The Fading Gaussian Broadcast Channel with Channel State Information and Output Feedback

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

***Abstract*—The fading broadcast channel (BC) with additive white Gaussian noise (AWGN) channel, channel output feedback (COF) and channel state information (CSI) is considered. Perfect CSI is available at the receivers, and unit delayed CSI along with COF at the transmitter. Under the assumption of memoryless fading, a posterior matching scheme that incorporates the additional CSI feedback into the coding scheme is presented. With COF, the achievable rates depend on the joint distribution of the fading process. Numerical examples show that the capacity region of two-user fading AWGN-BC is enlarged by COF. The coding scheme is however suboptimal since some parts of the achievable rate region are outperformed by superposition coding without COF.**

I. Introduction

In this paper we consider the problem of communication over wireless additive white Gaussian noise (AWGN) fading broadcast channels (BC). This builds upon work for the fading point-to-point (P2P) AWGN channel with various channel state information (CSI) assumptions. The capacity of P2P channels with memory and with imprecise or delayed CSI was investigated in [1]; the capacity of a fading channel with CSI at both the transmitter (CSIT) and receiver (CSIR), or at the receiver only, was derived in [2]; the Markov channel capacity with instantaneous CSIR and delayed CSIT was obtained in [3]; thus, with CSIR, the fading P2P capacity is known.

For the fading BC the capacity region is known when the channel state is fixed and time-invariant, and when the channel state is time-variant, for single antenna [4]–[7] and for multiple antennas under the assumption that perfect CSIT and CSIR [8]. It is shown in [8] that a superposition of Gaussian codes is optimal for the degraded vector broadcast channel and dirty-paper coding is optimal for the non-degraded case. However, the capacity is unknown once the assumption of perfect CSIT is removed. [9] obtains the capacity region for a certain class of single-antenna fading AWGN-BCs with CSIR only when the input distribution is Gaussian; [10] uses the entropy power inequality (EPI), along with an optimization framework, to derive upper bounds and superposition coding lower bounds on the sum-capacity of two-user fading BC where the CSI is unknown to the transmitter and one of the users has a constant (non-fading) channel; [11] provides inner and outer bounds to the capacity region of fading Gaussian BC without CSIT only, and demonstrates a binary expansion superposition (BES) scheme that achieves rates within 6.386 bits/s/Hz per user to the outer bound.

While channel output feedback (COF) cannot increase the capacity of memoryless P2P channels, it can significantly reduce the probability of error and even simplify capacity achieving schemes. Shayevitz and Feder propose a fundamental principle for P2P communication over general memoryless channels in the presence of COF, known as posterior matching (PM) [12]. For the P2P AWGN channel, it reduces to the Schalkwijk-Kailath (SK) scheme [13], [14], which has an error probability decaying double-exponentially with the block length, in contrast with the single exponential attained by schemes without feedback.

It is also known that COF can increase the capacity region for BCs, except for the physically degraded BC [15]. [16] constructs a coding scheme for AWGN-BC with two receivers and COF, and shows that the capacity region is enlarged when the received signals are conditionally independent given the transmitted signal but the noise variances are equal. [17] extended the P2P PM scheme in [12] to BCs and obtain the same exact region as in [16], [18] shows that 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 error probability decays doubly exponential in the block length.

The capacity region of the AWGN-BC with CSIR only is known only to within a constant gap [11], and the effect of COF on this practically relevant channel model has not been investigated to the best of our knowledge. Yet, all modern wireless systems have mechanisms to provide CSI at the transmitter through feedback. Towards understanding the performance enhancement due to CSI obtained at the transmitter through feedback, in our previous work [19], we use the Layered Packet Erasure BC (LPE-BC) model from [11] to gain insights into the high Signal to Noise Ratio (SNR) capacity of the fading Gaussian BC. Here we explore the advantages of incorporating both CSI and COF for single-antenna AWGN channels as a first step towards translating the high-SNR result of [19] to finite SNR. Similar ideas have been investigated in [3], [20], [21]. In particular, [20] proposes a capacity achieving linear feedback scheme for P2P i.i.d. fading and AWGN channel with CSIR and unit delayed CSIT, which is motivated by control theory and employs the SK scheme. However, it is not clear how to extended the scheme to BCs.

Contributions: All exact capacity results for AWGN-BCs are without COF [4]–[7], or without fading [16], [17]. We consider the AWGN-BC with current CSIR and unit delayed CSIT as well as COF. We extend the work in [17] so as to include channel fading and numerically compare the performance under different channel conditions. In particular, we derive an achievable region that incorporates the joint distribution of the fading and show that this region is larger than the capacity region without feedback.

Paper Organization: Section II introduces the channel model. Section III described our proposed scheme. Section IV gives some numerical evaluations. Section V concludes the paper.

