Stabilization of Nonlinear Dynamic Systems over Limited Capacity Communication Channels

Vahideh Sanjaroon, Alireza Farhadi, Abolfazl Seyed Motahari and Babak. H. Khalaj

***Abstract*—This paper addresses the stabilization of noiseless nonlinear dynamic systems over limited capacity communication channels. It is shown that the stability of nonlinear dynamic systems over memory-less communication channels implies an inequality condition between the Shannon channel capacity and the summation of the positive equilibrium Lyapunov exponents of the dynamic system or equivalently, the logarithms of the magnitude of the unstable eigenvalues of system Jacobian. Furthermore, we propose an encoder, decoder and a controller to prove that scalar nonlinear dynamic systems are stabilizable under the aforementioned inequality condition over the digital noiseless and the packet erasure channels, respectively, in sure and almost sure senses. The performance of the proposed coding scheme is illustrated by computer simulations.**

***Index Terms*—Nonlinear dynamic systems, stabilization, Limited bandwidth channel, The Lyapunov exponents.**

I. Introduction

A. Motivation and Backgrounds

A fundamental result from [1]–[6] shows that the stabilization of an unstable linear system over a communication channel is equivalent to the ability of transmitting $\sum_{\{i:|\lambda_i(A)\geq 1|\}} \log_2 |\lambda_i(A)|$ bits through the communication channel where $\sum_{\{i:|\lambda_i(A)\geq 1|\}} \log_2 |\lambda_i(A)|$ is the instability expansion rate of linear systems.

Results on linear systems have been extended to nonlinear networked control systems [7]–[22]. The observability problem of nonlinear dynamic systems is addressed over limited capacity communication channel [7], the digital noiseless channel [8], [9], the packet erasure channels [10], [11], the real erasure channels [12], and AWGN channel [11]. In [13], the authors considered a noiseless nonlinear dynamic system and obtained the tight bound for stabilization over the digital noiseless channel in terms of the so called feedback topological entropy of the system. This result is extended in [14] for uncertain dynamic systems. Minimal bit rate and stabilizing entropy is studied in [15] for continuous-time dynamic systems and exponential stabilization. The stabilization problem of continuous-time Lipschitz nonlinear dynamic systems is addressed in [16]. In [17], the authors presented a necessary condition in terms of the constant positive Lyapunov exponent for the stabilization of scalar nonlinear noiseless dynamic systems over the real erasure channel. Input to state stabilization with quantized measurements is studied in [18]. Invariance entropy is defined in [19] for continuous-time deterministic nonlinear system which measures the minimal amount of information to achieve invariance of a given subset of the state space. Subsequently, in [20], it is shown that the invariance entropy is equivalent to the feedback topological entropy. Also, in [21], the authors addressed the relation between these two quantities for nonlinear systems in networked control context. Finally, the linearized technique is used in [22] for control of nonlinear systems over the packet erasure channel.

V. Sanjaroon, A. Farhadi and B. H. Khalaj are with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran.

A. S. Motahari is with the Department of Computer Engineering, Sharif University of Technology, Tehran, Iran.

B. Paper Contributions

Key contributions of this work compared to aforementioned earlier works can be summarized as follows:

1) In the earlier works, e.g., [13], [17], necessary conditions for the stabilization of discrete-time nonlinear dynamic systems are presented in the case of special communication channels such as the digital noiseless or real erasure channels. However; this paper extends these results by using an information theoretic approach which shows that $C > \sum_i \max\{0, \kappa_i^* \Delta_i^*\}$ is a necessary condition for asymptotic stabilization of nonlinear dynamic systems over any memory-less communication channel (Theorem 1), where C is the Shannon channel capacity; and Δ_i^*s and κ_i^*s denote, respectively, distinct equilibrium Lyapunov exponents and their multiplicity numbers.

2) In Theorem 2, we propose encoder, decoder, and controller in order to achieve the asymptotic stability of scalar nonlinear dynamic systems over the digital noiseless channel under the aforementioned inequality condition. The most similar result to this theorem is presented in [13] for similar system dynamic in terms of the so called topological feedback entropy. However, we propose a different encoder-decoder and controller scheme in order to achieve the stability. It should be noted that for the case of nonlinear dynamic systems, unlike the linear time invariant systems, e.g., [1], [23], the transmitter sends packets with time varying length which is determined based on the value of the initial state and control inputs. In our proposed scheme, although, the parameters of encoder/decoder vary based on the initial state and input signal values, the minimum required capacity is independent of these parameters and only depends on the equilibrium Lyapunov exponents.

3) To the best of our knowledge, for the first time, the tight bound on almost sure asymptotic stabilization of scalar nonlinear dynamic systems over the packet erasure channel is presented in this paper (Theorem 3). The authors in [12] and [17], respectively, studied only the necessary condition on mean square exponential observability and stabilization of nonlinear dynamic systems over real erasure channel which has continuous alphabet. In this theorem, we consider transmitting quantized measurements through the channel and propose

an encoder-decoder scheme and controller in order to achieve asymptotic stabilization in almost sure sense. In [22], a coding and controller scheme is also proposed for the stabilization of nonlinear dynamic system over the packet erasure communication channel based on the linearization method which only works for sufficiently slow dynamic systems. Note that [2] only considers linear systems.

C. Paper Organization

The rest of the paper is organized as follows: The problem formulation is given in Section II. In Section III, a necessary condition on asymptotic stability is presented. In Section IV, sufficient conditions are presented. Simulation results are given in Section V and Section VI concludes the paper.

II. Notations and preliminaries

A. Notation

We use the following notations throughout the paper. $|\cdot|$ denotes the absolute value of a square matrix determinant and $||\cdot||$ the Euclidean norm, $\log(\cdot)$ is the binary logarithm, $E[\cdot]$ denotes the expected value function, $\mathbb{P}(\cdot)$ is the probability mass function, and $\mathcal{F}(\cdot)$ denotes the probability density function. v^t denotes the sequence $v_0, v_1 \ldots, v_t$. $\lambda_i(.)$s denote the eigenvalues of a matrix. $J_{f/x}(x^*, u^*)$ denotes the Jacobian matrix of the vector function $f(x,u)$ with respect to x, i.e., the elements of this matrix are as $\{J_{f/x}(x^*, u^*)\}_{ij} = \frac{\partial\big(f(x,u)\big)_i}{\partial(x)_j}\big|_{(x^*,u^*)}$, where the subscripts determine the index of elements such that the first subscript is the row number and the second is the column number. For the simplicity, if $g(\cdot)$ is a real function of a real variable, we use $g'(\cdot)$ for denoting its derivative. $\mathbb{N}$ denotes the set of positive integer numbers. The zero vector with dimension n is denoted by $\underline{0}_n$. $\triangleq$ and $\overset{p}{=}$ denote , respectively, the definition of a new variable and convergence in probability.

B. Preliminaries

The overall system model considered in this paper is shown in Fig. 1. In this system, the communication from plant to the distant controller is subject to communication imperfections; while the communication from distant controller to the plant is perfect. This structure of system has been considered in many research papers, e.g., [2], [1], [24]. This is the case, for example, in the tele-operation of micro autonomous vehicles, where the vehicle is supplied by limited capacity power supply; and hence, transmission from the vehicle to the remote base station, where the controller is located, is subject to communication imperfections; while as the base station can be supplied with high power, the communication from distant controller to vehicle is almost perfect. In what follows, we describe the input-output relation of each building block of this system.

Plant: The plant is described by the following noiseless nonlinear, time invariant and fully-observed dynamic system:

$$x_{k+1} = f(x_k, u_k), \quad y_k = x_k, \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state, $u_k \in \mathbb{R}^d$ is the input, and y_k is the observation signal. In fact, the physical output signal of

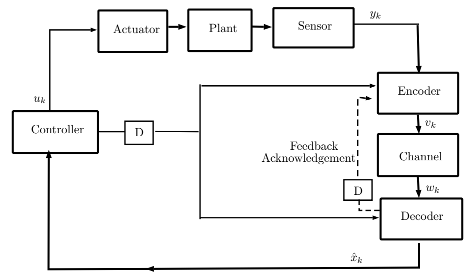


Fig. 1: Control of dynamic systems over the Discrete Memoryless Channels (DMCs)

the plant is transduced into measurable form y_k by a sensor. The initial state x_0 is known for the encoder; but unknown for decoder and controller. Hence, for them, it is assumed as a random variable with bounded entropy that takes values in a finite support set $\psi_0 \subset \mathbb{R}^n$.

