Control of Feedback Systems Subject to the Finite Rate Constraints via the Shannon Lower Bound

C. D. Charalambous
Department of Electrical and Computer Engineering
University of Cyprus
Nicosia, CYPRUS
E-mail: chadcha@ucy.ac.cy.

Alireza Farhadi
School of Information Technology and Engineering
University of Ottawa
Ottawa, Ontario, CANADA
Email: afarhadi@site.uottawa.ca.

***Abstract*— This paper is concerned with control of stochastic systems subject to limited feedback channel capacity. Specifically, the design of an encoder, decoder, and controller subject to the mean square observability and stabilizability is considered. It is shown that by transmitting information which is equal to the Shannon lower bound, mean square observability and stabilizability over Additive White Gaussian Noise (AWGN) channel is achieved. Furthermore, a modified definition for channel capacity and rate distortion is presented which is suitable for control applications and real time communication.**

I. Introduction

Recent advances in technology have created an increasingly demand on networks. Extensive research activity has been devoted to the question of how much bit rate must be allocated to each components of a network. This line of research is motivated by applications in which the communication data rates from the channel input to the controller input are limited and feedback is available from the output of the channel to the input of the channel. In such applications, due to limited capacity (Shannon capacity per source message) constraint, the main assumption is that the source outputs can not be represented with high precision at the end of communication system; and only a distorted version of source output is available. Therefore, the fundamental question in limited capacity applications is to find an encoding and/or stabilizibility schemes for reliable data reconstruction and/or stabilizability by transmitting information from source as minimum as possible.

The present paper tries to address this question. This paper is concerned with the control/communication system of Fig. 1. This block can be viewed as a basic model for the Networked Control Systems (NCS's) [1] in which the dynamical system (source) and the corresponding controller are connected through a shared communication media; while, there is unshared or high capacity communication link from controller to source. Thus, the connection from sensors to controller is subject to limited capacity constraint; while the connection from controller to the source is unconstraint (direct). The Control/communication system of Fig. 1 may correspond to tele-operation systems used for space exploration devices or wireless microsurgery. In such systems, the communication from dynamical system to remote controller is subject to limited capacity constraint due to limited power supply of such devices; while, the connection from controller to dynamical system is direct due to unlimited communication resource at the base station where the controller is located.

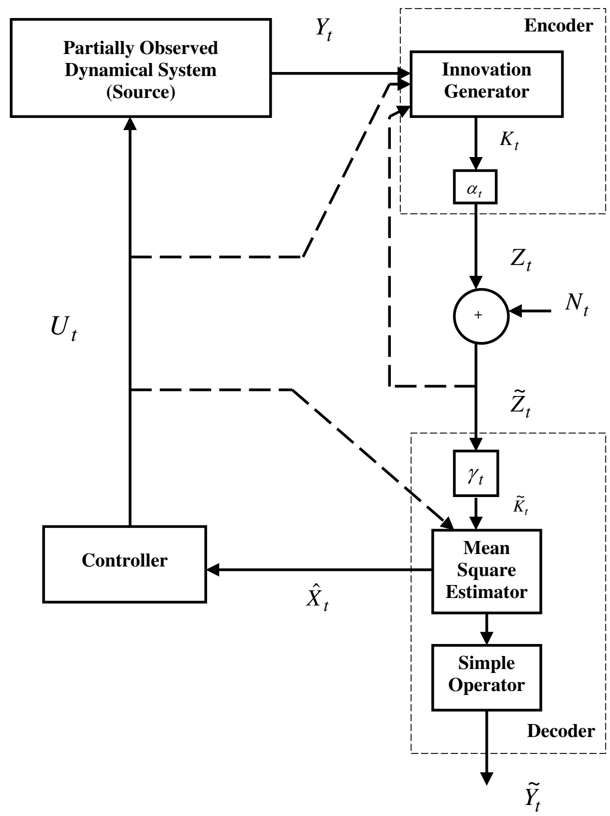

Fig. 1. A control/communication system

The control/communication system of Fig. 1 is defined on a complete probability space $(\Omega, \mathcal{F}(\Omega), P)$ with filtration $\{\mathcal{F}_t\}_{t\geq 0}$; $t \in \mathbf{N}_+ \triangleq \{0,1,2,...\}$, where , Y_t, K_t, Z_t $\tilde{Z}_t$, $\tilde{K}_t$, $\tilde{Y}_t$, $\hat{X}_t$ and U_t, $t \in \mathbf{N}_+$, are Random Variables (R.V.'s) denoting the source message, an innovation process associated with the source message (in the presence or absence of feedback), channel input codeword, channel output codeword, the reproduction of the innovation process, the reproduction of the source message, the estimated state variables in the mean square sense, and the control input to the source, respectively. Further, suppose the channel noise sequence $\{N_t; t \in \mathbf{N}_+\}$ is an orthogonal Gaussian process and $\{\alpha_t\}_{t\in\mathbf{N}_+}$ and $\{\gamma_t\}_{t\in\mathbf{N}_+}$ are deterministic real valued (to be defined when designing the encoding scheme).

The objective is to achieve reliable data reconstruction (known also as observability) in the sense that $\tilde{K}_t$ (resp. $\tilde{Y}_t$) follows K_t (resp. Y_t) measured by the mean square error, as well as, to achieve mean square stabilizability in the sense that the mean square value of the state variables is bounded for all times, when there is a limited data rate constraint.