Notation: Random variables (RVs) are denoted by upper-case letters, their realizations by corresponding lower-case letters. A real-valued RV X is associated with a distribution $P_X(\cdot)$ defined on the usual Borel σ-algebra over $\mathbb{R}$, and we write $X \sim P_X$. The cumulative distribution function (c.d.f.) of X is given by $F_X(x) = P_X\left((-\infty, x]\right)$, and the inverse c.d.f. is defined as $F_X^{-1}(t) := \inf\{x : F_X(x) > t\}$. The support of X is the intersection of all closed sets A for which $P_X(\mathbb{R}\backslash A) = 0$ and is denoted supp(X). We write $\mathbb{E}(\cdot)$ for expectation and $\mathbb{P}(\cdot)$ for the probability of a measurable event within the parentheses. The uniform probability distribution over $(0, 1)$ is denoted throughout by $\mathcal{U}$. The Gaussian distribution with mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$. Define $(f \circ g)(x) := f(g(x))$, $Y_i^{j(k)} := \{Y_m^{(k)}\}_{m=i}^{j}$ for $i \leq j$. We use $|\Delta|$ for the length of an interval $\Delta \subseteq \mathbb{R}$. We use $\operatorname{sgn}(x)$ to denote the sign function, where $\operatorname{sgn}(x) := 1$ if $x \geq 0$ and $\operatorname{sgn}(x) := -1$ if $x < 0$.

II. Channel Model

We consider a communication system where one transmitter and K receivers are connected via a fading AWGN channel, global CSIR is available at the receivers, and all channel outputs are noiselessly fed back to the transmitter. Let Θ_k be a random message point uniformly distributed over the unit interval that must be transmitted from the transmitter to receiver $k \in [\mathsf{K}]$. The received signal $Y_n^{(k)}$ for user $k \in [\mathsf{K}]$ at time $n \in \mathbb{N}$ is

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

where $X_n \in \mathbb{R}$ denotes the transmitted signal, $Z_n^{(k)} \sim \mathcal{N}(0, \sigma_k^2)$ is the real-valued AWGN for user k, and $S_n^{(k)} \in \mathcal{S}$ is the channel state of user k with alphabet $\mathcal{S}$. We assume that the random variables $(S^{(1)}, \ldots, S^{(\mathsf{K})})$ form a memoryless process over time, that the noise variables $(Z^{(1)}, \ldots, Z^{(\mathsf{K})})$ are independent across users and time, and that the input is subject to the average power constraint $\mathbb{E}[X^2] \leq P$.

In this paper, we focus on the case of one-unit-delayed transmitter-side and instantaneous receiver-side CSI. Hence, the channel state $S_n^{(k)}$ is available at the receiver k at time n and available at the transmitter at time $n+1$. The received signal $Y_n^{(k)}$ is sent back to the transmitter through the noiseless feedback channel. Thus, the transmitted symbol X_n at time n can depend on the messages $\{\Theta_k\}_{k=1}^{\mathsf{K}}$ to be conveyed to the receivers and the previous channel output sequences $\{S_1^{n-1(k)}, Y_1^{n-1(k)}\}_{k=1}^{\mathsf{K}}$.

A *transmission scheme* is a sequence of a-priori agreed upon measurable transmission functions $g_n : (0,1)^{\mathsf{K}} \times \mathbb{R}^{(n-1)\mathsf{K}} \times \mathbb{R}^{(n-1)\mathsf{K}} \to \mathbb{R}$, so that the input to the channel generated by the transmitter is given by

$$\begin{aligned} X_n =& g_n(\Theta_1, \cdots, \Theta_{\mathsf{K}}, \\ & S_1^{n-1(1)}, \cdots, S_1^{n-1(\mathsf{K})}, Y_1^{n-1(1)}, \cdots, Y_1^{n-1(\mathsf{K})}). \end{aligned}$$

A *decoding rule* is a sequence of measurable mappings $\{\Delta_n^{(k)} : \mathbb{R}^n \times \mathbb{R}^n \to \mathcal{E}\}_{n=1}^{\infty}$, where $\mathcal{E}$ is the set of all open intervals in $(0, 1)$. Without loss of generality, the transmitted message lies in some interval. We refer to $\Delta_n^{(k)}(y_1^{n(k)}, s_1^{n(k)})$ as the decoded interval of user k. For different users, $|\Delta_n^{(k)}(y_1^{n(k)}, s_1^{n(k)})|$ may be different. The correct decoded interval is the one that the transmitted message lies in. The *error probabilities* at time n associated with a transmission scheme and a decoding rule, are thus defined as

$$p_{n,e}^{(k)} := \mathbb{P}(\Theta_k \notin \Delta_n^{(k)}(S_1^{n(k)}, Y_1^{n(k)})),\ k \in [\mathsf{K}],$$

and the corresponding *rate* at time n is defined by

$$R_n^{(k)} := -\frac{1}{n} \log\left(|\Delta_n^{(k)}(S_1^{n(k)}, Y_1^{n(k)})|\right),\ k \in [\mathsf{K}]. \quad (1)$$

