# Tracking Nonlinear Noisy Dynamic Systems over Noisy Communication Channels

Alireza Farhadi and N. U. Ahmed

***Abstract*—This paper is concerned with tracking a vector of signal process generated by a family of distributed (geographically separated) nonlinear noisy dynamic subsystems over the binary symmetric channel. Nonlinear subsystems are subject to bounded external disturbances. Measurements are also subject to bounded noises. For this system and channel, subject to constraints on transmission rates, cross over probabilities and Lipschitz constants, a simple methodology is presented ensuring tracking with bounded mean absolute error.**

***Index Terms*—Tracking, the binary symmetric channel, the erasure channel, nonlinear systems.**

## I. Introduction

ONE of the emerging applications of sensor networks is in the tracking and automation of distributed systems. Some examples are civil infrastructure monitoring and control, intelligent solar farms, automated manufacturing of composite materials, etc. In these applications each sensor observes the state of a dynamic system under noisy environment and the observation signal $Y_t$ is then encoded, and transmitted to the fusion center where it is decoded all within one time step (i.e., the time duration between two successive time instants $t$ and $t+1$). Therefore, the commonly used techniques for data transmission are not suitable for tracking dynamical systems. In fact, these techniques are based on the separation principle and block encoding and decoding which result in a long decoding delay. This necessitates development of new techniques for real time communication of signals or messages produced by dynamic systems.

Dynamic systems can be viewed as continuous alphabet sources with memory. Consequently, many works in the literature are dedicated to the question of tracking and control over Additive White Gaussian Noise (AWGN) channel, which itself is naturally a continuous alphabet channel (e.g., [1] - [4]). However, addressing the question of tracking over digital links [5]-[17], which is naturally consistent with modern communication systems, is more interesting. Recently, there has been a significant progress in addressing this question only for linear dynamic systems over communication channels. Some references are [1]-[11]. A few of these publications (e.g., [12]-[17]) considered the problem of tracking nonlinear dynamic systems, in which the communication channel is a digital noiseless channel [12]-[17] and the dynamic system is noiseless [12]-[16].

Nonlinear dynamic systems subject to noise, and noisy communication channels are more realistic representations of complex network of dynamic and communication systems. Motivated by this, we study in this paper the problem of tracking nonlinear dynamic systems with distributed (geographically separated) nonlinear subsystems subject to both process noise and measurement noise, over the Binary Symmetric Channels (BSCs). A simple methodology for tracking this system is presented under the constraints on transmission rates, cross over probabilities, and Lipchitz constants. This gives satisfactory results on tracking with bounded mean absolute error. This paper complements the results of [1]-[11] by considering nonlinear dynamic systems rather than linear systems. It also complements the results of [12]-[17] by considering nonlinear noisy dynamic systems over the BSC.

The paper is organized as follows. In Section II, problem formulation is presented. In Section III, we find an equivalent erasure type communication channel for transmission via the BSC. Then we present a methodology for tracking a nonlinear system consisting of distributed subsystems. In Section IV, we present an example of a nonlinear noisy dynamic system; and we use computer simulations to demonstrate the promising performance of the proposed design technique. We conclude the paper in Section V.

## II. Problem Formulation

In this paper we are concerned with a cluster of $n$ distributed small sized sensor nodes. The $i$th node ($i \in \{1,2,...,n\}$) consists of sensors, a data processor, and a wireless communication unit which is equipped with a low capacity battery. Each sensor node uses a multimode transmission technique. The $i$th node uses low power transmission mode to transmit message bits and high power transmission mode to transmit flag bits. Multimode transmission technique is one of the adaptive communication methods and has been used in the literature [18]-[20]. In multimode transmission some of the system parameters such as rate or modulation are adapted to the changes of status of the transmitter, receiver, and transmission environment. Due to capacity limitation of the battery, most of the time, sensor nodes are tuned to transmit under low power transmission mode. The $i$th sensor node directly communicates with a fusion center located at a proper distance from all sensor nodes. For communication purposes, the modulator is the binary phase shift keying and the demodulator is the coherent maximum likelihood estimator. Since

Paper approved by S.-Y. Chung, the Editor for LDPC Coding and Information Theory of the IEEE Communications Society. Manuscript received August 18, 2009; revised January 14, 2010, July 17, 2010, and November 15, 2010.

This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) under grant no. A7109.

A. Farhadi is with INRIA, Grenoble, France (e-mail: afarhadi@alumni.uottawa.ca).

N. U. Ahmed is with the School of Information Technology and Engineering (SITE), University of Ottawa, Ontario, Canada (e-mail: ahmed@site.uottawa.ca).

Digital Object Identifier 10.1109/TCOMM.2011.020411.090488

the $i$th node sends message bits to the fusion center under low power transmission mode, the communication channel (for transmission) is modeled as the BSC. The BSC is a memoryless channel with the binary input $\{0,1\}$ and the binary output $\{0,1\}$ which flips the transmitted bit with the cross over probability $0 \le p_i \le \frac{1}{2}$. This channel is obtained by employing the binary phase shift keying modulator and coherent maximum likelihood demodulator. On the other hand, flag bits are noise free because they are transmitted under high power transmission mode. This noise free communication is equivalent to transmission of a sequence of bits consisting of the flag bit and other bits used for error correction, which are transmitted under the low power transmission mode over the BSC with a rate less than the channel capacity. Therefore, by Shannon's coding theorem [21] reliable communication is possible for this transmission.

The fusion center is equipped with a high capacity power supply. Therefore, the power of the transmitter of the fusion center can be chosen large enough so that the channel noise is negligible for transmission from the fusion center to the sensor nodes. Therefore, transmission from the fusion center to each of the sensor nodes can be considered noiseless.

