

# On the Capacity Enlargement of Gaussian Broadcast Channels with Passive Noisy Feedback

Aditya Narayan Ravi, Sibi Raj B. Pillai, Vinod M. Prabhakaran, Michèle Wigger

**Abstract**—It is well known that the capacity region of an average transmit power constrained Gaussian Broadcast Channel (GBC) with independent noise realizations at the receivers is enlarged by the presence of causal noiseless feedback. When the noise variances at the receivers are identical, even passive feedback via independent memoryless Gaussian links can lead to a capacity region enlargement. The last fact remains true even when the feedback noise variance is very high, and available only from one of the receivers. While such capacity enlargements are feasible for several other feedback models in the Gaussian BC setting, it is also known that feedback does not change the capacity region for physically degraded broadcast channels. In this paper, we consider a two user GBC with independent noise realizations at the receivers, where the feedback links from the receivers are corrupted by independent additive Gaussian noise processes. We investigate the set of four noise variances, two forward and two feedback, for which no capacity enlargement is possible. A sharp characterization of this region is derived, i.e., any quadruple outside the presented region will lead to a capacity enlargement, whereas quadruples inside will leave the capacity region unchanged. Our results lead to the conclusion that when the forward noise variances are different, too noisy a feedback from one of the receivers alone is not always beneficial for enlarging the capacity region, be it from the stronger user or the weaker one, in sharp contrast to the case of equal forward noise variances.

**Index Terms**—Gaussian Broadcast Channel, Noisy Feedback, Capacity Region, Feedback Coding, Capacity Enlargement.

## I. INTRODUCTION

**F**EEDBACK does not increase the capacity of a memoryless point-to-point channel, a result which goes back to C. E. Shannon [1]. However, feedback has a positive impact in simplifying coding schemes and boosting error exponents [2]. With the discovery of capacity regions for several multiuser

models in the '70s and '80s, it was of significant interest to find the impact of feedback on these models. In the absence of feedback, identifying suitable auxiliary variables which can lead to single letter expressions for the rate-region turned out to be a key step in discovering the capacity region of a degraded Broadcast Channel (BC) [3]. Notice that, without any feedback, one can turn a stochastically degraded BC in to an equivalent physical degraded setup, as only the marginal distributions to the individual receivers matter in the decay of error probability with blocklength. When perfect causal feedback is available in a two user BC, a single letter characterization of the capacity region in terms of one auxiliary variable is obtained by El Gamal in [4] and [5], where the physical degradation assumption is crucial. More specifically, [4] and [5] respectively show that the capacity regions of a physically degraded discrete memoryless BC and a physically degraded Gaussian BC (GBC) are unchanged by the presence of causal feedback. In contrast, Dueck [6] demonstrated a BC for which rate pairs outside its no-feedback capacity region can be attained using feedback. For a stochastically degraded two user GBC, Ozarow and Leung developed a feedback coding scheme to show that perfect noiseless feedback from both the receivers enlarges the capacity region [7], when there is no physical degradation. It was later shown that perfect noiseless feedback from the stronger receiver was sufficient to enlarge the capacity region [8]. The optimism of capacity enlargement using feedback did carry over to a variety of models. Recent works considered BCs with noisy feedback [9], [10], [11] and rate limited feedback [12]. In fact, for a two user GBC with equal receiver noise variances, passive noisy feedback from any one of the receivers enlarges the capacity region, even when the feedback noise is of very high variance [10].

It was shown in [13] that noisy feedback always enlarges the capacity region of a Gaussian Multiple Access Channel (MAC), a fact which remains true with the availability of feedback to only one of the transmitters. Furthermore, a MAC-BC duality while employing linear feedback coding schemes with noiseless feedback is known [14], allowing many of the MAC results to be relevant for the BC as well. In summary, capacity enlargement for Gaussian BCs using noisy feedback turned out to be true for several models which are not physically degraded. That this is not always the case when the receivers have different noise variances is shown in the present paper. The results here significantly expand some of the initial results in the conference versions [15], [16], which considered one sided feedback from the stronger receiver. The main interest here is in identifying the set of four noise variances, corresponding to two forward noise processes and

Parts of this paper were presented at the Information Theory Workshop (ITW 2015) and the International Symposium on Information Theory (ISIT 2020).

The work of Michèle Wigger has been supported by the ERC under grant agreement 715111. V. M. Prabhakaran's work was supported by a Ramanujan Fellowship from the Science and Engineering Research Board, India and the Department of Atomic Energy, Government of India, through project no. RT14001. S. R. B. Pillai acknowledges support from DST India under project EMR/2016/005847.

