A Control-Theoretic Linear Coding Scheme for the Fading Gaussian Broadcast Channel with Feedback

Siyao Li, Daniela Tuninetti, and Natasha Devroye
University of Illinois at Chicago, Chicago, IL 60607, USA, Email: {sli210, danielat, devroye}@uic.edu

2021 IEEE International Symposium on Information Theory (ISIT) | 978-1-5386-8209-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISIT45174.2021.9518060

***Abstract*—This paper proposes a linear coding scheme for the two-user fading additive white Gaussian noise broadcast channel, under the assumptions that: (i) perfect Channel State Information (CSI) is available at the receivers; and (ii) unit delayed CSI along with channel output feedback (COF) is available at the transmitter. The proposed scheme is derived from a control-theoretic perspective that generalizes the communication scheme for the point-to-point (P2P) fading Gaussian channel under the same assumptions by Liu *et al*. [1]. The proposed scheme asymptotically achieves the rates of a posterior matching scheme, from the same authors, for a certain choice of parameters.**

I. Introduction

This paper considers the problem of communication over an Additive White Gaussian Noise (AWGN) Broadcast Channels (BC) with fading and Channel Output Feedback (COF). In particular, the forward link of the channel, from the transmitter to the receivers, experiences i.i.d. time-varying fading assumed to be known to the receivers perfectly and without delay. The reverse link, also known as the feedback channel, enables the transmitter to access the exact channel outputs, including the channel state, with unit delay. Under the same Channel State Information (CSI) assumption, the capacity for the fading point-to-point (P2P) fading AWGN channels is known [2].

While COF cannot increase the capacity of memoryless P2P channels, it can significantly reduce the probability of error and even simplify capacity achieving schemes [3], [4]. The same is true for physically degraded BC [5]. In general, COF increases the capacity region of multi-user channels. For example, COF enlarges the capacity of the non-fading BC even when the feedback is available from only one of the receivers, and that in the degraded case the probability of error decays doubly exponential in the block length [6]. The coding scheme for non-fading AWGN-BC with two receivers and COF constructed in [7] extends the approach for P2P AWGN channel with feedback in [4], [8], and shows that the capacity region is enlarged except for the physically degraded BC, for which the capacity region is known. The P2P Posterior Matching (PM) scheme in [3] was extended to non-fading BCs in [9] to obtain the same exact region as in [7]. This region can be further enlarged by using *robust control theory* as proved in [10]. The Linear Quadratic Gaussian (LQG) control theory inspired the code design in [11], which performs the same as Elia's scheme in [10] for two users, and outperforms Kramer's scheme in [12] for more than two users. For symmetric non-fading AWGN-BCs with two users, the coding schemes in [10], [11] achieve the largest known sum-rate with COF. For non-symmetric non-fading AWGN-BCs with uncorrelated noises, the iterative coding scheme in [13] is sum-rate optimal among all linear-feedback schemes.

The capacity region of the fading AWGN-BC and COF is an open problem. In our previous work [14], we derived an achievable region for the two-user fading AWGN-BC with COF by integrating channel fading into design and presenting a PM-based coding scheme akin to [9]. Linear coding schemes from a control perspective have been only investigated for static BCs, to the best of our knowledge. These achievable schemes do not work for fading BCs, as the instantaneous channel fading is not available at the transmitter before signal transmission. The feedback scheme for P2P fading AWGN channel with receiver CSI and unit delayed COF proposed by Liu *et al*. in [1] is optimal. It is motivated by control theory and generalizes the Schalkwijk-Kailath (SK) scheme in [4], [8]. However, it is not immediately clear how to extend the scheme to a system with more than two users, since we need to refine the system state vector sequentially based on the feedback to stabilize each of the system state, which may be correlated for dependent channel fadings.

Contributions. All capacity results for the AWGN-BCs are without COF [15] or without fading [7], [9]–[13]. We demonstrate here an achievable rate region for the fading AWGN-BC with receiver CSI and COF. We propose a linear coding scheme over a two-user BC with fading and COF that generalizes (i) the work in [9] for static BCs with COF, and (ii) the scheme over P2P Gaussian channels with fading and COF in [1]. This linear coding scheme for a certain choice of parameters asymptotically achieves the rate region of the PM scheme we presented in [14]. This reveals yet again a tight connection between the feedback communication problem over an infinite-state fading channel with CSI available to the transmitter with a unit delay and a related feedback stabilization problem over the same channel. It is interesting to note that in the proposed scheme the encoder and decoder are causal and simple to implement. Moreover, the needed average transmission power can be determined by solving an optimal linear control problem known as *cheap control* [16]. Ongoing work includes the analysis of the proposed scheme for general parameters, in particular to understand whether, and if so by how much, one can outperform the PM-based scheme in [14].

Notation. In this paper, we represent time indices by subscripts, such as Y_i, and user indices by superscripts, such as $Y^{(k)}$. We use K to denote the number of receivers. We consider real-valued continuous random variables in discrete-time stochastic processes. Random Variables (RVs) are denoted by upper-case letters, their realizations by the corresponding lower-case letters, and sequences of RVs by boldface, such as $\mathbf{Y} := [Y^{(1)}, \cdots, Y^{(\mathsf{K})}]^T$. To denote explicitly the dimensionality of a vector, we use superscripts and subscripts in the following manner: $\mathbf{Y}_i^{j(k)} := [Y_i^{(k)}, \cdots, Y_j^{(k)}]$ for $i \le j$ and $\mathbf{Y}^{i-1} := [\mathbf{Y}^{i-1(1)}, \cdots, \mathbf{Y}^{i-1(\mathsf{K})}]^T$. $\mathbb{R}$ and $\mathbb{R}^n$ represent sets of real scalars and n-dimensional real column vectors, respectively. $\underline{A}^T$ denotes the transpose of matrix $\underline{A}$. A real-valued RV X is associated with a distribution $\mathbb{P}_X(\cdot)$ defined on the usual Borel σ-algebra over $\mathbb{R}$, and we write $X \sim \mathbb{P}_X$. We write $\mathbb{E}(\cdot)$ for expectation and $\mathbb{P}(\cdot)$ for the probability of a measurable event within the parentheses. The notation $\xrightarrow{\mathbb{P}}$ specifies convergence in probability. We use $\log$ in base 2, and $\operatorname{sgn}(x)$ to denote the sign function, where $\operatorname{sgn}(x) := 1$ if $x \ge 0$ and $\operatorname{sgn}(x) := -1$ if $x < 0$. Given a statement A, the indicator function $\mathbb{1}(A)$ is equal to one if A is true and zero otherwise. We let $\bar{k}$ denote $\bar{k} \ne k$ when $k \in [2]$.