Let $(\Omega, \mathcal{F}(\Omega), P)$ be a complete probability space and suppose all the random variables and random processes arising in our study of the network of systems and communication channels in this paper are based on this probability space. This system of multiple sensors defined on the probability space $(\Omega, \mathcal{F}(\Omega), P)$ is used to track a nonlinear dynamic system consisting of $n$ distributed coupled subsystems, as described below:

$$(s_i): \quad X_{t+1}^{(i)} = F_i(\beta_i^{(1)} X_t^{(1)}, ..., \beta_i^{(n)} X_t^{(n)}) + W_t^{(i)}, \quad X_0^{(i)} = \xi_i, \quad i = 1, 2, ..., n, \tag{1}$$

where $t \in \mathbf{N}_+ \triangleq \{0, 1, 2, ...\}$, $X_t^{(i)} \in \Re$ is the state of the $i$th subsystem, random variable $\xi_i \in \Re$ is the initial state, $W_t^{(i)} \in \Re$ is the process (system) noise of the $i$th subsystem, and $\beta_i^{(j)}$, $j \in \{1, 2, ..., n\}$ is either zero or one indicating the coupling from subsystem $s_j$ to subsystem $s_i$ (i.e., $\beta_i^{(j)} = 0$ means there is no coupling between the indicated nodes and therefore, the description of the nonlinear function $F_i(\cdot)$ does not include the corresponding variable $X_t^{(j)}$). Note that $\beta_i^{(i)}$ is always 1. Throughout this paper it is assumed that the system noise is essentially bounded, that is, $|W_t^{(i)}| \le \Omega_i$ ($\forall t \in \mathbf{N}_+$), P-a.s. It is also assumed that the probability measure of the initial state $X_0^{(i)}$ has a bounded support. That is, there exists a compact set $[-l_0^{(i)}, l_0^{(i)}] \subset \Re$ such that $P(X_0^{(i)} \in [-l_0^{(i)}, l_0^{(i)}]) = 1$. The nonlinear function $F_i(\cdot)$ is assumed to be Lipschitz continuous. That is, if the nonlinear function $F_i(\cdot)$ is a function of variables $\gamma_1$,...,$\gamma_m \in \Re$ ($m \le n$), there exists a positive scalar $K_i > 0$ such that for every $\gamma_1$,...,$\gamma_m \in \Re$ and $\eta_1$,...,$\eta_m \in \Re$, we have $|F_i(\gamma_1, ..., \gamma_m) - F_i(\eta_1, ..., \eta_m)| \le K_i\Big(|\gamma_1 - \eta_1| + ... + |\gamma_m - \eta_m|\Big)$.

The $i$th sensor node observes the state of the $i$th subsystem and provides a noisy output given by $Y_t^{(i)} = X_t^{(i)} + G_t^{(i)}$, where $G_t^{(i)} \in \Re$ is the measurement noise. It is assumed to be essentially bounded, i.e., $|G_t^{(i)}| \le \Upsilon_i$ ($\forall t \in \mathbf{N}_+$), P-a.s., and has zero mean, i.e., $E[G_t^{(i)}] = 0$, where $E[\cdot]$ denotes the expected value. The observation signal $Y_t^{(i)}$ is encoded and then transmitted via the BSC to the fusion center where it is reconstructed (decoded).

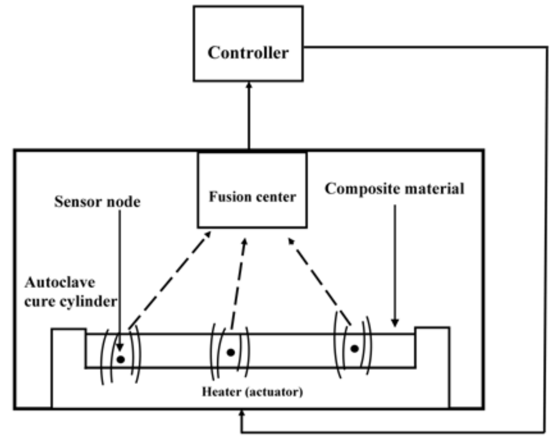


Fig. 1. Automated manufacturing of composite materials.

The objective of this paper is to find encoders and decoders to achieve mean absolute tracking for the nonlinear system (1) at the fusion center. This is formally defined as follows.

*Definition 2.1:* Consider the nonlinear system (1) with a cluster of sensor nodes, as described above. Let $\hat{X}_t^{(i)} \in \Re$ denote the reconstructed version of the state variable $X_t^{(i)} \in \Re$ at the fusion center at time $t$. Also, let $\mathcal{E}_t^{(i)} \triangleq |X_t^{(i)} - \hat{X}_t^{(i)}|$ denote the tracking error associated with the $i$th sensor node ($i \in \{1, 2, ..., n\}$). Then, we have mean absolute tracking if and only if for each $i \in \{1, 2, ..., n\}$ there exist an encoder, a decoder and a finite non-negative scalar $D_i$ such that $E[\mathcal{E}_t^{(i)}] \le D_i$, $\forall t \in \mathbf{N}_+$.

One potential application of the cluster of sensor nodes, as described above, is in the automated manufacturing of composite materials. One important step in manufacturing of composite materials is autoclave curing process, in which the blended materials are hardened inside an autoclave cure cylinder via heat and pressure. It is known that monitoring a composite's thermal or dielectric profile during autoclave curing process, and adjusting the heat or pressure accordingly, is important for having a high quality product. Traditionally, this is done by embedding a wired sensor network inside the composite material under cure. However, the wires brought out at the end of process is a source of damage to the material. To overcome this drawback, a wireless sensor network, as shown in Fig. 1, can be used.

## III. Tracking Nonlinear Systems

In this section, we address the tracking problem in the mean absolute sense, as described in Definition 2.1. Here, we first present an equivalent erasure type communication channel for transmission via the BSC. Then we present encoders and decoders for tracking the nonlinear dynamic system (1) over the equivalent erasure channels and therefore over the BSCs.

### A. Equivalent Communication Channel for Transmission via the BSC

Consider the $i$th sensor node. During each time step (i.e., the time period between two successive sampling times $t$ and $t+1$), the $i$th encoder encodes the observation signal $Y_t^{(i)}$ into $R_i$ message bits and transmits these $R_i$ bits one by one, under the low power transmission mode, via the BSC within the specific time period of $T$ seconds (which is less than the time step). After receiving each message bit the fusion center broadcasts the received bit via the digital noiseless feedback link to all nodes. Therefore all sensor nodes are aware of the received bit. If the bit is received correctly, then the $i$th encoder sends the next message bit. If the bit is flipped, then the $i$th transmitter is turned off for the rest of time period of $T$ seconds to conserve power. After the transmission time period of $T$ seconds, if all transmitted $R_i$ message bits are received correctly, then the $i$th encoder sends the flag bit "1", under the high power transmission mode, to inform the decoder that it can decode the transmitted message bits. If at least one of message bits is flipped, then the encoder sends the flag bit "0" under high power transmission mode. Recall that the flag bits are always received correctly. Similar to message bits, the fusion center also broadcasts the received flag bit to all nodes. At each time step, if the decoder receives the flag bit "1", it knows that the $R_i$ message bits were received correctly; and therefore, it decodes the transmitted bits. If it receives the flag bit "0", it disregards the transmitted message bits. Hence, the above protocol (consisting of transmission via the BSC and feedback acknowledgments) may be viewed as a transmission via a feedback erasure channel with rate $R_i$ and erasure probability $\alpha_i = 1-(1-p_i)^{R_i}$. This is a memoryless channel, which transmits $R_i$ bits in each channel use. The transmitted bits are received either correctly with probability $1-\alpha_i$ or erased with probability $\alpha_i$ [6]. In this case, the transmitter knows via the feedback acknowledgment whether the transmitted $R_i$ bits were received correctly or not.