A. N. Ravi is currently with the University of Illinois at Urbana-Champaign. The work reported is from his tenure at IIT Bombay. email: anravi2@illinois.edu

S. R. B. Pillai is with the Electrical Engineering Department, IIT Bombay, India. email: bsraj@ee.iitb.ac.in

V. M. Prabhakaran is with the School of Technology and Computer Science, TIFR Mumbai, India. email: vinodmp@tifr.res.in

M. Wigger is with LTCI Telecom Paris, IP Paris, France. email: michele.wigger@telecom-paristech.fr

Copyright (c) 2017 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org

two feedback noise processes, such that the capacity region differs from that without feedback. We wish to highlight the following aspects of the paper:

- noisy feedback from both the receivers are considered.
- the exact threshold behaviour is characterized, i.e., any set of four noise variances can be classified based on whether the capacity region of the GBC with feedback, having these parameters, is enlarged or not.
- that too noisy a feedback from the weaker receiver of a two user GBC does not enlarge the capacity region was hitherto unknown.
- a byproduct of our analysis is an outer bound which can be useful in some regimes where the conditions of no capacity enlargement are violated.

#### A. Notations:

For a positive integer  $n$ , we use  $U^n$  to denote the tuple  $U_1, U_2, \dots, U_n$ . A variable  $U_{ij}$  with multiple subscripts will be denoted as  $U_{i,j}$  in some contexts to make the index separation clear. The diagonal square matrix of size  $n \times n$  is denoted by  $\text{diag}(\vec{d})$ , where the diagonal elements are given by the vector  $\vec{d}$ . We write  $U \sim \mathcal{N}(\mu, \mathbf{K})$  to denote a random vector  $U$  having a Multivariate Gaussian Distribution with covariance matrix  $\mathbf{K}$ . The acronyms LHS and RHS stand for Left Hand Side and Right Hand Side respectively, of the mathematical equation in consideration. Logarithms are taken with respect to base 2.

#### B. Paper Organization:

The organization of the paper is as follows. We introduce the two user scalar GBC with passive noisy feedback from both the receivers in the next section, and present our objectives as well as the main results. Before proceeding to the detailed proofs, we take a detour in Section III to present results on the utility of feedback in some related channel models. Then, in order to prove the main results, we first develop some converse arguments in Section IV. In particular, for the two receiver GBC model with noisy feedback, we identify a regime where feedback does not help in capacity enlargement. The boundary of the regime thus characterized is then shown to be sharp, by constructing an achievable region which strictly enlarges the no feedback capacity for noise parameters exterior to the regime. This is presented in Section V. Finally, Section VI concludes the paper.

## II. MODEL AND RESULTS

Consider a memoryless two user scalar GBC, as shown in Fig. 1. Assume independent memoryless noisy feedback links from both receivers to the transmitter, in a causal manner. In this setup,  $X$  refers to the signal transmitted by the sender, and the additive noise processes  $(Z_1, Z_2)$  in the forward links are independent, distributed according to  $\mathcal{N}(0, \text{diag}(\sigma_1^2, \sigma_2^2))$ . Unless otherwise stated, we will assume

$$\sigma_1^2 \leq \sigma_2^2, \quad (1)$$

i.e., receiver 2's outputs are more noisy than receiver 1's. The passive causal feedback links are corrupted by additive

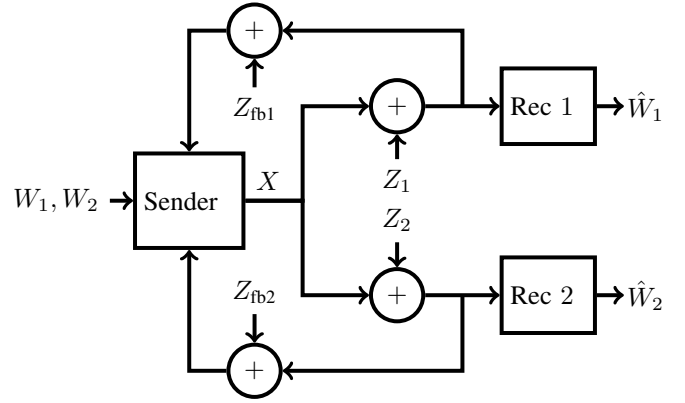


Fig. 1. GBC with causal passive noisy feedback from both the receivers.

noise  $(Z_{fb1}, Z_{fb2}) \sim \mathcal{N}(0, \text{diag}(\sigma_{fb1}^2, \sigma_{fb2}^2))$ , independent of the forward noise processes.