II. System Model

We consider a communication system where one transmitter and K receivers are connected via a fading AWGN channel. Global CSI is assumed available at the receivers. All channel outputs, including the CSI, are noiselessly fed back to the transmitter. The received signal $Y_n^{(k)}$ for user $k \in [\mathsf{K}]$ at time $n \in \mathbb{N}$ is

$$Y_n^{(k)} = \sqrt{S_n^{(k)}} U_n + Z_n^{(k)} \in \mathbb{R}, \tag{1}$$

where $U_n \in \mathbb{R}$ denotes the transmitted signal, $Z_n^{(k)}$ is the real-valued AWGN with unit power, and $S_n^{(k)} \in \mathcal{S}$ is the channel fading of user k with alphabet $\mathcal{S}$. We assume that $[S^{(1)}, \cdots, S^{(\mathsf{K})}]^T$ forms a memoryless process over time known at the receivers, that the noise $\mathbf{Z} := [Z^{(1)}, \ldots, Z^{(\mathsf{K})}]^T$ are independent across users and time, and that the input is subject to the average power constraint $\mathbb{E}(U^2) \le P$. Let $\mathbf{H} := [\sqrt{S^{(1)}}, \cdots, \sqrt{S^{(\mathsf{K})}}]^T$. We also assume that $U_0, \mathbf{H}, \mathbf{Z}$ are independent. A noiseless feedback channel transmits one-step delayed information $(\mathbf{H}_n, \mathbf{Y}_n)$ from the receivers to the transmitter.

Definition 1. A $(2^{nR^{(1)}}, \cdots, 2^{nR^{(\mathsf{K})}}, n)$ code for the fading AWGN-BC with COF consists of:

1) K independent and equally likely messages $M^{(k)} \in \{1, \cdots, 2^{nR^{(k)}}\}$, $k \in [\mathsf{K}]$;
2) an encoder that assigns a symbol $U_i(\mathbf{M}, \mathbf{Y}^{i-1}, \mathbf{H}^{i-1})$ to the message vector $\mathbf{M}$ and the previous channel output vectors $(\mathbf{Y}^{i-1}, \mathbf{H}^{i-1})$ for each $i \in [n]$; and
3) K decoders, where decoder $k \in [\mathsf{K}]$ assigns an estimate $\widehat{M}^{(k)}$ to each sequence $\mathbf{Y}_1^{n(k)} = (Y_1^{(k)}, \cdots, Y_n^{(k)})$.

The probability of error for receiver $k \in [\mathsf{K}]$ is defined as $p_{e,n}^{(k)} := \mathbb{P}(M^{(k)} \ne \widehat{M}^{(k)})$.

Definition 2. We say that $(R^{(1)}, \cdots, R^{(\mathsf{K})})$ is an achievable rate vector under (asymptotic block) power constraint P if there exists a sequence of $(2^{nR^{(1)}}, \cdots, 2^{nR^{(\mathsf{K})}}, n)$ codes such that $\lim_{n\to\infty} \sup \frac{1}{n} \sum_{i=1}^n \mathbb{E}(|U_i|^2) \le P$ and $\lim_{n\to\infty} \max_{k\in[\mathsf{K}]} p_{e,n}^{(k)} = 0$.

III. Main Result

Several connections between information theory and control theory, especially when COF is present, have been explored [1], [10], [11]. In this section, we present the main result of this paper, which relates the achievable rate region to the open-loop growth rate of a control system[1]. We introduce a linear feedback coding scheme for $\mathsf{K} = 2$ from a control perspective in Section III-A, analyze the achievable rate region in Section III-B, and derive the linear coding parameters based on cheap control in Section III-C.

A. A Control System and its Stabilization

A K-dimensional unstable dynamical system is stabilized by a controller having full state observation [17]. The system model at time $i \ge 0$ is

$$\text{State: } \mathbf{X}_{i+1} = \underline{A}(\mathbf{H}_{i-j+1}^i)\mathbf{X}_i - \underline{B}(\mathbf{H}_{i-j+1}^i)\mathbf{Y}_i, \tag{2a}$$

$$\text{Input: } U_i = \underline{C}(\mathbf{H}_{i-l+1}^{i-1})\mathbf{X}_i, \tag{2b}$$