Because each node communicates with the fusion center via the BSC, there is a possibility of collision of the received signals at the fusion center. In order to avoid such collision we need to employ an appropriate technique. Here, we use a Time Division Multiple Access (TDMA) scheme, in which each time step is divided into $n$ non-overlapping time slots of identical size and each slot is allocated to each sensor node. Each sensor node uses its allocated time slot to transmit its message bits and associated flag bit.

Having the above equivalent erasure model with rate $R_i$ for transmission via the BSC, we now present encoders and decoders for the mean absolute tracking.

### B. Encoders and Decoders for the Mean Absolute Tracking

In this section, we employ a differential coding technique for the mean absolute tracking of the nonlinear system (1). Differential coding techniques code the signal difference; and consequently, the amount of information to be sent is reduced. This type of coding has been used frequently in the literature, for instance in [6], [22]. In [6] using a differential quantization technique, tracking over the packet erasure channel is possible. Also using a differential coding scheme, reliable communication over AWGN channel is guaranteed as demonstrated in [22].

Now, consider the $i$th node ($i \in \{1,2,...,n\}$) in the cluster of sensor nodes, as described earlier. It is easy to verify that

$$|Y_0^{(i)}| = |X_0^{(i)} + G_0^{(i)}| \le l_0^{(i)} + \Upsilon_i \triangleq L_0^{(i)}.$$

At time $t=0$, the set $[-L_0^{(i)}, L_0^{(i)}]$ is partitioned into $2^{R_i}$ equal size, non-overlapping subintervals and the center of each subinterval is chosen as the index of that interval. Upon observing $Y_0^{(i)}$, the index of the subinterval which includes $Y_0^{(i)}$ is encoded into $R_i$ bits and transmitted via the channel. If erasure does not occur, the decoder can identify the index of the subinterval where $Y_0^{(i)}$ is located and the value of this index is chosen as $\hat{Y}_0^{(i)}$ which is the reconstructed (decoded) version of $Y_0^{(i)}$. Let $\hat{X}_0^{(i)}(=\hat{Y}_0^{(i)})$ denote the reconstructed (decoded) version of the initial state $X_0^{(i)}$. Then, the decoding error for this case is bounded above by

$$\begin{aligned}\mathcal{E}_0^{(i)} &= |X_0^{(i)} - \hat{X}_0^{(i)}| = |Y_0^{(i)} - G_0^{(i)} - \hat{Y}_0^{(i)}| \\ &\le |Y_0^{(i)} - \hat{Y}_0^{(i)}| + \Upsilon_i \le \frac{L_0^{(i)}}{2^{R_i}} + \Upsilon_i.\end{aligned}$$

If erasure occurs, then $\hat{Y}_0^{(i)} = \hat{X}_0^{(i)} = 0$; and therefore

$$\mathcal{E}_0^{(i)} = |X_0^{(i)} - \hat{X}_0^{(i)}| = |Y_0^{(i)} - G_0^{(i)} - \hat{Y}_0^{(i)}| \le L_0^{(i)} + \Upsilon_i.$$

Hence, we may write $\mathcal{E}_0^{(i)} \le V_0^{(i)} + \Upsilon_i$, where $V_0^{(i)} = L_0^{(i)} M_0^{(i)}$ and $M_0^{(i)}$ is a random variable satisfying $M_0^{(i)} = 1$ if erasure occurs, and $M_0^{(i)} = \frac{1}{2^{R_i}}$ if erasure does not occur.

At time instant $t=1$, using feedback acknowledgments from the fusion center, the $i$th sensor node can determine $\hat{X}_0^{(1)}(=\hat{Y}_0^{(1)})$, ..., $\hat{X}_0^{(n)}(=\hat{Y}_0^{(n)})$. Therefore, from Lipschitz continuity assumption, it follows that

$$\begin{aligned}&|Y_1^{(i)} - F_i(\beta_i^{(1)}\hat{X}_0^{(1)}, ..., \beta_i^{(n)}\hat{X}_0^{(n)})| \\ &= |F_i(\beta_i^{(1)}X_0^{(1)}, ..., \beta_i^{(n)}X_0^{(n)}) \\ &\quad - F_i(\beta_i^{(1)}\hat{X}_0^{(1)}, ..., \beta_i^{(n)}\hat{X}_0^{(n)}) + G_1^{(i)} + W_0^{(i)}| \\ &\le K_i \sum_{j=1}^{n} \beta_i^{(j)} |X_0^{(j)} - \hat{X}_0^{(j)}| + \Upsilon_i + \Omega_i \\ &\le K_i \sum_{j=1}^{n} \beta_i^{(j)} (V_0^{(j)} + \Upsilon_j) + \Upsilon_i + \Omega_i = L_1^{(i)}.\end{aligned}$$

Then, the interval $[-L_1^{(i)}, L_1^{(i)}]$ is partitioned into $2^{R_i}$ equal size, non-overlapping subintervals, and the center of each subinterval is chosen as the index of that interval. Upon observing $Y_1^{(i)}$, the index of the subinterval which includes $Y_1^{(i)} - F_i(\beta_i^{(1)}\hat{X}_0^{(1)}, ..., \beta_i^{(n)}\hat{X}_0^{(n)})$ is encoded into $R_i$ bits and transmitted to the fusion center. If erasure does not occur, then the decoder can identify the index of the subinterval which contains $Y_1^{(i)} - F_i(\beta_i^{(1)}\hat{X}_0^{(1)}, ..., \beta_i^{(n)}\hat{X}_0^{(n)})$ and the value of this index plus $F_i(\beta_i^{(1)}\hat{X}_0^{(1)}, ..., \beta_i^{(n)}\hat{X}_0^{(n)})$ is chosen as $\hat{Y}_1^{(i)}$, which is the reconstructed (decoded) version of $Y_1^{(i)}$. Let $\hat{X}_1^{(i)}(=\hat{Y}_1^{(i)})$ denote the reconstructed (decoded) version of $X_1^{(i)}$. Then, the decoding error for this case is bounded above by

$$\mathcal{E}_1^{(i)} = |X_1^{(i)} - \hat{X}_1^{(i)}| = |Y_1^{(i)} - G_1^{(i)} - \hat{Y}_1^{(i)}| \le \frac{L_1^{(i)}}{2^{R_i}} + \Upsilon_i.$$