For most part, research on reliable data reconstruction and stabilizability subject to limited capacity constraint has focused on the basic problem of Fig. 1, beginning with [2] and [3] and continuing with [4]-[15]. Various publications have introduced necessary and sufficient conditions for observability and stabilizability of Fig. 1 in various senses [2], [4]-[10], [12]-[15]. In most part, these conditions are given in the form of a lower bound on the capacity in terms of rate of change of the dynamical system. In particular, it is already known that the eigenvalues rate condition (i.e., the summation of the logarithms of the magnitude of the unstable eigenvalues of the open loop discrete time-invariant systems) is the minimum achievable capacity for observability and stabilizability of linear time-invariant dynamical systems. In most part, after finding a necessary condition in terms of a lower bound on the capacity, an encoder, decoder and controller are proposed which can achieve the lower bound found as a necessary condition. Subsequently in these publications, the minimum achievable capacity for observability and stabilizability and the encoder, decoder and controller which can achieve the minimum capacity have been proposed.

In this paper, we employ the Shannon lower bound and the rate distortion theory to obtain tight conditions on the channel capacity for mean square observability of partially observed dynamical systems subject to Gaussian measurement and process noises. This is obtained by relating the Shannon lower bound to the definitions of observability and stabilizability. The advantage of using Shannon lower bound to present conditions for observability and stabilizability is its relation to observability and stabilizability definition employed. Subsequently, it overcomes the drawbacks of using Shannon capacity in moment observability and stabilizability. Other advantage is finding a condition in terms of Shannon entropy rate which can be easily computed and can imply the eigenvalues rate condition.

Similar results for controlling fully observed linear dynamical systems subject to linear quadratic pay off have been reported in [9] in which the idea of source-channel matching is employed to obtain the results. Moreover, similar results to [9], previously have been reported in [16] where separation between communication and control system design is established under optimal transmission (encoder/decoder design) and Linear Quadratic Gaussian (LQG) pay-off (control design). Nevertheless, in the present paper using the idea of source-channel matching, the results of [9] are extended to the case of partially observed linear dynamical systems subject to measurement noise. Furthermore, since the separation principle between the design of communication and stabilizing controller holds, the communication and stabilizing controller is designed separately; while the whole system is optimal (with respect to linear quadratic pay off).

In control applications and real time communication, the causality of communication channel in the sense that the channel output does not anticipate the channel input as well as the causality of the rate distortion between message and the corresponding reproduction in the sense that the reproduction does not anticipate the message, are often desirable or even essential. Therefore, it is important to modify the classical definition of capacity and rate distortion by considering the causality constraint. In this paper a modified definition of channel capacity and rate distortion by considering causality constraint is also given [17].

This paper is organized as follows. In Section II, the problem formulation is given. In Section III, a modified definition for channel capacity and rate distortion is presented which is often desirable especially in control applications and real time communication. Section IV is concerned with mean square observability of the control/communication system of Fig. 1 and Section V deals with the mean square stabilizability of the control/communication system of Fig. 1.

II. Problem Formulation

Consider the control/communication system of Fig. 1, where $Y_t \in \mathcal{Y}_t$, $K_t \in \mathcal{K}_t$ $Z_t \in \mathcal{Z}_t$, $\tilde{Z}_t \in \tilde{\mathcal{Z}}_t$, $\tilde{K}_t \in \tilde{\mathcal{K}}_t$, $\tilde{Y}_t \in \tilde{\mathcal{Y}}_t$, $U_t \in \mathcal{U}_t$ are Random Variables (R.V.'s) denoting the source message, innovation, channel input, channel output, reproduced innovation, reproduced source message, and the control input to the source, respectively, at time $t \in \mathbf{N}_+$. It is assumed that $\mathcal{Y}_t$, $\mathcal{K}_t$, $\mathcal{Z}_t$, $\tilde{\mathcal{Z}}_t$, $\tilde{\mathcal{Z}}_t$, $\tilde{\mathcal{Y}}_t$, $\tilde{\mathcal{K}}_t$, and $\tilde{\mathcal{U}}_t$ are complete separable metric spaces and $(\mathcal{Y}_t, \mathcal{F}(\mathcal{Y}_t))$, $(\mathcal{K}_t, \mathcal{F}(\mathcal{K}_t))$, $(\mathcal{Z}_t, \mathcal{F}(\mathcal{Z}_t))$, $(\tilde{\mathcal{Z}}_t, \mathcal{F}(\tilde{\mathcal{Z}}_t))$, $(\tilde{\mathcal{Y}}_t, \mathcal{F}(\tilde{\mathcal{Y}}_t))$, $(\tilde{\mathcal{K}}_t, \mathcal{F}(\tilde{\mathcal{K}}_t))$, and $(\mathcal{U}_t, \mathcal{F}(\mathcal{U}_t))$, $\tilde{\mathcal{Z}}_t \subseteq \mathcal{Z}_t$ are measurable spaces (e.g., $\mathcal{F}(\mathcal{Y}_t)$ is an σ-algebra of subsets of the set $\mathcal{Y}_t$ generated by closed set). For $T, n \in \mathbf{N}_+$, sequences of the R.V.'s with length T and n of the source, innovation and channel, are identified with the product measurable spaces, $(\mathcal{Y}_{0,T-1}, \mathcal{F}(\mathcal{Y}_{0,T-1})) \triangleq \times_{k=0}^{T-1}(\mathcal{Y}_k, \mathcal{F}(\mathcal{Y}_k))$, and similarly for the innovation, channel input, channel output, reproductions and control input to the source, respectively.

Throughout, sequences of R.V.'s are denoted by $Y^T \triangleq (Y_0, Y_1, ..., Y_T)$ for $T \in \mathbf{N}_+$. $\log(.)$ denotes logarithm of base 2. A stochastic kernel, $P(dF; x)$, is a mapping $P : \hat{\mathcal{A}} \times A \rightarrow [0,1]$ which satisfies i) For every $x \in A$, the set function $P(:; x)$ is a probability measure on $\hat{\mathcal{A}}$, and ii) For every $F \in \hat{\mathcal{A}}$, the function $P(dF; .)$ is $\mathcal{A}$-measurable ($(A, \mathcal{A})$, $(\hat{A}, \hat{\mathcal{A}})$ are measurable spaces).

