[10] S. Mu, T. Chu, and L. Wang, "Coordinated collective motion in a motile particle group with a leader," *Physica A*, vol. 351, pp. 211–226, Jun. 2005.

[11] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 169–182, Feb. 2005.

[12] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, May 2005.

[13] L. Gao and D. Cheng, "Comment on 'Coordination of groups of mobile autonomous agents using nearest neighbor rules,'" *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 1913–1916, Nov. 2005.

[14] R. O. Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401–420, Mar. 2006.

[15] W. Wang and J. E. Slotine, "A theoretical study of different leader roles in networks," *IEEE Trans. Autom. Control*, vol. 51, no. 7, pp. 1156–1161, Jul. 2006.

[16] H. Shi, L. Wang, and T. Chu, "Virtual leader approach to coordinated control of multiple mobile agents with asymmetric interactions," *Physica D*, vol. 213, pp. 51–65, 2006.

[17] F. Xiao and L. Wang, "State consensus for multi-agent systems with switching topologies and time-varying delays," *Int. J. Control*, vol. 79, no. 10, pp. 1277–1284, 2006.

[18] G. Xie and L. Wang, "Consensus control for a class of networks of dynamic agents," *Int. J. Robust Nonlinear Control*, vol. 17, pp. 941–959, 2007.

[19] Y. Hong, J. Hua, and L. Gao, "Tracking control for multi-agent consensus with an active leader and variable topology," *Automatica*, vol. 42, pp. 1177–1182, 2006.

[20] Y. Hong, L. Gao, D. Cheng, and J. Hu, "Lyapunov-based approach to multiagent systems with switching jointly connected interconnection," *IEEE Trans. Autom. Control*, vol. 52, no. 5, pp. 943–948, May 2007.

[21] B. Liu, T. Chu, L. Wang, and G. Xie, "Controllability of a class of multi-agent systems with a leader," in *Proc. IEEE Conf. Amer. Control*, Minneapolis, MN, Jun. 2006, pp. 2844–2849.

[22] B. Liu, G. Xie, T. Chu, and L. Wang, "Controllability of interconnected systems via switching networks with a leader," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Taipei, Taiwan, Oct. 2006, pp. 3912–3916.

[23] C. W. Reynolds, "Flocks, birds, and school: A distributed behavioral model," *Comput. Graph.*, vol. 21, pp. 25–34, 1987.

[24] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Phys. Rev. Lett.*, vol. 675, pp. 1226–1229, 1995.

[25] V. Gazi and K. M. Passino, "Stability analysis of swarms," *IEEE Trans. Autom. Control*, vol. 48, no. 4, pp. 692–697, Apr. 2003.

[26] T. Chu, L. Wang, and T. Chen, "Self-organized motion in anisotropic swarms," *J. Control Theor. Appl.*, vol. 1, pp. 77–81, 2003.

[27] A. V. Savkin, "Coordinated collective motion of groups of autonomous mobile robots: Analysis of Vicsek's model," *IEEE Trans. Autom. Control*, vol. 49, no. 6, pp. 981–989, Jun. 2004.

[28] B. Liu, T. Chu, L. Wang, and Z. Wang, "Swarm dynamics of a group of mobile autonomous agents," *Chinese Phys. Lett.*, vol. 22, no. 1, pp. 254–257, 2005.

[29] T. Chu, L. Wang, T. Chen, and S. Mu, "Complex emergent dynamics of anisotropic swarms: Convergence vs. oscillation," *Chaos Solitons Fractals*, vol. 30, pp. 875–885, 2006.

[30] C. Godsil and G. Royle, *Algebraic Graph Theory*. New York: Springer-Verlag, 2001.

[31] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*. Berlin, Germany: Springer-Verlag, 1974.

[32] G. Xie and L. Wang, "Reachability realization and stabilizability of switched linear iscrete-time systems," *J. Math. Anal. Appl.*, vol. 280, pp. 209–220, 2003.

Control of Continuous-Time Linear Gaussian Systems Over Additive Gaussian Wireless Fading Channels: A Separation Principle

Charalambos D. Charalambous, Alireza Farhadi, and Stojan Z. Denic

***Abstract*—This note is concerned with the control of continuous-time linear Gaussian systems over additive white noise wireless fading channels subject to capacity constraints. Necessary and sufficient conditions are derived, for bounded asymptotic and asymptotic observability and stabilizability in the mean square sense, for controlling such systems. For the case of a noiseless time-invariant system controlled over a continuous-time additive white Gaussian noise channel, the sufficient condition for stabilizability and observability states that the capacity of the channel C must satisfy $C > \sum_{\{i;\mathrm{Re}(\lambda_i(A))\geq 0\}} \mathrm{Re}(\lambda_i(A))$, where A is the system matrix and $\lambda_i(A)$ denotes the eigenvalues of A. The necessary condition states that the channel capacity must satisfy $C \geq \sum_{\{i;\mathrm{Re}(\lambda_i(A))\geq 0\}} \mathrm{Re}(\lambda_i(A))$. Further, it is shown that a separation principle holds between the design of the communication and the control subsystems, implying that the controller that would be optimal in the absence of the communication channel is also optimal for the problem of controlling the system over the communication channel.**

***Index Terms*—Continuous time, mutual information, networked control system, stabilizability and observability.**

I. Introduction

In recent years, there has been a significant activity in addressing issues associated with the control of systems over limited capacity communication channels. A typical example is given in Fig. 1. The control/communication system of Fig. 1 can be used to describe a distributed control system in which the plant and the corresponding controller are connected through a shared communication media, while there is an unshared or high-capacity communication link from controller to plant. It can also be used to describe a teleoperation system in which the communication from the plant to the remote controller is subject to limited capacity constraint, while the connection from the controller to the plant is unconstrained. Since a discrete time model is more appropriate for today's digital communication links, previous work on this subject is focused on the observability and/or stabilizability of discrete-time systems, controlled over a discrete-time communication channel with finite capacity [1]–[8]. Nevertheless, in some applications, analog modulation schemes may be interesting due to the simplicity in realizing such schemes. On the other hand, having a complete theory that deals with continuous-time systems will help us gain additional insight and understanding into building control/communication systems.

It is already known that a necessary condition for observability and stabilizability of linear discrete-time invariant systems is given by

Manuscript received July 26, 2005; revised February 19, 2007, and November 1, 2007. Recommended by Associate Editor S. Dey. This work was supported in part by the European Commission under Project ICCCSYSTEMS and in part by the University of Cyprus via a medium-sized research grant.

C. D. Charalambous is with the Electrical and Computer Engineering Department, University of Cyprus, Nicosia 1678, Cyprus. He is also with the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: chadcha@ucy.ac.cy).

A. Farhadi is with the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: afarhadi@site.uottawa.ca).

S. Z. Denic is with Arizona State University, Tempe, Arizona 85287 USA (e-mail: denstol@yahoo.com).

Digital Object Identifier 10.1109/TAC.2008.919533

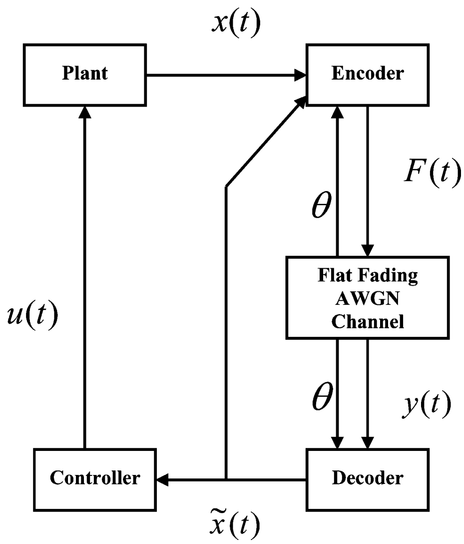

Fig. 1. Control/communication system over flat fading AWGN channel.

the condition $C \geq \sum_{\{i;|\lambda_i A)|\geq 1\}} \log|\lambda_i(A)|$, where C is the channel capacity and $\lambda_i(A)$'s are eigenvalues of system matrix A. Furthermore, when the inequality is strict, this condition is also sufficient.

