Mean Square Stabilization of Linear Discrete-time Systems over Power Constrained Fading Channels

Liang Xu, Yilin Mo, Lihua Xie* and Nan Xiao

Abstract—**This paper considers the mean square stabilization problem of discrete-time linear time-invariant (LTI) systems over a power constrained fading channel. Fundamental limitations on the mean square stabilizability are obtained via information theoretic arguments. For scalar systems and two-dimensional systems, necessary and sufficient conditions for the mean square stabilizability are provided. Moreover, an adaptive time division multiple access (TDMA) communication scheme is designed for high-dimensional systems, which achieves a larger stabilizability region than the conventional TDMA communication scheme, and is proved to be optimal under certain situations.**

I. Introduction

Control over communication channels/networks has been a hot research topic in the past decade [1], [2], motivated by the rapid development of wireless communication technologies that enable the connection of geographically distributed systems and devices. Until now, there have been plentiful results that reveal requirements on communication channels to ensure the stabilizability of networked control systems. For noiseless digital channels, the result is the celebrated data rate theorem [3]. For noisy channels, the problem is complicated by the fact that different channel capacities are required under different stability definitions. For almost sure stability, the Shannon capacity in relation to unstable dynamics of a system constitutes the critical condition for its stabilizability [4]. While for moment stability, the anytime capacity is introduced to characterize the stabilizability conditions [5]. The anytime capacity has more stringent reliability requirements than the Shannon capacity. Different from the fact that the Shannon capacity can be calculated via the maximization of mutual information, in general, there is no systematic method to calculate anytime capacities of channels. Currently, stabilizability results for networked control only exist for specific communication channels when moment stability is concerned. For example, [6], [7] characterize the mean square stabilizability conditions for discrete-time linear time-invariant (LTI) systems over fading channels with linear control policies. [8]–[11] study the mean square stabilization problem over additive white Gaussian noise (AWGN) channels/networks and characterize the conditions to ensure mean square stabilizability. Specifically, the results stated above deal with fading or AWGN separately. While in wireless communications, it is more practical to consider them as a whole.

In this paper, we are interested in a communication channel which is subjected to both fading and AWGN, and also has a channel input power constraint. We aim to find the critical condition on the channel to ensure the mean square stabilizability of LTI systems. For scalar systems, the problem lies in how to design encoders/decoders to render the closed-loop system mean square stable. For AWGN channels, [11] propose encoder/decoder designs based on the Schalkwijk coding scheme [12], which utilize the noiseless channel feedback to consecutively refine the estimation error. They show that such encoding/decoding schemes can stabilize scalar unstable systems with the minimal channel capacity requirement indicated in [9]. In this paper, we show that a modification of this coding scheme can stabilize scalar systems controlled over power constrained fading channels. For vector systems, the difficulty is how to optimally allocate channel resources among sub-systems. When the channel is only with Gaussian noise, [11] employs a time-invariant allocation with the time division multiple access (TDMA) strategy to solve this problem. The transmission through the channel is scheduled periodically. During every period, each sub-system is allocated a fixed portion of transmission slots proportional to the logarithm of the magnitude of the corresponding unstable eigenvalue. It is shown that such allocation together with proper encoder/decoder pairs can stabilize the vector system. Moreover, from the results in [9], we know that such TDMA strategy is optimal, which means that the fixed allocation with the TDMA strategy provides the exact channel resource required for stabilization of each sub-system. However, when fading exists, since the channel may have different capacity at different time due to the stochastic nature of the fading, the time-invariant allocation fails to provide the critical channel resource for stabilization of each sub-system. Similar issue is also encountered in networked control over rate limited communication channels. When the digital channel is with constant data rate, [3] shows that the time-invariant allocation achieved by time-sharing is optimal. When the digital channel is with stochastic data rate, the time-invariant allocation in [13] is only sufficient. The stabilizability region achieved in [13] is a convex hull, which can be conservative even for two-dimensional systems. Therefore, we propose to use time-varying allocations to achieve larger stabilizability regions in this paper.

The contributions of this paper are three folds. Firstly, information theoretic analysis is conducted for the networked control system, which implies the existence of fundamental limitations imposed by the power constrained fading channel on stabilizing unstable LTI systems. Secondly, a communication protocol with proper encoder/decoder/scheduler for two-dimensional systems with unstable eigenvalues having different magnitudes is proposed, which provides the optimal allocation of channel resources to each sub-system. Finally, an adaptive TDMA communication scheme is proposed for general high-dimensional systems, which is shown to achieve a larger stabilizability region than the conventional TDMA scheme.

This paper is organized as follows. The problem formulation is provided in Section II. The fundamental limitation of stabilizability over a power constrained fading channel is studied in Section III. In Section IV, conditions for the mean square stabilizability are provided. Section V provides numerical illustrations. This paper ends with concluding remarks in Section VI.

Throughout the paper, $\mathbb{R}, \mathbb{R}^n, \mathbb{N}, \mathbb{N}^+$ represent sets of real scalars, n-dimensional real column vectors, natural numbers and positive natural numbers, respectively. A sequence $\{\chi_i\}_{i=0}^t$ is denoted by χ^t. A' denotes the transpose of matrix A. $\mathrm{E}\{\cdot\}$ represents the expectation operator. $\mathrm{E}_y\{\cdot\}$ denotes the expectation conditioned on the event $Y = y$. $\ln(\cdot)$ denotes the natural logarithm.

* Author for correspondence.

This work was partially supported by National Research Foundation of Singapore under grant NRF-CRP8-2011-03.

Liang Xu, Yilin Mo, Lihua Xie and Nan Xiao are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore. Email:lxu006@e.ntu.edu.sg, {ylmo, elhxie}@ntu.edu.sg, xiao0023@e.ntu.edu.sg

II. Problem Formulation

This paper studies the following discrete-time linear system

$$x_{t+1} = Ax_t + Bu_t, \tag{1}$$

where $x \in \mathbb{R}^n$ is the system state; $u \in \mathbb{R}$ is the control input and (A, B) is controllable. The initial state $x_0 = [x_{1,0}, \ldots, x_{n,0}]'$ is randomly generated from a Gaussian distribution with zero mean and bounded covariance matrix. Without loss of generality, the following assumption is made as in [11], [13].

Assumption 1: All the eigenvalues of A are either on or outside the unit circle.

The configuration of the networked control system is depicted in Fig. 1. The system state x_t is observed and encoded by the sensor/encoder $f_t(\cdot)$ and transmitted to the controller/decoder $h_t(\cdot)$ through a slow fading channel. The encoder $f_t(\cdot)$ and the decoder $h_t(\cdot)$ are allowed to be of any causal form and can use all the available information till time t to generate their output. The fading channel is modeled as

$$r_t = g_t s_t + w_t, \tag{2}$$

where s_t denotes the channel input, which has an average power constraint, i.e., $\mathrm{E}\{s_t^2\} \leq P$; r_t represents the channel output; $\{g_t\}$ is the independent and identically distributed (i.i.d.) channel fading with bounded mean and variance; $\{w_t\}$ is an AWGN with zero-mean and variance σ_w^2. We also assume that $x_0, \{g_t\}, \{w_t\}$ are independent; after each transmission, the instantaneous fading g_t is known at the decoder side and there exists a channel feedback that transmits one-step delayed information of r_t, g_t from the decoder to the encoder.

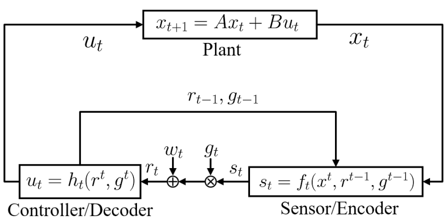

Fig. 1: Networked control over a power constrained fading channel

Remark 1: The knowledge of the fading level at the decoder side can be obtained for slow fading channels via receiver estimation [14]. Noiseless channel feedback may not be available in some settings. However, there are situations where this is a good assumption [12], [15]. Besides, channel feedback can be realized through the plant with suitably designed control policies in some scenarios [16]. Thus the assumption of noiseless channel feedback has been widely used in networked control research; see e.g., [5], [10], [15].

In this paper, for a given power constrained fading channel (2), we try to find requirements on the plant (1), such that there exists a causal encoder/decoder pair $\{f_t\}, \{h_t\}$ that can mean square stabilize the system, i.e., to render $\lim_{t\to\infty} \mathrm{E}\{x_t x_t'\} = 0$.

III. Fundamental Limitations

Since the entropy power provides a lower bound for the mean square value of the system state [9], we can treat the entropy power as a measure of the uncertainty of the system state and analyze its update, which provides a fundamental limitation of networked control over fading channels. The result is formalized in the following lemma, whose proof is given in Appendix A. The proof essentially follows the same steps as in [9], [10], [13], however, with some differences due to the channel structure.

Lemma 1: There exists a causal encoder/decoder pair $\{f_t\}$, $\{h_t\}$, such that the system (1) can be mean square stabilized over the channel (2), only if

$$(\det A)^{\frac{2}{n}} \mathrm{E}\left\{\mathrm{e}^{-\frac{2}{n}c_t}\right\} < 1, \tag{3}$$

where $c_t = \frac{1}{2}\ln(1 + \frac{g_t^2 P}{\sigma_w^2})$ is the instantaneous Shannon channel capacity of (2).