Communication channel: In this paper, we consider two types of channels as follows where channel input and output, respectively, are denoted by $v \in \mathbb{R}$ and $w \in \mathbb{R}$.

– The digital noiseless channel with average rate R_{av}, where for the kth transmission, the input is chosen from a set with 2^{R_k} members, and the channel output is the same as the channel input. In fact, R_k noiseless bits are transmitted at each channel use so that the average bit rate defined as $R_{av} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} R_k$ exists. Therefore, the channel is noiseless and memory-less and its capacity equals to $C = R_{av}$ bits in every channel use.

– The packet erasure channel with average rate R_{av} and erasure probability γ, where the space of channel input is a set with 2^{R_k} members, and the channel output is the same as the input symbol (with probability $1-\gamma$) or the erasure symbol (with probability γ). The capacity of this channel is $C = (1-\gamma)R_{av} = (1-\gamma)\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} E[R_k]$. Note that it is assumed that the limit of $\frac{1}{n}\sum_{k=0}^{n-1} E[R_k]$ exists.

Encoder: In general, given the feedback acknowledgment with one delay, an encoder maps $(y^k, u^{k-1}, w^{k-1}) \to v_k$ where u^{k-1} and w^{k-1} denote, respectively, the sequences $\{u_0, u_1, \ldots, u_{k-1}\}$ and $\{w_0, w_1, \ldots, w_{k-1}\}$. However, as there is no process noise in the system and all the ambiguity in the system is due to the initial state, we consider an special case of encoder which maps $(x_0, u^{k-1}, w^{k-1}) \to v_k$.

Decoder: The decoder is an operator that maps $(w^k, u^{k-1}) \to \hat{x}_{k|k}$, where $\hat{x}_{k|k}$ is the reconstruction of the state variable x_k at the time instant k given the observation $w_0, w_1, \ldots, w_k$. We assume $\hat{x}_{0|-1} = \underline{0}_n$.

Controller: The controller is defined as a nonlinear vector function ($g : \mathbb{R}^n \to \mathbb{R}^d$) which generates the input signal of the plant as $u_k = g(\hat{x}_{k|k})$. The controller commands are converted into changes in physical parameters by an actuator which is a hardware device. In this paper, we assume that the controller signal is sent back to the encoder and decoder and received with one delay.

In this paper, we consider the following definitions and assumptions to study the stabilization of system of Fig. 1.

Definition 1: System (1) over the given communication links is *asymptotic stabilizable at point x^* in sure, almost* *sure, and in probability senses if there exist an encoder and decoder pair and a controller such that, respectively;* $\lim_{k\to\infty} ||x_k - x^*|| = 0$*;* $\mathbb{P}[\lim_{k\to\infty} ||x_k - x^*|| = 0] = 1$*; for all* $\epsilon > 0$, $\lim_{k\to\infty} \mathbb{P}[||x_k - x^*|| \geq \epsilon] = 0$.

Assumption 1: The function $f(x, u)$ is assumed to be differentiable with respect to vector x.

Assumption 2: The Jacobian matrix $J_{f/x}(x, u)$ is assumed to be continuous with respect to both vectors x and u.

Assumption 3: Any control sequence that makes the state of the dynamic system (1) converge to x^*, has a unique steady state value which equals to u^*.

Assumption 4: The determinant of the Jacobian matrix $J_{f/x}$ at point (x^*, u^*) is non-zero.

Assumption 5: There is a differentiable control law $\mathscr{G}(\cdot)$ such that it

– makes the state of dynamic system (1) converge to point x^* for every $x_0 \in \psi_0$.

– makes the dynamic system $f(x_k, u_k + d_k)$ converge to point x^* for every $x_0 \in \psi_0$ and for every additive disturbance $d_k \in \mathbb{R}^n$ satisfying $\lim_{k\to\infty} ||d_k|| = 0$.

In the rest of the paper, without loss of generality, we assume that $(x^*, u^*) = (\underline{0}_n, \underline{0}_d)$ and the stability means the stability at point $x^* = \underline{0}_n$.

Remark 1: Note that x^* is the equilibrium point of the dynamic system $x_{k+1} = f(x_k, u^*)$. Therefore, based on the Lyapunov exponent definition in Theorem 4 of [25], it is concluded that the absolute value of the s distinct eigenvalues of the Jacobian matrix $J_{f/x}$ at the constant point (x^*, u^*) $(s \leq d)$ are equal to the Lyapunov exponents of the system $x_{k+1} = f(x_k, u^*)$ at point x^* which we referred them as the *equilibrium Lyapunov exponents*. The ith equilibrium Lyapunov exponent is denoted by Δ_i^* with the multiplicity κ_i^* for $1 \leq i \leq s$.

III. Necessary conditions for stabilization

In this section, we present a necessary condition for the stabilization of the system (1) over Discrete Memory-less Channels (DMCs) in the probability sense which also implies the necessity of this condition for the stabilization in sure and almost sure senses. It should be noted that the theorem holds in channels with feedback as well as without feedback. Before we present the main theorem of this section, we present the following preposition and lemma.

Preposition 1: For the Jacobian matrix $J_{f/x}(\underline{0}_n, \underline{0}_d) \in \mathbb{R}^{n\times n}$, there is a matrix P such that $PJ_{f/x}(\underline{0}_n, \underline{0}_d)P^{-1} = A$, where A is a block diagonal matrix in the form of $A = \begin{bmatrix} A^s & \\ & A^u \end{bmatrix}$ where A^s is a Jordan matrix of size $n_s \times n_s$ consisting all the Jordan blocks of stable eigenvalues of $J_{f/x}(\underline{0}_n, \underline{0}_d)$. The other Jordan blocks are in the matrix A^u of size $n_u \times n_u$ (i.e., $n = n_s + n_u$). See the Appendix B of [26].

Corollary 1: Let us define a new vector variable $z_k = Px_k$ as the linear transformation of x_k $\big(x_{k+1} = f(x_k, u_k)\big)$ which follows the following dynamic equation:

$$z_{k+1} = d(z_k, u_k) \triangleq Pf(P^{-1}z_k, u_k),\ z_0 = Px_0; \qquad (2)$$

Note that the Jacobian matrix $J_{d/z}(\underline{0}_n, \underline{0}_d)$ corresponding to the system (2), equals to the matrix A. Assume z_k^s and z_k^u denote, respectively, the first n_s and the last n_u elements of z_k. Consequently, corresponding to the stable and unstable subspace of A, we have: $z_{k+1} = \begin{bmatrix} z_{k+1}^s \\ z_{k+1}^u \end{bmatrix} = \begin{bmatrix} d^s(z_k^s, z_k^u, u_k) \\ d^u(z_k^s, z_k^u, u_k) \end{bmatrix}$.

Remark 2: The Assumptions 1-4 do hold for the system (2).

Lemma 1: There is some open set $\mathscr{N} = \mathscr{N}_{n_u} \times \mathscr{N}_{n_s} \times \mathscr{N}_d$ containing $(\underline{0}_n, \underline{0}_d)$ such that for every $z^s \in \mathscr{N}_{n_s}$ and $u \in \mathscr{N}_d$, the function $d^u(z^s, z^u, u)$ is a one to one function with respect to the vector variable z^u on domain $\mathscr{N}_{n_u}$. It should be noted that $\mathscr{N}_{n_u} \subset \mathbb{R}^{n_u}$, $\mathscr{N}_{n_s} \subset \mathbb{R}^{n_s}$, and $\mathscr{N}_d \subset \mathbb{R}^d$.