The different blocks of Fig. 1 are described below.

Information Source: The information source is defined by the marginal probability distribution $P(dY^T) = f_{Y^T} dY^T$, where

Y^T is the source output produced by the following linear discrete time stochastic partially observed control system.

$$\begin{cases} X_{t+1} = AX_t + NU_t + BW_t, \quad X_0 = X, \\ Y_t = CX_t + DG_t, \end{cases} \tag{1}$$

where $t \in \mathbf{N}_+$, $X_t \in \Re^q$ is the unobserved (state) process, $Y_t \in \Re^d$ is the observed process, $U_t \in \Re^o$ is the control signal, $W_t \in \Re^m$, $G_t \in \Re^l$ in which $\{W_t; t \in \mathbf{N}_+\}$ is Independent Identically Distributed (i.i.d.)$\sim N(0, I_{m\times m})$ and $\{G_t; t \in \mathbf{N}_+\}$ is i.i.d. $\sim N(0, I_{l\times l})$. Moreover, $X_0 \sim N(\bar{x}_0, \bar{V}_0)$ and $\{W_t, G_t, X_0; t \in \mathbf{N}_+\}$ are mutually independent.

Communication Channel: The communication channel at time $t \in \mathbf{N}_+$ is described by $\tilde{Z}_t = Z_t + N_t$ ($E[Z_t^{'} Z_t] \leq P_t$), where $Z_t, \tilde{Z}_t \in \Re^d$ are the channel input and output at time t and the stochastic process $\{N_t \in \Re^d; t \in \mathbf{N}_+\}$ is an orthogonal zero mean Gaussian process independent of Z_t.

Encoder: We define and discuss the following types of encoders.

Class A) The encoder at any time $t \in \mathbf{N}_+$ is modeled by a stochastic kernel $P(dZ_t; y^t, u^{t-1})$.

Class B) The encoder at any time $t \in \mathbf{N}_+$ is modeled by a stochastic kernel $P(dZ_t; y^t, u^{t-1}, \tilde{z}^{t-1})$.

Class C) The encoder at any time $t \in \mathbf{N}_+$ is modeled by a stochastic kernel $P(dZ_t; y^t, \tilde{z}^{t-1})$.

Decoder: We define and discuss the following types of decoders.

Class A) The decoder at any time $t \in \mathbf{N}_+$ is modeled by a stochastic kernel $P(d\tilde{K}_t; \tilde{z}^t)$ (resp. $P(d\tilde{Y}_t; \tilde{z}^t)$).

Class B) The decoder at any time $t \in \mathbf{N}_+$ is modeled by a stochastic kernel $P(d\tilde{K}_t; \tilde{z}^t, u^{t-1})$ (resp. $P(d\tilde{Y}_t; \tilde{z}^t, u^{t-1})$).

Controller: The control law at any time $t \in \mathbf{N}_+$ is modeled by a stochastic kernel $P(dU_t; \tilde{z}^t)$.

The objective of this paper is the mean square observability and stabilizability of the control/communication system of Fig. 1, when there is a limited amount of information constraint for describing source outputs, as defined as follows.

Definition 2.1: (Observability in the Mean Square Sense). Consider the block diagram of Fig. 1. The source is called observable in the mean square sense if there exists a finite $D_v \geq 0$, a control sequence, an encoder, and a decoder such that $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} E||K_t - \tilde{K}_t||^2 \leq D_v$ (resp. $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} E||Y_t - \tilde{Y}_t||^2 \leq D_v$), where $||.||$ is the Euclidian norm.

Definition 2.2: (Stabilizability in the Mean Square Sense). Consider the block diagram of Fig. 1. The source is called stabilizable in the mean square sense if there exists a finite $D_v^c \geq 0$, a controller, an encoder, and a decoder such that $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} E||X_t||^2_{C^{'}C} \leq D_v^c$.

Throughout, when we are concerned with observability, it is assumed that (A, C) is detectable and $(A, (BB^{'})^{\frac{1}{2}})$ is stabilizable; while, when we are also concerned with stabilizability, it is also assumed that $((C^{'}C)^{\frac{1}{2}}, A)$ is detectable and (A, N) is stabilizable [18].

III. Mathematical Preliminaries

In this section we give a modified definition for channel capacity and rate distortion which is often desirable in control applications and real time communication. We also recall the classical Shannon lower bound which is used in subsequent sections to describe the conditions for mean square observability and stabilizability.

For general feedback channels causality is often desirable especially in control or real time communication in which the channel output does not anticipate the channel input. The causality between stochastic kernel connecting Y^T to $\tilde{Y}^T$ is defined as follow.

Definition 3.1: (Causality)[17],[19] Given two sequences Y^T and $\tilde{Y}^T$, we shall say that the stochastic kernel connecting Y^T to $\tilde{Y}^T$ is causal if and only if $\{P(d\tilde{Y}_t; y^n, \tilde{y}^{t-1}) = P(d\tilde{Y}_t; y^t, \tilde{y}^{t-1})\}_{t=0}^T$ $\forall n > t$. The causality of the stochastic kernel connecting Y^T to $\tilde{Y}^T$ is denoted by $Y^T \xrightarrow{C} \tilde{Y}^T$.

Since the causality of the communication channel is often an essential future of communication channels used in control or real time communication, we have to modify the definition of the capacity by considering the causality constraint. This involves the restricted mutual information which is defined as follows.