A pair of messages  $(W_1, W_2)$ , independently and uniformly chosen from  $\{1, \dots, 2^{nR_1}\} \times \{1, \dots, 2^{nR_2}\}$  is to be conveyed in  $n$  channel uses of the GBC. The encoder chooses the symbol  $X_i$  at time  $i \in \{1, \dots, n\}$  as a function of both the messages as well as causal noisy feedback from both receivers:

$$X_i = g_i(W_1, W_2, Y_1^{i-1} + Z_{fb1}^{i-1}, Y_2^{i-1} + Z_{fb2}^{i-1}). \quad (2)$$

The transmissions are constrained to meet an average power constraint  $P$ , i.e.,  $\sum_{i=1}^n \mathbb{E}[X_i^2] \leq nP$ . After every block of  $n$  transmissions, each receiver attempts to decode the message intended to it. Let  $P_e(n)$  denote the average error probability that any of the two receivers makes a decoding error. We are interested in the capacity region  $\mathcal{C}^{\text{noisy-fb}}$  of this model. Using standard definitions [2], the capacity region can be concisely defined as follows.

*Definition 1:* The capacity region  $\mathcal{C}^{\text{noisy-fb}}$  is the convex closure of the set of all non-negative rate pairs  $(R_1, R_2)$  such that there exists a sequence of encoder-decoder pairs with  $P_e(n) \rightarrow 0$  as  $n \rightarrow \infty$ .

Let  $\mathcal{C}^{\text{wo-fb}}$  refer to the capacity region when there are no feedback links present, well known to be achieved by superposition coding [17].  $\mathcal{C}^{\text{wo-fb}}$  is the collection of non-negative rate pairs  $(R_1, R_2)$  such that

$$R_1 \leq \frac{1}{2} \log \left( 1 + \frac{\theta P}{\sigma_1^2} \right) \quad (3a)$$

$$R_2 \leq \frac{1}{2} \log \left( 1 + \frac{(1-\theta)P}{\theta P + \sigma_2^2} \right), \quad (3b)$$

for some  $\theta \in [0, 1]$ . By ignoring feedback, it is easy to see

$$\mathcal{C}^{\text{wo-fb}} \subseteq \mathcal{C}^{\text{noisy-fb}}. \quad (4)$$

For equal noise variances, i.e.  $\sigma_1^2 = \sigma_2^2$ , Venkataramanan and Pradhan [10] showed that the inclusion in (4) is strict, irrespective of the feedback noise variances. In fact, it is shown that noisy feedback from even one of the two receivers always enlarges the capacity region. This brings the following question to the fore: For what set of parameters  $(\sigma_1^2, \sigma_2^2, \sigma_{fb1}^2, \sigma_{fb2}^2)$  in  $(\mathbb{R}^+ \cup \{\infty\})^4$  will the inclusion in (4) stay strict? Alternately, are there regimes of noise variances

where feedback is rendered futile in enlarging the capacity region? Our main result is the following.

*Theorem 1:* For  $\sigma_1^2 \leq \sigma_2^2$ , the relation  $\mathcal{C}^{\text{wo-fb}} = \mathcal{C}^{\text{noisy-fb}}$  holds if and only if

$$\frac{\sigma_2^2}{\sigma_1^2} \geq \left( \frac{\sigma_1^2}{\sigma_{\text{fb}1}^2} + 1 \right) \left( \frac{\sigma_2^2}{\sigma_{\text{fb}2}^2} + 1 \right). \quad (5)$$

*Proof:* The proof involves two main steps. One is a converse argument to show that the capacity region is unchanged if the condition in (5) is satisfied. To this end, we employ a novel idea of equipping each receiver with suitable versions of the noise processes in the system. This enables the identification of a useful degradation structure in the broadcast model, in turn leading to a capacity region outer bound. The details are shown in Section IV-B. It further turns out that under the condition (5), the proposed outer bound is achieved by simple superposition coding schemes without any feedback. The sufficiency part of Theorem 1 is shown using a simple linear feedback coding scheme, which operates at rates outside the no feedback capacity region when (5) is not met. Section V details the coding scheme and achievable region. ■

The following corollary can be obtained by simple algebraic manipulations from Theorem 1.