$$\text{Output: } \mathbf{Y}_i = \mathbf{H}_i U_i + \mathbf{Z}_i, \tag{2c}$$

where $\mathbf{X}_i \in \mathbb{R}^{\mathsf{K}}$ is the system state vector; $U_i \in \mathbb{R}$ is the channel input; $\mathbf{Y}_i \in \mathbb{R}^{\mathsf{K}}$ is the channel output vector; $\mathbf{Z}_i \in \mathbb{R}^{\mathsf{K}}$ is the AWGN vector; and $\mathbf{H}_i \in \mathbb{R}^{\mathsf{K}}$ is the channel fading vector. As the controllers in (2a) and (2b) may depend on some previous channel fading parameters, for $\mathbf{H}_{i-j+1}^i$, j denotes the fixed memory length of matrices $\underline{A}$ and $\underline{B}$ defined next, and for $\mathbf{H}_{i-l+1}^{i-1}$, l denotes the fixed memory length of matrix $\underline{C}$. We assume $Z^{(k)} \sim \mathcal{N}(0,1)$ for all $k \in [\mathsf{K}]$ and mutually independent, and

$$\underline{A}(\mathbf{H}_{i-j+1}^i) = \operatorname{diag}(a_i^{(1)}(\mathbf{H}_{i-j+1}^i), \cdots, a_i^{(\mathsf{K})}(\mathbf{H}_{i-j+1}^i)) \in \mathbb{R}^{\mathsf{K}\times\mathsf{K}},$$

$$\underline{B}(\mathbf{H}_{i-j+1}^i) = \operatorname{diag}(b_i^{(1)}(\mathbf{H}_{i-j+1}^i), \cdots, b_i^{(\mathsf{K})}(\mathbf{H}_{i-j+1}^i)) \in \mathbb{R}^{\mathsf{K}\times\mathsf{K}},$$

$$|a_i^{(k)}(\mathbf{H}_{i-j+1}^i)| > 1, \forall k \in [\mathsf{K}], \tag{3}$$

$$\underline{C}(\mathbf{H}_{i-l+1}^{i-1}) \ne 0. \tag{4}$$

The term $\underline{B}(\mathbf{H}_{i-j+1}^i)$ in (2a) is the output feedback control gain, i.e., the output of the controller is $-\underline{B}(\mathbf{H}_{i-j+1}^i)\mathbf{Y}_i$. For convenience, in the rest of the paper we use $\underline{A}_i, \underline{B}_i$ and $\underline{C}_i$ to denote $\underline{A}(\mathbf{H}_{i-j+1}^i), \underline{B}(\mathbf{H}_{i-j+1}^i), \underline{C}(\mathbf{H}_{i-l+1}^{i-1})$ in (2); we use $a_i^{(k)}$ and $b_i^{(k)}$ to denote $a_i^{(k)}(\mathbf{H}_{i-j+1}^i)$ and $b_i^{(k)}(\mathbf{H}_{i-j+1}^i)$, respectively. We consider the stability of a system in the mean square sense [18].

Definition 3. The system in (2) is said to be *mean square stable* (MSS) if there exists a constant $\zeta < \infty$ such that $\mathbb{E}(\|\mathbf{X}_i\|^2) = \mathbb{E}(\sum_{k=1}^{\mathsf{K}} |X_i^{(k)}|^2) < \zeta$ for all i.

The open-loop system, namely $\mathbf{X}_{i+1} = \underline{A}_i\mathbf{X}_i$, is assumed to be unstable in (3) and thus the state $X_i^{(k)}$ in (2a) grows

[1]the open-loop growth rate is a measure of how unstable the open-loop control system is.

unboundedly at an asymptotic rate characterized by the Lyapunov exponent [19] and denoted by $R^{(k)\star}$, where

$$R^{(k)\star} := \lim_{n\to\infty} \frac{1}{n} \log \frac{|X_n^{(k)}|}{|X_0^{(k)}|} \xrightarrow{\mathbb{P}} \mathbb{E}\Big(\log |a_i^{(k)}|\Big). \qquad (5)$$

The closed-loop system dynamics is

$$\mathbf{X}_{i+1} = \underline{A}_{\text{cl},i}\mathbf{X}_i - \underline{B}_i\mathbf{Z}_i, \qquad (6)$$

where $\underline{A}_{\text{cl},i} := \underline{A}_i - \underline{B}_i\mathbf{H}_i\underline{C}_i \in \mathbb{R}^{\mathsf{K}\times\mathsf{K}}$. To have an MSS system, based on Definition 3, we need to have $\sum_{k=1}^{\mathsf{K}} \mathbb{E}(|X_i^{(k)}|^2)$ bounded.

Remark 1. *The information structure of the control system in* (2) *is classical [20], as the channel input U_i can access accurate information and histories of the state $\mathbf{X}_i$. This implies that the system model considered here is not relevant to the Witsenhausen Counterexample in [21]. In the LQG team decision problem, it is well-known that classical information structure allows optimal linear solutions [22]. This is why we focus on the linear representation in* (2)*.*

B. Reliable Communication Analysis

The control system model in (2) is a linear communication scheme for the fading BC in (1) with a power budget P, where each user k sees an entry of the output $Y_i^{(k)}$ and the CSI. The encoding function is a choice of $\underline{A}_i$, $\underline{B}_i$ and $\underline{C}_i$ in (2). Initially, divide the interval $[-\sqrt{P^{(k)}}, \sqrt{P^{(k)}}]$ into $2^{(n+1)(1-\epsilon)R^{(k)\star}}$ equal-length intervals, where $\sum_{k\in[\mathsf{K}]} P^{(k)} \le P$. $X_0^{(k)}$ is the midpoint of the interval with distance

$$D_n^{(k)} := \sqrt{P^{(k)}}/2^{(n+1)(1-\epsilon)R^{(k)\star}-1}, \qquad (7)$$

and encoder transmits U_0. The encoder recursively forms $\mathbf{X}_i$ and transmits U_i. Decoder k's estimate $\widehat{X}_i^{(k)}$ is generated as

$$\widehat{X}_i^{(k)} = \widehat{X}_{i-1}^{(k)} + \phi_i^{(k)} b_i^{(k)} Y_i^{(k)} \qquad (8)$$

where $\phi_i^{(k)} := \prod_{l=0}^{i}(a_i^{(k)})^{-1}$ and $\widehat{X}_0^{(k)} = 0$. Simple algebra shows the following invariance relationship regardless of the choice of B_i,

$$\widehat{X}_i^{(k)} = X_0^{(k)} - \phi_i^{(k)} X_{i+1}^{(k)}. \qquad (9)$$

Then, $\widehat{X}_i^{(k)}$ is mapped to the closest (in Euclidean distance) $X_0^{(k)}$ to obtain the decoded message. We next show that for the parameters that stabilize the control system in MSS yield an achievable rate region for the fading BC with COF. In particular, the code is obtained by minimizing the transmission power by using the tool of cheap control.