Let $\lambda_1, \ldots, \lambda_d$ be the distinct unstable eigenvalues (if λ_i is complex, we exclude from this list the complex conjugates $\bar{\lambda}_i$) of A in (1) with $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_d|$, and let m_i be the algebraic multiplicity of each λ_i. The real Jordan canonical form J of A then has the block diagonal structure $J = \mathrm{diag}(J_1, \ldots, J_d) \in \mathbb{R}^{n\times n}$ [3], where the block $J_i \in \mathbb{R}^{\mu_i \times \mu_i}$ and $|\det J_i| = |\lambda_i|^{\mu_i}$, with $\mu_i = m_i$ if $\lambda_i \in \mathbb{R}$, and $\mu_i = 2m_i$ otherwise. It is clear that we can equivalently study the following dynamical system instead of (1)

$$x_{t+1} = Jx_t + OBu_t, \tag{4}$$

for some similarity matrix O. Let $\mathcal{U} = \{1, \ldots, d\}$ denote the index set of unstable eigenvalues. Notice that each block J_i has an invariant real subspace $\mathcal{A}_{v_i}$ of dimension $a_i v_i$, for any $v_i \in \{0, \ldots, m_i\}$, where $a_i = 1$ if $\lambda_i \in \mathbb{R}$, and $a_i = 2$ otherwise. Consider the subspace $\mathcal{A}$ formed by taking the product of the invariant sub-spaces $\mathcal{A}_{v_i}$ for each real Jordan block. The total dimension of $\mathcal{A}$ is $v = \sum_{i\in\mathcal{U}} a_i v_i$. Denote by $x^{\mathcal{V}}$, the states of x belonging to $\mathcal{A}$. Then $x^{\mathcal{V}}$ evolves as

$$x_{t+1}^{\mathcal{V}} = J^{\mathcal{V}} x_{t+1}^{\mathcal{V}} + QOBu_t, \tag{5}$$

where Q is a transformation matrix and $|\det J^{\mathcal{V}}| = \prod_{i\in\mathcal{U}} |\lambda_i|^{a_i v_i}$. Since (4) is mean square stabilizable, (5) is also mean square stabilizable. In view of Lemma 1, the following fundamental limitations can be obtained [17].

Theorem 1: There exists a causal encoder/decoder pair $\{f_t\}, \{h_t\}$, such that the system (1) can be mean square stabilized over the channel (2) only if $[\ln|\lambda_1|, \ldots, \ln|\lambda_d|]' \in \mathbb{R}^d$ satisfy that for all $v_i \in \{0, \ldots, m_i\}$ and $i \in \mathcal{U}$ with $v = \sum_{i\in\mathcal{U}} a_i v_i$

$$\sum_{i\in\mathcal{U}} a_i v_i \ln|\lambda_i| < -\frac{v}{2} \ln \mathrm{E}\left\{\left(\frac{\sigma_w^2}{\sigma_w^2 + g_t^2 P}\right)^{\frac{1}{v}}\right\}. \tag{6}$$

Theorem 1 implies that even in the presence of a noiseless channel feedback, there still exists a fundamental limitation for the stabilizability of networked control over power constrained fading channels. Besides, for scalar systems where $A = \lambda_1$, $\ln|\lambda_1|$ should satisfy the following constraint to ensure mean square stabilizability

$$\ln|\lambda_1| < -\frac{1}{2} \ln \mathrm{E}\left\{\frac{\sigma_w^2}{\sigma_w^2 + g_t^2 P}\right\}. \tag{7}$$

Moreover, for two-dimensional systems with distinct eigenvalues λ_1, λ_2, the following requirement in addition to (7) should be satisfied

$$\ln|\lambda_1| + \ln|\lambda_2| < -\ln \mathrm{E}\left\{\left(\frac{\sigma_w^2}{\sigma_w^2 + g_t^2 P}\right)^{\frac{1}{2}}\right\}. \tag{8}$$

IV. Mean Square Stabilizability over Power Constrained Fading Channels

The existence of a noiseless channel feedback implies that there is no dual effect of control [18], i.e., separation between estimation and control holds, which will simplify the coding design. Indeed, we have the following lemma.

Lemma 2 ([10]): If (A, B) is controllable, and there exists an estimate $\hat{x}_t$ for the initial system state x_0, such that the estimation error $e_t = \hat{x}_t - x_0$ satisfies the following property,

$$\mathrm{E}\{e_t\} = 0, \tag{9}$$

$$\lim_{t\to\infty} A^t \mathrm{E}\left\{e_t e_t'\right\}(A')^t = 0, \tag{10}$$

then the system (1) can be mean square stabilized by the controller $u_t = K\left(A^t \hat{x}_t + \sum_{i=1}^t A^{t-i} B u_{i-1}\right)$ with K being selected such that $A + BK$ is stable.

Therefore, in the sequel we shall focus on the construction of communication/estimation algorithms which can achieve (9) and (10). To better convey our ideas, we start with scalar systems.

A. Scalar Systems

Theorem 2: Suppose $A = \lambda_1 \in \mathbb{R}$. There exists a causal encoder/decoder pair $\{f_t\}, \{h_t\}$, such that the system (1) can be mean square stabilized over the channel (2) if and only if (7) holds.

The necessity follows directly from Theorem 1. For the sufficiency, we can show that a variation of the Schalkwijk coding scheme [12] can stabilize the scalar system if (7) holds. The proof is similar to that of the AWGN case in [10] with some differences due to the existence of channel fading.

Proof: Suppose the estimation of x_0 formed by the decoder is $\hat{x}_t$ at time t and the estimation error is $e_t = \hat{x}_t - x_0$. The encoder is designed as

$$s_0 = \sqrt{\frac{P}{\sigma_{x_0}^2}} x_0, \quad s_t = \sqrt{\frac{P}{\sigma_{e_{t-1}}^2}}\left(\hat{x}_{t-1} - x_0\right),\ t \geq 1, \tag{11}$$

with $\sigma_{x_0}^2, \sigma_{e_{t-1}}^2$ representing the variance of x_0 and e_{t-1} respectively. The decoder is designed as

$$\hat{x}_0 = \sqrt{\frac{\sigma_{x_0}^2}{P}} r_0, \quad \hat{x}_t = \hat{x}_{t-1} - \frac{\mathrm{E}_{g_t}\{r_t e_{t-1}\}}{\mathrm{E}_{g_t}\{r_t^2\}} r_t,\ t \geq 1. \tag{12}$$

Since at time t, the encoder knows the one-step delayed channel output r_{t-1}, the fading g_{t-1} and the decoding law, it can thus simulate the decoder to obtain the estimate $\hat{x}_{t-1}$. With the designed encoder (11) and decoder (12), it is easy to show that $\mathrm{E}\left\{e_0\right\} = 0$ and $\mathrm{E}\left\{e_0^2\right\}$ is bounded. When $t \geq 1$, we have from (12) that

$$e_t = e_{t-1} - \frac{\mathrm{E}_{g_t}\{r_t e_{t-1}\}}{\mathrm{E}_{g_t}\{r_t^2\}} r_t. \tag{13}$$

By induction arguments, we have $\mathrm{E}\left\{e_t\right\} = 0$ for all $t \geq 1$. Thus (9) is satisfied. Denote $\hat{e}_{t-1} = \mathrm{E}_{g_t}\{r_t e_{t-1}\}/\mathrm{E}_{g_t}\{r_t^2\} r_t$. Since $\hat{e}_{t-1}$ is the minimal mean square error estimate (MMSE) of e_{t-1} based on r_t, from (13), we have $\mathrm{E}\{e_t^2\} = \mathrm{E}\{\mathrm{E}_{g_t}\{(e_{t-1} - \hat{e}_{t-1})^2\}\} \overset{(a)}{=} \mathrm{E}\{\frac{\sigma_w^2}{\sigma_w^2 + g_t^2 P} \mathrm{E}\{e_{t-1}^2\}\} = \mathrm{E}\{\frac{\sigma_w^2}{\sigma_w^2 + g_t^2 P}\}^t \mathrm{E}\{e_0^2\}$, where (a) is a direct consequence of the MMSE. Thus if $\lambda_1^2 \mathrm{E}\{\frac{\sigma_w^2}{\sigma_w^2 + g_t^2 P}\} < 1$, the designed encoder/decoder pair (11) and (12) can guarantee (10). In view of Lemma 2, the sufficiency is proved. ∎

Remark 2: Since g_t is known at the decoder side, we can show that a slight modification of the coding scheme in [11], where the expectation is replaced with the conditional expectation with respect to g_t, can stabilize the closed-loop system without channel feedback if (7) holds.

Remark 3: Theorem 2 indicates that the anytime capacity of the power constrained fading channel (2) corresponding to the anytime-reliability $2\ln|\lambda_1|$ is $C_a = -\frac{1}{2}\ln \mathrm{E}\{\frac{\sigma_w^2}{\sigma_w^2 + g_t^2 P}\}$. From Jensen's inequality, we know that $\mathrm{E}\{\mathrm{e}^{-2c_t}\} \geq \mathrm{e}^{-2\mathrm{E}\{c_t\}}$ and the equality holds if and only if c_t is a constant. Thus it follows that $C_a = \frac{1}{2}\ln\frac{1}{\mathrm{E}\left\{\mathrm{e}^{-2c_t}\right\}} \leq \frac{1}{2}\ln\frac{1}{\mathrm{e}^{-2\mathrm{E}\{c_t\}}} = \mathrm{E}\left\{c_t\right\} = C_{\mathrm{Shannon}}$, which means that the anytime capacity of the power constrained fading channel is no greater than its Shannon capacity. Besides, for AWGN channels, where c_t is a constant, we have that the anytime capacity is equal to its Shannon capacity, which coincides with the results in [5].

B. Two-Dimensional Systems

Theorem 3: Suppose $n = 2$. There exists a causal encoder/decoder pair $\{f_t\}, \{h_t\}$, such that the system (1) can be mean square stabilized over the channel (2) if and only if (6) holds.

In this subsection, we only provide the optimal communication scheme for two-dimensional systems with unstable eigenvalues having different magnitudes, i.e., $A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ with $\lambda_1, \lambda_2 \in \mathbb{R}$ and $|\lambda_1| > |\lambda_2| \geq 1$, and in view of Theorem 1, it suffices to show that a sufficient stabilizability condition is (7) and (8). For the case of two-dimensional systems with eigenvalues of equal magnitude, the communication scheme designed in the next subsection is shown to be optimal; see Corollary 1.