*Corollary 2 (One-sided feedback):* Consider a GBC with  $\sigma_1^2 \leq \sigma_2^2$ . When feedback is available only from receiver 2, the relation  $\mathcal{C}^{\text{wo-fb}} = \mathcal{C}^{\text{noisy-fb}}$  holds if and only if

$$\sigma_{\text{fb}2}^2 \geq \sigma_2^2 \left( \frac{\sigma_2^2}{\sigma_1^2} - 1 \right)^{-1}. \quad (6)$$

On the other hand, if feedback is only available from receiver 1, then the relation  $\mathcal{C}^{\text{wo-fb}} = \mathcal{C}^{\text{noisy-fb}}$  holds if and only if

$$\sigma_{\text{fb}1}^2 \geq \sigma_1^2 \left( \frac{\sigma_2^2}{\sigma_1^2} - 1 \right)^{-1}. \quad (7)$$

*Proof:* Take  $\sigma_{\text{fb}1}^2 \rightarrow \infty$  in Theorem 1 to get the first statement. Similarly, the second statement follows by taking  $\sigma_{\text{fb}2}^2 \rightarrow \infty$ . ■

The above corollary can be rephrased to show that one-sided feedback from receiver 1 increases capacity if and only if  $\frac{\sigma_2^2}{\sigma_{\text{fb}1}^2} < \frac{\sigma_1^2}{\sigma_{\text{fb}1}^2} \left( \frac{\sigma_1^2}{\sigma_{\text{fb}1}^2} + 1 \right)$ . Similarly, one-sided feedback from receiver 2 increases capacity if and only if  $\frac{\sigma_1^2}{\sigma_{\text{fb}2}^2} > \frac{\sigma_2^2/\sigma_{\text{fb}2}^2}{(\sigma_2^2/\sigma_{\text{fb}2}^2)+1}$ . Fig. 2 depicts the regimes of noise-variances where capacity is enlarged with one-sided feedback from receiver 1. The figure also allows for  $\sigma_1^2 > \sigma_2^2$ . In this case, the desired result is obtained by swapping indices 1 and 2 in the above corollary. The figure illustrates that for  $\sigma_1^2 = \sigma_2^2$ , one-sided feedback from receiver 1 is always helpful, no matter how noisy it is. Otherwise, if  $\sigma_1^2 \neq \sigma_2^2$ , then one-sided feedback from receiver 1 becomes useless for capacity enlargement when the feedback noise-variance exceeds a certain threshold. Similar statements hold also for one-sided feedback from receiver 2. In fact, the corresponding plot can be obtained from Fig. 2 by mirroring the plot on the  $\sigma_1^2 = \sigma_2^2$  line. Theorem 1 further provides the following corollary on the usefulness of feedback.

*Corollary 3:* If  $\sigma_1^2 = \sigma_2^2$ , feedback enlarges the capacity region whenever  $\sigma_{\text{fb}1}^2 < \infty$  or  $\sigma_{\text{fb}2}^2 < \infty$ . On the other hand,

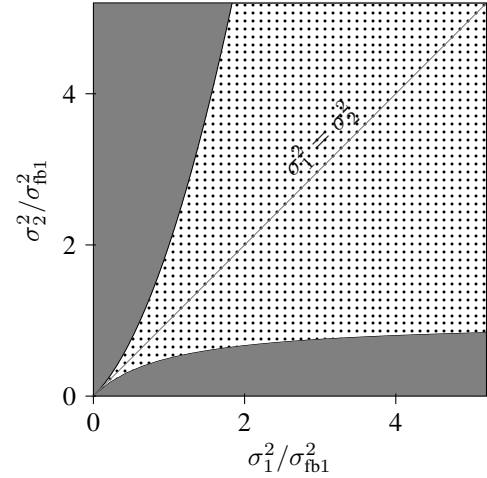


Fig. 2. The patterned region indicates where one-sided feedback from receiver 1 enlarges capacity, and the shaded region indicates where capacity remains unchanged.

if  $\sigma_1^2 \neq \sigma_2^2$ , then there are finite feedback noise variances  $\sigma_{\text{fb}1}^2 < \infty$  and  $\sigma_{\text{fb}2}^2 < \infty$  such that even feedback from both users does not enlarge capacity.

*Proof:* The proof follows by noting that for  $\sigma_1^2 = \sigma_2^2$ , the LHS of (5) evaluates to 1 and is thus dominated by its RHS whenever  $\sigma_{\text{fb}1}^2$  or  $\sigma_{\text{fb}2}^2$  are finite. For  $\sigma_1^2 < \sigma_2^2$ , the LHS of (5) exceeds 1 and it is possible to find finite  $\sigma_{\text{fb}1}^2$  and  $\sigma_{\text{fb}2}^2$  such that the RHS of (5) does not exceed its LHS. ■

While Theorem 1 is the main result of this paper, the converse and achievability techniques we employ in proving it can be used to infer outer and inner bounds for the Gaussian broadcast channel with feedback. We discuss these below.