Proposition 1. *The control system in* (2) *with* $\mathsf{K} = 2$*, unstable in open-loop with growth rates* $(R^{(1)\star}, R^{(2)\star})$ *in* (5) *and MSS in closed-loop, achieves any rate* $(R^{(1)}, R^{(2)})$ *arbitrarily close to* $(R^{(1)\star}, R^{(2)\star})$ *for a communication system.*

Proof: By (6), for a given $\mathbf{H}_0^n$, since $\mathbf{X}_{n+1} = \prod_{i=0}^{n} \underline{A}_{\text{cl},i}\mathbf{X}_0 - \sum_{i=0}^{n}\prod_{l=i}^{n} \underline{A}_{\text{cl},l}\underline{B}_i\mathbf{Z}_i$, where $\underline{A}_{\text{cl},i}, \underline{B}_i \in \mathbb{R}^{2\times 2}$ and $\mathbf{X}_0, \mathbf{Z}_i \in \mathbb{R}^2$, we have that $\mathbf{X}_{n+1}$ follows a Gaussian distribution conditioned on $(\mathbf{X}_0, \mathbf{H}_0^n)$ with mean $\mathbb{E}\left(\mathbf{X}_{n+1}|\mathbf{X}_0, \mathbf{H}_0^n\right) = \prod_{i=0}^{n} \underline{A}_{\text{cl},i}\mathbf{X}_0$ and variance

$$\mathbb{E}\Big(\big(\mathbf{X}_{n+1} - \mathbb{E}(\mathbf{X}_{n+1})\big)^2 |\mathbf{X}_0, \mathbf{H}_0^n\Big) = \Big(\sum_{i=0}^{n}\prod_{l=i}^{n} \underline{A}_{\text{cl},l}\underline{B}_i\Big)^2. \qquad (10)$$

Clearly, the variance depends on $\mathbf{H}_0^n$ but not on $\mathbf{X}_0$. From (9), $\widehat{\mathbf{X}}_n$ conditioned on $(\mathbf{X}_0, \mathbf{H}_0^n)$ follows the Gaussian distribution

$$\mathcal{N}\Big(\big(I_{2\times 2} - \prod_{i=0}^{n}\underline{A}_i^{-1}\prod_{i=0}^{n}\underline{A}_{\text{cl},i}\big)\mathbf{X}_0, (\prod_{i=0}^{n}\underline{A}_i^{-1})^2(\sum_{i=0}^{n}\prod_{l=i}^{n}\underline{A}_{\text{cl},l}\underline{B}_i)^2\Big).$$

where $\prod_{i=0}^{n}\underline{A}_i^{-1}\prod_{i=0}^{n}\underline{A}_{\text{cl},i}$ and $\left(\sum_{i=0}^{n}\prod_{l=i}^{n}\underline{A}_{\text{cl},l}\underline{B}_i\right)^2 \in \mathbb{R}^{2\times 2}$. Let $\begin{pmatrix} a_{\text{cl},11} & a_{\text{cl},12} \\ a_{\text{cl},21} & a_{\text{cl},22}\end{pmatrix} := \prod_{i=0}^{n}\underline{A}_{\text{cl},i}$ and let the diagonal elements of (10) from left to right be σ_1^2 and σ_2^2. Thus, the distribution of user k's estimate $\widehat{X}_n^{(k)}$ is

$$\mathcal{N}\Big((1-\phi_n^{(k)}a_{\text{cl},kk})X_0^{(k)} - \phi_n^{(k)}a_{\text{cl},k\bar{k}}X_0^{(\bar{k})}, (\phi_n^{(k)})^2\sigma_k^2\Big). \qquad (11)$$

In the limit $n\to\infty$, the probability of error $p_{e,n}^{(k)}$ of user k is upper bounded by $p_{e,n}^{(k)u}$, where

$$p_{e,n}^{(k)u} := 2Q\Big(\frac{1-(|\phi_n^{(k)}|/D_n^{(k)})\Big(|a_{\text{cl},kk}||X_0^{(k)}|+|a_{\text{cl},k\bar{k}}||X_0^{(\bar{k})}|\Big)}{(\sqrt{n+1}|\phi_n^{(k)}|/D_n^{(k)})(\sigma_k/\sqrt{n+1})}\Big),$$

and $\sum_{k\in[2]} P^{(k)} = P$. To show $p_{e,n}^{(k)} \to 0$, it is sufficient to show that $p_{e,n}^{(k)u} \xrightarrow{\mathbb{P}} 0$ or that the argument of the Q function tends to infinity, or equivalently $|\phi_n^{(k)}|/D_n^{(k)} \xrightarrow{\mathbb{P}} 0$, $\Big(|a_{\text{cl},kk}||X_0^{(k)}| + |a_{\text{cl},k\bar{k}}||X_0^{(\bar{k})}|\Big) \xrightarrow{\mathbb{P}} 0$, $\sqrt{n+1}|\phi_n^{(k)}|/D_n^{(k)} \xrightarrow{\mathbb{P}} 0$ and $\sigma_k/\sqrt{n+1} \xrightarrow{\mathbb{P}} 0$. Since $\frac{1}{n+1}\log|\phi_n^{(k)}| = -\frac{1}{n+1}\log\Big|\prod_{i=0}^{n} a_i^{(k)}\Big| \xrightarrow{\mathbb{P}} -R^{(k)\star}$, we have

$$\frac{1}{n+1}\log\frac{|\phi_n^{(k)}|}{D_n^{(k)}} = \frac{1}{n+1}\Big(\log|\phi_n^{(k)}| + (n+1)(1-\epsilon)R^{(k)\star} - \log\sqrt{P^{(k)}}\Big) \xrightarrow{\mathbb{P}} -\epsilon R^{(k)\star}.$$