1) Communication Structure: Since there are two sources $x_{1,0}$, $x_{2,0}$, we design two encoder/decoder pairs in the communication scheme and also design a scheduler to multiplex the channel use. The i-th encoder/decoder pair is used to transmit the information of $x_{i,0}$. The scheduler determines which encoder/decoder pair should use the channel. Suppose at time t, the i-th encoder/decoder pair has access to the channel. The Encoder i first generates a symbol $s_{i,t}$ and transmits it to the decoder through the communication channel. The Decoder i then forms an estimate $\hat{x}_{i,t}$ based on the channel output $r_{i,t}$. The controller maintains an array $\hat{x}_t = [\hat{x}_{1,t}, \hat{x}_{2,t}]'$ that represents the most recent estimate of x_0, which is set to 0 at $t = 0$. When the information about $x_{i,0}$ is transmitted, only $\hat{x}_{i,t}$ is updated at the controller side. The controller applies the control law in Lemma 2 to the plant at every step.

The structure of the communication protocol is illustrated in Fig. 2, where t_k^i is the time when the i-th encoder/decoder pair is scheduled to use the channel for its k-th transmission.

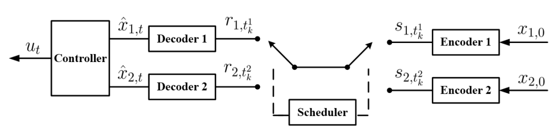

Fig. 2: Transmission protocol configuration

2) Encoder/Decoder Design: The following encoding/decoding strategy is used, which is modified from (11) and (12). The Encoder i is designed as

$$\begin{aligned} s_{i,t_0^i} &= \sqrt{\frac{P}{\sigma_{x_{i,0}}^2}} x_{i,0}, \\ s_{i,t_k^i} &= \sqrt{\frac{P}{\sigma_{e_{i,t_{k-1}^i}}^2}} (\hat{x}_{i,t_{k-1}^i} - x_{i,0}),\quad k \geq 1, \end{aligned} \tag{14}$$

where $\sigma_{x_{i,0}}^2$ and $\sigma_{e_{i,t}}^2$ represent the variance of $x_{i,0}$ and $e_{i,t}$, respectively, with $e_{i,t}$ being the i-th component of the estimation error e_t. The Decoder i satisfies

$$\begin{aligned} \hat{x}_{i,t_0^i} &= \sqrt{\frac{\sigma_{x_{i,0}}^2}{P}} r_{i,t_0^i}, \\ \hat{x}_{i,t_k^i} &= \hat{x}_{i,t_{k-1}^i} - \frac{\mathrm{E}_{g_{t_k^i}}\{r_{i,t_k^i} e_{i,t_{k-1}^i}\}}{\mathrm{E}_{g_{t_k^i}}\{r_{i,t_k^i}^2\}} r_{i,t_k^i},\quad k \geq 1. \end{aligned} \tag{15}$$

3) Scheduler Design: Throughout the paper, let $\delta = \frac{\sigma_w^2}{\sigma_w^2 + P}$. Define the scheduling indication vector as $\Phi(t) = [\phi_1(t), \phi_2(t)]'$ with $\phi_1(t), \phi_2(t) \in \{0, 1\}$ and $\phi_1(t) + \phi_2(t) = 1$. When the i-th encoder/decoder pair is scheduled to use the channel at time t, the variable $\phi_i(t)$ is set to 1, otherwise it is set to 0. Let $\Psi_p(i,j) = \prod_{k=i}^{j} (\frac{\sigma_w^2}{\sigma_w^2 + g_k^2 P})^{\phi_p(k)}$ with $p = 1, 2,\ i, j \in \mathbb{N}^+$

and $i \leq j$. Similar to the analysis for scalar systems, we can show that with the encoder (14) and the decoder (15), (9) always holds and $\mathrm{E}\{e_{i,t}^2\} = \mathrm{E}\{\Psi_i(t_0^i+1,t)\}\mathrm{E}\{e_{i,t_0^i}^2\}$ for $i=1,2$. Since $\phi_i(t)=0$ when $t<t_0^i$, to guarantee (10), we should design schedulers to ensure that, under the stochastic channel fading, $\lim_{t\to\infty}\mathrm{E}\left\{\lambda_1^{2t}\Psi_1(1,t)\right\}=0$ and $\lim_{t\to\infty}\mathrm{E}\left\{\lambda_2^{2t}\Psi_2(1,t)\right\}=0$, or equivalently $\lim_{t\to\infty}\mathrm{E}\{\lambda_1^{2t}\Psi_1(1,t)+\lambda_2^{2t}\Psi_2(1,t)\}=0$. Thus the scheduler should be designed to optimally allocate ϕ_1 and ϕ_2 to minimize $\lambda_1^{2t}\Psi_1(1,t)+\lambda_2^{2t}\Psi_2(1,t)$. The optimal allocation should satisfy $\sum_{j=1}^t\phi_2(j)\ln\frac{\sigma_w^2}{\sigma_w^2+g_j^2P}=2t\ln\frac{|\lambda_1|}{|\lambda_2|}+\sum_{j=1}^t\phi_1(j)\ln\frac{\sigma_w^2}{\sigma_w^2+g_j^2P}$, which is obtained by requiring $\lambda_1^{2t}\Psi_1(1,t)=\lambda_2^{2t}\Psi_2(1,t)$. To this end, Algorithm 1 is designed, which enforces ϕ_1 and ϕ_2 to meet the above requirement when t is sufficiently large.

In Algorithm 1, α_1 is the scheduler parameter to be defined latter; $T_k^{\mathrm{a}}=\sum_{j=1}^k T_j^{\mathrm{d}}$, $k\in\mathbb{N}^+$ is the time when k rounds of transmissions are completed and $T_0^{\mathrm{a}}=0$; T_k^{d} denotes the total time period to complete the k-th round of transmissions, i.e., $T_k^{\mathrm{d}}=T_k^1+T_k^2$. Here we assume that both the encoder and the decoder know the scheduling algorithm. Since the switching among transmissions in Algorithm 1 relies on the fading process, which is known to the encoder and the decoder, they are both aware of when to switch transmissions and what is the encoder/decoder pair currently using the channel. Thus we do not need to consider the coordination among encoders and decoders. The scheduled transmission periods are depicted in Fig. 3.

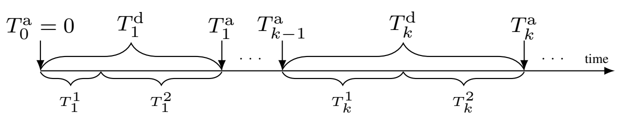

Fig. 3: Scheduled transmissions with Algorithm 1

Algorithm 1: Optimal Scheduler for Two-dimensional Systems

In the k-th round of transmissions

- The first encoder/decoder pair is scheduled to use the channel until
$$\sum_{t=T_{k-1}^{\mathrm{a}}+1}^{T_{k-1}^{\mathrm{a}}+T_k^1}\ln\frac{\sigma_w^2}{\sigma_w^2+g_t^2P}<\alpha_1\ln\delta \tag{16}$$
with T_k^1 being the minimal time period satisfying (16).
  - If
  $$\alpha_1\ln\delta+2T_k^1\ln\frac{|\lambda_1|}{|\lambda_2|}<0 \tag{17}$$
  the second encoder/decoder pair is scheduled to use the channel, until
  $$\sum_{t=T_{k-1}^{\mathrm{a}}+T_k^1+1}^{T_{k-1}^{\mathrm{a}}+T_k^1+T_k^2}\ln\frac{\sigma_w^2}{\sigma_w^2+g_t^2P}<2(T_k^1+T_k^2)\ln\frac{|\lambda_1|}{|\lambda_2|}+\alpha_1\ln\delta \tag{18}$$
  with T_k^2 being the minimal time period satisfying (18).
  - Otherwise, set $T_k^2=0$ and no transmission is carried out.
- Repeat this process.

It is clear from Algorithm 1 that T_i^{d} is independent of T_j^{d} and T_i^2 is independent of T_j^2 for any $i\neq j$, $i,j\in\mathbb{N}^+$. The switching condition (17) implies that if $T_k^1<T^{\mathrm{c}}:=\frac{\alpha_1\ln\delta}{2(\ln|\lambda_2|-\ln|\lambda_1|)}$, after the first encoder/decoder pair completes its transmission, the second encoder/decoder pair can use the channel. Otherwise, the first encoder/decoder pair continues to use the channel.

4) Scheduler Parameter Selection: If (7) holds, there exists θ_b with $0<\theta_b<1$ such that $\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\theta_b}\}=\lambda_1^{-2}$. Let $l(\theta_a)=2\theta_a\ln\frac{|\lambda_1|}{|\lambda_2|}-\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\theta_a}\}-2\ln|\lambda_1|$. If (8) holds, since $l(0)=-2\ln|\lambda_1|<0$, $l(\frac{1}{2})=-\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{1}{2}}\}-\ln|\lambda_1|-\ln|\lambda_2|>0$ and $l(\theta_a)$ is increasing in θ_a, there exists θ_a with $0<\theta_a<\frac{1}{2}$ such that $l(\theta_a)=0$, i.e., $\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\theta_a}\}=\lambda_1^{2(\theta_a-1)}\lambda_2^{-2\theta_a}$. The positive constant α_1 is then selected to satisfy
$$\alpha_1>\max\{\frac{-\ln(\lambda_1^{2(2-\theta_a)}\lambda_2^{2\theta_a})-\ln 4}{(1-2\theta_a)\ln\delta},\frac{-\ln\lambda_1^2-\ln 2}{(1-\theta_b)\ln\delta}\}. \tag{19}$$