In this note, we first consider the problem of stabilizability of a linear continuous-time invariant noiseless plant controlled over a continuous-time additive white Gaussian noise (AWGN) channel with memory. We then consider a linear stochastic Gaussian time-varying plant driven by Brownian motion controlled over a flat fading wireless channel. Here, we assume complete knowledge of the channel throughout the transmission, at the transmitter and the receiver ends [9]. Optimal encoding and decoding strategies that minimize the mean square decoding error as well as an optimal mean square stabilizability scheme are derived, and conditions for mean square observability and stabilizability are presented. It is also shown that a separation principle holds between the design of the control and the communication subsystems. Similar separation property for discrete-time systems is discussed in [6] and [10].

The note is organized as follows. In Section II, the problem formulation is given. In Section III, a necessary condition for stabilizability is presented. In Section IV, the optimal encoding/decoding scheme that ensures observability is given. In Section V, a stabilizing controller is designed by using a linear quadratic payoff, and it is shown that the encoder/decoder proposed achieves the capacity of the channel.

II. Problem Formulation

We shall give the precise definition of the signals and blocks associated with Fig. 1. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and time $t \in [0,T]$, $T > 0$. Let $x \triangleq \{x(s); 0 \leq s \leq T\}$, $x(t) \in \Re^n$ denote the output of the controlled plant (transmitted signal), $u \triangleq \{u(s); 0 \leq s \leq T\}$, $u(t) \in \Re^m$ the control signal, $y \triangleq \{y(s); 0 \leq s \leq T\}$, $y(t) \in \Re$ the output of the communication channel, $\theta \triangleq \{\theta(s); 0 \leq s \leq T\}$, $\theta(t) \in \Re^q$ the channel state information (CSI), $v \triangleq \{v(s); 0 \leq s \leq T\}$, $v(t) \in \Re$ the channel noise, $w \triangleq \{w(s); 0 \leq s \leq T\}$, $w(t) \in \Re^l$ the plant process noise, and $\tilde{x} \triangleq \{\tilde{x}(s); 0 \leq s \leq T\}$, $\tilde{x}(t) \in \Re^n$ the decoder output. Denote the fading process by $z \triangleq \{z(s, \theta(s)); 0 \leq s \leq T\}$, $z(t, \theta(t)) \in \Re$ that is assumed to be independent of the plant noise and the initial state. The plant noise w and the channel noise v are independent standard Brownain motions $(Ev(t)^2 = N_0 t, \mathrm{Cov}(w(t)) = I.t)$, which are independent of the initial state $x(0)$. Let $\{\mathcal{F}^x_{0,t}\}_{t\geq 0}$, $\{\mathcal{F}^{\tilde{x}}_{0,t}\}_{t\geq 0}$, $\{\mathcal{F}^y_{0,t}\}_{t\geq 0}$, and $\{\mathcal{F}^\theta_{0,t}\}_{t\geq 0}$ denote the complete filtration generated by $\mathcal{F}^x_{0,t} \triangleq \sigma\{x(s); 0 \leq s \leq t\}$, $\mathcal{F}^{\tilde{x}}_{0,t} \triangleq \sigma\{\tilde{x}(s); 0 \leq s \leq t\}$, $\mathcal{F}^y_{0,t} \triangleq \sigma\{y(s); 0 \leq s \leq t\}$, and $\mathcal{F}^\theta_{0,t} \triangleq \sigma\{\theta(s); 0 \leq s \leq t\}$, respectively, which are subsigma fields of $\{\mathcal{F}_t\}_{t\geq 0}$. Here, $\mathcal{F}^x_{0,t}$, $\mathcal{F}^{\tilde{x}}_{0,t}$, $\mathcal{F}^y_{0,t}$, and $\mathcal{F}^\theta_{0,t}$ are the Borel σ-algebras on the space of continuous functions $C([0,T]; \Re^n)$, $C([0,T]; \Re^n)$, $C([0,T]; \Re)$, and $C([0,T]; \Re^q)$ respectively. Next, the blocks of Fig. 1 are defined.

Plant: The state of the plant is described by the *Itô-controlled diffusion process*

$$dx(t) = A(t)x(t)dt + B(t)u(t)dt + G(t)dw(t), \quad x(0) \quad (1)$$

where $A : [0,T] \to \Re^{n\times n}$, $B : [0,T] \to \Re^{n\times m}$, and $G : [0,T] \to \Re^{n\times l}$, and $x(0)$ is Gaussian random variable $x(0) \sim N(\bar{x}_0, \bar{V}_0)$, which is independent of w. The control u is $\{\mathcal{F}_{0,t}\}_{t\geq 0}$ adapted, and $A(t)$, $B(t)$, and $G(t)$ are uniformly bounded.

Encoder: The encoder map $\{(x(s), \tilde{x}(s), z(s, \theta(s))); 0 \leq s \leq t\} \to F(t, x, \tilde{x}, \theta) (= F(t) \in \Re$ in compact notation) is adapted to $\{\mathcal{F}^{x,\tilde{x},\theta}_{0,t}\}_{t\geq 0}$ defined by $\mathcal{F}^{x,\tilde{x},\theta}_{0,t} \triangleq \mathcal{F}^x_{0,t} \bigvee \mathcal{F}^{\tilde{x}}_{0,t} \bigvee \mathcal{F}^\theta_{0,t}$, with power constraint $E[|F(t, x, \tilde{x}, \theta)|^2 | \mathcal{F}^\theta_{0,t}] \leq P$. The set of such admissible encoders is denoted by $\mathcal{F}_{\mathrm{ad}}$.

Channel: The communication channel is an AWGN, flat fading, wireless channel whose output y is defined by the following stochastic differential equation

$$dy(t) = z(t, \theta(t))F(t, x, \tilde{x}, \theta)dt + dv(t), \qquad 0 \leq t \leq T \quad (2)$$

where $y(0) = 0$ and v is a Brownian motion with $Ev^2(t) = N_0 t$. Throughout, we shall assume that (2) has a unique solution [11], and that $\int_0^T E[z(t, \theta(t))F(t, x, \tilde{x}, \theta)]^2 dt < \infty$, for finite T. Further, when computing the conditional mutual information over an infinite time horizon, we shall assume that $\limsup_{T\to\infty} \frac{1}{2T} \int_0^T E[z^2(t, \theta(t))]dt$ is finite. If $\lim_{t\to\infty} E[z^2(t, \theta(t))]$ exists, then $\limsup$ can be replaced by $\lim$.

Decoder: The decoder map $\{(y(s), z(s, \theta(s))); 0 \leq s \leq t\} \to \tilde{x}(t, y, \theta) (= \tilde{x}(t)$ in compact notation) is adapted to $\{\mathcal{F}^{y,\theta}_{0,t}\}_{t\geq 0}$, where $\mathcal{F}^{y,\theta}_{0,t} \triangleq \mathcal{F}^y_{0,t} \bigvee \mathcal{F}^\theta_{0,t}$. The set of admissible decoders is denoted by $\mathcal{D}_{\mathrm{ad}}$. The decoder plays the role of a state estimator.

Controller: The controller u is a nonanticipative functional of the output of the decoder and the CSI, e.g., u is $\{\mathcal{F}^{y,\theta}_{0,t}\}_{t\geq 0}$ adapted. The set of admissible controller is denoted by $\mathcal{U}_{\mathrm{ad}}$.

The objective of this note is to find necessary and sufficient conditions for bounded asymptotic and asymptotic observability and stabilizability of system (1), in the following sense.

Definition 2.1: Define $\mathcal{E}(t) \triangleq E[(x(t) - \tilde{x}(t, y, \theta))^{\mathrm{tr}}(x(t) - \tilde{x}(t, y, \theta)) | \mathcal{F}^{y,\theta}_{0,t}]$. System (1) and (2) is bounded asymptotically (respectively, asymptotically) observable, in the mean square sense, if there exist an encoder $F \in \mathcal{F}_{\mathrm{ad}}$ and decoder $\tilde{x} \in \mathcal{D}_{\mathrm{ad}}$, such that $\lim_{t\to\infty} \mathcal{E}(t) < \infty$, P-a.s. (respectively, $\lim_{t\to\infty} \mathcal{E}(t) = 0$, P-a.s.).