Specifically, the growth rate of $|\phi_n^{(k)}|/D_n^{(k)}$ converges to a strictly negative value in probability and thus $|\phi_n^{(k)}|/D_n^{(k)} \xrightarrow{\mathbb{P}} 0$. Similarly, $\frac{1}{n+1}\log\sqrt{n+1}|\phi_n^{(k)}|/D_n^{(k)} \xrightarrow{\mathbb{P}} -\epsilon R^{(k)\star}$, that is $\sqrt{n+1}|\phi_n^{(k)}|/D_n^{(k)} \xrightarrow{\mathbb{P}} 0$. By [23, Theorem 3.33], if the non-homogeneous system (6) is MSS, then the corresponding homogeneous system $\mathbf{X}_{i+1} = \underline{A}_{\text{cl},i}\mathbf{X}_i$ is MSS, which further leads to the almost sure convergence to zero of $\prod_{i=0}^{n}\underline{A}_{\text{cl},i}\mathbf{X}_0$ [23, Corollary 3.46]. Therefore, $\Big(|a_{\text{cl},kk}||X_0^{(k)}| + |a_{\text{cl},k\bar{k}}||X_0^{(\bar{k})}|\Big) \xrightarrow{\mathbb{P}} 0$.

As $\mathbb{E}((X_n^{(k)})^2) = \mathbb{E}((X_n^{(k)})^2|X_0^{(k)}, \mathbf{H}_0^n)$, by Markov's Inequality, for any $\epsilon > 0$,

$$\mathbb{P}\left(\mathbb{E}((X_n^{(k)})^2|X_0^{(k)}, \mathbf{H}_0^n) \ge (n+1)\epsilon\right) \le \frac{\mathbb{E}\Big((X_n^{(k)})^2|X_0^{(k)}, \mathbf{H}_0^n\Big)}{(n+1)\epsilon}.$$

Since the system is MSS in closed-loop, $\mathbb{E}((X_n^{(k)})^2)$ converges to some constant number. Hence $\mathbb{E}((X_n^{(k)})^2|X_0^{(k)}, \mathbf{H}_0^n)/(n+$

$1) \overset{\mathbb{P}}{\to} 0$. Also, as $\sigma_k^2 \le \mathbb{E}((X_n^{(k)})^2)$, it holds that $\sigma_k/\sqrt{n+1} \overset{\mathbb{P}}{\to} 0$. We conclude that $p_{e,n}^{(k)u} \overset{\mathbb{P}}{\to} 0$. Therefore, we have shown that closed-loop stability implies that the corresponding communication system can transmit reliably at rates arbitrarily close to $R^{(k)\star}$. ■

Remark 2. *The open-loop growth rates in* (5) *asymptotically achieve the rates of the PM-based scheme in [14] in all cases we tried numerically, and* $D_n^{(k)}$ *defined in* (7) *corresponds to the decoding interval of the message to user* k *at time* n *in [14]. Note that different choices of the stabilizing gain in* (2a) *can yield the same rate, but they may result in different average channel input powers.*

C. Cheap Control

Here we derive the coefficients of the linear feedback scheme in (2) that satisfy the conditions in Proposition 1, that is, we stabilize the control system with minimum power and enable reliable communication over the same channel with achievable rates arbitrarily close to the open-loop growth rate. With the help of cheap control, a method used in Linear Quadratic Regulator (LQR) control, the problem can be formulated as follows.

For the control gain $\underline{C}_i = [c_i^{(1)}, c_i^{(2)}]$ in (2), without loss of generality, we set $|c_i^{(1)}/c_i^{(2)}| = 1$, since the fraction can be compensated by optimizing over $a_i^{(k)}$ and $b_i^{(k)}$. Inspired by the parameters of the PM scheme in [14], in (2) we set

$$\underline{C}_i = \eta_i[1,\ \mathrm{sgn}(\rho_i)], \tag{12a}$$

where

$$\rho_i = \mathbb{E}\big(X_i^{(1)}X_i^{(2)}\big)\Big/\sqrt{\mathbb{E}\big(|X_i^{(1)}|^2\big)\mathbb{E}\big(|X_i^{(2)}|^2\big)}, \tag{12b}$$

$$\eta_i = \sqrt{\frac{\Big(\mathbb{E}\big(|X_i^{(1)}|^2\big) + \mathbb{E}\big(|X_i^{(2)}|^2\big)\Big)}{\mathbb{E}\Big(|X_i^{(1)} + \mathrm{sgn}(\rho_i)X_i^{(2)}|^2\Big)}}. \tag{12c}$$

Now the system dynamics in (2b) becomes

$$U_i = \eta_i X_i^{(1)} + \mathrm{sgn}(\rho_i)\eta_i X_i^{(2)}. \tag{13}$$

Here the goal is, for the given ergodic fading channel, to find the most efficient system in the form of (2) by finding $a_i^{(k)}$ and $b_i^{(k)}$ such that the open-loop is unstable with the growth rate satisfying (5), and the closed-loop is MSS with the least possible average channel input power. Thus, one needs to solve the following optimal control problem

$$P(R^{(1)}, R^{(2)}) := \min_{a_i^{(k)}, b_i^{(k)}} \lim_{n\to\infty} \frac{1}{n+1}\sum_{i=0}^{n} \mathbb{E}(|U_i|^2), \tag{14}$$

subject to (2) being MSS and satisfying (5). This is a type of cheap control problem as the objective function does not assign any direct penalty to the control effort $-\underline{B}_i\mathbf{Y}_i$. Intuitively, this translates to minimization of average channel input power subject to a rate constraint for its associated communication system. The solution is given as follows.