Let us recall the ingenious outer bound proposed by Ozarow and Leung [7] for a GBC with *perfect* causal feedback. This outer bound, which we will denote by  $\mathcal{O}^{\text{BC}}$ , is in fact the intersection of two regions. Specifically, let  $\mathcal{O}^{\text{sw,w}}$  denote the capacity region of a GBC, where the first receiver has access to  $(Y_1, Y_2)$ , while the second receiver observes  $Y_2$ . Since such a system is physically degraded, having perfect causal feedback of  $(Y_1, Y_2)$  at the encoder does not change the capacity region [5]. The region  $\mathcal{O}^{\text{sw,w}}$  is the collection of all rate pairs  $(R_1, R_2)$  such that  $R_1 \leq \frac{1}{2} \log \left( 1 + \left( \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \theta P \right)$  and  $R_2 \leq \frac{1}{2} \log \left( 1 + \frac{(1-\theta)P}{\sigma_2^2 + \theta P} \right)$  for some  $\theta \in [0, 1]$ . Reversing the roles of the receivers, we can give  $(Y_1, Y_2)$  to the second receiver, while the first receiver observes  $Y_1$ , resulting in the capacity region  $\mathcal{O}^{\text{s,ws}}$ . Now  $\mathcal{O}^{\text{BC}} \triangleq \mathcal{O}^{\text{sw,w}} \cap \mathcal{O}^{\text{s,ws}}$  is an outer bound to the capacity region of the original BC with feedback. Clearly, the region  $\mathcal{O}^{\text{BC}}$  will remain an outer bound with noisy feedback as well. An outer bound explicitly accommodating the feedback noise can be stated as follows.

*Theorem 4:* With  $\sigma_1^2 \leq \sigma_2^2$ , we have  $\mathcal{C}^{\text{noisy-fb}} \subseteq \mathcal{O}^{\text{noisy-fb}} \cap \mathcal{O}^{\text{BC}}$ , where  $\mathcal{O}^{\text{noisy-fb}}$  is the collection of all non-negative  $(R_1, R_2)$  such that

$$R_1 \leq \frac{1}{2} \log \left( 1 + \frac{\theta P}{\sigma^2} \right) \text{ and } R_2 \leq \frac{1}{2} \log \left( 1 + \frac{(1-\theta)P}{\theta P + \sigma_2^2} \right), \quad (8)$$

for some  $\theta \in [0, 1]$ , and

$$\sigma^2 = \min \left\{ \frac{1}{\frac{1}{\sigma_1^2 + \sigma_{\text{fb}1}^2} + \left( \frac{1}{\sigma_2^2} + \frac{1}{\sigma_{\text{fb}2}^2} \right)}, \sigma_1^2 \right\}. \quad (9)$$

*Proof:* The proof of the above statement is given in Section IV-C. ■

Notice that the outerbound  $\mathcal{O}^{\text{noisy-fb}}$  in (8) differs from the no-feedback capacity region  $\mathcal{C}^{\text{wo-fb}}$  of (3) in the bound on  $R_1$  where a new parameter  $\sigma^2$  (defined in (9)) replaces  $\sigma_1^2$ . The above theorem implies the ‘‘if part’’ (converse part) of Theorem 1. This follows from the fact that when (5) holds,  $\sigma^2$  in (9) is indeed  $\sigma_1^2$ . To see this, we may rewrite (5) as

$$\frac{1}{\sigma_2^2} + \frac{1}{\sigma_{\text{fb}2}^2} \leq \frac{\sigma_{\text{fb}1}^2/\sigma_1^2}{\sigma_1^2 + \sigma_{\text{fb}1}^2}, \quad (10)$$

and notice that using the upper bound of (10) on the bracketed term in (9) gives  $\sigma^2 = \sigma_1^2$ .

In order to show the required capacity enlargement claimed in Theorem 1, we will use a linear feedback coding scheme in conjunction with zero forcing decoders. Despite its seemingly simple nature, the scheme turns out to be powerful enough to show the capacity enlargement claimed in Theorem 1.

*Lemma 5:* For a pair of real valued variables  $(\gamma_1, \gamma_2)$ , and  $\beta \triangleq 2 + \frac{\gamma_1^2 \sigma_1^4}{P(\sigma_1^2 + \sigma_{\text{fb}1}^2)} + \frac{\gamma_2^2 \sigma_2^4}{P(\sigma_2^2 + \sigma_{\text{fb}2}^2)}$ , any non-negative rate pair  $(R_1, R_2)$  satisfying

$$R_1 \leq \frac{1}{4} \log \left( 1 + \frac{4\theta P}{\sigma_1^2 \beta + (\gamma_1^2 + 2\gamma_1) \frac{\sigma_1^4}{\sigma_1^2 + \sigma_{\text{fb}1}^2} + \gamma_2^2 \frac{\sigma_2^4}{\sigma_2^2 + \sigma_{\text{fb}2}^2}} \right) \quad (11a)$$

$$R_2 \leq \frac{1}{4} \log \left( 1 + \frac{4(1-\theta)P}{\sigma_2^2 \beta + (\gamma_2^2 - 2\gamma_2) \frac{\sigma_2^4}{\sigma_2^2 + \sigma_{\text{fb}2}^2} + \gamma_1^2 \frac{\sigma_1^4}{\sigma_1^2 + \sigma_{\text{fb}1}^2}} \right) \quad (11b)$$

for some  $0 \leq \theta \leq 1$ , is achievable for the GBC with noisy feedback.