5) Proof of Theorem 3: The necessity follows from Theorem 1. The gist of the sufficiency proof is to show that during one round of transmissions, the average value of $\lambda_p^{2t}\Psi_p(1,t)$ is smaller than one, i.e., $\mathrm{E}\{\lambda_p^{2T_1^{\mathrm{d}}}\Psi_p(1,T_1^{\mathrm{d}})\}<1$ for $p=1,2$. Since the transmission is scheduled periodically and $\{T_k^{\mathrm{d}}\}$ is i.i.d., we may expect that $\lim_{t\to\infty}\mathrm{E}\left\{\lambda_p^{2t}\Psi_p(1,t)\right\}=0$ holds in the long run, which together with Lemma 2 can guarantee the mean square stabilizability. The detailed proof is given as follows. Let $W_k=\ln\frac{\sigma_w^2}{\sigma_w^2+g_k^2P}$ and $L_t=\sum_{k=1}^t W_k$. Then it is immediate from (16) that T_1^1 is the first time such that $L_{T_1^1}<\varphi_1T_1^1+\gamma_1$ with $\varphi_1=0$ and $\gamma_1=\alpha_1\ln\delta$. Since there exist $0<\theta_a<\frac{1}{2}$, $0<\theta_b<1$ such that $\mathrm{E}\{e^{\theta_a(W_k-\varphi_1)}\}=\lambda_1^{2(\theta_a-1)}\lambda_2^{-2\theta_a}$, $\mathrm{E}\{\mathrm{e}^{\theta_b(W_k-\varphi_1)}\}=\lambda_1^{-2}$, from Lemma 4 in Appendix, we have
$$\mathrm{E}\{\lambda_1^{2(1-\theta_a)T_1^1}\lambda_2^{2\theta_aT_1^1}\}\leq\lambda_1^{2(1-\theta_a)}\lambda_2^{2\theta_a}\delta^{-\alpha_1\theta_a}, \tag{20}$$
$$\mathrm{E}\{\lambda_1^{2T_1^1}\}\leq\lambda_1^2\delta^{-\alpha_1\theta_b}. \tag{21}$$

Suppose T_1^1 is known and $T_1^1<T^{\mathrm{c}}$. Let $L_t=\sum_{k=T_1^1+1}^{T_1^1+t}W_k$. In view of the stopping condition (18), we know that T_1^2 is the first time starting from T_1^1 that $L_{T_1^2}<\varphi_2T_1^2+\gamma_2$ with $\varphi_2=2\ln\frac{|\lambda_1|}{|\lambda_2|}$ and $\gamma_2=2T_1^1\ln\frac{|\lambda_1|}{|\lambda_2|}+\alpha_1\ln\delta$. Since $\mathrm{E}\{\mathrm{e}^{\theta_a(W_k-\varphi_2)}\}=\lambda_1^{-2}$, in view of Lemma 4 in Appendix, we have
$$\mathrm{E}_\zeta\{\lambda_1^{2T_1^2}\}<\lambda_1^2\mathrm{e}^{-\theta_a\gamma_2}. \tag{22}$$
where ζ denote the event $T_1^1<T^{\mathrm{c}}$. Since $\theta_a<1$, when $T_1^1\geq T^{\mathrm{c}}$, we have $2T_1^1(\theta_a-1)\ln\frac{|\lambda_1|}{|\lambda_2|}\leq 2T^{\mathrm{c}}(\theta_a-1)\ln\frac{|\lambda_1|}{|\lambda_2|}<\alpha_1(1-\theta_a)\ln\delta+\ln 2+2\ln|\lambda_1|$. Rearranging both sides and applying the natural exponential function, we have $\Omega:=\lambda_2^{2T_1^1}-2\lambda_1^{2(1+T_1^1)}\mathrm{e}^{-\theta_a\gamma_2}\delta^{\alpha_1}<0$. In view of the conditional expectation, we have
$$\begin{aligned}
&\mathrm{E}\{\sum_{i=1}^2\lambda_i^{2T_1^{\mathrm{d}}}\Psi_i(1,T_1^{\mathrm{d}})\}\leq\mathrm{E}\{\lambda_1^{2T_1^{\mathrm{d}}}\delta^{\alpha_1}+\lambda_2^{2T_1^{\mathrm{d}}}\Psi_2(1,T_1^{\mathrm{d}})\}\\
&=\mathrm{E}\{\mathrm{E}_\zeta\{\lambda_1^{2T_1^{\mathrm{d}}}\delta^{\alpha_1}+\lambda_2^{2T_1^{\mathrm{d}}}\Psi_2(1,T_1^{\mathrm{d}})\}\}\\
&\quad+\mathrm{E}\{\mathrm{E}_\xi\{\lambda_1^{2T_1^{\mathrm{d}}}\delta^{\alpha_1}+\lambda_2^{2T_1^{\mathrm{d}}}\Psi_2(1,T_1^{\mathrm{d}})\}\}\\
&\overset{(a)}{\leq}\mathrm{E}\{\mathrm{E}_\zeta\{2\lambda_1^{2(T_1^1+T_1^2)}\delta^{\alpha_1}\}\}+\mathrm{E}\{\mathrm{E}_\xi\{\lambda_1^{2T_1^1}\delta^{\alpha_1}+\lambda_2^{2T_1^1}\}\}\\
&\overset{(b)}{\leq}\mathrm{E}\{\mathrm{E}_\zeta\{2\lambda_1^{2(1+T_1^1)}\mathrm{e}^{-\theta_a\gamma_2}\delta^{\alpha_1}\}\}+\mathrm{E}\{\mathrm{E}_\xi\{\lambda_1^{2T_1^1}\delta^{\alpha_1}+\lambda_2^{2T_1^1}\}\}\\
&=\mathrm{E}\{2\lambda_1^{2(1+T_1^1)}\mathrm{e}^{-\theta_a\gamma_2}\delta^{\alpha_1}\}+\mathrm{E}\{\mathrm{E}_\xi\{\lambda_1^{2T_1^1}\delta^{\alpha_1}+\Omega\}\}\\
&\overset{(c)}{\leq}2\lambda_1^2\delta^{\alpha_1(1-\theta_a)}\mathrm{E}\{\lambda_1^{2(1-\theta_a)T_1^1}\lambda_2^{2\theta_aT_1^1}\}+\mathrm{E}\{\mathrm{E}_\xi\{\lambda_1^{2T_1^1}\delta^{\alpha_1}\}\}\\
&\overset{(d)}{\leq}2\lambda_1^{2(2-\theta_a)}\lambda_2^{2\theta_a}\delta^{(1-2\theta_a)\alpha_1}+\lambda_1^2\delta^{(1-\theta_b)\alpha_1},
\end{aligned}$$
where ξ denotes the event $T_1^1\geq T^{\mathrm{c}}$; (a) follows from (18); (b) follows from (22); (c) follows from the fact that when $T_1^1\geq T^{\mathrm{c}}$, $\Omega<0$; (d) follows from (20) and (21). Since $\delta^{1-2\theta_a}<1$ and $\delta^{1-\theta_b}<1$, if α_1 is selected to satisfy (19), we have that

$\lambda_1^2\delta^{(1-\theta_b)\alpha_1} < \frac{1}{2}$, $2\lambda_1^{2(2-\theta_a)}\lambda_2^{2\theta_a}\delta^{(1-2\theta_a)\alpha_1} < \frac{1}{2}$, which guarantees $\mathrm{E}\{\lambda_1^{2T_1^{\mathrm{d}}}\Psi_1(1,T_1^{\mathrm{d}}) + \lambda_2^{2T_1^{\mathrm{d}}}\Psi_2(1,T_1^{\mathrm{d}})\} < 1$. Thus we have

$$\mathrm{E}\{\lambda_1^{2T_1^{\mathrm{d}}}\Psi_1(1,T_1^{\mathrm{d}})\} < 1, \quad \mathrm{E}\{\lambda_2^{2T_1^{\mathrm{d}}}\Psi_2(1,T_1^{\mathrm{d}})\} < 1. \tag{23}$$

Since $\Psi_p(1,T_k^{\mathrm{a}}) = \prod_{j=1}^{k}\Psi_p(T_{j-1}^{\mathrm{a}}+1, T_{j-1}^{\mathrm{a}}+T_j^{\mathrm{d}})$ and $\{\Psi_p(T_{j-1}^{\mathrm{a}}+1, T_{j-1}^{\mathrm{a}}+T_j^{\mathrm{d}})\}_{j=1}^{k}$ are i.i.d., we have

$$\begin{aligned}
&\sum_{k=0}^{\infty}\mathrm{E}\{\sum_{j=1}^{T_{k+1}^{\mathrm{d}}}\lambda_p^{T_k^{\mathrm{a}}+j}\Psi_p(1,T_k^{\mathrm{a}})\}\\
&=\sum_{k=0}^{\infty}\mathrm{E}\{\sum_{j=1}^{T_{k+1}^{\mathrm{d}}}\lambda_p^{T_0^{\mathrm{d}}+\cdots+T_k^{\mathrm{d}}+j}\prod_{j=1}^{k}\Psi_p(T_{j-1}^{\mathrm{a}}+1, T_{j-1}^{\mathrm{a}}+T_j^{\mathrm{d}})\}\\
&=\sum_{k=0}^{\infty}\mathrm{E}\{\frac{\lambda_p^{T_{k+1}^{\mathrm{d}}+2}-\lambda_p^2}{\lambda_p^2-1}\}\mathrm{E}\{\lambda_p^{T_1^{\mathrm{d}}}\Psi_p(1,T_1^{\mathrm{d}})\}^k,
\end{aligned} \tag{24}$$

for $p = 1,2$. In view of (23), we further have that $\mathrm{E}\{\sum_{t=1}^{\infty}(\lambda_1^{2t}\Psi_1(1,t) + \lambda_2^{2t}\Psi_2(1,t))\} = \sum_{k=0}^{\infty}\mathrm{E}\{\sum_{j=1}^{T_{k+1}^{\mathrm{d}}}(\lambda_1^{T_k^{\mathrm{a}}+j}\Psi_1(1,T_k^{\mathrm{a}}+j) + \lambda_2^{T_k^{\mathrm{a}}+j}\Psi_2(1,T_k^{\mathrm{a}}+j))\} \le \sum_{k=0}^{\infty}\mathrm{E}\{\sum_{j=1}^{T_{k+1}^{\mathrm{d}}}(\lambda_1^{T_k^{\mathrm{a}}+j}\Psi_1(1,T_k^{\mathrm{a}}) + \lambda_2^{T_k^{\mathrm{a}}+j}\Psi_2(1,T_k^{\mathrm{a}}))\} < \infty$, which implies that $\lim_{t\to\infty}\mathrm{E}\left\{\lambda_1^{2t}\Psi_1(1,t) + \lambda_2^{2t}\Psi_2(1,t)\right\} = 0$. The proof of sufficiency is completed. ■