Definition 2.2: Define $||x||^2 \triangleq x^{\mathrm{tr}}x, x \in \Re^n$. System (1) and (2) is bounded asymptotically (respectively, asymptotically) stabilizable, in the mean square sense, if there exit a controller, encoder, and decoder, such that $\lim_{t\to\infty} E[||x(t)||^2 | \mathcal{F}^\theta_{0,t}] < \infty$, P-a.s., (respectively, $\lim_{t\to\infty} E[||x(t)||^2 | \mathcal{F}^\theta_{0,t}] = 0$, P-a.s).

III. Bode's Integral Formula, Necessary Condition for Existence of Stabilizing Controller for Continuous-Time Systems

In this section, we consider the time-invariant noiseless analog of system (1) with $u(t) \in \Re$, that is, the plant is given by

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) \in \Re^n, \qquad u(t) \in \Re \tag{3}$$

$x(0) \sim N(\bar{x}_0, \bar{V}_0)$. The communication channel is an AWGN channel ($z = 1$) with memory given by

$$y(t) = o(t) + n(t), \qquad o(t) = h(t) * F(t), \qquad y(t) \in \Re \tag{4}$$

where o is a stochastic process with power spectral density (PSD) $S_o(\omega)$, n is a Gaussian white noise process with the PSD $S_n(\omega) = N_0$, $h(t)$ is a causal channel impulse response with corresponding transfer function $H(jw)$, "$*$" is the convolution operator, and $N(jw)$ and $Y(jw)$ are the Fourier transforms of the noise and output signals, respectively. Note that the Brownian motion and the white noise are related by $v(t) = \int_0^t n(s)ds$ with $Ev(t)v(s) = N_0 \min(t, s)$. In this section, it is assumed that the encoder, decoder, and controller are linear time-invariant with transfer functions $E(jw)$, $D(jw)$, and $C(jw)$, respectively.

The mutual information between the state of plant x and the channel output y is given by $I_T(x; y) \stackrel{\triangle}{=} E_{x,y}[\ln(dP_{x,y}(x, y)/(dP_x(x) \times dP_y(y)))]$, where $\ln(.)$ denotes the natural logarithm, $E_{x,y}[.]$ denotes expectation with respect to sample paths x and y, $P_{x,y}$ is the joint probability measure of x and y, P_x is the probability measure of x, and P_y is the probability measure of y. Subsequently, the finite-time channel capacity is defined by $C_h^T = \sup_{(x,F)\in\mathcal{X}\times\mathcal{F}_{\mathrm{ad}}}(1/T)I_T(x; y)$, and the infinite-horizon channel capacity by $C_h = \liminf_{T\to\infty} C_h^T$, where $\mathcal{X}$ is the set of all possible continuous sample paths x's, and $\mathcal{F}_{\mathrm{ad}}$ is the set of all admissible encoders [11].

Let $\lambda_i(A)$ denote the eigenvalues of A. We shall show that $C_h \geq \sum_{\{i;\mathrm{Re}(\lambda_i(A))\geq 0\}} \mathrm{Re}(\lambda_i(A))$ is a necessary condition for the existence of a stabilizing controller.

In this section, a controller is called stabilizable if the corresponding closed-loop sensitivity transfer function $S(jw) = Y(jw)/N(jw)$, from n to y, is strictly stable or alternatively $\lim_{t\to\infty} E|y(t)|^2 < \infty$ or $\lim_{t\to\infty} E||x(t)||^2 < \infty$.

The main result of this section is given in the following theorem.

Theorem 3.1: Consider the control/communication system of Fig. 1 described by (3) and (4) with linear time-invariant encoder, decoder, and controller in which there is no feedback from the output of the decoder to the input of the encoder. A necessary condition for the existence of a stabilizing controller is given by $C_h \geq \sum_{\{i:\mathrm{Re}(\lambda_i(A))\geq 0\}} \mathrm{Re}(\lambda_i(A))$, where C_h is the capacity of the AWGN channel with memory.

Proof: Assume there exist a stabilizing controller, and an encoder/decoder pair such that the control/communication system is stable. Define $l(t) \stackrel{\triangle}{=} \int_0^t y(s)ds$. Then, $l(t) = \int_0^t o(s)ds + \int_0^t n(s)ds$. Since, $v(t) = \int_0^t n(s)ds$ is Brownian motion, then we can apply [12] to get: $I_T(l; x) = \frac{1}{2N_0}E\{\int_0^T |o(t) - \hat{o}(t)|^2 dt\} = \frac{1}{2N_0}\int_0^T E|o(t) - \hat{o}(t)|^2 dt = \frac{1}{2N_0}\int_0^T \Sigma_t dt$ where $\hat{o}(t) \stackrel{\triangle}{=} E[o(t)|\mathcal{F}_{0,t}^y]$, and $\Sigma_t \stackrel{\triangle}{=} E|o(t) - \hat{o}(t)|^2$. Then, from the data processing inequality, it follows that $I_T(y; x) \geq I_T(l; x)$, and thus, $I_T(y; x) \geq \frac{1}{2N_0}\int_0^T \Sigma_t dt$. On the other hand, the mean square error is related to the PSDs via $\lim_{T\to\infty} \Sigma_T = \frac{N_0}{2\pi}\int_{-\infty}^{+\infty} \ln(1 + (S_o(\omega)/N_0))d\omega$ [12], where $S_o(\omega)$ is the PSD of o, and $S_n(\omega) = N_0$ is the PSD of n. Subsequently, $\lim_{T\to\infty} 1/T\int_0^T \Sigma_t dt = \lim_{T\to\infty} \Sigma_T = \frac{N_0}{2\pi}\int_{-\infty}^{+\infty} \ln(1 + (S_o(\omega)/N_0))d\omega$. Next, an application of Bode's integral formula [13] implies that

$$\begin{aligned}
\lim_{T\to\infty} \frac{1}{T} I_T(y; x) &\geq \lim_{T\to\infty} \frac{1}{2N_0}\frac{1}{T}\int_0^T \Sigma_t dt \\
&= \frac{1}{4\pi}\int_{-\infty}^{+\infty} \ln\left(1 + \frac{S_o(\omega)}{N_0}\right)d\omega \\
&= \frac{1}{4\pi}\int_{-\infty}^{+\infty} \ln \frac{S_y(\omega)}{S_n(\omega)}d\omega \\
&= \frac{1}{4\pi}\int_{-\infty}^{+\infty} \ln|S(\omega)|^2 d\omega \\
&= \sum_{\{i:\mathrm{Re}(\lambda_i(A))\geq 0\}} \mathrm{Re}(\lambda_i(A)).
\end{aligned} \tag{5}$$

Note that in order to derive the aforementioned result, the Bode integral formula is employed under the assumption that the stabilizing controller is chosen such that the corresponding open-loop transfer function (i.e., multiplications of channel, encoder, plant, controller, and decoder transfer functions) is strictly proper with degree at least 2. Subsequently, from (5), it follows that

$$\begin{aligned}
C_h &\stackrel{\triangle}{=} \liminf_{T\to\infty} \sup_{(x,F)\in\mathcal{X}\times\mathcal{F}_{\mathrm{ad}}} \frac{1}{T} I_T(y; x) \\
&\geq \lim_{T\to\infty} \frac{1}{T} I_T(y; x) \geq \sum_{\{i:\mathrm{Re}(\lambda_i(A))\geq 0\}} \mathrm{Re}(\lambda_i(A)).
\end{aligned} \tag{6}$$

Remark 3.2: Note that (6) is also a necessary condition for the case of AWGN channel [e.g., $h(t) = \delta(t)$]. That is, for AWGN channel, (6) is reduced to the following condition $C \geq \sum_{\{i:\mathrm{Re}(\lambda_i(A))\geq 0\}} \mathrm{Re}(\lambda_i(A))$, where C is the AWGN channel capacity. The equivalent condition of C for controlling discrete-time systems over a discrete-time noisy channel is given by [4] $C \geq \sum_{\{i;|\lambda_i(A)|\geq 1\}} \log|\lambda_i(A)|$. In Section V, we will achieve C by constructing an encoder, decoder, and controller that stabilize the plant.