Proof: To prove this lemma we first define a new function $f : \mathbb{R}^{n+d} \to \mathbb{R}^{n+d}$ as $\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = f(z^s, z^u, u) \triangleq \begin{bmatrix} z^s \\ d^u(z^s, z^u, u) \\ u \end{bmatrix}$. By the Assumptions 2 and 4, it is deduced that the Jacobian matrix of this new function is continuous with the determinant equals to $|J_{d^u/z^u}|$ which is non-zero at the origin. Therefore, we can apply the Inverse Function Theorem (Theorem 9.24 of [27]) on $f(z^s, z^u, u)$. This theorem results in the existence of some open set $\mathscr{N} = \mathscr{N}_{n_u} \times \mathscr{N}_{n_s} \times \mathscr{N}_d$ containing $(\underline{0}_n, \underline{0}_d)$ and an open set $\mathscr{P} = \mathscr{P}_{n_u} \times \mathscr{P}_{n_s} \times \mathscr{P}_d$ containing $f(\underline{0}_n, \underline{0}_d)$ such that $f : \mathscr{N} \to \mathscr{P}$ has a continuous inverse $f^{-1} : \mathscr{P} \to \mathscr{N}$ as $\begin{bmatrix} z^s \\ z^u \\ u \end{bmatrix} = f^{-1}(q_1, q_2, q_3) = \begin{bmatrix} q_1 \\ d^{u^{-1}}(q_1, q_2, q_3) \\ q_3 \end{bmatrix}$. In fact, for every $z^s \in \mathscr{N}_{n_s}$ and $u \in \mathscr{N}_d$, the function $q_2 = d^u(z^s, z^u, u)$ has an inverse function on domain $z^u \in \mathscr{N}_{n_u}$ denoted by $z^u = d^{u^{-1}}(z^s, q_2, u)$ where $q_2 \in \mathscr{P}_{n_u}$. Hence, for every $z^s \in \mathscr{N}_{n_s}$ and $u \in \mathscr{N}_d$, the function $d^u(z^s, z^u, u)$ is one to one with respect to the vector variable z^u on domain $\mathscr{N}_{n_u}$. ■

Now, we are ready to present the main result of this section.

Theorem 1: A necessary condition for the stabilization of the system (1) in the probability sense over DMCs with channel capacity C is that $C > \sum_i \max\{0, \kappa_i^* \Delta_i^*\}$ provided the Assumptions 1-4 hold.

Proof: We first note that necessary conditions for the stabilization of the systems (1) at point $x^* = \underline{0}_n$ and system (2) at point $z^* = \underline{0}_n$ are the same; therefore, we continue by proving the necessary condition of the system (2). To achieve this goal, we prove that $\sum_{i=1}^k I(z_i^u; w_i|w^{i-1}) \leq (k+1)C$ through the following steps where C is the Shannon channel capacity as defined in [28].

$$\begin{aligned}
\sum_{i=1}^k I(z_i^u; w_i|w^{i-1}) &\overset{(a)}{\leq} \sum_{i=1}^k I(z_i; w_i|w^{i-1}), \overset{(b)}{\leq} I(z^k; w^k), \\
&\overset{(c)}{=} h(w^k) - h(w^k|z^k, u^{k-1}), \\
&= h(w^k) - \sum_{i=0}^k h(w_i|w_0, \ldots, w_{i-1}, z^k, u^{k-1}), \\
&\overset{(d)}{=} h(w^k) - \sum_{i=0}^k h(w_i|w_0, \ldots, w_{i-1}, v_i, z^k, u^{k-1}), \\
&\overset{(e)}{=} h(w^k) - \sum_{i=0}^k h(w_i|v_i), \\
&\leq \sum_{i=0}^k I(w_i; v_i) \overset{(f)}{\leq} (k+1)C, \qquad (3)
\end{aligned}$$

where $I(\cdot;\cdot)$ is the mutual information function and $h(.)$ is the differential entropy for continuous variables as defined in [28]. (a) and (b) are obtained by writing the chain rule (Theorem 2.5.2 of [28]) and using the fact that the mutual information is a positive quantity. (c) follows since based on the structure of control/communication system of Fig. 1, u^{k-1} is a known sequence for both transmitter and receiver at the time instant k. (d) follows from the fact that v_i can be exactly obtained by knowing $w_0, \dots, w_{i-1}, u^{i-1}, z^i$. (e) is a direct result of memory-less channel since by the definition of a memory-less channel (Section 7.4 of [28]), w_i depends only on v_i and is conditionally independent of everything else. (f) is derived using the channel capacity definition.

Now, we choose the following function to measure distortion between any points x and x^* [2]:

$$d^\epsilon(x, x^*) = \left\{ \begin{array}{ll} 0, & ||x - x^*|| \leq \epsilon, \\ 1, & ||x - x^*|| > \epsilon. \end{array} \right. \tag{4}$$

Hence, $E[d^\epsilon(x, x^*)] = \mathbb{P}(||x - x^*|| > \epsilon)$. Note that we have assumed that (x^*, u^*) equals to $(\underline{0}_n, \underline{0}_d)$. Therefore, as the system (1) is stabilizable in the probability sense, there is a control sequence u_k such that $\lim_{k\to\infty} ||x_k|| \overset{p}{=} 0$. Moreover, based on the Assumption 3, it is deduced that $\lim_{k\to\infty} ||u_k|| \overset{p}{=} 0$. Therefore, for every $\epsilon > 0$, there is $k_1(\epsilon) \in \mathbb{N}$ such that for every $k \geq k_1$, we have $E[d^\epsilon\big((z_k, u_k), (\underline{0}_n, \underline{0}_d)\big)] \leq \epsilon$.

We choose the value of ϵ such that $\epsilon \leq \frac{1}{2}$ and the $n + d$ dimensional sphere with center $(\underline{0}_n, \underline{0}_d)$ and radius ϵ is placed inside the subspace $\mathscr{N}$ as introduced in Lemma 1.

In order to obtain a lower bound on $\sum_{i=1}^{k} I(z_i^u; w_i|w^{i-1})$, we define $\delta_i \triangleq d^\epsilon\big((z_i, u_i), (\underline{0}_n, \underline{0}_d)\big)$ as a Bernoulli random variable where for $i \geq k_1$, we have $\mathbb{P}(\delta_i = 1) \leq \epsilon$. Therefore, we can find a lower bound for $k > k_1$ as follows:

$$\begin{aligned}
I(z_k^u; w_k|w^{k-1}) =& h(z_k^u|w^{k-1}) - h(z_k^u|w^k), \\
=& h(z_k^u|w^{k-1}) - h(z_k^u, \delta_k|w^k) + H(\delta_k|z_k^u, w^k), \\
\overset{(a)}{\geq}& h(z_k^u|w^{k-1}) - h(z_k^u, \delta_k|w^k), \\
=& h(z_k^u|w^{k-1}) - H(\delta_k|w^k) - h(z_k^u|w^k, \delta_k), \\
=& h(z_k^u|w^{k-1}) - H(\delta_k|w^k) \\
& - h(z_k^u|w^k, \delta_k = 0)\mathbb{P}(\delta_k = 0) \\
& - h(z_k^u|w^k, \delta_k = 1)\mathbb{P}(\delta_k = 1), \\
\overset{(b)}{\geq}& (1 - \epsilon)h(z_k^u|w^{k-1}) - H(\delta_k) - h(z_k^u|w^k, \delta_k = 0),
\end{aligned} \tag{5}$$

where $H(.)$ is the entropy of discrete variables as defined in [28]. In the above equalities/inequalities, (a) follows since the entropy of a discrete variable is a positive quantity. Since conditioning on the entropy decreases the entropy, (b) holds.