C. High-Dimensional Systems

For general n-dimensional systems, the communication structure is designed similarly to that of the two-dimensional systems. There are n encoder/decoder pairs of the form (14) (15) to transmit the information of $x_{i,0}$, $i = 1,\ldots,n$. A scheduler is designed to multiplex the channel use. Define ϕ_i, $\Psi_i(\cdot,\cdot)$, $i = 1,\ldots,n$ analogously to the two-dimensional case. Similarly, we can prove that with such communication structure, (9) always holds and to guarantee (10), we only need to ensure that, $\lim_{t\to\infty}\mathrm{E}\left\{\lambda_i^{2t}\Psi_i(1,t)\right\} = 0$ for all $i = 1,\ldots,n$, or equivalently, $\lim_{t\to\infty}\mathrm{E}\left\{\sum_{i=1}^{n}\lambda_i^{2t}\Psi_i(1,t)\right\} = 0$. Thus the schedulers should be designed to optimally allocate ϕ_i to minimize $\sum_{i=1}^{n}\lambda_i^{2t}\Psi_i(1,t)$. The optimal choice of ϕ_i^* should satisfy $\sum_{j=1}^{t}\phi_i^*(j)\ln\frac{\sigma_w^2}{\sigma_w^2+g_j^2P} = (\sum_{j=1}^{t}\ln\frac{\sigma_w^2}{\sigma_w^2+g_j^2P} + 2t\sum_{i=1}^{n}\ln|\lambda_i|)/n - 2t\ln|\lambda_i|$. However ϕ_i^* is determined by $\sum_{j=1}^{t}\ln\frac{\sigma_w^2}{\sigma_w^2+g_j^2P}$, which is not causally available when transmitting $x_{i,0}$ at any time $k < t$. When $n = 2$, we can achieve the desired optimal allocation by first fixing ϕ_1 to be such that $\sum_{j=1}^{T_1^1}\phi_1(j)\ln(\frac{\sigma_w^2}{\sigma_w^2+g_j^2P}) < \alpha_1\ln\delta$ and then requiring ϕ_2 to achieve (18). However, this method is not applicable to the case of $n \ge 3$. In the following, we propose a scheduler design for general high-dimensional systems and show that such scheduling algorithm is optimal under certain situations.

Theorem 4: There exists a causal encoder/decoder pair $\{f_t\},\{h_t\}$, such that the system (1) can be mean square stabilized over the channel (2) if there exist β_i, $i = 1,\ldots,d$, with $0 < \beta_i \le 1$ and $\sum_{i=1}^{d}\beta_i = 1$, such that for all $i \in \mathcal{U}$,

$$\ln|\lambda_i| < -\frac{1}{2}\ln\mathrm{E}\left\{\left(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P}\right)^{\frac{\beta_i}{\mu_i}}\right\}. \tag{25}$$

The above stabilizability result is achieved via an adaptive TDMA scheduler. Different from the TDMA scheduler used in [17], the adaptive TDMA scheduler used here is adapted to the fading process. It switches the transmission only if certain stopping conditions are satisfied. By incorporating the information of the fading process, a larger stabilizability region is achieved. The detailed scheduler design and stability analysis is given as follows.

1) Scheduling Algorithm: The scheduler is described in Algorithm 2, where the parameters α_i, $i = 1,\ldots,n$ are defined in the sequel; $T_k^{\mathrm{a}} = \sum_{j=1}^{k}T_j^{\mathrm{d}}$, $k \in \mathbb{N}^+$ is the time when k rounds of transmissions are completed and $T_0^{\mathrm{a}} = 0$, and T_k^{d} denotes the total time period to complete the k-th round of transmissions, i.e. $T_k^{\mathrm{d}} = \sum_{i=1}^{n}T_k^i$. Since the fading $\{g_t\}$ is i.i.d., it is clear from Algorithm 2 that T_k^i is independent of T_k^j, for any $i \ne j$, $i,j \in \{1,2,\ldots,n\}$, $k \in \mathbb{N}^+$ and the random variables $\{T_1^{\mathrm{d}}, T_2^{\mathrm{d}}, \ldots\}$ are i.i.d..

Algorithm 2: Adaptive TDMA Scheduler for n-dimensional Systems

In the k-th round of transmissions

- The first encoder/decoder pair is scheduled to use the channel, until
$$\sum_{t=T_{k-1}^{\mathrm{a}}+1}^{T_{k-1}^{\mathrm{a}}+T_k^1}\ln\frac{\sigma_w^2}{\sigma_w^2+g_t^2P} < \alpha_1\ln\delta, \tag{26}$$
with T_k^1 being the minimal time period satisfying (26).
- …
- The j-th encoder/decoder pair is scheduled to use the channel, until
$$\sum_{t=T_{k-1}^{\mathrm{a}}+T_k^1+\cdots+T_k^{j-1}+1}^{T_{k-1}^{\mathrm{a}}+T_k^1+\cdots+T_k^{j-1}+T_k^j}\ln\frac{\sigma_w^2}{\sigma_w^2+g_t^2P} < \alpha_j\ln\delta, \tag{27}$$
with T_k^j being the minimal time period satisfying (27).
- …
- The n-th encoder/decoder pair is scheduled to use the channel, until
$$\sum_{t=T_{k-1}^{\mathrm{a}}+T_k^1+\cdots+T_k^{n-1}+1}^{T_{k-1}^{\mathrm{a}}+T_k^1+\cdots+T_k^{n-1}+T_k^n}\ln\frac{\sigma_w^2}{\sigma_w^2+g_t^2P} < \alpha_n\ln\delta, \tag{28}$$
with T_k^n being the minimal time period satisfying (28).
- Repeat this process.

---

2) Scheduler Parameter Selection: If (25) holds, there exist θ_i, $i = 1,\ldots,d$ with $0 \le \theta_i < \frac{\beta_i}{\mu_i}$, such that $\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\theta_i}\} = |\lambda_i|^{-2}$. The positive constants α_j, $j = 1,\ldots,n$ are selected as follows: if $x_{j,0}$ is the j-th component of x_0 in (4) that corresponds to the eigenvalue λ_i, $i \in \mathcal{U}$, then α_j is selected to be

$$\alpha_j = -\frac{2n\beta_i}{\mu_i\ln\delta}(\max_{k\in\mathcal{U}}\frac{\ln|\lambda_k|}{\beta_k/\mu_k-\theta_k}+\iota), \quad j = 1,\ldots,n, \tag{29}$$

with ι being an arbitrary positive constant.

3) Proof of Theorem 4: Here we only consider the case that $\lambda_1,\ldots,\lambda_d$ are real and $m_i = \mu_i = 1$. We can easily extend the analysis to other cases by combining the following analysis with the argument used in Chapter 2 of [2]. The sufficiency proof is focused on showing that $\lim_{t\to\infty}\mathrm{E}\left\{\lambda_i^{2t}\Psi_i(1,t)\right\} = 0$ for all $i = 1,\ldots,n$ under Algorithm 2. Similar to the derivation of (21), with Algorithm 2, we can show that $\mathrm{E}\{\lambda_i^{2T_1^j}\} \le \delta^{-\alpha_j\theta_i}\lambda_i^2$. Since $T_1^1, T_1^2, \ldots, T_1^n$ are independent with each other, we further have $\mathrm{E}\{\lambda_i^{2\sum_{j=1}^{n}T_1^j}\delta^{\alpha_i}\} \le \delta^{\alpha_i-\theta_i\sum_{j=1}^{n}\alpha_j}\lambda_i^{2n}$. If α_i is selected as (29), then $\sum_{i=1}^{n}\alpha_i = -\frac{2n}{\ln\delta}(\max_j\frac{\ln|\lambda_j|}{\beta_j-\theta_j}+\iota)$ and $\alpha_i/(\sum_{j=1}^{n}\alpha_j) = \beta_i$ for all $i = 1,\ldots,n$. Thus we have

$$\begin{aligned}
&\mathrm{E}\left\{\lambda_i^{2\sum_{j=1}^{n}T_1^j}\delta^{\alpha_i}\right\} \le (\delta^{\beta_i-\theta_i})^{\sum_{j=1}^{n}\alpha_j}\lambda_i^{2n}\\
&=(\delta^{\beta_i-\theta_i})^{-\frac{2n}{\ln\delta}(\max_j\frac{\ln|\lambda_j|}{\beta_j-\theta_j}+\iota)}(\delta^{\beta_i-\theta_i})^{\frac{2n}{\ln\delta}\frac{\ln|\lambda_i|}{\beta_i-\theta_i}}\\
&=(\delta^{\beta_i-\theta_i})^{\frac{2n}{\ln\delta}(\frac{\ln|\lambda_i|}{\beta_i-\theta_i}-\max_j\frac{\ln|\lambda_j|}{\beta_j-\theta_j}-\iota)}.
\end{aligned}$$