Definition 3.2: (Restricted Mutual Information)[17],[19] Consider the sequences of R.V.'s Y^{T-1} and $\tilde{Y}^{T-1}$ in which the stochastic kernel connecting Y^{T-1} to $\tilde{Y}^{T-1}$ is causal. Then, the mutual information from Y^{T-1} to $\tilde{Y}^{T-1}$ is defined by.

$$\begin{aligned} & I(Y^{T-1}; \tilde{Y}^{T-1})\Big|_R \\ & \triangleq E_{P(d\tilde{Y}^{T-1}, dY^{T-1})} \log \frac{P(d\tilde{Y}^{T-1}; y^{T-1})}{P(d\tilde{Y}^{T-1})}\Big|_R \\ & = \sum_{t=0}^{T-1} E_{P(d\tilde{Y}^t, dY^t)} \log \frac{P(d\tilde{Y}_t; y^t, \tilde{y}^{t-1})}{P(d\tilde{Y}_t; \tilde{y}^{t-1})} \\ & \triangleq \sum_{t=0}^{T-1} I(Y^t; \tilde{Y}_t|\tilde{Y}^{t-1}). \end{aligned} \tag{2}$$

Now, we are ready to define the information capacity for causal channels.

Definition 3.3: (Information Capacity for Causal Channels)[17],[19]. Consider a causal communication channel (with memory and feedback) $Z^{n-1} \xrightarrow{C} \tilde{Z}^{n-1}$ connecting input sequence Z^{n-1} to output sequence $\tilde{Z}^{n-1}$, and constraint set $\mathcal{M}_{CI}$ on the possible joint distribution functions, $P(dZ^{n-1})$. The causal information capacity for the time horizon n is defined as follow.

$$\mathcal{C}_{n|R} \triangleq \sup_{P(dZ^{n-1})\in\mathcal{M}_{CI}} I(Z^{n-1}; \tilde{Z}^{n-1})\Big|_R. \tag{3}$$

Subsequently, the capacity is defined by

$$\mathcal{C}_{|R} = \lim_{n\to\infty} \frac{1}{n}\mathcal{C}_{n|R} \tag{4}$$

provided the limit exists.

Please note that since Discrete Memoryless Channels (DMC's) as well as AWGN channel are causal, for such channels (4) is

equal to the classical Shannon capacity, $\mathcal{C}$, defined in [20].
Causality of the rate distortion between the source message and the corresponding reproduced message is also desirable in control applications and real time communication. Subsequently, we have the following modified definition for information rate distortion.

Definition 3.4: (Information Rate Distortion for Causal Systems)[17],[19] Let Y^{T-1} and $\tilde{Y}^{T-1}$ denote sequences of length T of the source output and the reproduction of the source output, respectively. Let also

$$\mathcal{M}_{DC}=\{P(d\tilde{Y}^{T-1};y^{T-1});\sum_{t=0}^{T-1}E\rho_t(Y^t,\tilde{Y}^t)\leq D_v\} \quad (5)$$

denotes the distortion constraint where $\{\rho_t(.,.);t\in\mathbf{N}_+\}$ is a sequence of distortion measures (which are continuous in the second argument and non-negative functions) and $D_v\geq 0$ is the distortion level.
Then, causal rate distortion function is defined by

$$\begin{aligned}&R_T(D_v)|_R\\&=\inf_{P(d\tilde{Y}^{T-1};y^{T-1})\in\mathcal{M}_{DC}} I(Y^{T-1};\tilde{Y}^{T-1})\Big|_R.\end{aligned} \quad (6)$$

Subsequently, the rate distortion is defined by

$$R(D_v)|_R=\lim_{T\to\infty}\frac{1}{T}R_T(D_v)|_R. \quad (7)$$

Moreover, the non-anticipative stochastic kernel which achieves the infimum of the rate distortion function is given by

$$\begin{aligned}P^*(d\tilde{Y}^{T-1};y^{T-1}) &= \prod_{t=0}^{T-1}P^*(d\tilde{Y}_t;y^t,\tilde{y}^{t-1}),\\ P^*(d\tilde{Y}_t;y^t,\tilde{y}^{t-1}) &= \frac{e^{s\rho_t(y^t,\tilde{y}^t)}P^*(d\tilde{Y}_t;\tilde{y}^{t-1})}{\int_{\tilde{\mathcal{Y}}_t}e^{s\rho_t(y^t,\tilde{y}^t)}P^*(d\tilde{Y}_t;\tilde{y}^{t-1})}\end{aligned} \quad (8)$$

where $s\leq 0$ is the solution of $s=\frac{d}{dD_v}R_T(D_v)|_R$.
From (8) follows that the causal rate distortion is the sequential rate distortion introduced in [22].
Please note that for sources producing independent asymptotic stationary sequences, the information rate distortion functions for causal systems (i.e., $R_T(D_v)|_R$ and $R(D_v)|_R$) are equal to the classical Shannon rate distortion functions (i.e., $R_T(D_v)$ and $R(D_v)$ defined in [20] and [21]).
Next, we give data processing inequalities for causal systems, using the restricted mutual information. The failure of symmetry associated with the restricted mutual information affects the derivation of data processing inequalities. Specifically, the derivation of classical data processing inequalities [20] is based on the symmetry of mutual information. Here, we need the Markovian property, which is equivalent to conditional independence assumption. We shall say $Y^T\to Z^n\to\tilde{Z}^n$ forms a Markov chain if $P(d\tilde{Z}_t;z^t,\tilde{z}^{t-1},y^T)=$ $P(d\tilde{Z}_t;z^t,\tilde{z}^{t-1})$ $(t\in\{0,1,...,n\})$ P-a.s.