IV. Optimal Encoding/Decoding Scheme for Observability

In this section, we design an optimal encoder/decoder pair for the time-varying system defined by (1) and (2) (in this section and the next section, we assume $N_0 = 1$) that guarantees the observability condition defined in the sense of Definition 2.1. The necessary and sufficient condition for the existence of such an encoder/decoder pair is given in terms of the capacity of the channel and the time-varying matrix $A(t)$. First, we consider the scalar case of system (1) because it is easier to present the idea developed that is based on [11] and [14], and then, we extend the result to the vector case.

A. *Optimal Encoding/Decoding Scheme for Observability: The Scalar Case*

1) Conditional Mutual Information and Capacity of Feedback Systems: We shall need the following equivalent expressions for conditional mutual information (which are variants of the mutual information found in [12] or [15]).

Theorem 4.1: Consider the model given by (1) and (2), shown in Fig. 1. The mutual information between the state of plant x and the channel output y, conditional on the state θ is given by

$$1)\ I_T(x; y|\theta) \stackrel{\triangle}{=} E_{x,y,\theta}\left[\ln \frac{dP_{x,y|\theta}(x, y|\theta)}{dP_{x|\theta}(x|\theta) \times dP_{y|\theta}(y|\theta)}\right] \tag{7}$$

$$2)\ I_T(x;y|\theta) = \frac{1}{2}E_\theta \int_0^T z^2(t,\theta(t))E[|F(t,x,\tilde{x},\theta)|^2 - |\hat{F}(t,\tilde{x},\theta)|^2|\mathcal{F}^\theta_{0,t}]dt \quad (8)$$

$$3)\ I_T(x;y|\theta) = \int_\theta \tilde{I}_T(\theta)dP_\theta(\theta) \quad (9)$$

where $\tilde{I}_T(\theta) = \int \ln(dP_{x,y|\theta}(x,y|\theta)/(dP_{x|\theta}(x|\theta) \times dP_{y|\theta}(y|\theta)))$ $dP_{x,y|\theta}(x,y|\theta)$, $E_{x,y,\theta}[.]$ represents expectation with respect to the sample paths x, y, θ, (similarly for $E_\theta[.]$) and $\hat{F}(t,\tilde{x},\theta) = E[F(t,x,\tilde{x},\theta)|\mathcal{F}^{y,\theta}_{0,t}]$. Here, P_θ is the probability measure induced by θ, while $P_{x,y|\theta}$ is a joint probability measure of x, and y conditional on θ (similarly for $P_{x|\theta}$ and $P_{y|\theta}$) (e.g., these are measures defined on the sample paths on the space of continuous functions).

Proof: 3) By invoking the definition of the conditional mutual information

$$I_T(x;y|\theta) \stackrel{\triangle}{=} \int \ln \frac{dP_{x,y|\theta}(x,y|\theta)}{dP_{x|\theta}(x|\theta) \times dP_{y|\theta}(y|\theta)} .dP_{x,y,\theta}(x,y,\theta) \quad (10)$$

where $dP_{x,y,\theta}(x,y,\theta) = dP_{x,y|\theta}(x,y|\theta)dP_\theta(\theta)$, we obtain (9).

Proof: 2) Using the methodology of [15], applied to mutual information, we deduce (8).

Next, the definition of the channel capacity for a Gaussian flat fading channel, when the CSI is fully known is given. Thereafter, an upper bound on the mutual information is introduced, and subsequently, it is shown that this upper bound is the channel capacity. Note that conditional mutual information $I_T(x;y|\theta)$ defined previously is different from conditional expectation [9].

Definition 4.2: Consider the model given by (1) and (2), when the fading process θ is completely known to the transmitter and receiver, subject to the instantaneous power constraint $E[|F(t,x,\tilde{x},\theta)|^2|\mathcal{F}^\theta_{0,t}] \leq P, \forall t \geq 0$, where P is fixed and positive. Then the finite-time and infinite-time channel capacity are defined by $C_f^T \stackrel{\triangle}{=} \sup_{(x,F)\in\mathcal{X}\times\mathcal{F}_{\rm ad}}(\frac{1}{T})I_T(x;y|\theta)$ and $C_f \stackrel{\triangle}{=} \liminf_{T\to\infty} C_f^T$, respectively. Here, the supremum is taken over all state processes $x \in \mathcal{X}$ that are solutions to (1) and over all encoding functions $F \in \mathcal{F}_{\rm ad}$ that satisfy the power constraint [9], [11], [16].

Since the encoder and decoder have access to the channel state information θ, then $\sup_{(x,F)\in\mathcal{X}\times\mathcal{F}_{\rm ad}}(1/T)I_T(x;y|\theta) = (1/T)\int_\theta \sup_{(x,F)\in\mathcal{X}\times\mathcal{F}_{\rm ad}} \tilde{I}_T(\theta)dP_\theta(\theta)$, and similarly for C_f. Hence, C_f^T and C_f are obtained via an average with respect to P_θ of the supremum over $(x,F) \in \mathcal{X}\times\mathcal{F}_{\rm ad}$ of $\tilde{I}_T(\theta)$. The selection of the instantaneous power constraint reflects the constraint on the channel input signal expressed in terms of the square of the amplitude. One may consider alternative power constraints to reflect a constraint on the average power over a time interval [0,T] per unit time, or even a constraint on the receiver. However, care must be taken to incorporate the alternative constraint in the main results.

Lemma 4.3: Consider the model given by (1) and (2) subject to instantaneous power constraint. Then, $(1/T)I_T(x;y|\theta) \leq \frac{1}{2T}P\int_0^T E[z^2(t,\theta(t))]dt$. Moreover, the infinite-time channel capacity is given by $C_f = \liminf_{T\to\infty} \frac{P}{2T}\int_0^T E[z^2(t,\theta(t))]dt$.

Proof: According to (8) and by considering the power constraint, we have $I_T(x;y|\theta) \leq (1/2)E_\theta \int_0^T z^2(t,\theta(t))$ $E[|F(t,x,\tilde{x},\theta)|^2|\mathcal{F}^\theta_{0,t}]dt \leq P/2\int_0^T E[z^2(t,\theta(t))]dt$. Following the methodology in [11, Sec. 16.4], it is shown that the aforementioned upper bound determines the channel capacity, that is, a white Gaussian noise signal x achieves the capacity.

2) Optimal Encoding and Decoding: In this section, we design an optimal (in mean square sense) encoding/decoding strategy that achieves the channel capacity C_f^T and C_f. We then employ the expression for the minimum mean square decoding error to obtain necessary and sufficient conditions for bounded asymptotic and asymptotic observability. In the subsequent development, only linear encoders are considered, because along the same lines of [11, Sec. 16.4], it can be shown that linear encoders achieve the channel capacity and the minimum mean square decoding error.

Definition 4.4: The set of linear admissible encoders $\mathcal{L}_{\rm ad}$, where $\mathcal{L}_{\rm ad} \subset \mathcal{F}_{\rm ad}$, is the set of linear nonanticipative functionals F with respect to $(x,\tilde{x},\theta)$, which have the following form $F(t,x,\tilde{x},\theta) = F_0(t,\tilde{x},\theta) + F_1(t,\tilde{x},\theta)x(t)$.