Note that the above inequality is also true for $k_1 \leq i \leq k-1$, which can be continued as follows:

$$\begin{aligned}
I(z_i^u; w_i|w^{i-1}) \geq& (1 - \epsilon)h(z_i^u|w^{i-1}) - H(\delta_i) - h(z_i^u|w^i, \delta_i = 0), \\
=& (1 - \epsilon)h(z_i^u|w^{i-1}) - H(\delta_i) - \epsilon h(z_i^u|w^i, \delta_i = 0) \\
& - (1 - \epsilon)h(z_i^u|w^i, \delta_i = 0), \\
\overset{(a)}{\geq}& (1 - \epsilon)h(z_i^u|w^{i-1}) - H(\delta_i) - \epsilon h(z_i^u|w^i, \delta_i = 0) \\
& - (1 - \epsilon)h(z_{i+1}^u|w^i, \delta_i = 0) + (1 - \epsilon)\log\big(|A^u| - \epsilon_m\big), \\
\overset{(b)}{\geq}& (1 - \epsilon)h(z_i^u|w^{i-1}) - H(\delta_i) - \epsilon h(z_i^u|\delta_i = 0) \\
& - (1 - \epsilon)h(z_{i+1}^u|w^i) + (1 - \epsilon)\log\big(|A^u| - \epsilon_m\big),
\end{aligned} \tag{6}$$

For (a), see the Appendix A, where ϵ_m is a constant depending on ϵ such that it will be small for a small value of ϵ. Since conditioning on entropy decreases the entropy, (b) is resulted.

By substituting (5) and (6), we can obtain the lower bound on $\sum_{i=k_1}^{k} I(z_i^u; w_i|z_i^s, w^{i-1})$ as follows:

$$\begin{aligned}
\sum_{i=k_1}^{k} I(z_i^u; w_i|w^{i-1}) \geq& (1 - \epsilon)h(z_{k_1}^u|w^{k_1-1}) - \sum_{i=k_1}^{k} H(\delta_i) \\
& - h(z_k^u|w^k, \delta_k = 0) - \epsilon\sum_{i=k_1}^{k-1} h(z_i^u|\delta_i = 0) \\
& + (1 - \epsilon)\sum_{i=k_1}^{k-1} \log\big(|A^u| - \epsilon_m\big), \\
\geq& (1 - \epsilon)h(z_{k_1}^u|w^{k_1-1}) - (k - k_1 + 1)H_b(\epsilon) \\
& - \log(K_{n_u}\epsilon^{n_u}) - (k - k_1)\epsilon\log(K_{n_u}\epsilon^{n_u}) \\
& + (1 - \epsilon)\sum_{i=k_1}^{k-1} \log\big(|A^u| - \epsilon_m\big),
\end{aligned} \tag{7}$$

where $H_b(\epsilon)$ is the binary entropy function defined as the entropy of a Bernoulli random variable with probability ϵ. In the aforementioned inequalities, the last inequality follows since δ_i, $i \geq k_1$ is a Bernoulli random variable whose probability is at most $\epsilon \leq \frac{1}{2}$; and therefore, $H(\delta_i) \leq H_b(\epsilon)$. Furthermore, note that the condition $\delta_i = 0$ implies that $||z_i^u|| \leq ||z_i|| \leq \epsilon$. On the other hand, an upper bound on the entropy of bounded random variables is obtained by the uniform distribution (Example 12.2.4 of [28]); and therefore, $h(z_i^u|\delta_i = 0) \leq \log(K_{n_u}\epsilon^{n_u})$ where K_{n_u} is the volume constant for the n_u dimensional sphere. In order to obtain the necessary condition in terms of capacity, we rewrite the inequality when k goes to infinity:

$$\begin{aligned}
C \geq& \lim_{k\to\infty} \sup_{\mathcal{F}(V^k)} \frac{1}{k}\left[\sum_{i=1}^{k_1} I(z_i^u; w_i|w^{i-1}) + \sum_{i=k_1}^{k} I(z_i^u; w_i|w^{i-1})\right], \\
\overset{(a)}{=}& \lim_{k\to\infty} \frac{1}{k}\sum_{i=k_1}^{k} I(z_i^u; w_i|z_i^s, w^{i-1}), \\
\overset{(b)}{\geq}& \lim_{k\to\infty} \frac{1}{k}\Big[(1 - \epsilon)h(z_{k_1}^u|z_{k_1}^s, w^{k_1-1}) - (k - k_1 + 1)H_b(\epsilon) \\
& - \log(K_{n_u}\epsilon^{n_u}) - (k - k_1)\epsilon\log(K_{n_u}\epsilon^{n_u})\Big] \\
& + \frac{1}{k}(1 - \epsilon)\sum_{i=k_1}^{k-1} \log\big(|A^u| - \epsilon_m\big), \\
\overset{(c)}{=}& - H_b(\epsilon) - \epsilon\log(K_{n_u}\epsilon^{n_u}) + (1 - \epsilon)\lim_{k\to\infty}\frac{1}{k}\sum_{i=k_1}^{k-1} \log\big(|A^u| - \epsilon_m\big),
\end{aligned}$$

where (a) follows from the fact that k_1 is finite; hence, $\sum_{i=1}^{k_1} I(z_i^u; w_i|w^{i-1})$ is finite. (b) is resulted from inequality (7). Since x_0 has bounded entropy and k_1 is finite, $h(x_{k_1}|w^{k_1-1})$ is also finite which results in the equality (c).

It should be noted that converging ϵ to zero results in the convergence of $H_b(\epsilon)$ to zero (Example 2.2,1 of [28]). Moreover, based on L'Hospital's Rule (Theorem 5.13 of [27]), it is deduced that $\epsilon\log(K_{n_u}\epsilon^{n_u})$ also converges to zero when ϵ goes to zero. Therefore, as ϵ can be chosen arbitrary small, $C > \log|A^u| = \sum_i \log\big|\lambda_i(A^u)\big| = \sum_i \log|\lambda_i(J_{d^u/z^u}(\underline{0}_n, \underline{0}_d))| = \sum_{i=1}^{n} \max\{0, \log\big|\lambda_i(J_{f/x}(\underline{0}_n, \underline{0}_d))\big|\}$ is a necessary condition for stabilization. Therefore, based on Remark 1, the necessary condition can be written as $C > \sum_{i=1}^{s} \max\{0, \kappa_i^*\Delta_i^*\}$. This completes the proof. ■

Note that in the above theorem it was proved that $C > \sum_{i=1}^{s} \max\{0, \kappa_i^* \Delta_i^*\}$ is a stabilization necessary condition generalizing the existing results for linear systems (e.g., [1], [2]) and noiseless channels (e.g., [13]).

IV. Sufficient conditions for stabilization

Considering the system model of Fig. 1, sufficient conditions for the stability of the system are presented for two distinct channels. First, we address the stability question over the digital noiseless channel and then over the packet erasure channel with feedback acknowledgment. For this purpose, we propose encoders and decoders that map $(x_0, u^{k-1}, w^{k-1}) \to v_k$ and $(w^k, u^{k-1}) \to \hat{x}_{k|k}$, respectively. In this section, we prove that using our proposed coding scheme and the controller law $\mathscr{G}(\cdot)$ satisfying the Assumptions 5, we can obtain a tight sufficient condition for the stabilization of scalar nonlinear systems over the digital noiseless and the packet erasure channels.

Encoder-Decoder scheme:

1) Let $\hat{x}_{0|-1} = 0$ and $\Psi_0 = [-L_0, L_0]$.

2) For the time instant $k = 0$:

a) Determination of the packet length R_0: We set $L_{0|-1} \triangleq L_0$, $\Omega_{0|-1} \triangleq \Psi_0$, and $m_0 \triangleq \max_{a\in\Omega_{0|-1}} |\mathscr{G}'(a)|$. Subsequently, the length of the packet is determined as $R_0 = \log(\max\{1, m_0\})$.

b) Transmission of the packet: The signal $e_{0|-1} \triangleq x_0 - \hat{x}_{0|-1} = x_0$ is quantized using the uniform quantizer with 2^{R_0} levels and finite range $2L_{0|-1}$ whose output is the center of the corresponding sub-interval. Subsequently, v_0 is obtained as the index of this quantization output and is sent through a packet of length R_0 over the communication channel. If the packet is received successfully, the acknowledgment is sent back to the transmitter and both transmitter and receiver which have access to the sub-interval index, update the initial state estimate $\hat{x}_{0|0}$ as the center of the sub-interval. Therefore, the ambiguity interval of initial state is reduced to $2L_{0|0}$ where $L_{0|0} = \frac{L_{0|-1}}{2^{R_0}}$. If this packet is not received correctly, it is erased and the initial state estimate and the ambiguity interval are not changed, i.e., $\hat{x}_{0|0} = \hat{x}_{0|-1}$ and $L_{0|0} = L_{0|-1}$. Finally, the receiver sends the value of $\hat{x}_{0|0}$ to the controller.