Proposition 2. *A linear coding scheme for the fading AWGN-BC, with two receivers and COF, obtained by solving the cheap control problem in* (14) *has the following parameters*

$$|a_i^{(k)}| = \sqrt{\frac{1 + S_i^{(k)}P}{1 + S_i^{(k)}(PP^{(\bar{k})}(1-\rho_i^2))/g_i}},\quad k, \bar{k} \in [2],\ \bar{k} \neq k, \tag{15a}$$

$$b_i^{(1)} = \frac{a_i^{(1)}\sqrt{S_i^{(1)}}\eta_i\Big(P^{(1)} + \sqrt{P^{(1)}P^{(2)}}|\rho_i|\Big)}{S_i^{(1)}\Big(P^{(1)} + P^{(2)}\Big) + 1}, \tag{15b}$$

$$b_i^{(2)} = \frac{a_i^{(2)}\sqrt{S_i^{(2)}}\eta_i\Big(\mathrm{sgn}(\rho_i)P^{(2)} + \sqrt{P^{(1)}P^{(2)}}\rho_i\Big)}{S_i^{(2)}\Big(P^{(1)} + P^{(2)}\Big) + 1}, \tag{15c}$$

and ρ_{i+1} *is updated as*

$$\rho_{i+1} = \sqrt{\frac{m_i}{f_i}}\Big(g_i\rho_i - \frac{l_iP}{m_i}(P\rho_i + \mathrm{sgn}(\rho_i)\sqrt{P^{(1)}P^{(2)}}(1+\rho_i^2))\Big), \tag{15d}$$

where

$$g_i := P + 2|\rho_i|\sqrt{P^{(1)}P^{(2)}}, \tag{15e}$$

$$m_i := (S_i^{(1)}P + 1)(S_i^{(2)}P + 1), \tag{15f}$$

$$f_i := (g_i + S_i^{(1)}P^{(2)}q_i)(g_i + S_i^{(2)}P^{(1)}q_i), \tag{15g}$$

$$q_i := P(1-\rho_i^2), \tag{15h}$$

$$l_i := S_i^{(1)}S_i^{(2)}P + S_i^{(1)} + S_i^{(2)}, \tag{15i}$$

for all i*, which shares the same encoding function as the PM-based scheme in [14].*

Proof: By (13), the average transmission power becomes

$$\lim_{n\to\infty} \frac{1}{n+1}\sum_{i=0}^{n} \mathbb{E}\left(\sum_{k\in[2]} |X_i^{(k)}|^2\right). \tag{16}$$

To have minimum average transmission power P, let $\lim_{i\to\infty} \mathbb{E}\Big(|X_i^{(k)}|^2\Big) = P^{(k)}$, and $\sum_{k\in[2]} P^{(k)} = P$.

From (6), $X_{i+1}^{(k)}$ are given by

$$X_{i+1}^{(1)} = \Big(a_i^{(1)} - \sqrt{S_i^{(1)}}b_i^{(1)}\eta_i\Big)X_i^{(1)} - \sqrt{S_i^{(1)}}b_i^{(1)}\,\mathrm{sgn}(\rho_i)\eta_i X_i^{(2)} - b_i^{(1)}Z_i^{(1)}, \tag{17a}$$

$$X_{i+1}^{(2)} = \Big(a_i^{(2)} - \sqrt{S_i^{(2)}}b_i^{(2)}\,\mathrm{sgn}(\rho_i)\eta_i\Big)X_i^{(2)} - \sqrt{S_i^{(2)}}b_i^{(2)}\eta_i X_i^{(1)} - b_i^{(2)}Z_i^{(2)}. \tag{17b}$$

By convergence, the following holds

$$P^{(1)} = \mathbb{E}\Big(\big(a_i^{(1)} - \sqrt{S_i^{(1)}}b_n^{(1)}\eta_i\big)^2\Big)P^{(1)} - 2\rho_i\sqrt{P^{(1)}P^{(2)}} \times \mathbb{E}\Big(\sqrt{S_i^{(1)}}b_i^{(1)}\,\mathrm{sgn}(\rho_i)\eta_i\big(a_i^{(1)} - \sqrt{S_i^{(1)}}b_i^{(1)}\eta_i\big)\Big) + \mathbb{E}\Big((\sqrt{S_i^{(1)}}b_i^{(1)}\,\mathrm{sgn}(\rho_i)\eta_i)^2\Big)P^{(2)} + \mathbb{E}\Big(\big(b_i^{(1)}\big)^2\Big), \tag{18a}$$

$$P^{(2)} = \mathbb{E}\Big((a_i^{(2)} - \sqrt{S_i^{(2)}}b_i^{(2)}\,\mathrm{sgn}(\rho_i)\eta_i)^2\Big)P^{(2)} - 2\rho_i$$

$$\times\sqrt{P^{(1)}P^{(2)}}\mathbb{E}\Big(\sqrt{S_i^{(2)}}b_i^{(2)}\eta_i\big(a_i^{(2)}-\sqrt{S_i^{(2)}}b_i^{(2)}\operatorname{sgn}(\rho_i)\eta_i\big)\Big)$$
$$+\mathbb{E}\Big((\sqrt{S_i^{(2)}}b_i^{(2)}\eta_i)^2\Big)P^{(1)}+\mathbb{E}\Big(\big(b_i^{(2)}\big)^2\Big). \tag{18b}$$

Hence, minimizing the transmission power (16) is equivalent to minimizing $P^{(1)}+P^{(2)}$ satisfying (18).