Using linear encoders, the received signal y is given by $dy(t) = z(t,\theta(t))[F_0(t,\tilde{x},\theta) + F_1(t,\tilde{x},\theta)x(t)]dt + dv(t),\ y(0) = 0.$

Decoding: Because the decoded signal $\tilde{x}$ is a function of the received signal y and the channel θ, the optimal decoder minimizing the mean square decoding error is $\tilde{x}_{\rm opt}(t,y,\theta) = \hat{x}(t,y,\theta) = E[x(t)|\mathcal{F}^{y,\theta}_{0,t}]$, which is the conditional mean. The conditional error variance for the decoder is $V(t,y,\theta) = E[(x(t) - \hat{x}(t,y,\theta))^2|\mathcal{F}^{y,\theta}_{0,t}]$. Moreover, they satisfy the following generalized Kalman filtering equations [11]: $d\hat{x}(t,y,\theta) = A(t)\hat{x}(t,y,\theta)dt + B(t)u(t)dt + z(t,\theta(t))V(t,y,\theta)F_1(t,\hat{x},\theta)[dy(t) - z(t,\theta(t))(F_0(t,\hat{x},\theta) + F_1(t,\hat{x},\theta)\hat{x}(t,y,\theta))dt]$, $\hat{x}(0) = \bar{x}_0$ and $\dot{V}(t,y,\theta) = 2A(t)V(t,y,\theta) - z^2(t,\theta(t))F_1^2(t,\hat{x},\theta)V^2(t,y,\theta) + G^2(t)$, $V(0) = \bar{V}_0$.

Encoding: From the point of view of the coding theorem, an encoder is optimal if it operates near the channel capacity, while ensuring a decoding error that tends to zero exponentially fast, as the codeword length tends to infinity. In our case, by choosing (F_0, F_1) appropriately, the conditional error variance is minimized, and the channel capacity C_f^T is achieved. The optimal encoder and decoder as well as the corresponding conditional error variance are given by the following theorem. The methodology is similar to one found in [11, Sec. 16.4], except that the channel state information θ has to be taken into account, by working with the conditional mutual information instead of the unconditional.

Theorem 4.5 (Coding theorem): Suppose the received signal is defined by (2) and the source by (1). Then, the encoder, which achieves the finite-time channel capacity C_f^T, the optimal decoder, and the corresponding error variance are, respectively, given by

$$F^*(t,x,\hat{x}^*,\theta) = \sqrt{\frac{P}{V^*(t,y,\theta)}}(x(t) - \hat{x}^*(t,y,\theta)) \quad (11)$$

$$d\hat{x}^*(t,y,\theta) = A(t)\hat{x}^*(t,y,\theta)dt + B(t)u(t)dt + z(t,\theta(t))\sqrt{PV^*(t,y,\theta)}dy(t), \quad \hat{x}^*(0) = \bar{x}_0 \quad (12)$$

$$V^*(t,y,\theta) = V^*(0)\exp\left\{2\int_0^t A(s)ds - \int_0^t z^2(s,\theta(s))Pds\right\} + \int_0^t G^2(s)\exp\left\{2\int_s^t A(u)du - \int_s^t z^2(u,\theta(u))Pdu\right\}ds, \quad V^*(0) = \bar{V}_0. \quad (13)$$

Proof: Substituting (11)–(13) into the conditional mutual information of Theorem 4.1, we deduce that for finite time, the upper bound of Lemma 4.3 is achieved (the proof is a variant of the one given in [11, Sec. 16.4]).

From (13), it follows that by employing the proposed optimal encoding/decoding scheme the mean square estimation error $V^*(t,y,\theta)$ is independent of a control signal. This suggests that the encoder and decoder can be designed independent of the controller; thus, a separation principle holds between the control and the communication designs. Optimality of the separated design will be shown in the last section.

Theorem 4.6: 1) When $G(t)\neq 0$, a sufficient condition for bounded asymptotic observability in the form of: $\limsup_{t\to\infty}\mathcal{E}(t)<\infty$ is $\inf_{t\in[0,\infty)}(Pz^2(t,\theta(t))-2[A(t)]^+)>0$, P-a.s., where $[a]^+=a$ if $a\geq 0$ and $[a]^+=0$ otherwise, and $C_f=\liminf_{T\to\infty}\frac{P}{2T}\int_0^T E[z^2(t,\theta(t))]dt$.

2) When $G(t)=0$, the aforementioned condition is also a sufficient condition for asymptotic observability in the mean square sense [i.e., $\lim_{t\to\infty}\mathcal{E}(t)=0$].

3) For the case of a time-invariant analog system and when $G\neq 0$ and $z=1$, a sufficient condition for bounded asymptotic observability [i.e., $\lim_{t\to\infty}\mathcal{E}(t)<\infty$] is given by $C=C_f=(P/2)>[A]^+$. Moreover, a necessary condition for such observability is given by $C\geq [A]^+$. Furthermore, the aforementioned conditions are also sufficient and necessary conditions (if $\bar{V}_0\neq 0$) for asymptotic observability [i.e., $\lim_{t\to\infty}\mathcal{E}(t)=0$], respectively, for the case of $G=0$.

Proof (*sufficient part*): 1) Suppose $\inf_{t\in[0,\infty)}(Pz^2(t,\theta(t))-2[A(t)]^+)>0$, P-a.s. holds. Then, from (13), it follows that by using the optimal encoder and decoder of Theorem 4.5, $\limsup_{t\to\infty}V^*(t,y,\theta)<\infty$ P-a.s. 2) This follows along the same lines of 1). 3) Follows easily from (13).

Necessary part: Consider the case of $G\neq 0$. If condition $C=P/2\geq [A]^+$ is not satisfied, then $A>P/2$ and consequently, $V^*(t,y,\theta)\to\infty$ as $t\to\infty$, P-a.s. Therefore, since among all admissible encoding/decoding schemes (including nonlinear ones), the proposed encoding/decoding scheme is optimal [i.e., $\mathcal{E}(t)\geq V^*(t,y,\theta)$, P-a.s.], and the mean square estimation error $\mathcal{E}(t)$ associated with all other admissible encoding schemes is going to be unbounded asymptotically, P-a.s. This implies that condition $C\geq [A]^+$ is a necessary condition for bounded asymptotic observability. The result for the case of $G=0$ follows similarly.

Remark 4.7: When $G(t)=0$, $\limsup_{t\to\infty}(1/t)\int_0^t A(s)ds$ is bounded, and the channel is the continuous-time AWGN channel ($z=1$), for which the channel capacity is $C=P/2$, it is easily shown that another sufficient condition for asymptotic observability is $C=P/2>\limsup_{t\to\infty}(1/t)\int_0^t A(s)ds$. Moreover, a necessary condition for asymptotic observability is $C\geq\liminf_{t\to\infty}1/t\int_0^t A(s)ds$.

By looking at the transmitted signal $F^*(t,x,\hat{x}^*,\theta)=\sqrt{\frac{P}{V^*(t,y,\theta)}}(x(t)-\hat{x}^*(t,y,\theta))$, it can be concluded that the encoding scheme is equivalent to a controller that uses the weighted error between the transmitted signal and its estimate. This kind of solution in which only the error is transmitted resembles the usual tracking strategy of feedback control theory. The goal of the output of the decoder $\hat{x}$ is to track the information signal.

B. Optimal Encoding/Decoding Scheme for Observability: The Vector Case

In this section, we extend the previous results to the vector case. From the classical control literature (e.g., [17]), we know that if $A(t)$ is time invariant [i.e., $A(t)=A$], it is diagonalizable, if there exists a similarity transformation S, such that $SAS^{-1}=\Lambda=\mathrm{diag}(\lambda_1(A),\ldots,\lambda_n(A))$, where $\lambda_i(A)$'s are eigenvalues of A. Next, we assume A is diagonalizable and $\lambda_i(A)$'s are real numbers (e.g., A is a symmetric matrix), and we apply such a similarity transformation $\gamma(t)\stackrel{\triangle}{=}Sx(t)$ to transform system (1) into the following system

$$d\gamma(t)=\Lambda\gamma(t)dt+SB(t)u(t)dt+SG(t)dw(t) \tag{14}$$

where $\gamma(0)=S\bar{x}_0$ and

$$\Lambda=\begin{pmatrix}\Lambda_s & 0\\ 0 & \Lambda_{\mathrm{us}}\end{pmatrix}$$

in which Λ_s block corresponds to the stable subspace and Λ_{us} block corresponds to the unstable subspace. From the previous part, we noticed that the stable eigenvalues do not contribute to the capacity requirement for observability. Thus, for the transformed system (14), without loss of generality, we can restrict our attention to A matrices that contains only unstable eigenvalues (e.g., A is positive semi definite). A similar idea for discrete-time systems is used in [4] and [5].