3) For the time instant $k \geq 1$:

a) Determination of the packet length R_k: Given $\hat{x}_{k-1|k-1}$ and control sequence u^{k-1}, the following parameters are obtained through the following equations:

$$L_{k-1|k-1} = \frac{L_{k-1|k-2}}{2^{\beta_{k-1} R_{k-1}}}, \tag{8}$$

$$\Omega_{k-1|k-1} = \{s \in \mathbb{R}; |s - \hat{x}_{k-1|k-1}| \leq L_{k-1|k-1}\}, \tag{9}$$

$$s_k = \max_{b\in\Omega_{k-1|k-1}} \left|J_{f/x}(b, u_{k-1})\right|, \tag{10}$$

$$\hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1}, u_{k-1}), \tag{11}$$

$$L_{k|k-1} = s_k L_{k-1|k-1}, \tag{12}$$

$$\Omega_{k|k-1} = \{s \in \mathbb{R}; |s - \hat{x}_{k|k-1}| \leq L_{k|k-1}\}, \tag{13}$$

$$m_k = \max_{a\in\Omega_{k|k-1}} |\mathscr{G}'(a)|, \tag{14}$$

$$r_k = \log(\max\{1, (m_k/m_{k-1})s_k\}), \tag{15}$$

where β_{k-1} equals to zero if the packet has been erased at the time instant $k-1$; otherwise, $\beta_{k-1} = 1$. Finally, the value of the packet length R_k is determined at both transmitter and receiver using the following equation,

$$R_k = (1 - \beta_{k-1})R_{k-1} + r_k + \epsilon, \tag{16}$$

where $\epsilon > 0$ is an arbitrary value. Note that for the fractional R_k, the time sharing technique should be used as explained in the Appendix A of [11].

b) Transmission of the packet: The signal $e_{0|k-1} \triangleq x_0 - \hat{x}_{0|k-1}$ is quantized using the uniform quantizer with 2^{R_k} levels and finite range $2L_{0|k-1}$ whose output is the center of the corresponding sub-interval. Subsequently, v_k is obtained as the index of this quantization output and sent through a packet of length R_k over the communication channel. If the packet is received successfully, the acknowledgment is sent back to the transmitter and both transmitter and receiver which have access to the sub-interval index, update the initial state estimate $\hat{x}_{0|k}$ using the center of the sub-interval. In fact, by receiving the packet, the ambiguity interval of initial state is reduced to $2L_{0|k}$ where $L_{0|k} = \frac{L_{0|k-1}}{2^{R_k}}$. If this packet is not received, the initial state estimate and the ambiguity interval is not changed, i.e., $\hat{x}_{0|k} = \hat{x}_{0|k-1}$ and $L_{0|k} = L_{0|k-1}$. After reconstruction of $\hat{x}_{0|k}$, $\hat{x}_{k|k}$ is obtained through the recursive equation $\hat{x}_{t|k} = f(\hat{x}_{t-1|k}, u_{t-1}),\ t = 1, \ldots, k$, and sent to the controller.

Applying this coding scheme leads to the following results:

- The value of the packet length R_k is adaptively changed based on the estimate of the current states, the input signals, previous packets length, and the channel erasure information; all of them being available at both transmitter and receiver at the time instant k. Hence, such coding structure is causal.
- Finite range $2L_{0|k}$ is determined through the recursive equation at any time instant $k \geq 0$, as follows: $L_{0|k} = \frac{L_{0|k-1}}{2^{\beta_k R_k}}, L_{0|-1} = L_0$, where β_is are independent indicator random variables with the distribution $p(\beta_i = 0) = \gamma$ and $p(\beta_i = 1) = 1 - \gamma$ which model the erasure of packets. Therefore, we have: $|e_{0|k}| \leq L_{0|k} = \frac{L_{0|-1}}{2^{\sum_{i=0}^{k} \beta_i R_i}}$.
- The set $\Omega_{k|k}$ is a subset of $\Omega_{k|k-1}$ (i.e., $\Omega_{k|k} \subset \Omega_{k|k-1}$).
- It should be noted that $x_0 \in \Omega_{0|0}$ and therefore, all future estimates of x_0 are in $\Omega_{0|0}$ (i.e., $\hat{x}_{0|t} \in \Omega_{0|0}\ , t \geq 0$). By the definition of $L_{1|1}$ and $\Omega_{1|1}$, it is deduced that $x_1 \in \Omega_{1|1}$, since

$$\begin{aligned} |x_1 - \hat{x}_{1|t}| &= |f(x_0, u_0) - f(\hat{x}_{0|t}, u_0)|, \\ &= |e_{0|t}|\left|J_{f/x}(b, u_0)\right|, \\ &\overset{(a)}{\leq} \max_{b\in\Omega_{0|0}} |e_{0|t}|\left|J_{f/x}(b, u_0)\right|, \\ &\leq s_1|e_{0|t}| \leq s_1 \frac{L_{0|-1}}{2^{\sum_{i=0}^{t} \beta_i R_i}} = \frac{L_{1|1}}{2^{\sum_{i=2}^{t} \beta_i R_i}}, \end{aligned} \tag{17}$$

where the existence of point b between x_0 and $\hat{x}_{0|0}$ is guaranteed by the Mean Value Theorem (Theorem 5.10 of [27]). As points x_0 and $\hat{x}_{0|0}$ belong to the set $\Omega_{0|0}$, the inequality (a) is resulted.

Therefore, $|x_1 - \hat{x}_{1|1}| \leq L_{1|1}$; i.e., $x_1 \in \Omega_{1|1}$ and the future estimates ($\hat{x}_{1|t}\ , t \geq 1$) also belong to $\Omega_{1|1}$. If we repeat this procedure for $|x_k - \hat{x}_{k|t}|$, $t \geq k$, we can see that $|x_k - \hat{x}_{k|t}| \leq \frac{L_{k|k}}{2^{\sum_{i=k+1}^{t} R_i}}$; and therefore, $x_k, \hat{x}_{k|t} \in \Omega_{k|k}$ for $t \geq k$.

Now, using the above coding scheme, we have the following sufficient condition on the stabilization of control/communication system over the digital noiseless channel.

Theorem 2: Consider the scalar control/communication system described by the system (1) ($x_0 \in [-L_0, L_0]$), the digital noiseless channel, the proposed encoder and decoder pair and a controller that satisfies the Assumption 5. This system is surely asymptotic stabilizable at the point $x^* = 0$ provided the Assumptions 1,2 and 4 hold and $C > \max\{0, \Delta^*\}$.