Now we assume all $a_i^{(k)}\geq 1$ are given and look for the optimal control $\{b_i^{(k)}\}$. Then, we look for $\{a_i^{(k)}\}$ by searching over all $\{a_i^{(k)}\}$ satisfying (5). Summing up the right-hand side terms of (18a) and (18b), we obtain $\lambda(b_i^{(1)},b_i^{(2)})$, a function of $b_i^{(1)}$ and $b_i^{(2)}$, where

$$\lambda(b_i^{(1)},b_i^{(2)}):=\Big(a_i^{(1)}-\sqrt{S_i^{(1)}}b_i^{(1)}\eta_i\Big)^2P^{(1)}+\sum_{k=1}^{2}\Big(b_i^{(k)}\Big)^2$$
$$+\Big(a_i^{(2)}-\sqrt{S_i^{(2)}}b_i^{(2)}\operatorname{sgn}(\rho_i)\eta_i\Big)^2P^{(2)}+\eta_i^2\sum_{k=1}^{2}S_i^{(k)}(b_i^{(k)})^2P^{(\bar{k})}$$
$$-2\sqrt{S_i^{(1)}}b_i^{(1)}\operatorname{sgn}(\rho_i)\eta_i\Big(a_i^{(1)}-\sqrt{S_i^{(1)}}b_i^{(1)}\eta_i\Big)\sqrt{P^{(1)}P^{(2)}}\rho_i$$
$$-2\sqrt{S_i^{(2)}}b_i^{(2)}\eta_i\Big(a_i^{(2)}-\sqrt{S_i^{(2)}}b_i^{(2)}\operatorname{sgn}(\rho_i)\eta_i\Big)\sqrt{P^{(1)}P^{(2)}}\rho_i.$$

Recall that $b_i^{(k)}$ are functions of $S_i^{(k)},k\in[2]$. For functional optimization, each of the expressions inside the expectation in the right-hand side of (18) needs to be minimized, in other words, we need to solve $\min_{b_i^{(1)},b_i^{(2)}}\lambda(b_i^{(1)},b_i^{(2)})$ for each $S_i^{(k)}$. Letting $\frac{\partial\lambda(b_i^{(1)},b_i^{(2)})}{\partial b_i^{(k)}}=0,\ \forall k\in[2]$, the minimum value of $\lambda(b_i^{(1)},b_i^{(2)})$ is achieved when $b_i^{(1)}$ and $b_i^{(2)}$ satisfy (15b) and (15c) respectively. Taking g_i defined in (15e) and substituting (15b) and (15c) into (18), we attain that

$$1=\mathbb{E}\left(\frac{(a_i^{(k)})^2\big(1+S_i^{(k)}(PP^{(\bar{k})}(1-\rho_i^2))/g_i\big)}{S_i^{(k)}P+1}\right).$$

Denote $d_i^{(k)}:=\frac{(a_i^{(k)})^2\left(1+S_i^{(k)}(PP^{(\bar{k})}(1-\rho_i^2))/g_i\right)}{S_i^{(k)}P+1}$. Now the cheap control problem is reduced to

$$\min_{d_i^{(1)},\ d_i^{(2)}} P$$
$$\text{s.t.}\begin{cases}\mathbb{E}\left(d_i^{(k)}\right)=1,\forall k,\bar{k}\in[2],\ \bar{k}\neq k,\\ \mathbb{E}\left(\log d_i^{(k)}+\log\left(\frac{1+S_i^{(k)}P}{1+S_i^{(k)}PP^{(\bar{k})}(1-\rho_i^2)/g_i}\right)\right)=2R^{(k)\star}.\end{cases}$$

With $P^{(1)}=\alpha P$ and $P^{(2)}=(1-\alpha)P$ for $\alpha\in[0,1]$, we have

$$\log\left(\frac{1+S_i^{(k)}P}{1+S_i^{(k)}PP^{(\bar{k})}(1-\rho_i^2))/g_i}\right) \tag{19}$$
$$=\log\left(\frac{1+S_i^{(k)}P}{1+S_i^{(k)}P(1-\rho_i^2))(1-\alpha)/(1+2|\rho_i|\sqrt{\alpha(1-\alpha)})}\right).$$

As the right-hand side term of (19) is monotonically increasing with P, to minimize P, we need to minimize $\mathbb{E}\left(\log\left(\frac{1+S_i^{(k)}P}{1+S_i^{(k)}PP^{(\bar{k})}(1-\rho_i^2)/g_i}\right)\right)$ or equivalently maximize $\mathbb{E}\big(\log d_i^{(k)}\big)$ for fixed $R^{(k)}$ subject to $\mathbb{E}\big(d_i^{(k)}\big)=1$. From the concavity of the logarithm, it follows that the maximization is achieved if all $d_i^{(k)}=1$, which gives $|a_i^{(k)}|$ as in (15a) with $|a_i^{(k)}|>1$.