Since $\theta_i < \beta_i$ and $0 < \delta < 1$, we have

$$\mathrm{E}\left\{\lambda_i^{2T_1^{\mathrm{d}}}\delta^{\alpha_i}\right\} = \mathrm{E}\left\{\lambda_i^{2\sum_{j=1}^{n} T_1^j}\delta^{\alpha_i}\right\} < 1, \quad (30)$$

for all $i = 1, \ldots, n$. Since $\Psi_i(1, T_k^{\mathrm{a}}) = \prod_{j=1}^{k} \Psi_i(T_{j-1}^{\mathrm{a}} + 1, T_{j-1}^{\mathrm{a}} + T_j^{\mathrm{d}})$ and $\Psi_i(T_{j-1}^{\mathrm{a}} + 1, T_{j-1}^{\mathrm{a}} + T_j^{\mathrm{d}}) < \delta^{\alpha_i}$ for any $j \in \mathbb{N}^+$, in view of (30), we have $\mathrm{E}\{\sum_{t=1}^{\infty}\lambda_i^{2t}\Psi_i(1,t)\} = \sum_{k=0}^{\infty}\mathrm{E}\{\sum_{j=1}^{T_{k+1}^{\mathrm{d}}}\lambda_i^{2(T_k^{\mathrm{a}}+j)}\Psi_i(1, T_k^{\mathrm{a}} + j)\} < \sum_{k=0}^{\infty}\mathrm{E}\{\sum_{j=1}^{T_{k+1}^{\mathrm{d}}}\lambda_i^{2(T_k^{\mathrm{a}}+j)}\prod_{j=1}^{k}\Psi_i(T_{j-1}^{\mathrm{a}} + 1, T_{j-1}^{\mathrm{a}} + T_j^{\mathrm{d}})\} < \sum_{k=0}^{\infty}\mathrm{E}\{\sum_{j=1}^{T_{k+1}^{\mathrm{d}}}\lambda_i^{2(T_k^{\mathrm{a}}+j)}\delta^{k\alpha_i}\} = \sum_{k=0}^{\infty}\mathrm{E}\{\lambda_i^{2T_1^{\mathrm{d}}}\delta^{\alpha_i}\}^k\mathrm{E}\{(\lambda_i^{2T_{k+1}^{\mathrm{d}}+2} - \lambda_i^2)/(\lambda_i^2 - 1)\} < \infty$. Thus $\lim_{t\to\infty}\mathrm{E}\left\{\lambda_i^{2t}\Psi_i(1,t)\right\} = 0$ for all $i = 1, \ldots, n$. The proof of sufficiency is completed. ∎

Remark 4: The stabilizability conditions in the derived theorems involve the calculation of the expectation $\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\beta}\}$ for some β. For some fading distributions, we can give the closed form of this term. For example, when $g_t \sim \text{Bernoulli}(\epsilon)$[1], this term is given by $(1-\epsilon)(\frac{\sigma_w^2}{\sigma_w^2+P})^{\beta} + \epsilon$. For other fading distributions that are not possible to calculate the closed forms, this term can be evaluated numerically via MATLAB or Mathematica.

Remark 5: In Theorem 4, the stabilizability condition is expressed in terms of of parameters $\{\beta_i\}_{i\in\mathcal{U}}$. β_i has the physical interpretation that it represents the fraction of channel resources that is allocated to the sub-dynamics corresponding to the eigenvalue λ_i. For the given communication channel and system matrix, the existence of $\{\beta_i\}_{i\in\mathcal{U}}$ can be checked via the following feasibility problem

$$\exists\beta_i > 0, i = 1, \ldots, d$$

$$\text{s.t.} \sum_{i=1}^{d}\beta_i = 1 \quad (31)$$

$$|\lambda_i|^2 < l_i(\beta_i), \quad i = 1, \ldots, d \quad (32)$$

with $l_i(\beta_i) := \mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{\beta_i}{\mu_i}}\}^{-1}$. Since $l_i(\beta_i)$ is increasing in β_i and $l_i(0) = 1 \le |\lambda_i|^2$, there exists $\beta_i^* \ge 0$ such that $l_i(\beta_i^*) = |\lambda_i|^2$(binary search can be used to find equation roots to obtain β_i^*). In view of (32), any feasible β_i must satisfy that $\beta_i > \beta_i^*$. If $\sum_i\beta_i^* \ge 1$, there exists no feasible solution since (31) is violated. Otherwise, one feasible solution is given by $\beta_i = \frac{\beta_i^*}{\sum_i\beta_i^*}$.

Remark 6: Theorem 4 indicates that the stabilzable region of $[\ln|\lambda_1|, \ldots, \ln|\lambda_d|]' \in \mathbb{R}^d$ for a given power constrained fading channel achieved with Algorithm 2 is

$$\mathcal{S} = \cup_{\beta_i>0,\sum_i\beta_i=1} \times_{i\in\mathcal{U}} [0, -\frac{1}{2}\ln\mathrm{E}\left\{\left(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P}\right)^{\frac{\beta_i}{\mu_i}}\right\}),$$

where $\times$ denotes the Cartesian product. We can prove that $\mathcal{S}$ is convex. Suppose $\mathsf{x} = [\mathsf{x}_1, \ldots, \mathsf{x}_d]' \in \mathcal{S}$ and $\mathsf{y} = [\mathsf{y}_1, \ldots, \mathsf{y}_d]' \in \mathcal{S}$. Then there exist $[\varrho_1, \ldots, \varrho_d]'$ with $\varrho_i > 0$, $\sum_i\varrho_i = 1$ and $[\eta_1, \ldots, \eta_d]'$ with $\eta_i > 0$, $\sum_i\eta_i = 1$ such that $\mathsf{x}_i < -\frac{1}{2}\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{\varrho_i}{\mu_i}}\}$, $\mathsf{y}_i < -\frac{1}{2}\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{\eta_i}{\mu_i}}\}$ for $i = 1, \ldots, d$. Let $\mathsf{z} = [\mathsf{z}_1, \ldots, \mathsf{z}_d]' = c\mathsf{x} + (1-c)\mathsf{y}$ with $0 < c < 1$, then $z_i = cx_i + (1-c)y_i$ and

$$z_i < -\frac{c}{2}\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{\varrho_i}{\mu_i}}\} - \frac{1-c}{2}\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{\eta_i}{\mu_i}}\}$$
$$= -\frac{1}{2}\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{\varrho_i}{\mu_i}}\}^c\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{\eta_i}{\mu_i}}\}^{1-c}$$
$$\overset{(a)}{\le} -\frac{1}{2}\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{c\varrho_i+(1-c)\eta_i}{\mu_i}}\},$$

where (a) follows from the Hölder's inequality. Thus there exist β_is with $\beta_i = c\varrho_i + (1-c)\eta_i > 0$ and $\sum_i\beta_i = 1$ such that $\mathsf{z}_i < -\frac{1}{2}\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{\beta_i}{\mu_i}}\}$ for all $i = 1, \ldots, d$, which means $\mathsf{z} \in \mathcal{S}$. Thus $\mathcal{S}$ is convex.

Remark 7: The sufficiency achieved via the TDMA scheduler in [17] can be alternatively formulated as follows: if there exist β_is with $0 < \beta_i \le 1$ and $\sum_{i=1}^{d}\beta_i = 1$, such that

$$\ln|\lambda_i| < -\frac{\beta_i}{2\mu_i}\ln\mathrm{E}\left\{\frac{\sigma_w^2}{\sigma_w^2+g_t^2P}\right\}, \quad (33)$$

for all $i = 1, 2, \ldots, d$, then the system (1) can be mean square stabilized. Since $l(z) = z^{\frac{\beta_i}{\mu_i}}$ with $0 < \frac{\beta_i}{\mu_i} \le 1$ is concave, from the Jensen's inequality, we have $-\frac{\beta_i}{2\mu_i}\ln\mathrm{E}\{\frac{\sigma_w^2}{\sigma_w^2+g_t^2P}\} \le -\frac{1}{2}\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{\beta_i}{\mu_i}}\}$. Thus any λ_i that satisfies (33) must also satisfy (25) with the same β_i, which implies that the adaptive TDMA scheduler in this paper achieves a stabilizability region no smaller than the TDMA scheduler in [17].

Remark 8: If $g_t = 1$, channel (2) degenerates to an AWGN channel and the necessary and sufficient condition to ensure mean square stabilizability, following from (6) and (25), is $\sum_{i=1}^{d}\mu_i\ln|\lambda_i| < \frac{1}{2}\ln(1 + \frac{P}{\sigma_w^2})$, which recovers the results in [8], [9]. If $g_t \sim \text{Bernoulli}(\epsilon)$, by taking the limit $\sigma_w^2 \to 0$ and $P \to \infty$, we can obtain that the stabilizability condition over an erasure channel is $\lambda_1^2 < \frac{1}{\epsilon}$, which degenerates to the results in [6], [19].

When all the strictly unstable eigenvalues have the same magnitude, we can show that the sufficient condition (25) coincides with the necessary condition (6), as shown in the following corollary.

Corollary 1: Suppose $|\lambda_1| = \cdots = |\lambda_{d_u}| = \tilde{\lambda} > 1$ and $|\lambda_{d_u+1}| = \cdots = |\lambda_d| = 1$ with $1 \le d_u \le d$. There exists an encoder/decoder pair $\{f_t\}, \{h_t\}$, such that the system (1) can be mean square stabilized over the channel (2) if and only if $\ln\tilde{\lambda} < -\frac{1}{2}\ln\mathrm{E}\{(\frac{\sigma_w^2}{\sigma_w^2+g_t^2P})^{\frac{1}{\mu_1+\cdots+\mu_{d_u}}}\}$.

Remark 9: The results derived in this paper for the power constrained fading channel (2) can be easily extended to the following channel model

$$r_t = g_t(s_t + w_t), \quad (34)$$

which is suitable for modeling the digital erasure channel with $\{w_t\}$ denoting the quantization error and $\{g_t\}$ representing the erasure process. If $g_t = 0$, the communication channel cannot transmit any information. Otherwise, we can always multiply the received signal r_t by $1/g_t$ at the decoder side, and thus the resulted channel is equivalent to an AWGN channel. From this perspective, channel (34) is essentially the power constrained lossy channel studied in [20]. Thus the results derived in [20] apply directly to the channel (34).