Proof: In the following, it is shown that by applying the proposed coding scheme and controller law $\mathscr{G}(.)$ introduced in Assumption 5, the system is asymptotically stable in sure sense at point $x^* = 0$. Towards this goal, we study the convergence behavior of the sequence $G_{k|k} \triangleq \mathscr{G}(x_k) - \mathscr{G}(\hat{x}_{k|k})$. Note that in the case of the digital noiseless channel, the value of β_k, $k \geq 0$, equals to 1 and the packet length is determined as $R_k = r_k + \epsilon$. Therefore, we can write:

$$\begin{aligned} |G_{k|k}| \triangleq |\mathscr{G}(x_k) - \mathscr{G}(\hat{x}_{k|k})| &= |\mathscr{G}'(a)||x_k - \hat{x}_{k|k}|, \\ &\leq \max_{a \in \Omega_{k|k}} |\mathscr{G}'(a)||L_{k|k}|, \\ &\leq \max_{a \in \Omega_{k|k-1}} |\mathscr{G}'(a)||L_{k|k}|, \\ &\leq \frac{m_k}{2^{R_k}} |L_{k|k-1}|, \\ &\leq \frac{m_k \prod_{i=1}^{k} s_i}{2^{\sum_{i=0}^{k} R_i}} L_0, \end{aligned} \tag{18}$$

where the existence of the point a in $\Omega_{k|k}$ is proved by the Mean Value Theorem. The values of the packet length R_i leads to:

$$\begin{aligned} \sum_{i=0}^{k} R_i = \sum_{i=0}^{k} r_i + k\epsilon &\leq \log(m_0 \prod_{i=1}^{k} \frac{m_i}{m_{i-1}} s_i) + k\epsilon, \\ &= \log(m_k \prod_{i=1}^{k} s_i) + k\epsilon. \end{aligned} \tag{19}$$

Substituting (19) in (18) results in the convergence of the sequence $G_{k|k}$ to zero. Therefore, the system $x_{k+1} = f(x_k, \mathscr{G}(\hat{x}_{k|k}))$ can be seen as the system $x_{k+1} = f(x_k, \mathscr{G}(x_k) + G_{k|k})$ which based on the Assumption 5, converges to $x^* = 0$ for every $x_0 \in \psi_0 = [-L_0, L_0]$. Consequently, the system is asymptotically stable in sure sense at point $x^* = 0$. Note that the equation (19) also results in the convergence of the sequence $L_{k|k}$ to zero, since: $L_{k|k} = \frac{\prod_{i=1}^{k} s_i}{2^{\sum_{i=0}^{k} R_i}} L_0 \leq \frac{2^{-k\epsilon}}{m_k}$. Now, we compute the average length of the packets used in this stabilizing scheme. We know both sequences x_k and $L_{k|k}$ converge to zero; and therefore $\hat{x}_{k|k}$ also converges to zero. Therefore, equations (9)-(15) result in the convergence of s_k, m_k and R_k, respectively, to $\left|J_{f/x}(0,0)\right|$, $|\mathscr{G}'(0)|$, and $\max\{0, \log(\left|J_{f/x}(0,0)\right|)\} + \epsilon$.

Consequently, from the Cesaro Mean (Theorem 4.2.3 of [28]), it follows that the required average bit rate for such a transmission, which is equal to the average of packet lengths, is given by: $R_{av} = \lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} R_i = \max\{0, \log(\left|J_{f/x}(0,0)\right|)\} + \epsilon$. From Remark 1, it follows that $\log(\left|J_{f/x}(0,0)\right|)$ is the equilibrium Lyapunov exponent of the system (1) $\big($i.e., $\Delta^* = \log(\left|J_{f/x}(0,0)\right|)\big)$. This proves that $C = R_{av} > \max\{0, \Delta^*\}$ is a sufficient condition for the system (1) to be surely asymptotic stabilizable over the digital noiseless channel, since having this condition guarantees the existence of a coding scheme with the packet length $R_k = r_k + \epsilon$ which stabilizes the system. ■

The final theorem is concerned with the sufficient condition for the almost sure stabilization of the dynamic system (1) over the packet erasure channel. In order to achieve this condition, we consider the following additional assumption:

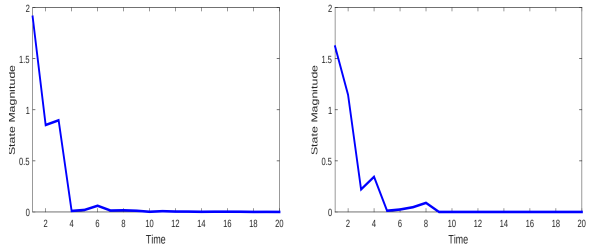
Assumption 6: The sequence r_k is assumed to be bounded.

Theorem 3: Consider the scalar control/communication system described by the system (1) ($x_0 \in [-L_0, L_0]$), the packet erasure channel with feedback acknowledgment and erasure probability γ, the proposed encoder and decoder pair and a controller satisfying the Assumption 5. Then, the point $x^* = 0$ is almost surely asymptotic stabilizable provided the Assumptions 1,2,4 and 6 hold and $C > \max\{0, \Delta^*\}$.

Proof: We show that using the proposed coding scheme and the controller $\mathscr{G}(\cdot)$ as proposed in the Assumption 5, the system is almost surely stable at the point $x^* = 0$. Towards this goal, we study the convergence behavior of the sequence $G_{k|k}$. By following the similar steps as we used in the proof of Theorem 2, we have: $|G_{k|k}| \leq \frac{m_k \prod_{i=0}^{k-1} s_i}{2^{\sum_{i=0}^{k} \beta_i R_i}} L_0$. Applying the proposed coding scheme with the transmitted packets length R_k, leads to the almost sure convergence of $|G_{k|k}|$ to zero (see the Appendix B). Therefore, based on the Assumption 5, the state of system $x_{k+1} = f(x_k, \mathscr{G}(\hat{x}_{k|k})) = f(x_k, \mathscr{G}(x_k) + G_{k|k})$ almost surely converges to point $x^* = 0$. Consequently, the system is stable in almost sure sense. Following the similar steps as used in the proof of Theorem 1 and Appendix B, we have almost sure convergence of $L_{k|k}$ to zero. Consequently, s_k, m_k, and r_k converge to, respectively, $\left|J_{f/x}(0,0)\right|$, $|\mathscr{G}'(0)|$, and $r^* \triangleq \max\{0, \log(\left|J_{f/x}(0,0)\right|)\}$ in almost sure sense. Now, taking the expected value from (16) leads to the equation $E[R_i] = \gamma E[R_{i-1}] + E[r_i] + \epsilon$ due to the in-dependency of β_{i-1} and R_{i-1}. Note that r_i is a bounded random variable which converges to r^*; therefore, by Bounded Convergence Theorem (Theorem 1.5.3 of [29]), $E[r_i]$ converges to r^*. Furthermore, as γ is less than 1, it can be deduced that $E[R_i]$ has a limit point as $\lim_{i\to\infty} E[R_i] = \frac{1}{(1-\gamma)}\big(r^* + \epsilon\big)$.

Therefore, based on the Cesaro Mean, the average bit rate and required channel capacity for such transmission, respectively, are obtained as $R_{av} = \lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} E[R_i] = \frac{1}{(1-\gamma)}\big(\max\{0, \log\left|J(\frac{f}{x})\right|_{(0,0)}\} + \epsilon\big)$ and $C = (1-\gamma) R_{av} = \max\{0, \log\left|J(\frac{f}{x})\right|_{(0,0)}\} + \epsilon$. Note that the system equilibrium Lyapunov exponent Δ^* equals to $\log(\left|J_{f/x}(0,0)\right|)$. Hence, $C > \max\{0, \Delta^*\}$ is a tight sufficient condition for almost sure stabilization of the control systems over the packet erasure channel. ■

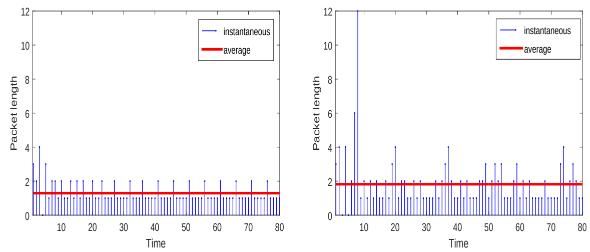
Hence, unlike [22], the tight bound on almost sure asymptotic stabilization of scalar nonlinear dynamic systems over the packet erasure channel is obtained. It should be noted that [22] studies the sufficient condition whose tightness was not proved. Finally, we study these theorems for the special case of discrete - time linear system of $x_{k+1} = f(x_k, u_k) = Ax_k + Bu_k$. By Remark 1, the equilibrium Lyapunov exponents for this system are the logarithm of the magnitudes of the eigen-



(a) For the digital noiseless channel

(b) For the packet erasure channel with $\gamma = 0.3$

Fig. 2: The state magnitude of the dynamic system



(a) For the digital noiseless channel

(b) For the packet erasure channel with $\gamma = 0.3$

Fig. 3: The length of transmitted packets

values of the matrix $J_{f/x}(0,0) = A$. Hence, the necessary and sufficient condition is reduced to the well-known eigenvalues-rate condition $C > \sum_i \max\{0, \log(|\lambda_i(A)|)\}$ (also appearing in e.g., [1], [2]).

