\infty}log(1+\frac{\lambda_1^*}{\lambda_2^*}|H(f)|^2)df, \tag{II.35}$$

where the Lagrange multiplier λ_1^* and λ_2^* can be found from the following equation

$$\int_{-\infty}^{+\infty}S^*_Z(f)df=P,\qquad \int_{-\infty}^{+\infty}S^*_V(f)df=P_n, \tag{II.36}$$

where

$$S^*_Z(f)=\frac{\lambda_1^*}{\lambda_2^*}S^*_V(f),\ S^*_V(f)=\frac{|H(f)|^2}{2(\lambda_1^*|H(f)|^2+\lambda_2^*)} \tag{II.37}$$

III. Robust Information Transmission Theorem

In this section, we invoke the data processing inequality to derive a robust version of the information transmission theorem. This theorem provides a necessary condition for end to end transmission up to distortion level D, (e.g. $E\rho(X,\widetilde{X})\leq D$), when there is uncertainty on the source as well as communication channel.

Theorem 3.1: (Robust Information Transmission Theorem) A necessary condition for reproducing the source output X up to distortion level D by $\widetilde{X}$ at the output of the decoder, when there is uncertainty on the source and the communication channel is

$$R_{robust}(D) \leq C_{robust}. \tag{III.38}$$

Proof [16].
Theorem 3.1 is given without any restriction on source and channel symbols, consequently this theorem must be applicable when we use a sequences of source and channel symbols with length m and n $(m \leq n)$, respectively. The next corollary shows that this theorem is applicable when we use sequences of source and channel symbols with length m and n and when there are feedbacks between input and output of encoder, channel and decoder, namely when we deal with communication block diagram given in Figure $IV.4$.

Corollary 3.2: (Robust Information Transmission Theorem in the Presence of Feedback) A necessary condition for reproducing the source output $X^m = (X_0, X_1, X_2, ..., X_{m-1})$ up to distortion level D by $\widetilde{X}^m = (\widetilde{X}_0, \widetilde{X}_1, \widetilde{X}_2, ..., \widetilde{X}_{m-1})$ at the output of the decoder for n-times channel use $(m \leq n)$, when there is uncertainty on the source and communication channel, and when there are feedbacks between input and output of encoder, channel and decoder is

$$R_{m,robust}(D) \leq C_{n,robust}, \tag{III.39}$$

where $C_{n,robust}$ denotes the robust information channel capacity for n-times channel use and $R_{m,robust}(D)$ denotes the robust rate distortion for sequence of length m of source symbols.
Proof. [16].

Remark 3.3: Corollary 3.2 extends the result given in [9] to the case when source and communication link are subject to uncertainty. Also this corollary covers the result given in [9] for T-times channel use (e.g. $n = T$), and sequences with length 1 (e.g. $m = 1$) of source symbols, when there is no uncertainty in source and communication link.

IV. Necessary Condition for Observability and Stabilizability

In this section a necessary condition for observability and stabilizability of fully observed, finite dimensional, noiseless, uncertain linear systems over uncertain channels is introduced.
Suppose for each $\Delta A \in S \subset \Re^{n\times n}$ the linear time-invariant system defined on a filter probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \{P_X^{\Delta A}, \Delta A \in S\})$, is given by

$$\begin{aligned}&(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \{P_X^{\Delta A}, \Delta A \in S\}) :\\ &X_{t+1} = (A + \Delta A)X_t + BU_t, \quad t \geq 0,\ X_0 \in \Re^n.\end{aligned} \tag{IV.40}$$

where $\{X_t\}$ is $\Re^n$ value state process, and $\{U_t\}$ is a $\Re^m$-valued control process. We have $A \in \Re^{n\times n}$, and $B \in \Re^{n\times m}$. The initial position, X_0, is distributed according to the probability density p_{X_0} with finite differential entropy $H(p_{X_0})$. Moreover system $(IV.40)$ is supposed to be detectable and stabilizable for each $\Delta A \in S$.
In $(IV.40)$, ΔA is an unknown matrix, which belongs to the uncertainty set $\Delta A \in S$. Let $P_X^{\Delta A}$ be the probability measure induced by the dynamic $(IV.40)$, when $\Delta A \in S$. Then $P_X^{\Delta A}$ is a function of ΔA, and it is induced by $\{X_t\}_{t\geq 0}$. Since ΔA belongs to the uncertainty set $\Delta A \in S$, then $P_X^{\Delta A}$ belongs to the set of uncertainty $P_X^{\Delta A} \in \mathcal{M}_{SU} = \{P_X^{\Delta A}; \Delta A \in S, p_X^{\Delta A} = \frac{dP_X^{\Delta A}}{dx}\}$.
The nominal model corresponding to the true model given in $(IV.40)$ is

$$\begin{aligned}&(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P_X) :\\ &X_{t+1} = AX_t + BU_t,\ t \geq 0,\ X_0 \sim p(X_0), X_0 \in \Re^n.\end{aligned} \tag{IV.41}$$

The dynamic $(IV.40)$ is cascaded with an uncertain communication channel (See Figure $IV.5$) and the objective is to find necessary condition for almost surely uniform asymptotically observability and stabilizability which are defined as follows.