*Proof:* The lemma is proved in Section V, along with the achievability proof of Theorem 1. ■

To exemplify the utility of this lemma, let us consider the case with  $\sigma_1^2 = \sigma_2^2$ , and  $\sigma_{\text{fb}2}^2 = \infty$ , i.e., no feedback from the second receiver. Then, taking  $\theta' = 0.5$  and  $\gamma_1 = \gamma_2 = 0$  will recover the equal rate point on the no-feedback capacity region. However, some minimal algebra suffices to show that small negative values of  $\gamma_1$  will cause the region given by (11) to contain rate-pairs outside the no-feedback capacity region. Therefore, this scheme almost immediately suggests a capacity enlargement using passive noisy feedback in a GBC with  $\sigma_1^2 = \sigma_2^2$ . Generalizing the above idea to different noise variances at the receivers needs more analytical effort, which is done in Section V after the proof of Lemma 5. We remark that when compared to the scheme in [10], which achieves potentially superior rate regions, the scheme yielding Lemma 5 is remarkably simpler. This simplicity (and the consequent simplicity of the rate region) allows us to perform the analysis needed to prove the capacity enlargement in Theorem 1.

### III. MORE ON THE USEFULNESS OF FEEDBACK

In this section we consider two related BC setups where the feedback links to the transmitter do not increase the capacity

region. The results here may be of interest on their own. Moreover, Corollary 2 can be proved using the first result that we present in this section, stated as Theorem 6 below. However, the results and tools presented in this section seem not enough to prove the converse to our main result presented in Theorem 1.

#### A. A Vector-BC With Partial Feedback

We start with a slight extension of El Gamal’s negative result on the usefulness of feedback for physically degraded BCs, to vector channels with partial feedback. Consider a general (not necessarily Gaussian) BC where the first receiver observes  $Y_1 = (\tilde{Y}_1, \hat{Y}_1)$ , and the second receiver observes  $Y_2 = (\tilde{Y}_2, \hat{Y}_2)$ . Let the transition law of the BC be described by  $P_{\tilde{Y}_1, \hat{Y}_1, \tilde{Y}_2, \hat{Y}_2|X}$ . Feedback is only from outputs  $\tilde{Y}_1$  and  $\tilde{Y}_2$  to the transmitter, and can be noisy or perfect. The following negative result holds.

*Theorem 6:* Consider a memoryless BC of transition law  $P_{\tilde{Y}_1, \hat{Y}_1, \tilde{Y}_2, \hat{Y}_2|X}$ . If there exists a conditional law  $P'_{\tilde{Y}_2, \hat{Y}_2|\tilde{Y}_1, \hat{Y}_1}$  such that the joint law

$$P'_{\tilde{Y}_1, \hat{Y}_1, \tilde{Y}_2, \hat{Y}_2|X} \triangleq P'_{\tilde{Y}_2, \hat{Y}_2|\tilde{Y}_1, \hat{Y}_1} \cdot P_{\tilde{Y}_1, \hat{Y}_1|X} \quad (12)$$

satisfies

$$P'_{\tilde{Y}_1|X, \tilde{Y}_1, \tilde{Y}_2} = P_{\tilde{Y}_1|X, \tilde{Y}_1, \tilde{Y}_2}, \quad (13a)$$

$$P'_{\tilde{Y}_2|X, \tilde{Y}_1, \tilde{Y}_2} = P_{\tilde{Y}_2|X, \tilde{Y}_1, \tilde{Y}_2}, \quad (13b)$$

$$P'_{\tilde{Y}_1, \hat{Y}_2|X} = P_{\tilde{Y}_1, \hat{Y}_2|X}, \quad (13c)$$

then causal feedback of the two outputs  $\tilde{Y}_1$  and  $\tilde{Y}_2$  to the transmitter does not enlarge the capacity region.