Definition 3.5: (Data Processing Inequalities For Causal Systems)[17],[19] Let $Y^T\to Z^n\to\tilde{Z}^n\to\tilde{Y}^T$ forms a Markov chain, and $Z^n\xrightarrow{C}\tilde{Z}^n$, $Y^T\xrightarrow{C}\tilde{Y}^T$.
Then

$$I(Z^n;\tilde{Z}^n)\Big|_R\geq I(Y^T;\tilde{Z}^n)\geq I(Y^T;\tilde{Y}^T)\Big|_R, \quad (9)$$

where $I(.;.)$ is the classical mutual information [20].
Next, we recall the classical Shannon lower bound which is a tight approximation of the classical rate distortion function; and subsequently, a tight lower bound for causal rate distortion when the classical rate distortion and causal rate distortion are the same.

Lemma 3.6: (Shannon Lower Bound). Let Y^{T-1}, $Y_t\in\Re^d$, $0\leq t\leq T-1$ be a sequence with length T produced by the source $P(dY^{T-1})=f_{Y^{T-1}}dY^{T-1}$. Consider the following distortion constraint $\mathcal{M}_{DC}=\{P(d\tilde{Y}^{T-1};y^{T-1});\frac{1}{T}\sum_{t=0}^{T-1}\rho(y_t,\widetilde{y}_t)\leq D_v\}$, where $\rho(y_t,\widetilde{y}_t)=\rho(y_t-\widetilde{y}_t):\Re^d\to[0,\infty)$ is continuous.
Then
i) A lower bound for $\frac{1}{T}R_T(D_v)$ is given by

$$\frac{1}{T}R_T(D_v)\geq\frac{1}{T}H_S(Y^{T-1})-\max_{h\in G_D}H_S(h), \quad (10)$$

where $H_S(.)$ is the Shannon differential entropy [20] and G_D is defined by $G_D\triangleq\{h:\Re^d\to[0,\infty);\int_{\Re^d}h(\xi)d\xi=1,\ \int_{\Re^d}\rho(\xi)h(\xi)d\xi\leq D_v,\quad \xi\in\Re^d\}$. Moreover, when $\int_{\Re^d}e^{s\rho(\xi)}d\xi<\infty$ for all $s<0$, then $h^*(\xi)\in G_D$ that maximizes $H_S(h)$ is

$$h^*(\xi)=\frac{e^{s\rho(\xi)}}{\int_{\Re^d}e^{s\rho(\xi)}d\xi},\quad \int_{\Re^d}\rho(\xi)h^*(\xi)d\xi=D_v. \quad (11)$$

Subsequently, when $R(D_v)$ and $\mathcal{H}_S(\mathcal{Y})\triangleq\lim_{T\to\infty}\frac{1}{T}H_S(Y^{T-1})$ exist, the Shannon lower bound, $R_S(D_v)$, is given by

$$R(D_v)\geq\mathcal{H}_S(\mathcal{Y})-\max_{h\in G_D}H_S(h)\triangleq R_S(D_v). \quad (12)$$

ii) Suppose the difference distortion measure $\rho(.)$ satisfies the conditions a,b,d of ([24], pp. 2029), $\int_{\Re^d}e^{s\rho(\xi)}d\xi<\infty$ for all $s<0$, $\mathcal{H}_S(\mathcal{Y})>-\infty$ and there exists an $y^*\in\Re^d$ such that $E\rho(y-y^*)<\infty,\forall y\in\Re^d$.
Then, in the limit as $D_v\to 0$, the lower bound is asymptotically exact. That is, for the case when $R(D_v)$ and $\mathcal{H}_S(\mathcal{Y})$ exist, $\lim_{D_v\to 0}\Big[R(D_v)-\Big(\mathcal{H}_S(\mathcal{Y})-H_S(h^*)\Big)\Big]=0$. Furthermore, for independent asymptotic stationary sources $R(D_v)|_R=R(D_v)=R_S(D_v)$.
Proof: Follows from [24] by considering the method proposed in ([21], pp. 140) or [25].

IV. Mean Square Observability over AWGN Channel

In this section, it is shown that if the capacity is at least equal to the Shannon lower bound, there exists an encoding scheme which guarantees the mean square observability over AWGN channel. This result together with the results of [15] implies that the Shannon lower bound is the minimum achievable capacity for the mean square observability. For simplicity, we consider the case of $Y_t \in \Re$ (the case of $Y_t \in \Re^d$ follows similarly).

Consider the control/communication system of Fig. 1. Assume the encoder and decoder is of Class A, and the innovation generator block of Fig. 1 produces an orthogonal Gaussian innovations process $K_t = Y_t - E[Y_t|\sigma\{Y^{t-1}, U^{t-1}\}] \sim N(0, \Lambda_t)$, where $\sigma\{.\}$ denotes the σ-algebra and $\Lambda_t \triangleq CV_tC^{'} + DD^{'}$ in which

$$\begin{aligned} V_{t+1} &= AV_tA^{'} - AV_tC^{'}(CV_tC^{'} + DD^{'})^{-1}CV_tA^{'} \\ &\quad +BB^{'},\ V_0 = \bar{V}_0. \end{aligned} \tag{13}$$

Next, for the distortion constraint $\mathcal{M}_{DC} = \{P(d\tilde{K}^{T-1}; k^{T-1}); \frac{1}{T}\sum_{t=0}^{T-1} E||K_t - \tilde{K}_t||^2 \leq D_v\}$, consider the rate distortion function of Definition 3.4. For $D_v < \min_{t\in\mathbf{N}_+}\Lambda_t$, the minimizing kernel is given by

$$\begin{aligned} &P^*(d\tilde{K}_t; k^t, \tilde{k}^{t-1}) = q^*(\tilde{K}_t|k_t)d\tilde{K}_t, \\ &q^*(\tilde{K}_t|k_t) \sim N(\beta_t K_t, \beta_t D_v), \quad \beta_t \triangleq 1 - \frac{D_v}{\Lambda_t}. \end{aligned} \tag{14}$$