If erasure occurs, then $\hat{Y}_1^{(i)} = \hat{X}_1^{(i)} = F_i(\beta_i^{(1)}\hat{X}_0^{(1)}, ..., \beta_i^{(n)}\hat{X}_0^{(n)})$; and therefore, $\mathcal{E}_1^{(i)} = |X_1^{(i)} - \hat{X}_1^{(i)}| \leq L_1^{(i)} + \Upsilon_i$. Hence, we may write $\mathcal{E}_1^{(i)} \leq V_1^{(i)} + \Upsilon_i$, where $V_1^{(i)} = L_1^{(i)} M_1^{(i)}$ and $M_1^{(i)}$ is a random variable satisfying $M_1^{(i)} = 1$ if erasure occurs, and $M_1^{(i)} = \frac{1}{2^{R_i}}$ if erasure does not occur.

Following the procedure, as described above, we construct the following sequence $\{\hat{X}_1^{(i)}, \hat{X}_2^{(i)}, \hat{X}_3^{(i)}, \cdots\}$.

In the following proposition we show that using the above methodology we have the mean absolute tracking.

*Proposition 3.1:* Consider the nonlinear system (1) and cluster of sensor nodes, as described earlier, over the BSCs. Using flag bits and feedback from the receiver, the transmission over the BSC is equivalent to the transmission over the erasure channel with feedback, as described in Section III-A. For this transmission, suppose that the rates $\{R_i\}_{i=1}^n$ are chosen such that the following matrix is stable (i.e., all its eigenvalues are in the open unit circle).

$$\underline{\mathcal{A}} = \begin{pmatrix} K_1\beta_1^{(1)}(\frac{1-\alpha_1}{2^{R_1}} + \alpha_1) & ... & K_1\beta_1^{(n)}(\frac{1-\alpha_1}{2^{R_1}} + \alpha_1) \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ K_n\beta_n^{(1)}(\frac{1-\alpha_n}{2^{R_n}} + \alpha_n) & ... & K_n\beta_n^{(n)}(\frac{1-\alpha_n}{2^{R_n}} + \alpha_n) \end{pmatrix}. \quad (2)$$

Then, there exist encoders and decoders that guarantee the mean absolute tracking in the sense of Definition 2.1 .

*Proof:* Choose any rates $\{R_i\}_{i=1}^n$ under which the matrix $\underline{\mathcal{A}}$ is stable. Recall that for the $i$th node ($i \in \{1, 2, ..., n\}$), the random variable $M_t^{(i)}$ satisfies: $P(M_t^{(i)} = \frac{1}{2^{R_i}}) = 1 - \alpha_i$ and $P(M_t^{(i)} = 1) = \alpha_i$. This random variable indicates the successful transmission (if $M_t^{(i)} = \frac{1}{2^{R_i}}$); or failed transmission (if $M_t^{(i)} = 1$) at time $t$. Therefore, it is independent of the other variables $M_{\tilde{t}}^{(i)}$ for $\tilde{t}(\in \mathbf{N}_+) \neq t$. It is also clear that the process $\{M_t^{(i)}\}_{t\in\mathbf{N}_+}$ is identically distributed. So, the random process $\{M_t^{(i)}\}_{t\in\mathbf{N}_+}$ is an i.i.d. process. It is also evident that the process $M_t^{(i)}$ and $M_t^{(j)}$ are mutually independent for any $j(\neq i) \in \{1, 2, ..., n\}$.

Using the methodology described in Section III-B, we have $\mathcal{E}_t^{(i)} = |X_t^{(i)} - \hat{X}_t^{(i)}| \leq V_t^{(i)} + \Upsilon_i$, P-a.s., where

$$V_t^{(i)} = K_i M_t^{(i)} \sum_{j=1}^{n} \beta_i^{(j)} (V_{t-1}^{(j)} + \Upsilon_j) + M_t^{(i)}(\Upsilon_i + \Omega_i),$$
$$V_0^{(i)} = L_0^{(i)} M_0^{(i)}, \ t = 1, 2, 3, ... .$$

Let $diag(\cdot)$ denote the diagonal matrix and $\prod_{n\geq j\geq 1} a_j \triangleq a_n a_{n-1}...a_1$ denote the ordered product. By a straightforward computation, one can easily verify that

$$\underline{V}_t = \underline{A}_t \underline{V}_{t-1} + \underline{A}_t \underline{\Upsilon} + \underline{M}_t(\underline{\Upsilon} + \underline{\Omega}), \underline{V}_0 = \begin{pmatrix} V_0^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ V_0^{(n)} \end{pmatrix},$$
$$\forall t \geq 1 \quad (3)$$

where

$$\underline{V}_t = \begin{pmatrix} V_t^{(1)} \\ \cdot \\ \cdot \\ \cdot \\ V_t^{(n)} \end{pmatrix}, \ \underline{M}_t = diag(M_t^{(1)}, ..., M_t^{(n)}),$$
$$\underline{\Upsilon} = \begin{pmatrix} \Upsilon_1 \\ \cdot \\ \cdot \\ \cdot \\ \Upsilon_n \end{pmatrix}, \ \underline{\Omega} = \begin{pmatrix} \Omega_1 \\ \cdot \\ \cdot \\ \cdot \\ \Omega_n \end{pmatrix},$$
$$\underline{A}_t = \begin{pmatrix} K_1\beta_1^{(1)}M_t^{(1)} & . \ . \ . & K_1\beta_1^{(n)}M_t^{(1)} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ K_n\beta_n^{(1)}M_t^{(n)} & . \ . \ . & K_n\beta_n^{(n)}M_t^{(n)} \end{pmatrix}.$$