We say that a transmission scheme together with a decoding rule achieves a rate tuple $(R_1, \cdots, R_{\mathsf{K}})$ over the fading AWGN channel if for all $k \in [\mathsf{K}]$, it satisfies

$$\lim_{n\to\infty} \mathbb{P}\left(R_n^{(k)} < R_k\right) = 0,\ \lim_{n\to\infty} p_{n,e}^{(k)} = 0. \quad (2)$$

The rate tuple is achieved with an input power constraint P for $P > 0$, if it also satisfies

$$\lim_{n\to\infty} \sup \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i^2] \leq P. \quad (3)$$

Note that in the standard framework [22], a uniformly distributed message set $\left[2^{\lceil n\hat{R}_k \rceil}\right]$ is conveyed to receiver k with coding rate $\hat{R}_k$. By [12, Lemma 3] and incorporating the feedback of channel fading, the achievability as defined in (2) and (3) implies achievability in the standard framework.

Remark: For the single-user case, we incorporate the additional CSI feedback into the PM scheme as follows, which now takes on a recursive time-varying form by standard manipulations. The channel input at time $n+1$ is

$$\begin{aligned} X_{n+1} &= g_{n+1}(\Theta_0, y_1^n, s_1^n) \\ &= F_X^{-1} \circ F_{\Theta_0|Y_1^n, S_1^n}(\Theta_0 | y_1^n, s_1^n), \\ &= F_X^{-1} \circ F_{X_n|Y_n, S_n}(X_n | Y_n, S_n), \end{aligned} \quad (4)$$

for any n, where Θ_0 is uniform over the unit interval, and with the initial condition $X_1 = g_1(\Theta_0) = F_X^{-1}(\Theta_0)$. The equality in (4) indicates that each signal component at time $n+1$ is statistically independent of the signal feedback from the corresponding user at time n by [17, Section III]. One can show that the scheme achieves the ergodic capacity $C = \frac{1}{2}\mathbb{E}\left[\log(1 + S\frac{P}{\sigma^2})\right]$, where the expectation is over the CSI variable S, and that the probability of error decays doubly-exponentially in n, i.e., $p_{n,e} \leq 4\exp\left(-\frac{\pi}{8} 2^{2n(C-R)}\right)$ – details are not reported here for sake of space.

III. The Two-User Fading AWGN-BC

In this section, we consider the fading AWGN-BC with $\mathsf{K} = 2$ receivers. Since the fading $S_n^{(k)}$ is available at receiver k at time n, we can equivalently express the channel model as

$$Y_n^{(k)} = X_n + \frac{Z_n^{(k)}}{\sqrt{S_n^{(k)}}},\ k \in [\mathsf{K}], \tag{5}$$

where the variance of the AWGN in (5) varies over time. We propose an extension of the PM scheme presented in [17] and of the single-user case in the remark in Section II that achieves the rates in Theorem 1.

Theorem 1. *The rate-tuple (R_1, R_2) is achievable if it satisfies*

$$R_k < R_k^\star := \lim_{n\to\infty} -\frac{1}{n}\log \mathbb{E}\left[\prod_{i=1}^{n} a_i^{(k)}\right],\ k \in [2], \tag{6}$$

and $W_n^{(k)} := \mathbb{E}[V_n^{(k)}]^2$ is bounded, where the expectation is with respect to the channel fading and where $a_i^{(k)}$ is defined in (10).

The rest of the section is dedicated to proving Theorem 1.

Transmission scheme: In the first two transmissions, the transmitter broadcasts messages $V_1^{(1)} = F_{V_1}^{-1}(\Theta_1)$ and $V_2^{(2)} = F_{V_2}^{-1}(\Theta_2)$ separately, where $V_k \sim \mathcal{N}(0, P_k), P_1 + P_2 = P$. Set $V_2^{(1)} = V_1^{(1)}$ after the second transmission. For $n \geq 3$, we let

$$X_n = \beta_n \left(V_n^{(1)} + \mathrm{sgn}(\mu_n) V_n^{(2)}\right),$$
$$V_n^{(k)} = \frac{1}{a_{n-1}^{(k)}}\left(V_{n-1}^{(k)} - b_{n-1}^{(k)} Y_{n-1}^{(k)}\right), k \in [2], \tag{7}$$