Definition 4.1: Let the error be $E_t = X_t - \hat{X}_t$, where $\hat{X}_t$ is the state estimate. System $(IV.40)$ is almost surely uniform asymptotically observable iff the exists an encoder and decoder such that

$$\sup_{P_X^{\Delta A} \in \mathcal{M}_{SU}} \Pr(\lim_{t\to\infty} ||E_t||_2 \neq 0) = 0. \tag{IV.42}$$

System $(IV.40)$ is almost surely uniform asymptotically stabilizable iff there exist an encoder, decoder and controller such that

$$\sup_{P_X^{\Delta A} \in \mathcal{M}_{SU}} \Pr(\lim_{t\to\infty} ||X_t||_2 \neq 0) = 0. \tag{IV.43}$$

Remark 4.2: The above definition for observability is a natural generalization of classical definition of asymptotic observability to the case when there is uncertainty in the information source.

Proposition 4.3: Given $(IV.40)$, a necessary condition on the operational robust channel capacity ($C_{robust}^{cap} = \lim_{n\to\infty} \frac{1}{n} C_{n,robust}$) for almost surely uniform asymptotically observability is

$$C_{robust}^{cap} \geq \sum_{\lambda(A+\Delta A_{max})} max\{0, log|\lambda(A + \Delta A_{max})|\}. \tag{IV.44}$$

where $\Delta A_{max} \in S$ is chosen such that

$$|det(A + \Delta A)| = \prod_{i=1}^{d} |\lambda_i(A + \Delta A)| \tag{IV.45}$$

is maximum.
Proof [16].

Proposition 4.4: Given $(IV.40)$, a necessary condition on the operational robust channel capacity ($C^{cap}_{robust} = \lim_{n\to\infty} \frac{1}{n} C_{n,robust}$)for almost surely uniform asymptotically stabilizability is

$$C^{cap}_{robust} \geq \sum_{\lambda(A+\Delta A_{max})} max\{0, log|\lambda(A+\Delta A_{max})|\}. \tag{IV.46}$$

Proof. [16].

Remark 4.5: In proving the necessary conditions shown previously, we did not need to explicitly describe the encoder, decoder and channel, and we did not use an assumption of separation between the observer and controller. Hence, the conditions hold independently of the choice of these components. Moreover, Propositions 4.3 and 4.4 extend the result derived in [9] under Propositions 3.2 and 3.3.

References

[1] D. F. Delchamps. Stabilizing a linear system with quantized state feedback. IEEE Transaction on Automatic Control,Vol. 35, pp. 916-924, 1990.

[2] S. Tatikonda and S. Mitter. Control of LQG systems under communication constraints. In Proc. 37th IEEE Conf. Dec. Contr. , pp. 1165-1170, 1998.

[3] W. S. Wong and R. W. Brockett. Systems with finite communication bandwidth constraints II: stabilization with limited information feedback. IEEE Transaction on Automatic Control,Vol. 44, pp. 1049-1053, 1999.

[4] G. N. Nair and R. J. Evans. Stabilization with data-rate-limited feedback: tightest attainable bounds. Sys. Contr. Lett., 41(1), pp. 49-56, Sep 2000.

[5] N. Elia and S. K. Mitter. Stabilization of linear systems with limited information. IEEE Transaction on Automatic Control, Vol. 46, pp. 1384-1400, 2001.

[6] G. N. Nair and R. J. Evans. Mean square stabilisability of stochastic linear systems with data rate constraints. In Proc. 41st IEEE Conf. Dec. Contr., pp. 1632-1637, Las Vegas, USA, Dec. 2002.

[7] G. N. Nair and R. J. Evans. Exponential stabilisability of finite-dimensional linear systems with limited data rates. In Automatica,Vol. 39 pp. 585-593, 2003.

[8] K. Tsumura and J. Maciejowski. Stabilizability of SISO control systems under constraints of channel capacities. In Proc. 42nd IEEE Conf. Dec. Contr., pp. 193-198, 2003.

[9] S. Tatikonda, and S. Mitter. Control over noisy channels. IEEE Transaction on Automatic Control, vol. 49, No. 7,pp. 1196-1201, July 2004.