*Proof:* Consider the BC of transition law  $P'_{\tilde{Y}_1, \hat{Y}_1, \tilde{Y}_2, \hat{Y}_2|X}$  as defined in (12). This channel is indeed physically degraded because its law satisfies the Markov chain  $X \rightarrow (\tilde{Y}_1, \hat{Y}_1) \rightarrow (\tilde{Y}_2, \hat{Y}_2)$ . Moreover, by assumptions (13), under any encoding strategy, and both in the presence and absence of feedback, the joint law of the tuples  $(W_1, \tilde{Y}_1^n, \hat{Y}_1^n)$  and  $(W_2, \tilde{Y}_2^n, \hat{Y}_2^n)$  is the same in both the original BC  $P_{\tilde{Y}_1, \hat{Y}_1, \tilde{Y}_2, \hat{Y}_2|X}$  and the physically degraded version  $P'_{\tilde{Y}_1, \hat{Y}_1, \tilde{Y}_2, \hat{Y}_2|X}$ . To see this, consider

$$\begin{aligned} & P(w_1, \tilde{y}_1^n, \hat{y}_1^n) \\ &= \sum_{w_2, \tilde{y}_2^n} P(w_1, w_2) \prod_{i=1}^n P(\tilde{y}_{1i}, \tilde{y}_{2i}, \hat{y}_{1i} | w_1, w_2, \tilde{y}_1^{i-1}, \tilde{y}_2^{i-1}, \hat{y}_1^{i-1}) \\ &= \sum_{w_2, \tilde{y}_2^n} P(w_1, w_2) \prod_{i=1}^n P(\tilde{y}_{1i}, \tilde{y}_{2i} | w_1, w_2, \tilde{y}_1^{i-1}, \tilde{y}_2^{i-1}, \hat{y}_1^{i-1}) \\ &\quad \times P(\hat{y}_{1i} | w_1, w_2, \tilde{y}_1^i, \tilde{y}_2^i, \hat{y}_1^{i-1}) \\ &\stackrel{(a)}{=} \sum_{w_2, \tilde{y}_2^n} P(w_1, w_2) \prod_{i=1}^n P(\tilde{y}_{1i}, \tilde{y}_{2i} | x_i) P(\hat{y}_{1i} | x_i, \tilde{y}_{1i}, \tilde{y}_{2i}) \\ &\stackrel{(b)}{=} \sum_{w_2, \tilde{y}_2^n} P(w_1, w_2) \prod_{i=1}^n P'(\tilde{y}_{1i}, \tilde{y}_{2i} | x_i) P'(\hat{y}_{1i} | x_i, \tilde{y}_{1i}, \tilde{y}_{2i}). \end{aligned}$$

Here (a) follows since  $x_i = g_i(w_1, w_2, \tilde{y}_1^{i-1}, \tilde{y}_2^{i-1})$  and the channel is memoryless, and (b) follows from the conditions (13a) and (13c). Similar arguments apply at receiver 2 as well.

Since the error probability at each receiver only depends on its own observations, but not on the observations at the other receiver, the capacity regions of the original and the physically degraded BCs coincide. On any physically degraded BC, the capacity region, even with full causal feedback, remains the same as that without feedback [4], [5]. By the above observations, the same must also hold for our original BC, thus concluding the proof. ■

### B. Feedback From the Weak Receiver to all Terminals

Consider a stochastically degraded GBC where  $Y_1 = X + Z_1$  and  $Y_2 = X + Z_2$  are observed at the respective receivers. Let us assume that  $Y_c \triangleq Y_2 + Z_{fb2}$  is given as feedback from receiver 2 to the transmitter in a causal manner. In addition, suppose that the same feedback values are observed at receiver 1 as well. Thus, receiver 1 observes its own channel outputs  $Y_1^n$  as well as the noisy feedback outputs  $Y_c^n$ , before decoding. The model is illustrated in Fig. 3a. We take  $(Z_1, Z_2, Z_{fb2}) \sim \mathcal{N}(0, \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_{fb2}^2))$ , with  $\sigma_1^2 \leq \sigma_2^2$ . Let us denote the capacity region of this model by  $\mathcal{C}^{\text{weak-fb}}$ . Notice that the described setup is physically degraded only if the feedback noise  $Z_{fb2}^n$  is deterministic and thus  $\sigma_{fb2}^2 = 0$ . Interestingly, such a degraded setup was employed in [7] to derive an outer bound to the GBC capacity region with perfect causal feedback.

The following theorem, which is the main result of this subsection, shows that the feedback capacity region  $\mathcal{C}^{\text{weak-fb}}$  is unaffected by the presence of the feedback link *to the transmitter*. In other words, once receiver 1 learns the feedback signal from receiver 2, the capacity region does not depend on whether also the transmitter observes the feedback or not.