where $\mathrm{sgn}(\mu_n)$ is used to align the signal of user 1 with the signal of user 2; where $V_n^{(k)}$ is the encoded message of user k at time n and X_n is the message to be broadcast to all users; where

$$\mu_n = \frac{\mathbb{E}\left[V_n^{(1)} V_n^{(2)} | S_{n-1}^{(1)} = s_{n-1}^{(1)}, S_{n-1}^{(2)} = s_{n-1}^{(2)}\right]}{\sqrt{P_1 P_2}}, \tag{8}$$

$$\beta_n = \sqrt{\frac{P}{P + 2|\mu_n|\sqrt{P_1 P_2}}}, \tag{9}$$

$$a_n^{(k)} = \sqrt{\frac{\mathrm{var}(V_n^{(k)} | Y_n^{(k)} = y_n^{(k)}, S_n^{(k)} = s_n^{(k)})}{P_k}}, \tag{10}$$

$$b_n^{(k)} = \frac{\mathbb{E}[V_n^{(k)} Y_n^{(k)} | S_n^{(k)} = s_n^{(k)}]}{\mathrm{var}(Y_n^{(k)} | S_n^{(k)} = s_n^{(k)})}. \tag{11}$$

In particular, μ_n in (8) is the normalized correlation coefficient of the messages to two users, β_n in (9) is used to satisfy the input power constraint. User $k \in [2]$ receives the signal

$$Y_n^{(k)} = \beta_n \left(V_n^{(1)} + \mathrm{sgn}(\mu_n) V_n^{(2)}\right) + Z_n^{(k)} / \sqrt{s_n^{(k)}}, \tag{12}$$

where $s_n^{(k)}$ is the realization of $S_n^{(k)}$ and is available at the receiver, and will be available at the transmitter through the feedback. Given the output $Y_n^{(k)}$ and fading $S_n^{(k)}$, the conditional variance $\mathrm{var}(V_{n-1}^{(k)} | Y_{n-1}^{(k)} = y_{n-1}^{(k)}, S_{n-1}^{(k)} = s_{n-1}^{(k)})$ can be interpreted as the noise power. Then, $a_n^{(k)}$ in (10) may be interpreted as the square root of the noise-to-signal ratio of user k at time n. The parameters $a_n^{(k)}$ in (10) and $b_n^{(k)}$ in (11) are chosen so that

$$V_{n+1}^{(k)} = F_V^{-1} \circ F_{V_n^{(k)} | Y_n^{(k)}, S_n^{(k)}}(V_n^{(k)} | Y_n^{(k)}, S_n^{(k)}). \tag{13}$$

Decoding rule: In the first two transmissions, user $k \in [2]$ receives $Y_1^{(k)} = V_1^{(1)} + Z_1^{(k)} / \sqrt{s_1^{(k)}}$, and $Y_2^{(k)} = V_2^{(2)} + Z_2^{(k)} / \sqrt{s_2^{(k)}}$, separately. Each user $k \in [2]$ selects a fixed interval $J_1^{(k)} = (l_k, t_k) \subset \mathbb{R}$ with respect to $V_n^{(k)}, n \in [2]$.

For $n \geq 3$, each user k receives $Y_n^{(k)}$ given by (12). Each user k updates the interval $J_n^{(k)}$ for the $V_n^{(k)}$ as

$$J_n^{(k)} := \left(T_n^{(k)}(l_k), T_n^{(k)}(t_k)\right), \tag{14}$$

$$T_n^{(k)}(x) := w_1^{(k)} \circ w_2^{(k)} \circ \cdots \circ w_n^{(k)}(x), \tag{15}$$

$$w_n^{(k)}(x) := a_n^{(k)} x + b_n^{(k)} Y_n^{(k)}. \tag{16}$$

According to (10), $a_n^{(k)} > 0$ which ensures that $w_n^{(k)}(x)$ in (16) and $T_n^{(k)}(x)$ in (15) are monotonically increasing in x for any realization of $Y_1^{n(k)}$. The decoded interval $\Delta_n^{(k)}(Y_1^{n(k)}, S_1^{n(k)})$ of the message Θ_k is determined as

$$\Delta_n^{(k)}(Y_1^{n(k)}, S_1^{n(k)}) = F_{V_k}\left(J_n^{(k)}\right), \tag{17}$$

where $F_{V_k}\left(J_n^{(k)}\right)$ is the c.d.f. of an interval and we define

$$F_{V_k}\left((a, b)\right) := \left(\int_{-\infty}^{a} f_{V_k}(x)dx, \int_{-\infty}^{b} f_{V_k}(x)dx\right). \tag{18}$$