It follows easily from (3) that

$$\underline{V}_t = \Big(\prod_{t\geq j\geq 1} \underline{A}_j\Big)\underline{V}_0 + \Big(\sum_{j=1}^{t}(\prod_{t\geq m\geq j} \underline{A}_m)\Big)\underline{\Upsilon} + \Big(\sum_{j=1}^{t}(\prod_{t\geq m\geq j+1} \underline{A}_m)\underline{M}_j\Big)(\underline{\Upsilon} + \underline{\Omega}), \ \forall t \in \mathbf{N}_+. \quad (4)$$

Hence,

$$E[\underline{V}_t] = \Big(\prod_{t\geq j\geq 1} E[\underline{A}_j]\Big)E[\underline{V}_0] + \Big(\sum_{j=1}^{t}(\prod_{t\geq m\geq j} E[\underline{A}_m])\Big)\underline{\Upsilon} + \Big(\sum_{j=1}^{t}(\prod_{t\geq m\geq j+1} E[\underline{A}_m])E[\underline{M}_j]\Big)(\underline{\Upsilon} + \underline{\Omega})$$
$$= \underline{\mathcal{A}}^t\underline{\mathcal{V}}_0 + \Big(\sum_{j=1}^{t}\underline{\mathcal{A}}^{t-j+1}\Big)\underline{\Upsilon} + \Big(\sum_{j=1}^{t}(\underline{\mathcal{A}}^{t-j}\underline{\mathcal{M}})\Big)(\underline{\Upsilon} + \underline{\Omega})$$

where

$$\underline{\mathcal{V}}_0 = \begin{pmatrix} L_0^{(1)}(\frac{1-\alpha_1}{2^{R_1}} + \alpha_1) \\ \cdot \\ \cdot \\ \cdot \\ L_0^{(n)}(\frac{1-\alpha_n}{2^{R_n}} + \alpha_n) \end{pmatrix},$$
$$\underline{\mathcal{M}} = diag(\frac{1-\alpha_1}{2^{R_1}} + \alpha_1, ..., \frac{1-\alpha_n}{2^{R_n}} + \alpha_n).$$

Let $[V_t]_i$ denote the $i$th component of the vector $\underline{V}_t \in \Re^n$, i.e., $[V_t]_i = V_t^{(i)}$. It follows from the above equation that

$$E[V_t^{(i)}] = \Big[E[\underline{V}_t]\Big]_i = [\underline{\mathcal{A}}^t\underline{\mathcal{V}}_0 + \Big(\sum_{j=1}^{t}\underline{\mathcal{A}}^{t-j+1}\Big)\underline{\Upsilon} + \Big(\sum_{j=1}^{t}(\underline{\mathcal{A}}^{t-j}\underline{\mathcal{M}})\Big)(\underline{\Upsilon} + \underline{\Omega})]_i, \ \forall t \in \mathbf{N}_+.$$

Following our assumptions, the matrix $\underline{\mathcal{A}}$ is stable. Hence, for each $i \in \{1, 2, ..., n\}$, there exists a non-negative scalar $P_i$

defined as follows:

$$P_i = \sup_{t\in\mathbf{N}_+}[\underline{\mathcal{A}}^t\underline{\mathcal{V}}_0 + \Big(\sum_{j=1}^{t}\underline{\mathcal{A}}^{t-j+1}\Big)\underline{\Upsilon} + \Big(\sum_{j=1}^{t}(\underline{\mathcal{A}}^{t-j}\underline{\mathcal{M}})\Big)(\underline{\Upsilon}+\underline{\Omega})]_i,$$

such that $E[V_t^{(i)}] \leq P_i, \quad \forall t \in \mathbf{N}_+$. This completes the proof because $\mathcal{E}_t^{(i)} \leq V_t^{(i)} + \Upsilon_i$, P-a.s.; and therefore, $E[\mathcal{E}_t^{(i)}] \leq E[V_t^{(i)}] + \Upsilon_i \leq P_i + \Upsilon_i \stackrel{\triangle}{=} D_i, \forall i \in \{1,2,...,n\}$.

*Remark 3.2:* i) If the system is completely decoupled (i.e., $\forall i \in \{1,2,...,n\}$ we have $\beta_i^{(i)} = 1$ and $\beta_i^{(j)} = 0, \forall j \neq i$), the matrix $\underline{\mathcal{A}}$ is diagonal (i.e., $\underline{\mathcal{A}} = diag\Big(K_1(\frac{1-\alpha_1}{2^{R_1}} + \alpha_1),...,K_n(\frac{1-\alpha_n}{2^{R_n}} + \alpha_n)\Big)$); and therefore, it is stable if and only if $K_i(\frac{1-\alpha_i}{2^{R_i}} + \alpha_i) < 1$, $\forall i \in \{1,2,...,n\}$.
ii) For the special case of the digital noiseless channel (i.e., $\alpha_i = 0, \forall i = \{1,2,...,n\}$), we have tracking in the almost sure sense. From (4) it follows that for this case using the proposed technique we have $\mathcal{E}_t^{(i)} \leq V_t^{(i)} + \Upsilon_i$ ($\forall i \in \{1,2,...,n\}$), P-a.s., where $V_t^{(i)}$ is the $i$th component of the vector $\underline{V}_t \in \Re^n$, as described below:

$$\underline{V}_t = \underline{A}^t\underline{V}_0 + (\sum_{j=1}^{t}\underline{A}^{t-j+1})\underline{\Upsilon} + (\sum_{j=1}^{t}\underline{A}^{(t-j)}\underline{M})(\underline{\Upsilon}+\underline{\Omega}),$$

$$\underline{A} = \begin{pmatrix} \frac{K_1\beta_1^{(1)}}{2^{R1}} & \cdot & \cdot & \cdot & \frac{K_1\beta_1^{(n)}}{2^{R_1}} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \frac{K_n\beta_n^{(1)}}{2^{R_n}} & \cdot & \cdot & \cdot & \frac{K_n\beta_n^{(n)}}{2^{R_n}} \end{pmatrix}, \underline{V}_0 = \begin{pmatrix} \frac{L_0^{(1)}}{2^{R_1}} \\ \cdot \\ \cdot \\ \cdot \\ \frac{L_0^{(n)}}{2^{R_n}} \end{pmatrix},$$

$$M = diag(\frac{1}{2^{R_1}},...,\frac{1}{2^{R_n}}).$$

Now, following our assumptions, the rates $\{R_i\}_{i=1}^n$ are chosen such that the matrix $\underline{A}$ is stable. Hence, for each $i \in \{1,2,...,n\}$, there exists a non-negative scalar $P_i$, as given below:

$$P_i = \sup_{t\in\mathbf{N}_+}[\underline{A}^t\underline{V}_0 + (\sum_{j=1}^{t}\underline{A}^{t-j+1})\underline{\Upsilon} + (\sum_{j=1}^{t}\underline{A}^{t-j}\underline{M})(\underline{\Upsilon}+\underline{\Omega})]_i,$$

such that $V_t^{(i)} \leq P_i$, $\forall t \in \mathbf{N}_+$, P-a.s.; and therefore $\mathcal{E}_t^{(i)} \leq P_i + \Upsilon_i$, $\forall t \in \mathbf{N}_+$, P-a.s.
iii) The results of this section can be extended to the multi-dimensional case $X_t^{(i)} \in \Re^{q_i}$, $q_i \geq 1$ without much difficulty. For the special case, $n = 1$, the results of Proposition 3.1 yield the following result.