Thus, the convergence of $\mathbb{E}((X_n^{(k)})^2)$ holds by choosing $b_i^{(k)}$ and $a_i^{(k)}$ as in (15b)–(15a), which also ensures the closed-loop MSS. From (12b), after some simple algebra, the correlation coefficient ρ_{i+1} is updated as (15d) with notations defined in (15e)–(15i). Furthermore, one can see that the choice of parameters of the encoder in [14] are consistent with the coefficients chosen here. Therefore, the PM scheme for fading AWGN-BC in [14] can be obtained from a control-oriented perspective. ■

Remark 3. *We note that the linear coding scheme in Proposition 2 is obtained based on the choice of C_i in* (12a) *and it recovers the result in [1] when the number of user $\mathsf{K}=1$. Also, the cheap control formulation here differs from the PM scheme in [14] in that the cheap control performs the same operation at every step, whereas the PM scheme's startup phase differs from later phases. The cheap control approach has the advantage of unifying the operations for all steps (which simplifies the control-oriented analysis), but it has to wait long enough until that exponentially vanishing bias (i.e., the term $\phi_n^{(k)}a_{cl,kk}X_0^{(k)}-\phi_n^{(k)}a_{cl,k\bar{k}}X_0^{(\bar{k})}$ in* (11)*) becomes negligible [10]. In contrast, the PM scheme is unbiased since the special startup operation provides an unbiased estimation error and the action taken in this first step has no effect on the average transmission power used by later steps, which depends on the steady state behavior. Therefore, except for these minor differences, we show that the PM-based scheme in [14] is a special case of the cheap control method introduced here.*

IV. Conclusions and Future Work

This paper presented a linear code for the Gaussian two-user fading broadcast channel with feedback by using tools from control theory. We showed that the problem of feedback stabilization over control channels is closely related to the problem of communication over the same channel, when the encoder has access to noiseless feedback from the decoders. The growth rate of the open-loop system becomes the rate at which information is communicated over the channel. Based on this equivalence, feedback control methods can be applied to obtain communication schemes, i.e. one can translate the problem of maximizing the achievable rates for fixed power constraint in a communication system to the problem of minimizing the transmission power for fixed rates in a control system. The achievable rates of this linear coding scheme are not optimal as we choose the controller based on the PM scheme, which sometimes is worse than non-feedback schemes. Future work includes broadening and optimizing the form of the controller, incorporating other control methods to derive an achievable region that uniformly outperforms the region without feedback.

References

[1] J. Liu, N. Elia, and S. Tatikonda, "Capacity-achieving feedback schemes for gaussian finite-state markov channels with channel state information," *IEEE Trans. Inf. Theor.*, vol. 61, no. 7, pp. 3632–3650, July 2015.

[2] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theor.*, vol. 43, no. 6, pp. 1986–1992, Nov 1997.

[3] O. Shayevitz and M. Feder, "Optimal Feedback Communication Via Posterior Matching," *IEEE Trans. Inf. Theor.*, vol. 57, no. 3, pp. 1186–1222, March 2011.

[4] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–i: No bandwidth constraint," *IEEE Trans. Inf. Theor.*, vol. 12, no. 2, pp. 172–182, April 1966.

[5] A. E. Gamal, "The feedback capacity of degraded broadcast channels (corresp.)," *IEEE Trans. Inf. Theor.*, vol. 24, no. 3, pp. 379–381, May 1978.

[6] S. R. Bhaskaran, "Gaussian Broadcast Channel With Feedback," *IEEE Trans. Inf. Theor.*, vol. 54, no. 11, pp. 5252–5257, Nov 2008.

[7] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback (Corresp.)," *IEEE Trans. Inf. Theor.*, vol. 30, no. 4, pp. 667–671, July 1984.

[8] J. Schalkwijk, "A coding scheme for additive noise channels with feedback–ii: Band-limited signals," *IEEE Trans. Inf. Theor.*, vol. 12, no. 2, pp. 183–189, 1966.

[9] L. V. Truong and H. Yamamoto, "Posterior Matching for Gaussian Broadcast Channels with Feedback," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. 100-A, pp. 1165–1178, 2017.

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

[11] E. Ardestanizadeh, P. Minero, and M. Franceschetti, "LQG Control Approach to Gaussian Broadcast Channels With Feedback," *IEEE Trans. Inf. Theor.*, vol. 58, no. 8, pp. 5267–5278, 2012.

[12] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. Inf. Theor.*, vol. 48, no. 6, pp. 1423–1438, 2002.

[13] S. B. Amor and M. Wigger, "Linear-feedback MAC-BC duality for correlated BC-noises, and iterative coding," in *Annu. Allerton Conf. Commun., Contr., Comput.*, 2015, pp. 1502–1509.

[14] S. Li, D. Tuninetti, and N. Devroye, "The Fading Gaussian Broadcast Channel with Channel State Information and Output Feedback," in *IEEE Int. Symp. Inf. Theor.*, 2020, pp. 1474–1479.

[15] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. USA: Cambridge University Press, 2012.

[16] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York, NY, USA: Wiley, 1972.

[17] K. J. Astrom and R. M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers*. USA: Princeton University Press, 2008.

[18] A. A. Zaidi, T. J. Oechtering, S. Yuksel, and M. Skoglund, "Stabilization of linear systems over gaussian networks," *IEEE Trans. Automat. Contr.*, vol. 59, no. 9, pp. 2369–2384, 2014.

[19] Y. Fang, "Stability analysis of linear control systems with uncertain parameters," PhD dissertation, Case Western Reserve University, 1994.

[20] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proc. IEEE*, vol. 59, no. 11, pp. 1557–1566, 1971.

[21] H. Witsenhausen, "A counterexample in stochastic optimum control," *Siam Journal on Control*, vol. 6, pp. 131–147, 1968.

[22] Yu-Chi Ho and K'ai-Ching Chu, "Team decision theory and information structures in optimal control problems–part i," *IEEE Trans. Automat. Contr.*, vol. 17, no. 1, pp. 15–22, 1972.

[23] O. Costa, M. Fragoso, and R. Marques, *Discrete-Time Markov Jump Linear Systems*. London, U.K.: London Springer, 2005.