[10] C. D. Charalambous and F. Rezaei. Robust Coding for Uncertain Sources. Preprint, 2004.

[11] T. M Cover and Joy A. Thomas. Elements of Information Theory John Wiley and Sons, 1991.

[12] R. E. Blahut.Principles and Practice of Information Theory. Addison- Wesely Publishing Company, 1987.

[13] W. L. Root, and P. P. Varaya. Capacity of classes of Gaussian channels, SIAM J. Appl. Math, vol. 16, No. 6, November 1968.

[14] C. D. Charalambous, S. Z. Denic and S. M. Djouadi. Robust capacity for uncertain additive Gaussian channels. Preprint, 2004.

[15] C. D. Charalambous and F. Rezaei. A general formulation for rate distortion. Preprint, 2004.

[16] C. D. Charalambous, Alireza Farhadi, Stojan Denic, and Farzad Rezaei. Mathematical theory for robust control and communication subject to uncertainty. Preprint, 2004.

[17] A. Lapidoth and P. Narayan. Reliable communication under channel uncertainty. IEEE Transaction on Information Theory, vol. 44, No. 6,pp. 2148-2177, October 1998.

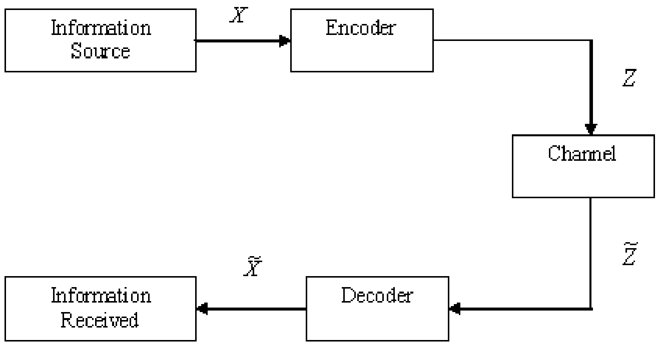

Fig. IV.1. Block diagram of communication system

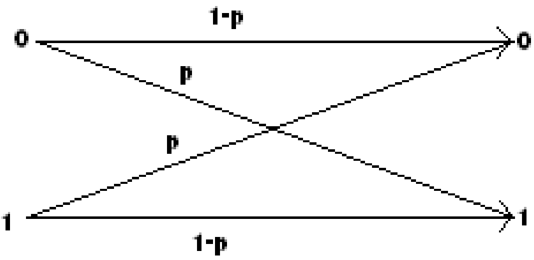

Fig. IV.2. BSC channel(memoryless channel)

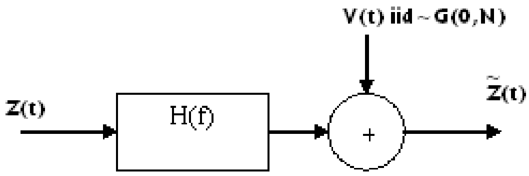

Fig. IV.3. Additive white Gaussian channel (channel with memory)

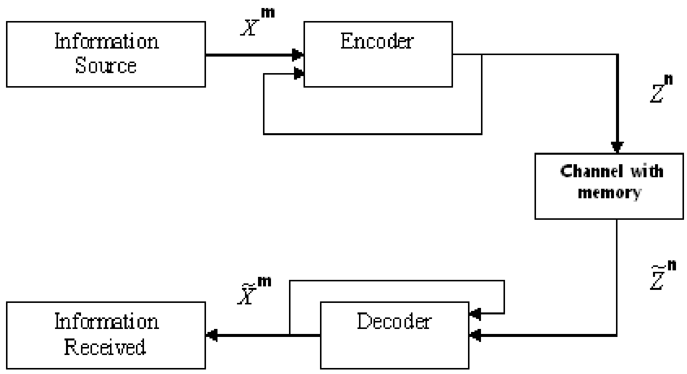

Fig. IV.4. Block diagram of communication system with feedback

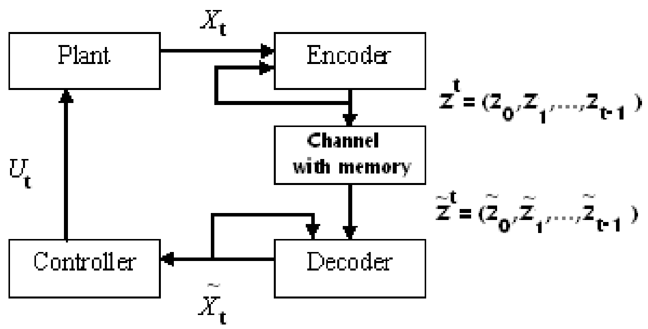

Fig. IV.5. Control system with communication channel