Subsequently, the solution to the rate distortion is given by $R_T(D_v)|_R = \sum_{t=0}^{T-1}\frac{1}{2}\log\frac{\Lambda_t}{D_v}$. Next, under assumption of (C, A) is detectable, $(A, (BB^{'})^{\frac{1}{2}})$ is stabilizable, and $D \neq 0$, $\lim_{T\to\infty}\Lambda_T = \Lambda_\infty$ [18], where $\Lambda_\infty = CV_\infty C^{'} + DD^{'}$ and V_∞ is the solution of the Algebraic Riccati equation corresponding to the Riccati equation (13). Therefore, for $D_v < \min_{t\in\mathbf{N}_+}\Lambda_t$, we have

$$R(D_v)|_R = \frac{1}{2}\log\frac{\Lambda_\infty}{D_v}. \tag{15}$$

On the other hand, for the same distortion measure, as above, the Shannon lower bound, $R_S(D_v)$, for the innovations process $\{K_t\}_{t\in\mathbf{N}_+}$ is given by

$$\begin{aligned} R(D_v)|_R = R(D_v) \geq R_S(D_v) &\triangleq \mathcal{H}_s(\mathcal{K}) - \max_{h\in G_D} H_S(h) \\ &= \frac{1}{2}\log\frac{\Lambda_\infty}{D_v}. \end{aligned} \tag{16}$$

Subsequently, from (15) and (16) it follows that for $D_v < \min_{t\in\mathbf{N}_+}\Lambda_t$, the Shannon lower bound is exact.

Next, a matched channel (e.g., a communication channel in which the channel input-to-channel output behaves like the rate distortion infimizing stochastic kernel [22]) corresponding to (14) is the following AWGN channel

$$\begin{aligned} &\tilde{Z}_t = Z_t + N_t, \quad Z_t \in \Re \\ &E[Z_t^2] \leq P_t,\ N_t \text{ orthogonal} \sim N(0, \frac{D_v}{\beta_t}),\ t \in \mathbf{N}_+. \end{aligned} \tag{17}$$

Next, if we choose $\alpha_t = 1$ and $\gamma_t = \beta_t$, the power constraint associated to this encoding scheme is $P_t = E[K_t^2] = \Lambda_t$ and for $D_v < \min_{t\in\mathbf{N}_+}\Lambda_t$, it can be easily shown that the capacity is $\mathcal{C}_{|R} = R_S(D_v)$; also $E||K_t - \tilde{K}_t||^2 = D_v$, $\forall t \in \mathbf{N}_+$. That is, using this encoding scheme the mean square observability of innovations process over the matched channel (17) with capacity $R_S(D_v)$ can be obtained.

On the other hand, from ([15], Theorem 3.3), we conclude that $\mathcal{C}_{|R} \geq R_S(D_v)$ is also a necessary condition for existence of an encoding scheme (e.g., for the existence of α_t and γ_t) for the mean square observability. Subsequently, following this result and the above result, we can conclude that for $D_v < \min_{t\in\mathbf{N}_+}\Lambda_t$, $\mathcal{C}_{|R} = R_S(D_v)$ is the minimum achievable capacity for the mean square observability of innovations process over the matched AWGN channel (17), where this capacity is obtained for $\alpha_t = 1$ and $\gamma_t = \beta_t$.

In the above analysis we have shown that over the matched AWGN channel (17), reliable data reconstruction is possible by transmitting information from the source at least equals to $R_S(D_v)$. Nevertheless, in rare situations the communication channel is of the form (17) and it is normally of the following form.

$$\begin{aligned} &\tilde{Z}_t = Z_t + N_t, \quad Z_t \in \Re \\ &E[Z_t^2] \leq P_t, \quad N_t \text{ orthogonal } \sim N(0, W). \end{aligned} \tag{18}$$

However, over this communication channel by choosing $\alpha_t = \sqrt{\frac{\beta_t W}{D_v}}$ and $\gamma_t = \sqrt{\frac{D_v\beta_t}{W}}$, we can still have the mean square observability of innovations process by transmitting $\mathcal{C}_{|R} = R_S(D_v)$ bits in each time step. In other words, $\mathcal{C}_{|R} = R_S(D_v)$ is also the minimum achievable capacity over the AWGN channel (18) in which this capacity is obtained by choosing $\alpha_t = \sqrt{\frac{\beta_t W}{D_v}}$ and $\gamma_t = \sqrt{\frac{D_v\beta_t}{W}}$.

Please note that from classical information theoretic results (e.g., information transmission theorem [21] [23]), we already know that for the (asymptotic) stationary ergodic sources over DMC's or AWGN channels, the rate distortion and subsequently the Shannon lower bound is the minimum achievable capacity for observability. This result is obtained following the random coding argument in which it does not address a specific encoding scheme for reliable data reconstruction of a given source over a given channel. Nevertheless, in this section, we investigated the validity of this theorem for the Gaussian system (1) over AWGN channel (18) by proposing a specific encoding scheme. Further, we relate this rate to the parameters of the Gaussian system (1), in particular to the Shannon entropy rate.

V. Mean Square Stabilizability

In the previous section when the encoder is of Class A, we proposed an encoding scheme that reliably transmits information for the partially observed system (1) over AWGN channel (18). Next, we can use this transmitted information to stabilize the dynamical system (1). Since we are interested in mean square stabilizability, we shall first provide the mean square estimation of the states of the dynamical system.