*Corollary 3.3:* Consider the system (1) with only one sub-system ($n = 1$) over the BSC. Using flag bits and feedback from the receiver, the transmission over the BSC is equivalent to the transmission over the erasure channel with feedback, as described in Section III-A. For this equivalent transmission,

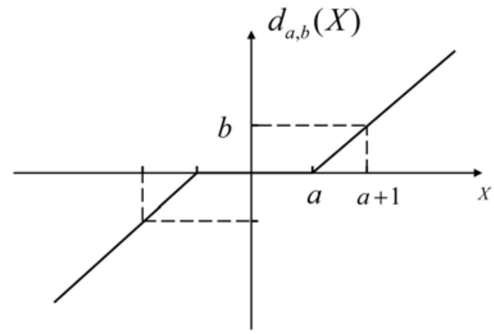


Fig. 2. Dead - zone nonlinearity function.

suppose that the rate $R_1$ is chosen such that the following condition holds:

$$\frac{K_1(1-\alpha_1)}{2^{R_1}} + K_1\alpha_1 < 1. \quad (5)$$

Then, the methodology of Section III-B guarantees mean absolute tracking.

*Remark 3.4:* We have the following remarks regarding the above result.
i) It is evident that the results of Corollary 3.3 also holds for linear systems (i.e., $X_{t+1}^{(1)} = K_1X_t^{(1)} + W_t^{(1)}$). For this special case, the condition (5) reduces to the conditions presented in [8], [9], [23].
ii) It was shown in ([8], Chapter 7, Theorem 7.2.1) that a necessary and sufficient condition for the mean absolute tracking of a fully observed linear system subject to bounded external disturbance (i.e., $X_{t+1}^{(1)} = K_1X_t^{(1)} + W_t^{(1)}$, $|W_t^{(1)}| \leq \Omega_1$) over the binary erasure channel (i.e., $R_1 = 1$) with erasure probability $\alpha_1$ is $K_1 < \frac{2}{1+\alpha_1}$. It is interesting to notice that for the special case of $R_1 = 1$, the condition (5), which guarantees the mean absolute tracking over the equivalent erasure channel, reduces to the above condition.
iii) It was shown in [23] that a necessary and sufficient condition for mean square tracking of the scalar partially observed linear system with mode $K_1$ over the digital noiseless channel (i.e., $\alpha_1 = 0$) is $R_1 > \log K_1$. Again, it is interesting to notice that for this case, the condition (5), which guarantees the mean absolute tracking over the equivalent erasure channel, reduces to the above condition.
iv) It was shown in [9] that a necessary and sufficient condition for mean square tracking of the scalar partially observed linear system with mode $K_1$ over an erasure channel with high rate (i.e., $R_1 \to \infty$) and erasure probability $\alpha_1$ is $K_1^2\alpha_1 < 1$. Again for this high rate case, it is clear that the condition (5), which guarantees the mean absolute tracking over the equivalent erasure channel, reduces to $K_1\alpha_1 < 1$, which is similar to the above condition.
From the above discussions it follows that the proposed technique presents some optimal features for the mean absolute tracking over the equivalent erasure channel.

In general, we can protect the message bits over the binary symmetric channel by adding some extra bits for error correction. Using such techniques, the probability of receiving message bits correctly will increase. However, this

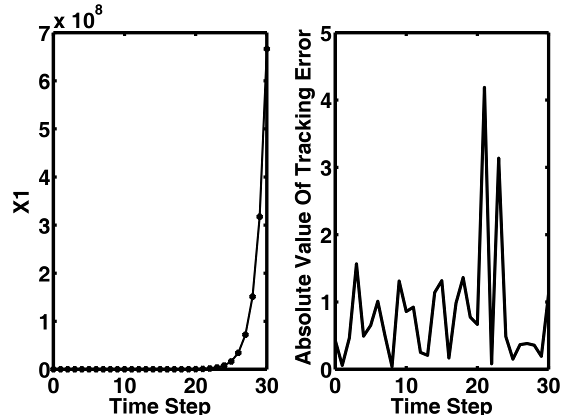


Fig. 3. Left figure: $X_t^{(1)}$ (solid line) and $\hat{X}_t^{(1)}$ (dotted) versus time step. Right figure: $\mathcal{E}_t^{(1)}$ versus time step.

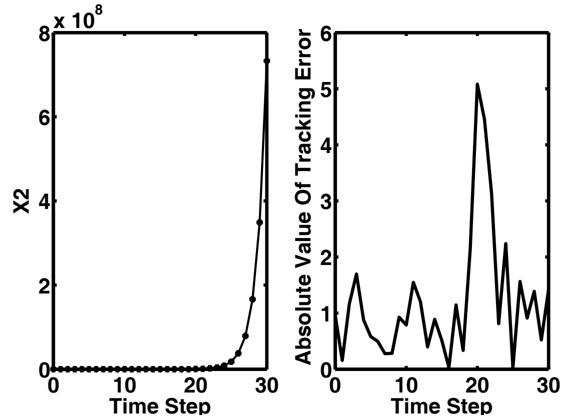


Fig. 4. Left figure: $X_t^{(2)}$ (solid line) and $\hat{X}_t^{(2)}$ (dotted) versus time step. Right figure: $\mathcal{E}_t^{(2)}$ versus time step.

will increase the number of transmitted bits. Therefore, we can just use few extra bits for error correction in each transmission. Consequently, the receiver may not be able to correct all the flipped message bits. But, using feedback acknowledgments from the receiver, the transmitter will know if the receiver is able to correct all the flipped message bits. Thus, if there are still some flipped message bits, which can not be corrected at the receiver using the implemented error correction technique, the transmitter sends the flag bit "0" to inform the receiver that the received message bits contain errors, which can not be corrected. Since the decoding law is recursive, if the receiver decodes the message bits containing errors, the decoding error may grow with time. Therefore, the decoder disregards the message bits if it receives the flag bit "0" and reconstructs the state using the available message bits received correctly.