Analysis: By definition, we need to show (2). Let $R_n^{(k)}$ be the instantaneous rate of transmission of message Θ_k to user k defined in (1). For any rate R_k, we have

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}(R_n^{(k)} < R_k) \\
&= \lim_{n\to\infty} \mathbb{P}\left(-\frac{1}{n}\log\left|\Delta_n^{(k)}(Y_1^{n(k)}, S_1^{n(k)})\right| < R_k\right) && (19)\\
&= \lim_{n\to\infty} \mathbb{P}\left(\left|\Delta_n^{(k)}(Y_1^{n(k)}, S_1^{n(k)})\right| > 2^{-nR_k}\right) \\
&\leq \lim_{n\to\infty} \mathbb{P}\left(\left|J_n^{(k)}\right| > \frac{2^{-nR_k}}{l}\right) && (20)\\
&\leq \lim_{n\to\infty} l \cdot 2^{nR_k}\mathbb{E}\left[|J_n^{(k)}|\right] && (21)\\
&= \lim_{n\to\infty} l \cdot 2^{nR_k}\mathbb{E}\Big[\Big|w_1^{(k)} \circ w_2^{(k)} \circ \cdots \circ w_n^{(k)}(t_k) \\
&\quad - w_1^{(k)} \circ w_2^{(k)} \circ \cdots \circ w_n^{(k)}(l_k)\Big|\Big] && (22)
\end{aligned}$$

$$\text{where}\quad l = \sup_{x\in\mathbb{R}}\{f_V(x)\}; \tag{23}$$

where (19) follows from (1); where (20) follows from (17), (18) and (23); where (21) follows from Markov's inequality; where (22) follows from (14) and (15). From (16), we have

$|w_n^{(k)}(t) - w_n^{(k)}(l)| = a_n^{(k)}|t-l|$. Recall that $R_k^\star$ is defined in (6). Also, here $a_n^{(k)}$ is a function of the channel fading instead of a sequence of real numbers that concentrates to some a_k as $n \to \infty$ in [17]. Continuing with (22), we have

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}(R_n^{(k)} < R_k) \\
&= \lim_{n\to\infty} l \cdot 2^{nR_k} \mathbb{E}\left[a_1^{(k)} a_2^{(k)} \cdots a_n^{(k)} |t_k - l_k|\right] \\
&= \lim_{n\to\infty} l \cdot 2^{nR_k} \mathbb{E}\left[\prod_{i=1}^n a_i^{(k)} \left|J_1^{(k)}\right|\right] \\
&= \lim_{n\to\infty} l \cdot 2^{nR_k} \mathbb{E}\left[\prod_{i=1}^n a_i^{(k)}\right] \left|J_1^{(k)}\right| \\
&= \lim_{n\to\infty} l \cdot 2^{-n(R_k^\star - R_k)} \left|J_1^{(k)}\right|.
\end{aligned}$$

Hence, for some $\epsilon \in (0,1)$, any rate $R_k < R_k^\star$, $\lim_{n\to\infty} \mathbb{P}(R_n^{(k)} < R_k) \to 0$ holds if

$$\left|J_1^{(k)}\right| = o\left(2^{\epsilon n(R_k^\star - R_k)}\right). \tag{24}$$

Similar to the proof of [17, Theorem 1], by selecting $J_1^{(k)}$ satisfying (24) and $|J_1^{(k)}| \to \infty$ as $n \to \infty$, we can show $\lim_{n\to\infty} p_{n,e}^{(k)} = 0$. Next, we want to evaluate the achievable rate in (6). According to [12, Lemma 1], we have

$$F_{V_n^{(k)}|Y_n^{(k)}, S_n^{(k)}}(V_n^{(k)}|Y_n^{(k)} = y_n^{(k)}, S_n^{(k)} = s_n^{(k)}) \sim \mathcal{U}.$$

Since $V_k \sim \mathcal{N}(0, P_k)$ and the transmitted signal for user $k, k \in [2]$ is (13), then for any $n \geq 1$, $V_n^{(k)} \sim \mathcal{N}(0, P_k)$. Due to space limitations, the computation details are omitted. Following similar steps as in [17, Section V], we obtain for all n

$$a_n^{(k)} = \sqrt{\frac{\frac{\sigma_k^2}{s_n^{(k)}} + P_{\bar{k}} P(1-\mu_n^2)/(P + 2|\mu_n|\sqrt{P_1 P_2})}{P + \frac{\sigma_k^2}{s_n^{(k)}}}},\ k \neq \bar{k},$$

$$b_n^{(1)} = \frac{\beta_n(P_1 + |\mu_n|\sqrt{P_1P_2})}{P + \frac{\sigma_1^2}{s_n^{(1)}}},$$

$$b_n^{(2)} = \frac{\beta_n(\operatorname{sgn}(\mu_n)P_2 + \mu_n\sqrt{P_1P_2})}{P + \frac{\sigma_2^2}{s_n^{(2)}}}.$$