When the decoder is of Class B, the optimal mean square state estimator is given by $\hat{X}_t = E[X_t|\tilde{K}^{t-1}, U^{t-1}]$. Nevertheless, for the unstable system (i.e., when some of the eigenvalues of the system matrix A in (1) are outside or on the unit circle), the mean square estimation error associated to the innovations process produced by the encoder of Class A is going to be unbounded. Subsequently, in the control/communication system of Fig. 1, we shall use an innovations encoder which uses channel with feedback. That is, we use an encoder of Class B to produce the orthogonal Gaussian innovations process $K_t = Y_t - CE[X_t|\tilde{K}^{t-1}, U^{t-1}]$. This encoder scales K_t by $\alpha_t = \sqrt{\frac{\eta_t W}{D_v}}$ where $D_v \leq \min_{t\in\mathbf{N}_+} \Upsilon_t$ ($\Upsilon_t \stackrel{\triangle}{=} C\Pi_t C^{'} + DD^{'}$), and $\eta_t \stackrel{\triangle}{=} 1 - \frac{D_v}{\Upsilon_t}$ in which Π_t is obtained from the following recursive equation

$$\begin{aligned}\Pi_{t+1} &\stackrel{\triangle}{=} A\Pi_t A^{'} - A\Pi_t C^{'}(C\Pi_t C^{'} + DD^{'} + \frac{W}{\alpha_t^2})^{-1} \\ &\quad .C\Pi_t A^{'} + BB^{'}, \quad \Pi_0 = \bar{V}_0. \end{aligned} \tag{19}$$

The decoder, on the other hand, scales the output of the channel by $\gamma_t = \sqrt{\frac{D_v \eta_t}{W}}$ and produces $\tilde{K}_t = \sqrt{\frac{D_v \eta_t}{W}} \tilde{Z}_t$. Consequently, using this encoding scheme, it can be easily shown that under the assumption of (A, C) is detectable and $(A, (B^{'}B)^{\frac{1}{2}})$ is stabilizable, $\mathcal{C}_{|R} = \frac{1}{2}\log\frac{\Upsilon_\infty}{D_v} = R_S(D_v)$, where $R_S(D_v)$ is the Shannon lower bound associated to the innovations process $K_t = Y_t - CE[X_t|\tilde{K}^{t-1}, U^{t-1}]$ and $\Upsilon_\infty = C\Pi_\infty C^{'} + DD^{'}$ where Π_∞ is the solution to the Algebraic Riccati equation corresponding to the Riccati equation (19). Further, following the expression for the innovation process and since $E[X_t|\tilde{K}^{t-1}, U^{t-1}]$ is known for both the encoder and decoder, the reproduction of the source message at the decoder end, is $\tilde{Y}_t \stackrel{\triangle}{=} \tilde{K}_t + CE[X_t|\tilde{K}^{t-1}, U^{t-1}]$. Subsequently, it can be easily shown that $E||Y_t - \tilde{Y}_t||^2 = E||K_t - \tilde{K}_t||^2 = D_v$, $\forall t \in \mathbf{N}_+$. Next, we can use the mean square estimator to estimate the state variable (even for the unstable system). This estimator is given by the following recursive equation (Kalman filter)

$$\begin{aligned}\hat{X}_{t+1} &= A\hat{X}_t + \frac{1}{\alpha_t\gamma_t} A\Pi_t C^{'}(C\Pi_t C^{'} + DD^{'} + \frac{W}{\alpha_t^2})^{-1}\tilde{K}_t \\ &\quad + NU_t, \quad \hat{X}_0 = \bar{x}_0. \end{aligned} \tag{20}$$

Next, consider the following quadratic pay off functional

$$\lim_{T\to\infty} \frac{1}{T} E \sum_{t=0}^{T-1} \left(||X_t||^2_{C^{'}C} + ||U_t||^2_H \right) \quad (H > 0). \tag{21}$$

From classical separation principle [18] follows that the stabilizing controller that minimizes the pay off functional (21) subject to AWGN communication constraint is given by

$$\begin{aligned} U_t^* &= -\Delta \hat{X}_t \\ \Delta &\stackrel{\triangle}{=} (H + N^{'}P_\infty N)^{-1} N^{'} P_\infty A \\ P_\infty &= A^{'}P_\infty A - A^{'}P_\infty N (H + N^{'}P_\infty N)^{-1} N^{'} P_\infty A \\ &\quad + C^{'}C \end{aligned} \tag{22}$$

provided $((C^{'}C)^{\frac{1}{2}}, A)$ is detectable and (A, N) is stabilizable. That is, for $D_v < \min_{t\in\mathbf{N}_+} \Upsilon_t$ over AWGN channel (18), using the control policy (22), we can have

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} E||X_t||^2_{C^{'}C} < D_v^c, \tag{23}$$

where D_v^c is the minimum quadratic cost (i.e., $D_v^c = \lim_{T\to\infty} \frac{1}{T} E \sum_{t=0}^{T-1} \left(||X_t||^2_{C^{'}C} + ||U_t^*||^2_H \right)$), when the transmission data rate is $\mathcal{C}_{|R} = R_S(D_v)$.

Please note that since $\sigma\{U^t\} \subseteq \sigma\{\tilde{K}^t\} \subseteq \sigma\{\tilde{Z}^t\}$, then by knowing the output of the communication channel, the encoder and decoder can also specify the control sequence. Consequently, the encoder can be of Class C, while the decoder of Class A. Moreover, from the above construction it is evident that a separation principle exists between the design of the control and the communication systems.

Moreover, the above analysis shows that mean square observability of innovations process over an AWGN channel with minimum achievable capacity is possible in the present of channel with feedback. That is, if the encoder is of Class C and decoder is of Class A. Furthermore, one can consider K^{T-1} as an orthogonal version of Y^{T-1}. That is, $K^{T-1} = \Gamma^{-1} Y^{T-1}$ where $Cov((K_0, K_1, ..., K_{T-1})^{'}) = \Gamma^{-1} Cov((Y_0, Y_1, ..., Y_{T-1})^{'})(\Gamma^{-1})^{'}$ (i.e., Γ^{-1} is the unitary matrix that diagonalize $Cov((Y_0, Y_1, ..., Y_{T-1})^{'})$). Subsequently, from ([21], pp.110) rate distortion function between K^{T-1} and $\tilde{K}^{T-1}$ is identical to the rate distortion function between Y^{T-1} and $\tilde{Y}^{T-1}$. Thus, in the presence of channel with feedback mean square observability of the observed process over an AWGN channel with minimum achievable capacity is possible.