In order to extend the previous result, the following assumptions are introduced.

Assumption 4.8: In (14), it is assumed that $SG(t)G^{\mathrm{tr}}(t)S^{\mathrm{tr}}$ and $S\bar{V}_0S^{\mathrm{tr}}$ are diagonal matrices.

Notice that Assumptions 4.8 are satisfied if $G(t)G^{\mathrm{tr}}(t)=I$ (e.g., G is orthogonal), A is a symmetric system matrix (note that for A symmetric, $S^{\mathrm{tr}}=S^{-1}$), and the initial condition $x(0)$ has a corresponding covariance matrix of the form $\bar{V}_0=\alpha.I,\alpha\geq 0$. Under Assumptions 4.8, it can be shown that the optimal mean square decoding error, obtained by transmitting $\gamma(t)$, is diagonal. Notice that $\gamma(t)^{\mathrm{tr}}\gamma(t)=x^{\mathrm{tr}}(t)S^{\mathrm{tr}}Sx(t)$ ($=x^{\mathrm{tr}}(t)x(t)$ for A symmetric). Thus, stabilizability of $\gamma(t)$ is equivalent to the stabilizability of $x(t)$ and vice versa. Further, observability of $\gamma(t)$ is equivalent to observability of $x(t)$, particularly for A symmetric. Therefore, without loss of generality, we can consider the transformed system (14) instead of system (1) in our analysis. By replacing $\gamma(t)$ with $x(t)$, the results obtained in Section IV-A1 do not change, and therefore, the finite-time capacity of the flat fading AWGN channel is given by $C_f^T=\frac{P}{2T}\int_0^T E[z^2(t,\theta(t))]dt$.

By defining $F_1(t,\tilde{\gamma},\theta)\stackrel{\triangle}{=}[f_{11}(t,\tilde{\gamma},\theta),\ldots,f_{nn}(t,\tilde{\gamma},\theta)]$, the received signal is

$$dy(t)=z(t,\theta(t))[F_0(t,\tilde{\gamma},\theta)+F_1(t,\tilde{\gamma},\theta)\gamma(t)]dt+dv(t) \tag{15}$$

$y(0)=0$. The optimal mean square error decoder is $\tilde{\gamma}_{\mathrm{opt}}(t,y,\theta)=\hat{\gamma}(t,y,\theta)=E[\gamma(t)|\mathcal{F}_{0,t}^{y,\theta}]$ with error covariance $V(t,y,\theta)=E[(\gamma(t)-\hat{\gamma}(t,y,\theta))(\gamma(t)-\hat{\gamma}(t,y,\theta))^{\mathrm{tr}}|\mathcal{F}_{0,t}^{y,\theta}]$. Moreover [11], the decoder $\hat{\gamma}(t,y,\theta)$ and the corresponding error covariance $V(t,y,\theta)$ satisfy the following generalized Kalman filter equation:

$$\begin{aligned}d\hat{\gamma}(t,y,\theta)=&\,\Lambda\hat{\gamma}(t,y,\theta)dt+SB(t)u(t)dt\\&+V(t,y,\theta)F_1^{\mathrm{tr}}(t,\hat{\gamma},\theta)z(t,\theta(t))[dy(t)\\&-z(t,\theta(t))(F_0(t,\hat{\gamma},\theta)\\&+F_1(t,\hat{\gamma},\theta)\hat{\gamma}(t,y,\theta))dt],\quad \hat{\gamma}(0)=S\bar{x}_0\end{aligned} \tag{16}$$

$$\begin{aligned}\dot{V}(t,y,\theta)=&\,2\Lambda V(t,y,\theta)-z^2(t,\theta(t))V(t,y,\theta)\\&.F_1^{\mathrm{tr}}(t,\hat{\gamma},\theta)F_1(t,\hat{\gamma},\theta)V(t,y,\theta)+SG(t)G^{\mathrm{tr}}(t)S^{\mathrm{tr}},\\V(0)=&\,S\bar{V}_0S^{\mathrm{tr}}.\end{aligned} \tag{17}$$

Applying Assumptions 4.8, $V(t,y,\theta)$ is diagonal, e.g., $V(t,y,\theta)=\mathrm{diag}(V_{11}(t,y,\theta),\ldots,V_{nn}(t,y,\theta))$, and the ith diagonal element of (17) is given by $\dot{V}_{ii}(t,y,\theta)=2\lambda_i(A)V_{ii}(t,y,\theta)-z^2(t,\theta(t))V_{ii}^2(t,y,\theta)f_{ii}^2(t,\tilde{\gamma},\theta)+[SG(t)G^{\mathrm{tr}}(t)S^{\mathrm{tr}}]_{ii}$, where $[SG(t)G^{\mathrm{tr}}(t)S^{\mathrm{tr}}]_{ii}$ is the ith diagonal element of $SG(t)G(t)^{\mathrm{tr}}S^{\mathrm{tr}}$. Consequently, following the same methodology used to prove Theorem 4.5, we have the following theorem.

Theorem 4.9: Suppose Assumptions 4.8 hold, the received signal is defined by (15) and the source by (14). Then, the encoder, which achieves the finite-time channel capacity $C_f^T = \frac{P}{2T}\int_0^T E[z^2(t,\theta(t))]dt$, the optimal decoder, and the corresponding error covariance are given by $F^*(t,\gamma,\hat{\gamma}^*,\theta) = F_0^*(t,\hat{\gamma}^*,\theta) + \sum_{i=1}^n f_{ii}^*(t,\hat{\gamma}^*,\theta)\gamma_i(t), F_0^*(t,\hat{\gamma}^*,\theta) = -\sum_{i=0}^n f_{ii}^*(t,\hat{\gamma}^*,\theta)\hat{\gamma}_i^*(t,y,\theta)$, $f_{ii}^*(t,\hat{\gamma}^*,\theta) = \sqrt{\frac{\alpha_i P}{V_{ii}^*(t,y,\theta)}}, d\hat{\gamma}^*(t,y,\theta) = \Lambda\hat{\gamma}^*(t,y,\theta)dt + SB(t)u(t)dt + z(t,\theta(t))[\sqrt{\alpha_1 PV_{11}^*(t,y,\theta)},\ldots,\sqrt{\alpha_n PV_{nn}^*(t,y,\theta)}]^{\mathrm{tr}}dy(t)$,

$$V_{ii}^*(t,y,\theta) = [S\bar{V}_0S^{\mathrm{tr}}]_{ii}\exp\left\{2\int_0^t \lambda_i(A)ds - \int_0^t \alpha_i z^2(s,\theta(s))Pds\right\} + \int_0^t [SG(s)G^{\mathrm{tr}}(s)S^{\mathrm{tr}}]_{ii} \cdot \exp\left\{2\int_s^t \lambda_i(A)du - \int_s^t \alpha_i z^2(u,\theta(u))Pdu\right\}ds \tag{18}$$

where $\hat{\gamma}_i^*(t,y,\theta)$ is the ith element of $\hat{\gamma}^*(t,y,\theta)$ and $0 \le \alpha_i \le 1$, $\sum_{i=1}^n \alpha_i = 1$.

Next, we have the following theorem that extends Theorem 4.6 to the vector case.

Theorem 4.10: Suppose Assumptions 4.8 hold and the encoder/decoder of Theorem 4.10 is employed.