It can be easily checked that $a_n^{(k)} \in [\frac{1}{1 + s_n^{(k)}P/\sigma_k^2}, 1), k \in [2]$. By (7), for two successive time slots, we need to have

$$\begin{aligned}
\mu_{n+1} = &\frac{\sqrt{A_n}}{\sqrt{L_n(\mu_n)}} \Big(D(\mu_n)\mu_n \\
&-PB_n\Big(P\mu_n + \operatorname{sgn}(\mu_n)\sqrt{P_1P_2}(1+\mu_n^2)\Big)/A_n\Big)
\end{aligned} \tag{25}$$

where $A_n = \left(P + \frac{\sigma_1^2}{s_n^{(1)}}\right)\left(P + \frac{\sigma_2^2}{s_n^{(2)}}\right)$, $B_n = P + \frac{\sigma_1^2}{s_n^{(1)}} + \frac{\sigma_2^2}{s_n^{(2)}}$, $D(x) = P + 2|x|\sqrt{P_1P_2}$, $G(x) = P(1-x^2)$, and $L_n(x) = (D(x)/s_n^{(1)} + P_2G(x))(D(x)/s_n^{(2)} + P_1G(x))$.

Different from [17], we cannot show that $|\mu_n|$ is unchanged by using the trick that setting $\mu_n = (-1)^{n+1}\mu$, due to the time-variant i.i.d. channel fading in (25). However, since the transmitter sends a linear combination of $V_n^{(1)}$ and $V_n^{(2)}$ from time $n = 3$ onwards, we can assign some number in $(-1,1)$ to μ_2 and update μ_n based on the realization of fading states by (25). Finally, we obtain for $k \in [2], k \neq \bar{k}$

$$\begin{aligned}
R_k < R_k^\star &= \lim_{n\to\infty} -\frac{1}{n}\log\mathbb{E}\left[\prod_{i=1}^n a_i^{(k)}\right] \\
&= \lim_{n\to\infty} -\frac{1}{n}\log\mathbb{E}\left[\prod_{i=1}^n \sqrt{\frac{\frac{\sigma_k^2}{s_i^{(k)}} + \frac{P_{\bar{k}}P(1-\mu_i^2)}{(P+2|\mu_i|\sqrt{P_1P_2})}}{P + \frac{\sigma_k^2}{s_i^{(k)}}}}\right].
\end{aligned} \tag{26}$$

We have not been able to rigorously prove that the limits in (26) exist. We note that in all cases we tried numerically, the achievable rates converge to appropriate values, regardless of the starting point μ_2. The numerical examples in Section IV obtained the rates by iterating (25).

Finally, we note that for the AWGN-BC with COF, the achievable region should depend on the joint distribution of the channel fading, which is confirmed by (26) since $(s_n^{(1)}, s_n^{(2)})$ jointly determine μ_n in (25).

Remark: When fading is a constant (i.e., $S_n^{(k)} = s_k, \forall n$), we can set $\mu_{n+1} = -\mu_n$ and $|\mu_n|$ is a constant in $(0,1)$ obtained by solving (25), following the same idea as in [17]. Let $\gamma_k = \frac{s_kP}{\sigma_k^2}, |\mu_n| = \mu$, then (26) simplifies to

$$R_k < \frac{1}{2}\log\left(\frac{1+\gamma_k}{1 + \gamma_k\frac{P_{\bar{k}}(1-\mu^2)}{P+2\mu\sqrt{P_1P_2}}}\right),\quad \begin{matrix} k \in [2] \\ k \neq \bar{k}. \end{matrix}$$

For the non-fading case, since the achievable rate is a function of μ_n given in (25), based on our experiments, μ depends on the difference of γ_1 and γ_2, i.e., when $\gamma_1 > \gamma_2$, the larger γ_1/γ_2 is, the larger μ is and the opposite holds when $\gamma_1 < \gamma_2$. In particular, when $\gamma_1 = \gamma_2$, the achievable region of the PM scheme with COF always exceeds the capacity region without COF and the two regions get closer as γ_1 (or γ_2) increases. Note that it is not trivial to find parameters of the PM scheme (i.e., $a_n^{(k)}, b_n^{(k)}, \beta_n$) to achieve the region in [23] for the static channel, since they send messages with powers that vary at each transmission.

The case with fading is even more complicated and experiments show that the achievable region highly relates to the probability of the difference of the fading pairs when the capacity of P2P transmission for each user is fixed. That is, the higher probability of fading pairs with small difference, the smaller variance of $|\mu_n|$, the better the performance of the coding scheme. Detailed examples are given in the following section.