## IV. Simulation Results

In this section, we illustrate the performance of the proposed design technique. Let $d_{a,b}(X)$ denote the dead-zone nonlinearity function as shown in Fig. 2. Here, we are concerned with a simple nonlinear dynamic system consisting of two coupled subsystems $s_1$ and $s_2$, as described below:

$$(s_1): \begin{cases} X_{t+1}^{(1)} = F_1(X_t^{(1)}, X_t^{(2)}) + W_t^{(1)}, \ X_0^{(1)} = \xi_1, \\ Y_t^{(1)} = X_t^{(1)} + G_t^{(1)}, \end{cases}$$

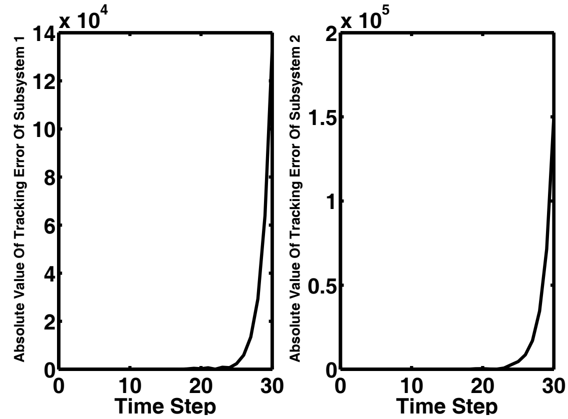


Fig. 5. Left figure: $\mathcal{E}_t^{(1)}$ versus time step for $R_1 = R_2 = 1$. Right figure: $\mathcal{E}_t^{(2)}$ versus time step for $R_1 = R_2 = 1$.

$$(s_2): \begin{cases} X_{t+1}^{(2)} = F_2(X_t^{(1)}, X_t^{(2)}) + W_t^{(2)}, \ X_0^{(2)} = \xi_2, \\ Y_t^{(2)} = X_t^{(2)} + G_t^{(2)}, \end{cases}$$

where $F_1(\gamma_1, \gamma_2) = d_{\frac{1}{2},1.1}(\gamma_1 + \gamma_2)$, $F_2(\gamma_1, \gamma_2) = d_{1,1}(\gamma_1 + \gamma_2)$, $X_t^{(i)} \in \Re$, $Y_t^{(i)} \in \Re$, $W_t^{(i)} \in \Re$, $G_t^{(i)} \in \Re$ ($i \in \{1, 2\}$), random variables $\xi_1$ and $\xi_2$ are uniformly distributed with distribution $\xi_1 \sim U(-1, 1)$ and $\xi_2 \sim U(-2, 2)$, respectively, $W_t^{(1)}$ and $W_t^{(2)}$ are i.i.d. with distribution $W_t^{(1)} \sim U(-1, 1)$ and $W_t^{(2)} \sim U(-2, 2)$, respectively, and $G_t^{(i)}$ is i.i.d. with distribution $G_t^{(i)} \sim U(-1, 1)$. Here, $\xi_1$ and $\xi_2$ are independent of system noises and measurement noises. Furthermore, $\{\xi_1, \xi_2\}$, $\{W_t^{(1)}, G_t^{(1)}\}$, $\{W_t^{(2)}, G_t^{(2)}\}$, $\{W_t^{(1)}, G_t^{(2)}\}$ and $\{W_t^{(2)}, G_t^{(1)}\}$ are mutually independent.

Note that the nonlinear functions $F_1(\cdot)$ and $F_2(\cdot)$ are Lipschitz because for each $\gamma_1$, $\gamma_2$, $\eta_1$, $\eta_2 \in \Re$, we have

$$|F_1(\gamma_1, \gamma_2) - F_1(\eta_1, \eta_2)| \leq 1.1(|\gamma_1 - \eta_1| + |\gamma_2 - \eta_2|),$$

$(i.e., \ K_1 = 1.1)$, and

$$|F_2(\gamma_1, \gamma_2) - F_2(\eta_1, \eta_2)| \leq |\gamma_1 - \eta_1| + |\gamma_2 - \eta_2|,$$

$(i.e., \ K_2 = 1)$. For subsystem $s_i$ we use the methodology presented in Section III to transmit observations to the fusion center via the BSC with cross over probability $p_i = 0.1$. For this transmission, we have $1 - \alpha_1 = (1 - p_1)^{R_1} = 0.9^{R_1}$ and $1 - \alpha_2 = (1 - p_2)^{R_2} = 0.9^{R_2}$. Hence, the matrix $\underline{\mathcal{A}}$ is given by

$$\underline{\mathcal{A}} = \begin{pmatrix} 1.1(\frac{0.9^{R_1}}{2^{R_1}} + 1 - 0.9^{R_1}) & 1.1(\frac{0.9^{R_1}}{2^{R_1}} + 1 - 0.9^{R_1}) \\ (\frac{0.9^{R_2}}{2^{R_2}} + 1 - 0.9^{R_2}) & (\frac{0.9^{R_2}}{2^{R_2}} + 1 - 0.9^{R_2}) \end{pmatrix}.$$

Table I summarizes the pair $(R_1, R_2)$ under which the matrix $\underline{\mathcal{A}}$ is stable. It also includes the corresponding upper bounds on the tracking errors, i.e., $D_1$ and $D_2$. From Table I, it follows that the rates $R_1 = 3$ and $R_2 = 3$ (which are relatively small) correspond to the smallest values for $D_1$ and $D_2$.