V. Simulation results

There are many dynamic systems satisfying all the assumptions considered in this paper, e.g., the class of scalar nonlinear dynamic systems: $x_{k+1} = f(x_k) + u_k$ where $f(x)$ is continuously differentiable and its derivative is non-zero at the origin. For this system by choosing a control law as $u_k = -f(x_k)$, the Assumption 5 is also satisfied.

In order to illustrate the performance of the proposed coding scheme, we consider $f(x) = 2x^2 + 2x$ and apply the coding scheme on this system where $x_0 \in [-2, 2]$ is a random variable. It is assumed that $\hat{x}_{0|-1} = 0$. We simulate the magnitude of the state x_k over the packet erasure channel with erasure probabilities $\gamma = 0$ and $\gamma = 0.3$ using the proposed encoding/decoding scheme for the case of $\epsilon = 0.2$. The magnitude of the states presented in Fig. 2 illustrate the stability of the nonlinear system. Moreover, the instantaneous and average length of transmitted packets are presented in Fig. 3. Based on the simulation results, the average bit rate equals to 1.29(bit/time unit) and 1.82(bit/time unit), respectively, in the case of $\gamma = 0$ and $\gamma = 0.3$. It should be noted that $\Delta^* = 1$(bit/time unit) for this system; and therefore, the equality $(1-\gamma)R_{av} = \Delta^* + \epsilon$ approximately holds. In fact, the simulation results are consistent with Theorem 2 and Theorem 3, which present the condition $C = (1-\gamma)R_{av} > \Delta^*$ as a sufficient stabilization condition. In this simulation, we use the time sharing technique to obtain the integer length for transmitted packets.

VI. Conclusion

The stabilization problem of the noiseless nonlinear dynamic systems with random initial state over limited capacity communication channels were studied in this paper. It was proved that the stabilization of such systems over a memoryless channel is impossible if the channel capacity is less than the summation of the positive equilibrium Lyapunov exponents (i.e. $C < \sum_i \max\{0, \Delta_i^*\}$). For scalar nonlinear systems, it was shown that this condition is tight. It means that under the condition $C > \max\{0, \Delta^*\}$, there exist an encoder, decoder and a controller which make the system asymptotically stable in sure and almost sure senses, respectively, over the digital noiseless and the packet erasure channels.

Appendix A

In this appendix, we obtain the inequality between $h(z_{i+1}^u | w^i, \delta_i = 0)$ and $h(z_i^u | w^i, \delta_i = 0)$. Based on the structure for the control/communication system of Fig. 1, the sequence of control signal u^i is known at both transmitter and receiver sides at the time instant $i+1$. Therefore, given the variable z_i^s, the value of random vector z_{i+1}^u is determined based on the value of random vector z_i^u through the nonlinear function $z_{i+1}^u = d^u(z_i^s, z_i^u, u_i)$ which is a one to one function based on Lemma 1 with respect to z_i^u on the domain $\mathscr{N}_{n_u}$ for every $z_i^s \in \mathscr{N}_{n_s}$ and $u_i \in \mathscr{N}_d$.

Note that ϵ has been chosen so that $n+d$ dimensional sphere with center $(\underline{0}_n, \underline{0}_d)$ and radius ϵ is placed inside the subspace $\mathscr{N}$. Therefore, the condition $\delta_i = 0$ implies that given the variable z_i^s, the function $d^u(z_i^s, z_i^u, u_i)$ is one to one with respect to z_i^u and based on equation (1.25) in [30], we can write the following equation : $\mathcal{F}(z_{i+1}^u | z_i^s, w^i, \delta_i = 0) = \mathcal{F}(z_i^u | z_i^s, w^i, \delta_i = 0). \left|\frac{dz_{i+1}^u}{dz_i^u}\right|^{-1}$. We use the total derivative rule for $z_{i+1}^u = d^u(z_i^u, z_i^s, u_i)$ in order to obtain $\left|\frac{dz_{i+1}^u}{dz_i^u}\right|$ as $\frac{dz_{i+1}^u}{dz_i^u} = J_{d^u/z^u}(z_i, u_i) + J_{d^u/z^s}(z_i, u_i)\frac{dz_i^s}{dz_i^u}$. Note that $J_{d^u/z^u}(z,u)$ and $J_{d^u/z^s}(z,u)$ are continuous function with respect to both variables by the Assumption 2; and therefore, $\lim_{||(z,u)||\to 0} ||J_{d^u/z^u}(z,u) - A^u|| = 0$ and $\lim_{||(z,u)||\to 0} ||J_{d^u/z^s}(z,u)|| = 0$. Similarly, we can use the total derivative rule to prove that the sequence $\frac{dz_i^s}{dz_i^u}$ converges to zero vector. Therefore, by continuity of determinant function, we have, $\lim_{||(z,u)||\to 0} \left|\frac{dz_{i+1}^u}{dz_i^u}\right| = |A^u|$. Hence, the condition $\delta_i = 0$ (or equivalently, $||(z,u)|| \le \epsilon$) implies that there exists a constant value ϵ' such that $\left| \left|\frac{dz_{i+1}^u}{dz_i^u}\right| - |A^u| \right| \le \epsilon'$. We define ϵ_m as the minimum value of ϵ'. Consequently,

$$\mathcal{F}(z_{i+1}^u | z_i^s, w^i, \delta_i = 0) \ge \mathcal{F}(z_i^u | z_i^s, w^i, \delta_i = 0), (|A^u| + \epsilon_m)^{-1},$$
$$\mathcal{F}(z_{i+1}^u | z_i^s, w^i, \delta_i = 0) \le \mathcal{F}(z_i^u | z_i^s, w^i, \delta_i = 0), (|A^u| - \epsilon_m)^{-1}. \tag{20}$$

By multiplying (20) in $\mathcal{F}(z_i^s | w^i, \delta_i = 0)$ and taking integral with respect to z^s, we have:

$$\mathcal{F}(z_{i+1}^u | w^i, \delta_i = 0) \ge \mathcal{F}(z_i^u | w^i, \delta_i = 0), (|A^u| + \epsilon_m)^{-1},$$
$$\mathcal{F}(z_{i+1}^u | w^i, \delta_i = 0) \le \mathcal{F}(z_i^u | w^i, \delta_i = 0), (|A^u| - \epsilon_m)^{-1}. \tag{21}$$

Subsequently, the recursive relation for $h(z^u_{i+1}|w^i,\delta_i=0)$ is obtained as $h(z^u_{i+1}|w^i,\delta_i=0)\overset{(a)}{=}E\big[-\log\big(\mathcal{F}(z^u_{i+1}|w^i,\delta_i=0)\big)\big|\delta_i=0\big]\overset{(b)}{\geq}E\big[-\log\big(\mathcal{F}(z^u_i|w^i,\delta_i=0)(\big|A^u\big|-\epsilon_m)^{-1}\big)\big|\delta_i=0\big]=E\big[-\log\big(\mathcal{F}(z^u_i|w^i,\delta_i=0)\big)+\log(\big|A^u\big|-\epsilon_m)\big|\delta_i=0\big]=E\big[-\log\big(\mathcal{F}(z^u_i|w^i,\delta_i=0)\big)\big|\delta_i=0)\big]+\log(\big|A^u\big|-\epsilon_m)\overset{(c)}{=}h(z^u_i|w^i,\delta_i=0)+\log(\big|A^u\big|-\epsilon_m)$. The definition of conditional entropy results in (a) and (c). (b) is resulted from (21).