IV. Numerical Evaluations

The capacity region of the fading AWGN-BC without feedback is known when the fading is constant [4]. In this section,

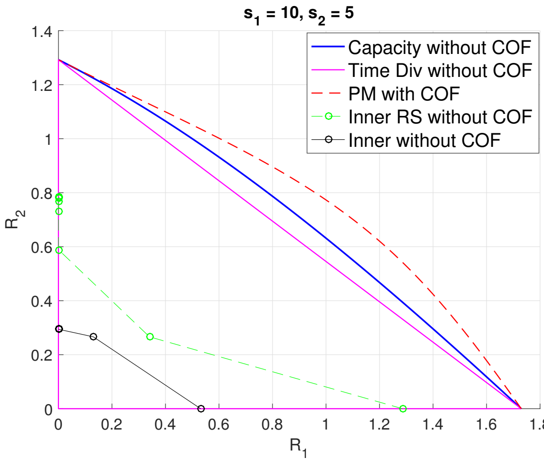

Fig. 1: Non-fading AWGN-BC: receivers 1 and 2 have SNRs of 10dB and 7dB. The inner bounds without COF are from [11].

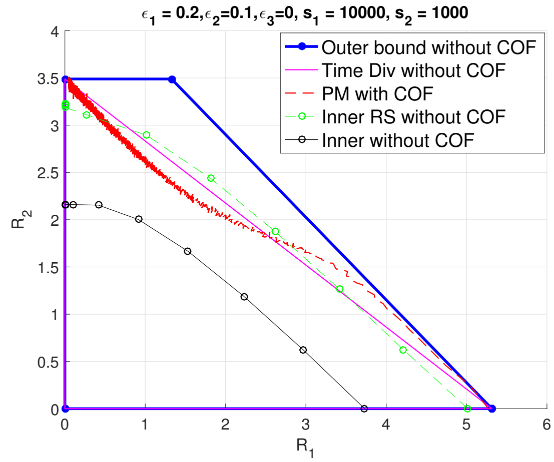

Fig. 2: Fading AWGN-BC: receivers 1 and 2 have time-varying SNRs of 40 and 30 dB with probability 0.8 and 0.7 respectively.

we set $S_n^{(k)} = s_k^\star, k \in [2]$ with some probability p_k and zero otherwise, and compare the achievable rates in (26) with the achievable rates in [11] and the capacity region without feedback. Note that the channel model in [11] considers both real and complex components; here we only consider the real component, thus we scale the rates in [11] by 1/2. In all figures, blue solid line represents the capacity region without COF (achieved by superposition coding) or the outer bound to the capacity region without COF from [11]; magenta solid line indicates the achievable region of time division without COF with power control; red dashed line demonstrates the proposed PM scheme with COF; green and black marked lines illustrate the inner bound without COF achieved by reverse stripping (RS) in [11, Theorem 5] and by [11, Theorem 4], respectively.

For numerical comparisons, we set $P = 1, \sigma_1^2 = \sigma_2^2 = 1$. For the non-fading channel, Fig. 1 considers $s_1^\star = 10, s_2^\star = 5$, i.e., user 1 has SNR 10dB and user 2 has SNR 7dB. We note that COF indeed enlarges the capacity region of (non-physically degraded) fading AWGN-BC, as part of the achievable rate pairs are outside the capacity region without feedback.

Next, we consider the fading channel with four states, where $(S_n^{(1)}, S_n^{(2)})$ takes value $(0,0)$ with probability ϵ_1, $(s_1^\star, 0)$ with probability ϵ_2, $(0, s_2^\star)$ with probability ϵ_3, and $(s_1^\star, s_2^\star)$ with probability ϵ_4 such that $\epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4 = 1$. In this case the capacity region without COF is unknown, thus we plot the outer bound in [11, Theorem 3]. For Fig. 2 and Fig. 3 we set $s_1^\star = 10000, s_2^\star = 1000$. Specifically, $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) = (0.2, 0.1, 0, 0.7)$ for Fig. 2 and $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) = (0, 0.3, 0.2, 0.5)$ for Fig. 3. Comparing the two plots with fading, the achievable region of our PM scheme depends on the joint distribution of the channel fading instead of the marginal ones. When $S_1 = 0, S_2 = 0$, nobody receives messages, thus it does not contribute to the achievable rates. Here the smallest difference between $S^{(1)}$ and $S^{(2)}$ is $S_1 = s_1^\star, S_2 = s_2^\star$, except for $S_1 = 0, S_2 = 0$ and $\mathbb{P}(S_1 = s_1^\star, S_2 = s_2^\star)$ is larger in Fig. 2 than the one in