V. Numerical Illustrations

A. Scalar Systems

The authors in [21] derive the necessary and sufficient condition for mean square stabilization of scalar LTI systems over a power constrained fading channel with linear encoders/decoders as $\frac{1}{2}\ln(1 + \frac{\mu_g^2P}{\sigma_g^2P+\sigma_w^2}) > \ln|\lambda|$ with μ_g and σ_g^2 being the mean and variance of g_t. We can similarly define the mean square capacity of the power constrained fading channel achieved with linear encoders/decoders as $C_m = \frac{1}{2}\ln(1 + \frac{\mu_g^2P}{\sigma_g^2P+\sigma_w^2})$. Assume that the fading follows the Bernoulli distribution, i.e., $g_t \sim \text{Bernoulli}(\epsilon)$, and let $P = 1$ and $\sigma_w^2 = 1$, the channel capacities in relation to the erasure probability

[1]$\Pr(g_t = 0) = \epsilon$, $\Pr(g_t = 1) = 1 - \epsilon$, where $g_t = 0$ represents the appearance of fading and $g_t = 1$ means that the channel is free of fading.

are plotted in Fig. 4. It is clear that $C_{\text{Shannon}} \geq C_a \geq C_m$ at any erasure probability ϵ. This result is obvious since we have proved that the Shannon capacity is no smaller than the anytime capacity. Besides, we have more freedom in designing the causal encoder/decoder pair compared with the linear encoder/decoder pair, thus allowing to achieve a higher capacity. The three kinds of capacity degenerate to the same value when $\epsilon = 0$ and $\epsilon = 1$, which represent the AWGN channel case and the disconnected case respectively. This fact is trivial for the disconnected case and is consistent for the AWGN channel case in [5], [8], [9], in which the authors show that the anytime capacity is equal to the Shannon capacity for AWGN channels and causal encoder/decoder pair cannot provide any benefits in increasing the channel capacity.

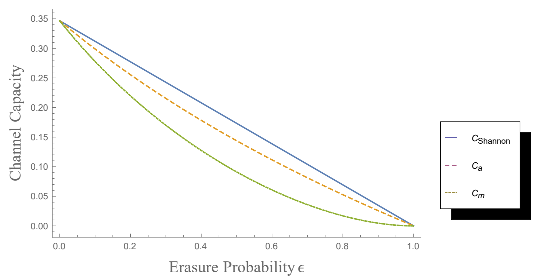

Fig. 4: Comparison of different channel capacities for scalar systems

B. Vector Systems

Consider a two-dimensional system (4) with $J = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, and the fading in (2) follows the Rayleigh distribution with probability density function $l(z;\sigma) = \frac{z}{\sigma^2} \mathrm{e}^{-\frac{z^2}{2\sigma^2}}$, $z \geq 0$. Let $P = 1$, $\sigma_w^2 = 1$, $\sigma = 2$, then the necessary stabilizability region, the sufficient stabilizability regions achieved with the optimal scheduler in Algorithm 1, the adaptive TDMA scheduler in Algorithm 2, the conventional TDMA scheduler in [17] and with linear encoders/decoders in [21], in terms of $(\ln |\lambda_1|, \ln |\lambda_2|)$ are plotted in Fig. 5. We can observe that the region of $(\ln |\lambda_1|, \ln |\lambda_2|)$ that can be stabilized with the designed causal encoders/decoders is much larger than that by linear encoders/decoders in [21]. Thus by extending encoder/decoders from linear settings to causal requirements, we can tolerate more unstable systems. It is clear from the figure that the optimal scheduler proposed in Algorithm 1 covers the whole necessary stabilizability region. Besides, as noted in Remark 7, the adaptive TDMA scheduler achieves a larger stabilizability region than that of the conventional TDMA scheduler. Moreover, the adaptive TDMA scheduler is optimal at three points, i.e., $|\lambda_1| = |\lambda_2|$, $|\lambda_1| = 1$ and $|\lambda_2| = 1$. This is consistent with Corollary 1.

VI. Conclusions

This paper has characterized the requirement for a power constrained fading channel to allow the existence of a causal encoder/decoder pair that can mean square stabilize a discrete-time LTI system. Fundamental limitations have been provided in terms of the system dynamics and channel parameters. Optimal communication designs have been provided for scalar systems and two-dimensional systems. For high-dimensional systems, a new communication scheme has also been provided, which can achieve larger stabilizability regions than existing results. What would the results be when there is no channel feedback or the channel fading is correlated is still under investigation. Further work will also be devoted to shrinking the gap between the necessary condition and the sufficient condition for high-dimensional systems. How to design communication/control schemes that only use finite memory is also of interest.

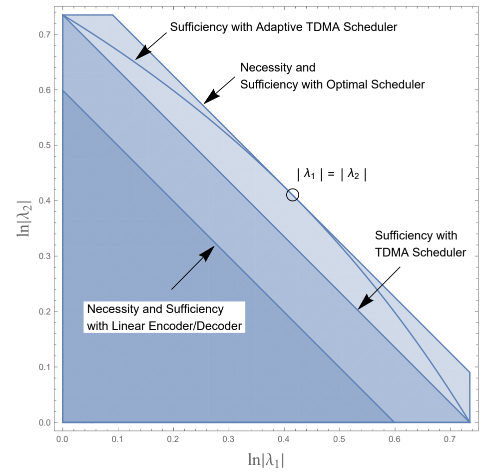

Fig. 5: Comparison of stabilizability regions for two-dimensional systems

Appendix

The following definitions are needed in the proof of Lemma 1 and are stated first, which are borrowed from [9]. All random variables are assumed to exist on a common probability space. The probability density of a random variable X is denoted by p_X, and the probability density of X conditioned on the event $Y = y$ is denoted by $p_{X|y}$. The differential entropy of X is defined by $H(X) = -\mathrm{E}\{\ln p_X\}$, provided that the defining integral exists. Denote the entropy of X given the event $Y = y$ by $H_y(X) = H(X|Y = y) = -\mathrm{E}_y\{\ln p_{X|y}\}$, and the random variable associated with $H_y(X)$ by $H_Y(X)$. The conditional entropy of X given the event $Y = y$ and averaged over Y is defined by $H(X|Y) = \mathrm{E}\{H_Y(X)\}$, and the conditional entropy of X given the events $Y = y$ and $Z = z$ and averaged only over Y by $H_z(X|Y) = \mathrm{E}_z\{H_{Y,Z}(X)\}$. The mutual information between two random variables X and Y given the event $Z = z$ is defined by $I_z(X;Y) = H_z(X) - H_z(X|Y)$. Given a random variable $X \in \mathbb{R}^n$ with entropy $H(X)$, the entropy power of X is defined by $N(X) = \frac{1}{2\pi \mathrm{e}} \mathrm{e}^{\frac{2}{n} H(X)}$. Denote the entropy power of X given the event $Y = y$ by $N_y(X) = \frac{1}{2\pi \mathrm{e}} \mathrm{e}^{\frac{2}{n} H_y(X)}$, and the random variable associated with $N_y(X)$ by $N_Y(X)$. The conditional entropy power of X given the event $Y = y$ and averaged over Y is defined by $N(X|Y) = \mathrm{E}\{N_Y(X)\}$. For any encoding strategy, the following lemma shows that the amount of information that the channel output contains about the source equals to that of the channel output contains about the channel input.

Lemma 3: Let X be an n-dimensional random variable, $f(X)$ be a function of X, and $Y = f(X) + N$ with N being a random variable that is independent of X. Then $I(X;Y) = I(f(X);Y)$.

Proof: Since $H(Y|X) = H(Y|X, f(X)) \leq H(Y|f(X))$, we have $H(Y) = I(X;Y) + H(Y|X) \leq I(X;Y) + H(Y|f(X))$. Thus $H(Y) - H(Y|f(X)) = I(Y; f(X)) \leq I(X;Y)$. Besides, since $X \to f(X) \to Y$ forms a Markov chain, $Y \to f(X) \to X$ also forms a Markov chain. The data processing inequality [22] then implies that $I(X;Y) \leq I(f(X);Y)$. Combining the two facts, we have $I(X;Y) = I(f(X);Y)$. ∎

Proof of Lemma 1: Here we use the uppercase letters X, S, R, G to denote random variables of the system state, the channel input, the channel output and the channel fading. We use the lowercase letters x, s, r, g to denote their realizations. The average entropy power of X_t conditioned on (R^t, G^t) is $N(X_t|R^t, G^t) = \mathrm{E}\{N_{R^t,G^t}(X_t)\} \overset{(a)}{=} \mathrm{E}\{\mathrm{E}_{R^{t-1},G^t}\{N_{R^t,G^t}(X_t)\}\} \overset{(b)}{=} \frac{1}{2\pi e}\mathrm{E}\{\mathrm{E}_{R^{t-1},G^t}\{\mathrm{e}^{\frac{2}{n}H_{R^t,G^t}(X_t)}\}\}$ where (a) follows from the law of total expectation and (b) from the definition of entropy power. Since

$$\begin{aligned}
&\mathrm{E}_{r^{t-1},g^t}\{\mathrm{e}^{\frac{2}{n}H_{R^t,G^t}(X_t)}\} \overset{(c)}{\geq} \mathrm{e}^{\frac{2}{n}\mathrm{E}_{r^{t-1},g^t}\{H_{R^t,G^t}(X_t)\}}\\
&\overset{(d)}{=} \mathrm{e}^{\frac{2}{n}H_{r^{t-1},g^t}(X_t|R_t)} = \mathrm{e}^{\frac{2}{n}\left(H_{r^{t-1},g^t}(X_t)-I_{r^{t-1},g^t}(X_t;R_t)\right)}\\
&\geq \mathrm{e}^{\frac{2}{n}\left(H_{r^{t-1},g^t}(X_t)-I_{r^{t-1},g^t}(X^t;R_t)\right)}\\
&\overset{(e)}{=} \mathrm{e}^{\frac{2}{n}\left(H_{r^{t-1},g^t}(X_t)-I_{r^{t-1},g^t}(S_t;R_t)\right)}\\
&\overset{(f)}{\geq} \mathrm{e}^{\frac{2}{n}\left(H_{r^{t-1},g^t}(X_t)-c_t\right)} \overset{(g)}{=} \mathrm{e}^{-\frac{2}{n}c_t}\mathrm{e}^{\frac{2}{n}H_{r^{t-1},g^{t-1}}(X_t)},
\end{aligned}$$