APPENDIX B

In this appendix, we prove the almost sure convergence of $G_{k|k}$ to zero. Since $r_i\leq\log(\frac{m_i}{m_{i-1}}s_i)$, the following equality holds:

$$|G_{k|k}|\leq\frac{m_k\prod_{i=0}^{k-1}s_i}{2^{\sum_{j=0}^{k}\beta_jR_j}}\leq 2^{\sum\limits_{i=0}^{k-1}r_i-\beta_jR_j}=2^{k\left(\frac{1}{k}\sum\limits_{i=1}^{k}r_i-\beta_iR_i\right)}. \tag{22}$$

From the proposed packet length $R_i=(1-\beta_{i-1})R_{i-1}+r_i+\epsilon$, we have:

$$\begin{aligned}\frac{1}{k}\sum_{i=1}^{k}r_i-\beta_iR_i=&\frac{1}{k}\sum_{i=1}^{k}(r_i+\epsilon)-\beta_iR_i-\epsilon\\=&\frac{1}{k}\sum_{i=1}^{k}\prod_{j=i}^{k}(1-\beta_j)(r_i+\epsilon)-\epsilon\\\leq&\frac{1}{k}\sum_{i=1}^{k}\prod_{j=i}^{k}(1-\beta_j)\bar{Q}-\epsilon.\end{aligned} \tag{23}$$

where the Assumption 6 results in the existence of the constant number $\bar{Q}$. Substituting (23) in (22) leads to the conclusion that if $\frac{1}{k}\mathcal{R}_k\triangleq\frac{1}{k}\sum_{i=1}^{k}\prod_{j=i}^{k}(1-\beta_j)\bar{Q}$ converges to zero almost surely, then the sequence $|G_{k|k}|$ also converges almost surely to zero. Note that $\mathcal{R}_k$ takes the values $0,\bar{Q},\dots,k\bar{Q}$, respectively, with probability $(1-\gamma),\gamma(1-\gamma),\dots,\gamma^k$. Hence, we can compute its second moment as $E[\mathcal{R}_k^2]=(1-\gamma)\bar{Q}^2\sum_{i=0}^{k-1}i^2\gamma^i+k^2\bar{Q}^2\gamma^k\leq\bar{Q}^2\sum_{i=0}^{k}i^2\gamma^i$. Using the Chebyshevś inequality (Theorem 1.6.4 of [29]), we have: $\mathbb{P}(\mathcal{R}_k>k\delta)<\frac{1}{k^2\delta^2}E[\mathcal{R}_k^2]\leq\frac{1}{k^2\delta^2}\bar{Q}^2\sum_{i=0}^{k}i^2\gamma^i$. Since the condition $\sum_{k=1}^{\infty}\mathcal{P}(\mathcal{R}_k>k\delta)\leq\infty$ is sufficient for almost sure convergence of $\frac{1}{k}\mathcal{R}_k$ to zero (Theorem 7.5 of [31]), we compute $\sum_{k=1}^{\infty}\frac{1}{k^2}\sum_{i=0}^{k}i^2\gamma^i=\frac{\gamma}{(1-\gamma)^2}\big(\sum_{k=1}^{\infty}\frac{1+\gamma}{k^2}-\sum_{k=1}^{\infty}\frac{(k+1)^2}{k^2}\gamma^k+\sum_{k=1}^{\infty}\frac{2k^2+2k-1}{k^2}\gamma^{k+1}-\sum_{k=1}^{\infty}\gamma^{k+2}\big)$.

Note that $\gamma<1$ and hence all terms in the above equation are finite. Therefore, almost sure convergence of $\frac{1}{k}\mathcal{R}_k$ to zero is proved and the sequence $G_{k|k}$ converges to zero almost surely.

REFERENCES

[1] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1056–1068, 2004.
[2] ——, "Control over noisy channels," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1196–1201, 2004.
[3] A. Zaidi, T. Oechtering, S. Yüksel, and M. Skoglund, "Stabilization of linear systems over Gaussian networks," *IEEE Trans. Autom. Control*, vol. 59, no. 9, pp. 2369–2384, 2014.
[4] P. Minero, L. Coviello, and M. Franceschetti, "Stabilization over Markov feedback channels: the general case," *IEEE Trans. Autom. Control*, vol. 58, no. 2, pp. 349 –362, 2013.
[5] V. Sanjaroon, A. Farhadi, B. H. Khalaj, and A. S. Motahari, "Estimation and stability over awgn channel in the presence of fading, noisy feedback channel and different sample rates," *Systems and Control Letters*, vol. 123, pp. 75–84, 2019.
[6] N. Ellia, "When bode meets shannon: control-oriented feedback communication schemes," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1477–1488, 2004.
[7] A. Matveev and A. Pogromsky, "Observation of nonlinear systems via finite capacity channels: constructive data rate limits," *Automatica*, vol. 70, pp. 217–229, 2016.
[8] A. Pogromsky and A. Matveev, "A topological entropy approach for observation via channels with limited data rate," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 14 416–14 421, 2011.
[9] D. Liberzon and S. Mitra, "Entropy and minimal data rates for state estimation and model detection," in *Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control*, 2016, pp. 247–256.
[10] A. Farhadi and N. U. Ahmed, "Tracking nonlinear noisy dynamic systems over noisy communication channels," *IEEE Trans. Commun.*, vol. 59, no. 4, pp. 955–961, 2011.
[11] V. Sanjaroon, A. Farhadi, A. S. Motahari, and B. H. Khalaj, "Estimation of nonlinear dynamic systems over communication channels," *IEEE Transactions on Automatic Control*, vol. 63, pp. 3024 – 3031, 2018.
[12] A. Diwadkar and U. Vaidya, "Limitation for nonlinear observation over erasure channel," *IEEE Trans. Autom. Control*, vol. 58, no. 2, pp. 454–459, 2013.
[13] 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, 2004.
[14] A. V. Savkin, "Analysis and synthesis of networked control systems: topological entropy, observability, robustness and optimal control," *Automatica*, vol. 42, no. 1, pp. 51–62, 2006.
[15] F. Colonius, "Minimal bit rates and entropy for exponential stabilization," *SIAM Journal on Control and Optimization*, vol. 50, no. 5, pp. 2988–3010, 2012.
[16] C. De Persis, "On stabilization of nonlinear systems under data rate constraints using output measurements," *International Journal of Robust and Nonlinear Control*, vol. 16, no. 6, pp. 315–332, 2006.
[17] U. Vadia and N. Elia, "Stabilization of nonlinear systems over packet-drop links: scalar case," *Systems and Control Letters*, pp. 959–966, 2012.
[18] Y. Sharon and D. Liberzon, "Input to state stabilizing controller for systems with coarse quantization," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 830–844, 2012.
[19] F. Colonius and C. Kawan, "Invariance entropy for control systems," *SIAM Journal on Control and Optimization*, vol. 48, no. 3, pp. 1701–1721, 2009.
[20] F. Colonius, C. Kawan, and G. Nair, "A note on topological feedback entropy and invariance entropy," *Systems & Control Letters*, vol. 62, no. 5, pp. 377–381, 2013.
[21] C. Kawan, "Data rate of nonlinear control systems and feedback entropy," *Encyclopedia of Systems and Control*, pp. 259–264, 2015.
[22] A. Parsa and A. Farhadi, "Measurement and control of nonlinear dynamic systems over the internet (iot): applications in remote control of autonomous vehicles," *Automatica*, vol. 95, pp. 93–103, 2018.
[23] K. You and L. Xie, "Minimum data rate for mean square stabilizability of linear systems with markovian packet losses," *IEEE Transactions on Automatic Control*, vol. 56, no. 4, pp. 772–785, 2011.
[24] 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. Autom. Control*, vol. 53, no. 4, pp. 1013–1019, 2008.
[25] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, "Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. part 1: theory," *Meccanica*, vol. 15, no. 1, pp. 9–20, 1980.
[26] G. Strang, *Linear Algebra and Its Applications*. Cengage Learning US, 2006.
[27] W. Rudin *et al.*, *Principles of mathematical analysis*. McGraw-hill New York, 1964, vol. 3.
[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 2006.
[29] R. Durrett, *Probability: Theory and Examples*. Cambridge university press, 2010.
[30] J. V. Michalowicz, J. M. Nichols, and F. Bucholtz, *Handbook of Differential Entropy*. CRC Press., 2013.
[31] H. Pishro-Nik, *Introduction to Probability, Statistics and Random processes*. Kappa Research, LLC, 2014.