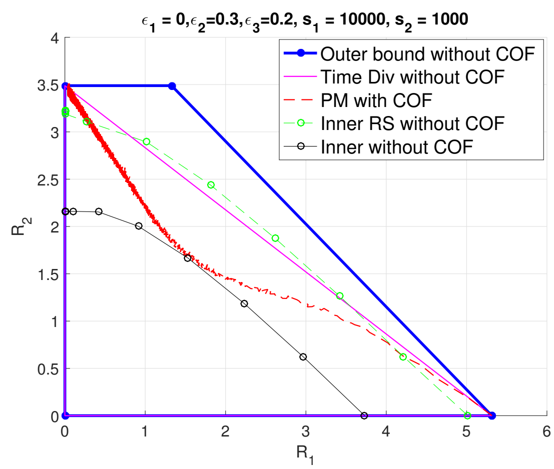

Fig. 3: Receivers 1 and 2 have time-varying SNRs of 40 and 30 dB with probability 0.8 and 0.7 respectively, but a different joint distribution from Fig. 2

Fig. 3, which might be the reason that the performance in Fig. 2 is better than Fig. 3 based on our analysis in the end of Section III.

V. Conclusions

We presented a feedback scheme for the fading AWGN two-user broadcast channel. The analysis of the proposed posterior matching scheme was modified and compared to past work for static channels by considering the dynamic nature of the channel fading in the feedback. Numerical evaluations show that the achievable region of the two-user scheme exceeds the schemes without feedback in some cases. In general, our proposed PM scheme is sub-optimal because it sometimes achieves rates below those without feedback. The reason might be that the messages are transmitted with a fixed power at each time. Future work includes determining outer bounds, deriving new strategies for an achievable region that uniformly outperforms the region without feedback, and incorporating a two-phase scheme from our past work [19].

REFERENCES

[1] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 2007–2019, Sep. 1999.

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

[3] H. Viswanathan, "Capacity of markov channels with receiver csi and delayed feedback," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 761–771, March 1999.

[4] P. Bergmans, "A simple converse for broadcast channels with additive white gaussian noise (corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 279–280, March 1974.

[5] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 2–14, January 1972.

[6] Lifang Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels .i. ergodic capacity," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1083–1102, March 2001.

[7] D. N. Tse, "Optimal power allocation over parallel gaussian broadcast channels," in *Proceedings of IEEE International Symposium on Information Theory*, June 1997, pp. 27–.

[8] H. Weingarten, Y. Steinberg, and S. S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[9] D. Tuninetti and S. Shamai, "On two-user fading gaussian broadcast channels with perfect channel state information at the receivers," in *IEEE International Symposium on Information Theory, 2003. Proceedings.*, June 2003, pp. 345–345.

[10] A. Jafarian and S. Vishwanath, "On the capacity of one-sided two user gaussian fading broadcast channels," in *2008 Information Theory and Applications Workshop*, Jan 2008, pp. 428–432.

[11] D. N. C. Tse and R. D. Yates, "Fading broadcast channels with state information at the receivers," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3453–3471, June 2012.

[12] O. Shayevitz and M. Feder, "Optimal feedback communication via posterior matching," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1186–1222, March 2011.

[13] 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, April 1966.

[14] J. Schalkwijk, "A coding scheme for additive noise channels with feedback–ii: Band-limited signals," *IEEE Transactions on Information Theory*, vol. 12, no. 2, pp. 183–189, April 1966.

[15] A. E. Gamal, "The feedback capacity of degraded broadcast channels (corresp.)," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 379–381, May 1978.

[16] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the gaussian broadcast channel with feedback (corresp.)," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 667–671, July 1984.

[17] L. V. Truong and H. Yamamoto, "Posterior Matching for Gaussian Broadcast Channels with Feedback," *arXiv e-prints*, p. arXiv:1404.2520, Apr 2014.

[18] S. R. Bhaskaran, "Gaussian broadcast channel with feedback," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5252–5257, Nov 2008.

[19] S. Li, D. Tuninetti, and N. Devroye, "On the capacity region of the layered packet erasure broadcast channel with feedback," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, May 2019, pp. 1–6.

[20] Jialing Liu, N. Elia, and S. Tatikonda, "Capacity-achieving feedback scheme for flat fading channels with channel state information," in *Proceedings of the 2004 American Control Conference*, vol. 4, June 2004, pp. 3593–3598 vol.4.

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

[22] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York:Wiley, 2006.

[23] S. Belhadj Amor, Y. Steinberg, and M. Wigger, "Mimo mac-bc duality with linear-feedback coding schemes," *IEEE Transactions on Information Theory*, vol. 61, no. 11, pp. 5976–5998, 2015.