where (c) follows from Jensen's inequality; (d) from the definition of conditional entropy; (e) from Lemma 3; (f) from the definition of channel capacity, i.e., $I_{r^{t-1},g^t}(S_t;R_t) \leq c_t$ with $c_t = \frac{1}{2}\ln(1+\frac{g_t^2 P}{\sigma_w^2})$ being the instantaneous Shannon channel capacity and (g) from the fact that G_t is independent of X_t, we have $N(X_t|R^t, G^t) \geq \frac{1}{2\pi e}\mathrm{E}\{\mathrm{e}^{-\frac{2}{n}c_t}\mathrm{e}^{\frac{2}{n}H_{R^{t-1},G^{t-1}}(X_t)}\} = \mathrm{E}\{\mathrm{e}^{-\frac{2}{n}c_t}\}N(X_t|R^{t-1}, G^{t-1})$. Since $\mathrm{e}^{\frac{2}{n}H_{r^t,g^t}(X_{t+1})} = \mathrm{e}^{\frac{2}{n}H_{r^t,g^t}(AX_t+BU_t)} \overset{(h)}{=} \mathrm{e}^{\frac{2}{n}H_{r^t,g^t}(AX_t)} \overset{(i)}{=} \mathrm{e}^{\frac{2}{n}H_{r^t,g^t}(X_t)+\frac{2}{n}\ln|\det A|} = (\det A)^{\frac{2}{n}}\mathrm{e}^{\frac{2}{n}H_{r^t,g^t}(X_t)}$, where (h) follows from the fact that $u_t = h_t(r^t, g^t)$ and (i) from Theorem 8.6.4 in [22], we have $N(X_{t+1}|R^t, G^t) = \mathrm{E}\{\frac{1}{2\pi e}(\det A)^{\frac{2}{n}}\mathrm{e}^{\frac{2}{n}H_{R^t,G^t}(X_t)}\} = (\det A)^{\frac{2}{n}}N(X_t|R^t, G^t)$. In view of above results, we have $N(X_{t+1}|R^t, G^t) \geq (\det A)^{\frac{2}{n}}\mathrm{E}\{\mathrm{e}^{-\frac{2}{n}c_t}\}N(X_t|R^{t-1}, G^{t-1})$. In light of Proposition II.1 in [9], to ensure mean square stability, $N(X_{t+1}|R^t, G^t)$ should converge to zero asymptotically, which requires $(\det A)^{\frac{2}{n}}\mathrm{E}\{\mathrm{e}^{-\frac{2}{n}c_t}\} < 1$. The proof is completed. ∎

Lemma 4: Suppose $\{W_i\}$ with $W_i \leq 0$ is i.i.d. with bounded nonzero mean, define $L_t = \sum_{i=1}^t W_i$ and let T be the first time such that $L_T < \varphi T + \gamma$ with $\varphi \geq 0$, $\gamma < 0$. If there exists $\theta \geq 0$ such that $\mathrm{E}\{\mathrm{e}^{\theta(W_i-\varphi)}\} = \lambda^{-2}$, then $\mathrm{E}\{\lambda^{2T}\} \leq \lambda^2\mathrm{e}^{-\theta\gamma}$.

Proof: When $\varphi > 0$, since L_t is non-increasing and $\varphi t + \gamma$ is increasing, the stopping time T is bounded. When $\varphi = 0$, T is unbounded if and only $\gamma \leq \lim_{t\to\infty}\sum_{i=1}^t W_i \leq 0$. Since $\{W_i\}$ is i.i.d., in view of the law of large numbers, we have $\Pr(\lim_{t\to\infty}\sum_{i=1}^t W_i/t = \mathrm{E}\{W_i\}) = 1$. Thus $\Pr(\lim_{t\to\infty}\sum_{i=1}^t W_i = \infty) = 1$, which implies $\Pr(\gamma \leq \lim_{t\to\infty}\sum_{i=1}^t W_i \leq 0) = 0$. Thus T is almost surely bounded.

Define $Y_t = \mathrm{e}^{\theta L_t + bt}$ with $b = 2\ln|\lambda| - \theta\varphi$, then $\mathrm{E}\{Y_{t+1}|Y_t, \ldots, Y_1\} = Y_t\mathrm{E}\{\mathrm{e}^{\theta W_{t+1}+b}\} = Y_t$. Thus Y_t is a martingale. Since T is either a bounded or an almost surely bounded stopping time, in view of the optional stopping theorem [23], we have $\mathrm{E}\{Y_T\} = \mathrm{E}\{Y_1\} = 1$.

Define $\eta = \varphi T + \gamma - L_T$. Since $L_T < \varphi T + \gamma$ and $L_{T-1} \geq \varphi(T-1) + \gamma$, we have $\eta > 0$. When $\varphi = 0$, since $L_{T-1} \geq \gamma$, we have $\eta = \gamma - L_T = \gamma - L_{T-1} - W_T \leq -W_T$. When $\varphi > 0$ and $\varphi(T-1) + \gamma \leq L_{T-1} \leq \varphi T + \gamma$, we have $\eta = \varphi T + \gamma - L_T = \varphi(T-1) + \gamma - L_{T-1} + \varphi - W_T \leq \varphi - W_T$. When $\varphi > 0$ and $\varphi T + \gamma < L_{T-1} \leq 0$, we have $\eta = \varphi T + \gamma - L_T = \varphi T + \gamma - L_{T-1} - W_T < -W_T$. Thus in general, $\eta \leq \varphi - W_T$.

Since $\mathrm{E}\{Y_T\} = \mathrm{E}\{\mathrm{e}^{\theta(\varphi T+\gamma-\eta)+bT}\} = \mathrm{e}^{\theta\gamma}\mathrm{E}\{\mathrm{e}^{(\theta\varphi+b)T}\mathrm{e}^{-\theta\eta}\} = \mathrm{e}^{\theta\gamma}\mathrm{E}\{\lambda^{2T}\mathrm{e}^{-\theta\eta}\} = 1$, and $\mathrm{E}\{\lambda^{2T}\mathrm{e}^{-\theta\eta}\} \geq \mathrm{E}\{\lambda^{2T}\mathrm{e}^{\theta(W_T-\varphi)}\} = \mathrm{E}\{\mathrm{E}_T\{\lambda^{2T}\mathrm{e}^{\theta(W_T-\varphi)}\}\} = \mathrm{E}\{\lambda^{2T}\mathrm{E}_T\{\mathrm{e}^{\theta(W_T-\varphi)}\}\} \overset{(a)}{=} \lambda^{-2}\mathrm{E}\{\lambda^{2T}\}$ where (a) follows from the definition of θ, we have $\mathrm{E}\{\lambda^{2T}\} \leq \lambda^2\mathrm{e}^{-\theta\gamma}$. ∎

References

[1] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, 2007.

[2] G. Como, B. Bernhardsson, and A. Rantzer, *Information and control in networks*. Cham: Springer International Publishing, 2014.

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

[4] A. S. Matveev and A. V. Savkin, "An analogue of Shannon information theory for detection and stabilization via noisy discrete communication channels," *SIAM Journal on Control and Optimization*, vol. 46, no. 4, pp. 1323–1367, 2007.

[5] 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 Transactions on Information Theory*, vol. 52, no. 8, pp. 3369–3395, 2006.

[6] N. Elia, "Remote stabilization over fading channels," *Systems & Control Letters*, vol. 54, no. 3, pp. 237–249, 2005.

[7] N. Xiao, L. Xie, and L. Qiu, "Feedback stabilization of discrete-time networked systems over fading channels," *IEEE Transactions on Automatic Control*, vol. 57, no. 9, pp. 2176–2189, 2012.

[8] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, "Feedback stabilization over signal-to-noise ratio constrained channels," *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1391–1403, 2007.

[9] J. S. Freudenberg, R. H. Middleton, and V. Solo, "Stabilization and disturbance attenuation over a Gaussian communication channel," *IEEE Transactions on Automatic Control*, vol. 55, no. 3, pp. 795–799, 2010.

[10] U. Kumar, J. Liu, V. Gupta, and J. N. Laneman, "Stabilizability across a Gaussian product channel: Necessary and sufficient conditions," *IEEE Transactions on Automatic Control*, vol. 59, no. 9, pp. 2530–2535, 2014.

[11] A. A. Zaidi, T. J. Oechtering, S. Yuksel, and M. Skoglund, "Stabilization of linear systems over Gaussian networks," *IEEE Transactions on Automatic Control*, vol. 59, no. 9, pp. 2369–2384, 2014.

[12] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–i: No bandwidth constraint," *IEEE Transactions on Information Theory*, vol. 12, no. 2, pp. 172–182, 1966.

[13] P. Minero, M. Franceschetti, S. Dey, and G. N. Nair, "Data rate theorem for stabilization over time-varying feedback channels," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 243–255, 2009.

[14] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge: Cambridge University Press, 2005., 2005.

[15] E. I. Silva, G. C. Goodwin, and D. E. Quevedo, "Control system design subject to SNR constraints," *Automatica*, vol. 46, no. 2, pp. 428–436, 2010.

[16] S. Tatikonda and S. Mitter, "Control over noisy channels," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1196–1201, 2004.

[17] L. Xu, L. Xie, and N. Xiao, "Mean square capacity of power constrained fading channels with causal encoders and decoders," in *Proceedings of the 54th IEEE Conference on Decision and Control*, Osaka, Japan, 2015, Conference Proceedings, pp. 2465–2470.

[18] S. Yuksel and T. Basar, *Stochastic networked control systems: stabilization and optimization under information constraints*. New York, NY : Springer New York : Imprint: Birkhauser, 2013., 2013.

[19] V. Gupta, B. Hassibi, and R. M. Murray, "Optimal LQG control across packet-dropping links," *Systems & Control Letters*, vol. 56, no. 6, pp. 439–446, 2007.

[20] L. Xu, Y. Mo, and L. Xie, "Mean square stabilization of vector LTI systems over power constrained lossy channels," in *Proceedings of the 2016 American Control Conference*, Boston, MA, USA, 2016, Conference Proceedings, pp. 7129–7134.

[21] N. Xiao and L. Xie, "Analysis and design of discrete-time networked systems over fading channels," in *Proceedings of the 30th Chinese Control Conference*, Yantai, China, 2011, Conference Proceedings, pp. 6562 – 6567.

[22] T. M. Cover and J. A. Thomas, *Elements of information theory*. Hoboken, N.J.: Wiley-Interscience, 2006.

[23] R. B. Ash and C. A. Dolans-Dade, *Probability and measure theory*. San Diego: Harcourt/Academic Press, 2000.