VI. Conclusion

The present paper complements the results of [15] by designing encoder, decoder and controller which can guarantee observability and stabilizability. Further a modified definition for channel capacity and rate distortion which is suitable for control applications and real time communication has been presented. In this paper the aim is to show the achievablity over wireless communication channels. Therefore, AWGN channel which is a basic model for wireless communication channels, was considered. Nevertheless, for future direction it would be interesting to consider the effects of fading and interference since wireless communication channels are normally subject to fading and interference.

Acknowledgment

This work was supported by Cyprus Research Promotion Foundation under Grant $\Pi\Lambda$HPO$/05/0506/20$.

References

[1] Wei Zhang, *Stability Analysis of Networked Control Systems*, Ph.D. Thesis, Department of Electrical Engineering and Computer Science, Case Western Reserve University, Auguest 2001.

[2] D. F. Delchamps, Stabilizing a Linear System with Quantized State Feedback, *IEEE Transactions on Automatic Control*, vol. 35, No. 8, pp. 916-924, Auguest 1990.

[3] W. S. Wong and R. W. Brockett, Systems with Finite Communication Bandwidth Constraints-Part I: State Estimation Problems, *IEEE Transactions on Automatic Control*, vol. 42, No. 9, pp. 1294-1299, September 1997.

[4] A. V. Savkin and I. R. Petersen, Set-Valued State Estimation via a Limited Capacity Communication Channel, *IEEE Transactions on Automatic Control*, vol. 48, No. 4, pp. 676-680, Appril 2003.

[5] V. Malyavej and A. V. Savkin, The problem of Optimal Robust Kalman State Estimation via Limited Capacity Digital Communication Channels, *System and Control Letters*, vol. 45, No. 3, pp. 283-292, March 2005.
[6] G. N. Nair and R. J. Evans, Stabilizability of Stochastic Linear Systems With Finite Feedback Data Rates, *SIAM Journal of Control and Optimization*, vol. 43, No. 2, pp. 413-436, 2004.
[7] S. Tatikonda and S. Mitter, Control over Noisy Channels, *IEEE Transactions on Automatic Control*, vol. 49, No. 7, pp. 1196-1201, July 2004.
[8] S. Tatikonda and S. Mitter, Control under Communication Constraints, *IEEE Transactions on Automatic Control*, vol. 49, No. 7, pp. 1056-1068, July 2004.
[9] S. Tatikonda, A. Sahai, and S. Mitter, Stochastic Linear Control Over a Communication Channel, *IEEE Transactions on Automatic Control*, vol. 49, No. 9, pp. 1549-1561, Sempetmber 2004.
[10] Nicola Elia, When Bode Meets Shannon: Control-Oriented Feedback Communication Schemes, *IEEE Transactions on Automatic Control*, vol. 49, No. 9, pp. 1477-1488, September 2004.
[11] K. Li and J. Baillieal, Robust Quantization for Digital Finite Communication Bandwidth (DFCB) Control, *IEEE Transactions on Automatic Control*, vol. 49, No. 9, pp. 1573-1584, September 2004.
[12] G. N. Nair, R. J. Evans, I. M. Y. Mareels and W. Moran, Topological Feedback Entropy and Nonlinear Stabilization, *IEEE Transations on Automatic Control*, vol. 49, No. 9, pp. 1585-1597, September 2004.
[13] D. Liberzon and J. P. Hespanha, Stabilization of Nonlinear Systems with Limited Information Feedback, *IEEE Transactions on Automatic Control*, vol. 50, No. 6, pp. 910-915, June 2005.
[14] N. C. Martins, A. Dahleh, and N. Elia, Feedback Stabilization of Uncertain Systems in the Presence of a Direct Link, *IEEE Transactions on Automatic Control*, vol. 51, No. 3, March 2006.
[15] C. D. Charalambous and Alireza Farhadi, “A Mathematical Framework for Robust Control over Uncertain Communication Channels”, in the *Proceedings of the 44th IEEE Conference on Decision and Control and 2005 European Control Conference*, pp. 2530-2535, Seville, December 12-15, 2005.
[16] Rajesh Bansal and Tamer Basar, Simultaneous Design of Measurement and Control Strategies for Stochastic Systems with Feedback, *Automatica*, vol. 25, No. 5, pp. 679-694, 1989.
[17] C. D. Charalambous,“Information Theory for Control Systems: Causality and Feedback”, *the Lecture Presented at the Symposium on Towards a Science of Networks, Communication Networks and Complexity*, Athens, Greece, Auguest 31st, 2006.
[18] P. E. Caines, *Linear Stochastic Systems*, John Wiley and Sons, 1988.
[19] C. D. Charalambous, Alireza Farhadi, and F. Rezaei, Information Theory for Control Systems: Causality, Feedback and Separation Principle, (*Preprint*).
[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John willey and Sons, 1991.
[21] Toby Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*, Printice-Hall, 1971.
[22] Sekhar Tatikonda, *Control Under Communication Constraint*, Ph.D. Thesis, Department of Electrical Engineering and Computer Science, MIT., September 2000.
[23] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley and Sons, INC., 1968.
[24] T. Linder and R. Zamir, On the Asymptotic Tightness of the Shannon Lower Bound, *IEEE Transactions on Information Theory*, vol. 40, No. 6, pp. 2026-2031, November 1994.
[25] J. T. Pinkston, *Information Rates of Independent Sample Sources*, M.S. Thesis, Department of Electerical Engineering, MIT., Cambridge, Mass. 1966.