**Simulation Results.** The results shown in Fig. 3 and Fig. 4 illustrate the performance of the proposed technique for tracking the above nonlinear system at the fusion center, with the rates $R_1 = 3$ and $R_2 = 3$. It is clear from the figures that the reconstructed signals coincide with the actual signals. This certainly indicates the promising performance of the design technique presented in this paper. Fig. 5 illustrates the

TABLE I
ADMISSIBLE RATES $(R_1, R_2)$ AND THE CORRESPONDING $(D_1, D_2)$

| $(R_1, R_2)$ ; $(D_1, D_2)$ | $(R_1, R_2)$ ;$(D_1, D_2)$ |
|---|---|
| (2,2) ; (11.51, 11.02) | (2,3) ; (9.89, 8.88) |
| (2,4) ; (11.06, 10.42) | (2,5) ; (14.3, 14.69) |
| (2,6) ; (21.7, 24.42) | (2,7) ; (45.7, 56) |
| (3,1) ; (35.43, 49.2) | (3,2) ; (9.15, 9.49) |
| (3,3) ; (8.06, 7.84) | (3,4) ; (8.85, 9.04) |
| (3,5) ; (10.91, 12.15) | (3,6) ; (14.93, 18.3) |
| (3,7) ; (24.19, 32.22) | (3,8) ; (59.22, 85.16) |
| (4,1) ; (70.88, 92.42) | (4,2) ; (10.84, 10.59) |
| (4,3) ; (9.38, 8.59) | (4,4) ; (10.43, 10.03) |
| (4,5) ; (13.31, 13.95) | (4,6) ; (19.6, 22.51) |
| (4,7) ; (37.72, 47.19) | (5,2) ; (15.74, 13.75) |
| (5,4) ; (14.93, 12.85) | (5,5) ; (21.09, 19.77) |
| (5,6) ; (40.36, 41.44) | (6,2) ; (28.07, 21.72) |
| (6,3) ; (20.65, 14.99) | (6,4) ; (25.75, 19.62) |
| (6,5) ; (48.73, 40.44) | (7,2) ; (83.92, 57.77) |
| (7,3) ; (41.97, 27.1) | (7,4) ; (67.32, 45.62) |

performance of the technique when $R_1 = R_2 = 1$. It is clear from the figure that for these rates the tracking errors explode. Note that according to Table I, for the rates $R_1 = R_2 = 2$, which are very close to the rates $R_1 = R_2 = 1$, the mean absolute tracking error is bounded.

## V. CONCLUSION

In this paper, we have developed a simple technique for design of encoders and decoders for tracking nonlinear noisy dynamic systems over the binary symmetric channel. The promising performance of this technique has been illustrated by an example. For future it should be of great interest to consider more relaxed assumptions for dynamic system, e.g., local Lipschitz continuity instead of global one. Also, it may be interesting to consider a Lipschitz condition based on $p$-norm instead of 1-norm as considered in this note.

## REFERENCES


[1] J. Liu and N. Elia, "Writing on dirty paper with feedback," *Commun. Inf. Syst.*, vol. 5, no. 4, pp. 401–422, 2005.

[2] N. Elia, "When Bode meets Shannon: control-oriented feedback communication schemes," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1477–1488, 2004.

[3] C. D. Charalambous and A. Farhadi, "LQG optimality and separation principle for general discrete time partially observed stochastic systems over finite capacity communication channels," *Automatica*, vol. 44, no. 12, pp. 3181–3188, 2008.

[4] C. D. Charalambous, A. Farhadi, and S. Z. Denic, "Control of continuous-time linear Gaussian systems over additive Gaussian wireless fading channels: a separation principle," *IEEE Trans. Automat. Contr.*, vol. 53, no. 4, pp. 1013–1019, 2008.

[5] A. S. Matveev and A. V. Savkin, "Shannon zero error capacity in the problems of state estimation and stabilization via noisy communication channels," *International J. Control*, vol. 80, pp. 241–255, Feb. 2007.

[6] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Trans. Automat. Contr.*, vol. 49, no. 7, pp. 1196–1201, July 2004.

[7] T. Simsek, R. Jain, and P. Varaiya, "Scalar estimation and control with noisy binary observations," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1598–1603, Sep. 2004.

[8] A. Sahai, "Anytime information theory," Ph.D. thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Feb. 2001.

[9] P. Minero, M. Franceschetti, S. Dey, and N. Nair, "Data rate theorem for stabilization over time-varying feedback channels," *IEEE Trans. Automat. Contr.*, vol. 54, no. 2, pp. 243–255, Feb. 2009.

[10] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link–part I: scalar systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.

[11] S. Yuksel and T. Basar, "Minimum rate coding for LTI systems over noiseless channels," *IEEE Trans. Automat. Contr.*, vol. 51, pp. 1878–1887, Dec. 2006.

[12] A. V. Savkin and T. M. Cheng, "Detectability and output feedback stabilizability of nonlinear networked control systems," *IEEE Trans. Automat. Contr.*, vol. 52, no. 4, pp. 730–735, Apr. 2007.

[13] T. M. Cheng and A. V. Savkin, "Output feedback stabilisation of nonlinear networked control systems with non-decreasing nonlinearities," in *Proc. American Control Conf.*, 2006, pp. 2795–2800.

[14] C. De Persis and A. Isidori, "Stabilizability by state feedback implies stabilizability by encoded state feedback," *Syst. Contr. Lett.*, vol. 53, pp. 249–258, 2004.

[15] D. Liberzon and J. P. Hespanha, "Stabilization of nonlinear systems with limited information feedback," *IEEE Trans. Automat. Contr.*, vol. 50, no. 6, pp. 910–915, June 2005.

[16] G. N. Nair and R. J. Evans, "State estimation via a capacity-limited communication channel," in *Proc. 36th IEEE Conf. Decision Contr.*, 1997, pp. 866–871.

[17] T. M. Cheng, "Robust stabilisation of nonlinear systems using output measurements via finite data-rate communication channels," in *Proc. 44th IEEE Conf. Decision Contr.*, 2005, pp. 866–871.

[18] E. Okamoto and H. Ogawa, "A block-coded modulation method for one-way multimode data transmission," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2124–2135, Dec. 2002.

[19] E. Okamoto, Y. Iwanami, and T. Ikegami, "Multimode transmission using wavelet packet modulation and OFDM," in *Proc. 58th IEEE Conf. Veh. Technol.*, Oct. 2003, pp. 1458–1462.

[20] R. Chen, Z. Shen, J. G. Andrews, and R. W. Heath, Jr., "Multimode transmission for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3294–3302, July 2008.

[21] T. M. Cover and J. A. Tomas, *Elements of Information Theory*. Wiley Series in Telecommunications, 1991.

[22] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–I: no bandwidth constraint," *IEEE Trans. Inf. Theory*, vol. 12, pp. 172–182, 1966.

[23] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM J. Control Optimization*, vol. 43, no. 3, pp. 413–436, 2004.