1) If there exists a set of $\{\alpha_i\}_{i=1}^n$ such that $0 \le \alpha_i \le 1, \sum_{i=1}^n \alpha_i = 1$ in which $\forall i = 1,2,\ldots,n,\ \ P-a.s.$,

$$\inf_{t\in[0,\infty)} (\alpha_i Pz^2(t,\theta(t)) - 2[\mathrm{Re}(\lambda_i(A))]^+) > 0 \tag{19}$$

then, for the case of $G(t) \neq 0$, we have bounded asymptotic observability in the sense $\limsup_{t\to\infty} E[(\gamma(t)-\tilde{\gamma}(t))^{\mathrm{tr}}(\gamma(t)-\tilde{\gamma}(t))|\mathcal{F}_{0,t}^{y,\theta}] < \infty$. Moreover, for the case of $G(t) = 0$, we have asymptotic observability in the mean square sense [i.e., $\lim_{t\to\infty} E[(\gamma(t)-\tilde{\gamma}(t))^{\mathrm{tr}}(\gamma(t)-\tilde{\gamma}(t))|\mathcal{F}_{0,t}^{y,\theta}] = 0$].

2) For the time-invariant analog of (14) and when $G \neq 0$ and $z = 1$, a sufficient condition for bounded asymptotic observability [i.e., $\lim_{t\to\infty} E[(\gamma(t)-\tilde{\gamma}(t))^{\mathrm{tr}}(\gamma(t)-\tilde{\gamma}(t))|\mathcal{F}_{0,t}^{y,\theta}] < \infty$] is given by

$$C = C_f = \frac{P}{2} > \sum_{\{i;\mathrm{Re}(\lambda_i(A))\ge 0\}} \mathrm{Re}(\lambda_i(A)). \tag{20}$$

Moreover, a necessary condition for such observability is given by

$$C = C_f = \frac{P}{2} \ge \sum_{\{i;\mathrm{Re}(\lambda_i(A))\ge 0\}} \mathrm{Re}(\lambda_i(A)). \tag{21}$$

Furthermore, the aforementioned conditions are also sufficient and necessary conditions for asymptotic observability [i.e., $\lim_{t\to\infty} E[(\gamma(t)-\tilde{\gamma}(t))^{\mathrm{tr}}(\gamma(t)-\tilde{\gamma}(t))|\mathcal{F}_{0,t}^{y,\theta}] = 0$], respectively, for $G = 0$.

Proof (sufficient part): 1) Follows from (18).

2) If condition (20) holds, there exists a set $\{\alpha_i\}_{i=1}^n$ such that $0 \le \alpha_i \le 1$ and $\sum_{i=1}^n \alpha_i = 1$ in which $\alpha_i P/2 > [\mathrm{Re}(\lambda_i(A))]^+$. Subsequently, from (18), the result is obtained.

Necessary part: 2) Consider the case of $G \neq 0$. If condition (21) is not satisfied, then for each set $\{\alpha_i\}_{i=1}^n$ such that $0 \le \alpha_i \le 1$, and $\sum_{i=1}^n \alpha_i = 1$, there exists one element $\alpha_j \in \{\alpha_i\}_{i=1}^n$ such that $\alpha_j P/2 < [\mathrm{Re}(\lambda_j(A))]^+$. This implies that $V_{jj}^*(t,y,\theta) \to \infty$ as $t \to \infty$. Subsequently, there is no other encoding scheme with asymptotic bounded mean square estimation error. The result for the case of $G = 0$ follows similarly.

V. Optimal Controller, Sufficient Condition for Stabilizability

In this section, we propose a state feedback controller that minimizes a quadratic payoff while stabilizing the time-invariant analog of system (1) [e.g., when $A(t) = A, B(t) = B, G(t) = G$, for all t]. Similarly to the previous section, we assume that there exists a similarity transformation S under which $SAS^{-1} = \Lambda$, in which we can consider the time-invariant analog of transformed system (14). Since, it is assumed that the channel state information is known, for a fixed channel sample path θ, the state feedback controller is chosen to minimize the following quadratic payoff

$$J^T = \frac{1}{T}E\left\{\int_0^T [\gamma^{\mathrm{tr}}(t)\gamma(t) + u^{\mathrm{tr}}(t)Ru(t)]dt\right\} \tag{22}$$

where $R > 0$ is symmetric weighting matrix. Moreover, we also consider the infinite horizon $\bar{J} = \lim_{T\to\infty} J^T$. For the infinite horizon, we assume the controllability rank condition $\mathrm{Rank}(\mathcal{C}) = n, \mathcal{C} \stackrel{\triangle}{=} (B\ \ AB\ \cdots\ A^{n-1}B)$.

According to the classical separation theorem of estimation and control (see [17, pp. 389–395] and [11]), the optimal controller that minimizes (22) subject to a flat fading AWGN communication channel and the linear encoder $F(t,\gamma,\tilde{\gamma},\theta) = F_0(t,\tilde{\gamma},\theta) + F_1(t,\tilde{\gamma},\theta)\gamma(t)$ is separated into a state estimator and a certainty equivalent controller given by

$$u^*(t) = -K(t)\hat{\gamma}(t,y,\theta), \qquad K(t) = R^{-1}B^{\mathrm{tr}}P(t) \tag{23}$$

where the state estimator $\hat{\gamma}(t,y,\theta)$ is the solution of (16) with the corresponding observer Riccati equation (17), and $P(t)$ is the solution of the following regulator Riccati equation

$$-\dot{P}(t) = I - P(t)SBR^{-1}B^{\mathrm{tr}}S^{\mathrm{tr}}P(t) + 2\Lambda P(t) \tag{24}$$

where $P(T) = 0$. For a fixed sample path of the channel, it follows that if the observer Riccati equation has a steady-state solution $\bar{V}$ and the regulator Riccati equation has a steady-state solution $\bar{P}$, then the averaged criterion $\bar{J} = \lim_{T\to\infty}\frac{1}{T}E\{\int_0^T[\gamma^{\mathrm{tr}}(t)\gamma(t) + u^{*,\mathrm{tr}}(t)Ru^*(t)]dt\}$ can be expressed in the alternative form [17, pp. 395]

$$\bar{J} = \lim_{T\to\infty}\frac{1}{T}E\left\{\int_0^T [\gamma^{\mathrm{tr}}(t)\gamma(t) + u^{*,\mathrm{tr}}(t)Ru^*(t)]dt\right\} = Trac[\bar{P}SGG^{\mathrm{tr}}S^{\mathrm{tr}} + \bar{V}\bar{K}^{\mathrm{tr}}R\bar{K}] \tag{25}$$

where $\bar{K} = R^{-1}B^{\mathrm{tr}}\bar{P}$. From [17, Th. 3.5, p. 231], it follows that if the controllability rank condition holds, the Riccati equation (24) has steady-state solution $\lim_{t\to\infty} P(t) = \bar{P}$. Under Assumptions 4.8, and assuming the optimal encoding/decoding scheme of Theorem 4.9 is used, then the observer Riccati equation (17) is reduced to the linear time-varying equation $V(t,y,\theta) = V^*(t,y,\theta) \stackrel{\triangle}{=} \mathrm{diag}(V_{11}^*(t,y,\theta),\ldots,V_{nn}^*(t,y,\theta))$, $\dot{V}_{ii}^*(t,y,\theta) = \tilde{A}_i(t,\theta(t))V_{ii}^*(t,y,\theta) + [SGG^{\mathrm{tr}}S^{\mathrm{tr}}]_{ii}, V_{ii}^*(0,y,\theta) = [S\bar{V}_0S^{\mathrm{tr}}]_{ii}, i = 1,\ldots,n, \tilde{A}_i(t,\theta(t)) \stackrel{\triangle}{=} 2\lambda_i(A) - \alpha_i z^2(t,\theta(t))P$, which shows the solution is independent of the sample path y. Moreover, if in addition, $|\tilde{A}_i(t,\theta)| \le k, k > 0, 1 \le i \le n$ for all $\theta \in \Re^q, 0 \le t \le T$, the channel fading process converges to a random variable for large t, which implies the following limit exists, $\lim_{t\to\infty} z^2(t,\theta(t)) = z_\infty^2$, P-a.s., and $\tilde{A}_{i,\infty} \stackrel{\triangle}{=} 2\lambda_i(A) - \alpha_i z_\infty^2 P < 0$, P-a.s., then $\lim_{t\to\infty} V(t,y,\theta) = \bar{V}$, P-a.s., where $\bar{V} = \mathrm{diag}(\bar{V}_{11},\ldots,\bar{V}_{nn})$, $\bar{V}_{ii} = [SGG^{\mathrm{tr}}S^{\mathrm{tr}}]_{ii}/(\alpha_i z_\infty^2 P - 2\lambda_i(A)), i = 1,\ldots,n$. Such a fading process is encountered in practice when the Doppler spread of the channel is zero (e.g., there is no relative motion between the transmitter and the receiver [9]). Notice that the statements of Theorems 4.10 and 4.9 do not assume a random variable fading process. Clearly, when $G = 0$, then $\bar{V} = 0$, P-a.s.

Proposition 5.1: Consider the time-invariant analog of system (14) and assume $\mathrm{Rank}(\mathcal{C}) = n$.

Then, for a fixed sample path of the channel, we have the following (P-a.s.).

1) Assuming $G \neq 0$ and $V(t, y, \theta) \to \bar{V}$, as $t \to \infty$, by using the optimal policy (23), we have $E||\gamma(t)||^2 < \infty$ and $E||u(t)||_R^2 < \infty$, as $t \to \infty$, where $||u(t)||_R^2 \stackrel{\triangle}{=} u^{\mathrm{tr}}(t) R u(t)$.

2) Assuming $G = 0$ and $V(t, y, \theta) \to 0$, as $t \to \infty$, by using the optimal policy (23), we have $E||\gamma(t)||^2 \to 0$ and $E||u(t)||_R^2 \to 0$, as $t \to \infty$.

Proof: This follows from (25).

Next, by using Theorem 4.10 and Proposition 5.1, the following theorem for bounded asymptotic and asymptotic stabilizability in the mean square sense is derived.

Theorem 5.2: Consider the time-invariant analog of the system (14), assume $\mathrm{Rank}(\mathcal{C}) = n$ and Assumptions 4.8 hold.

Then, the following cases occur.

1) For the case when $G \neq 0$ and $z = 1$, a sufficient condition for bounded asymptotic stabilizability in the mean square sense is given by

$$C = \frac{P}{2} > \sum_{\{i; \mathrm{Re}(\lambda_i(A)) \geq 0\}} \mathrm{Re}(\lambda_i(A)). \tag{26}$$

2) For the case when $G = 0$ and $z = 1$, (26) is also a sufficient condition for asymptotic stabilizability in the mean square sense.

3) For the case when $G = 0$ and in the presence of fading (i.e., $z \neq 1$), (19) is a sufficient condition for asymptotic stabilizability in the mean square sense.

Proof: 1) Suppose the condition (26) is satisfied. Then, by using the optimal encoding/decoding scheme proposed in Theorem 4.9, $\lim_{t\to\infty} V(t, y, \theta) = \bar{V}$. Consequently, from Proposition 5.1, it follows that by using the optimal control policy (23) and the optimal encoding/decoding scheme proposed in Theorem 4.9, the stabilizability condition is guaranteed. 2) It is shown along the same lines of 1). 3) It is shown along the same lines of 1).

Remark 5.3: As it was shown in this section, a separation principle holds between the design of the communication and the control subsystems. The efficient encoding/decoding scheme that minimizes the mean square estimation error and achieves the channel capacity is given in Section IV, while the optimal certainty equivalent controller that optimizes a quadratic cost functional is given by (23). Although we designed an optimal encoder/decoder pair and controller separately, the whole system is optimal since the separation principle holds and the communication system sends information at capacity.

REFERENCES

[1] A. V. Savkin and I. R. Petersen, "Set-valued state estimation via a limited capacity communication channel," *IEEE Trans. Autom. Control*, vol. 48, no. 4, pp. 676–680, Apr. 2003.

[2] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM J. Control Optim.*, vol. 43, no. 2, pp. 413–436, 2004.

[3] G. N. Nair, R. J. Evans, I. M. Y. Mareels, and W. Moran, "Topological feedback entropy and nonlinear stabilization," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1585–1597, Sep. 2004.

[4] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1196–1201, Jul. 2004.

[5] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.

[6] S. Tatikonda, A. Sahai, and S. Mitter, "Stochastic linear control over a communication channel," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1549–1561, Sep. 2004.

[7] Nicola Elia, "When Bode meets Shannon: Control-oriented feedback communication schemes," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1477–1488, Sep. 2004.

[8] K. Li and J. Baillieul, "Robust quantization for digital finite communication bandwidth (DFCB) control," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1573–1584, Sep. 2004.

[9] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[10] C. D. Charalambous and A. Farhadi, "LQG optimality and separation principle for general discrete time partially observed systems over finite capacity communication channels," *Automatica,* to be published.

[11] R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes-Applications I, II*. New York: Springer-Verlag, 1977, 1978.

[12] T. T. Kadota, M. Zakai, and J. Ziv, "Mutual information of the white Gaussian channel with and without feedback," *IEEE Trans. Inf. Theory*, vol. 17, no. 4, pp. 368–371, Nov. 1971.

[13] B. F. Wu and E. A. Jonckheere, "A simplified approach to Bode's theorem for continuous-time and discrete-time systems," *IEEE Trans. Autom. Control*, vol. 37, no. 11, pp. 1797–1802, Nov. 1992.

[14] C. D. Charalambous and Stojan Denic, "On the channel capacity of wireless fading channels," in *Proc. 41st IEEE Conf. Decis. Control*, Las Vegas, Nevada, Dec. 2002, pp. 4036–4041.

[15] T. Duncan, "On the calculation of mutual information," *SIAM J. Appl. Math*, vol. 19, pp. 215–220, Jul. 1970.

[16] S. Ihara, "Coding theorems for a continuous-time Gaussian channel with feedback," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 2041–2045, Nov. 1994.

[17] H. Kwakernnak and R. Sivan, *Linear Optimal Control System*. New York: Wiley, 1972.

Adaptive Control for the Systems Preceded by Hysteresis

Xinkai Chen, Chun-Yi Su, and Toshio Fukuda

***Abstract*—Hysteresis hinders the effectiveness of smart materials in sensors and actuators. It is a challenging task to control the systems with hysteresis. This note discusses the adaptive control for discrete time linear dynamical systems preceded with hysteresis described by the Prandtl–Ishlinskii model. The time delay and the order of the linear dynamical system are assumed to be known. The contribution of the note is the fusion of the hysteresis model with adaptive control techniques without constructing the inverse hysteresis nonlinearity. Only the parameters (which are generated from the parameters of the linear system and the density function of the hysteresis) directly needed in the formulation of the controller are adaptively estimated online. The proposed control law ensures the global stability of the closed-loop system, and the output tracking error can be controlled to be as small as required by choosing the design parameters. Simulation results show the effectiveness of the proposed algorithm.**

***Index Terms*—Adaptive control, discrete time linear systems, hysteresis, Prandtl-Ishlinskii model.**

Manuscript received February 23, 2006; revised September 22, 2006 and March 23, 2007. Recommended by Associate Editor D. Dochain. The work of X. Chen was supported by the Research Foundation for the Electrotechnology of Chubu (Japan).

X. Chen is with the Department of Electronic and Information Systems, Shibaura Institute of Technology, Saitama 337-8570, Japan (e-mail: chen@sic.shibaura-it.ac.jp).

C.-Y. Su is with the Department of Mechanical and Industrial Engineering, Concordia University, Montreal, QC H3G 1M8, Canada (e-mail: cysu@alcor.concordia.ca).

T. Fukuda is with the Department of Mechano-Informatics and Systems, Nagoya University, Nagoya 464-8603, Japan (e-mail: fukuda@mein.nagoya-u.ac.jp).

Digital Object Identifier 10.1109/TAC.2008.919551