*Theorem 7:* The capacity region  $\mathcal{C}^{\text{weak-fb}}$  is unchanged by the presence of feedback to the transmitter. In particular,

$$\mathcal{C}^{\text{weak-fb}} = \bigcup_{\theta \in [0,1]} (R_1, R_2) \text{ such that}$$

$$R_1 \leq \frac{1}{2} \log \left( 1 + \frac{\theta P}{\sigma_{\text{comb}}^2} \right) \text{ and } R_2 \leq \frac{1}{2} \log \left( 1 + \frac{(1-\theta)P}{\theta P + \sigma_2^2} \right), \quad (14)$$

where

$$\sigma_{\text{comb}}^2 \triangleq \frac{\sigma_1^2(\sigma_2^2 + \sigma_{fb2}^2)}{\sigma_1^2 + \sigma_2^2 + \sigma_{fb2}^2}. \quad (15)$$

*Proof:* The achievability of  $\mathcal{C}^{\text{weak-fb}}$  follows by superposition coding, without using any feedback to the transmitter. Receiver 1 then performs maximum ratio combining of  $Y_1$  and  $Y_c$ .

The key steps in the converse are to show that the capacity region  $\mathcal{C}^{\text{weak-fb}}$  is included in the capacity region of the *augmented* BC depicted in Fig. 3b, and to invoke Theorem 6 to show that feedback does not increase the capacity of this augmented BC. The final step is then to argue that the capacity region without feedback of the augmented BC coincides with the region defined in Theorem 7.

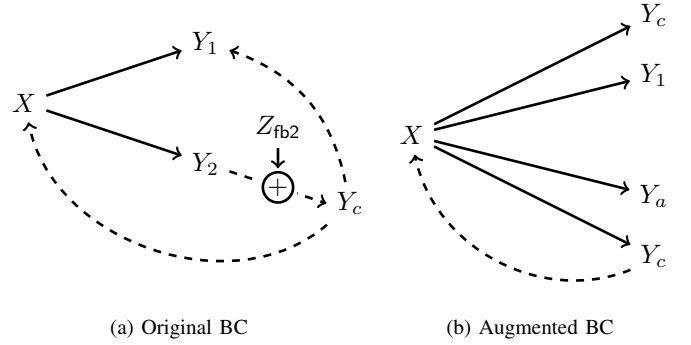


Fig. 3. BC models with feedback signal available at both the receivers

Consider the augmented BC depicted in Fig. 3b. Receiver 1 observes outputs  $(Y_1^n, Y_c^n)$  and receiver 2 observes outputs  $(Y_a^n, Y_c^n)$ , where

$$Y_a^n \triangleq Y_2^n - \alpha Z_{fb2}^n = X^n + \underbrace{Z_2^n - \alpha Z_{fb2}^n}_{\triangleq Z_a^n} \quad (16)$$

and

$$\alpha \triangleq \frac{\sigma_2^2}{\sigma_{fb2}^2}. \quad (17)$$

The transmitter observes the feedback output  $Y_{ci}$  after each channel use  $i$ . The capacity region  $\mathcal{C}_{\text{aug}}^{\text{noisy-fb}}$  of this augmented BC with noisy feedback includes the original capacity region  $\mathcal{C}^{\text{weak-fb}}$ , i.e.,

$$\mathcal{C}^{\text{weak-fb}} \subseteq \mathcal{C}_{\text{aug}}^{\text{noisy-fb}} \quad (18)$$

holds, because receiver 1 in the augmented BC observes the same outputs as in the original BC, and receiver 2 can compute  $Y_2^n$  from its observations  $(Y_c^n, Y_a^n)$ .

Notice next that feedback does not increase the capacity of the augmented BC, because this BC satisfies the assumptions in Theorem 6. To see this, specialize the theorem to  $\tilde{Y}_1 = Y_c$ ,  $\hat{Y}_1 = Y_1$ ,  $\tilde{Y}_2 = Y_c$ , and  $\hat{Y}_2 = Y_a$ . Now set  $P'_{\tilde{Y}_2|\hat{Y}_1, \tilde{Y}_1} = P_{\tilde{Y}_2|\hat{Y}_1, \tilde{Y}_1} = P_{\tilde{Y}_2|\hat{Y}_1}$  (i.e.  $\tilde{Y}_1 = \hat{Y}_1 = \tilde{Y}_2$  under both  $P$  and  $P'$ ), and choose  $P'_{\hat{Y}_2|\hat{Y}_1, \tilde{Y}_1, \tilde{Y}_2} = P_{\hat{Y}_2|\hat{Y}_1}$  in such a way that

$$\hat{Y}_2 = \hat{Y}_1 + \hat{Z}_2 \quad (19)$$

for  $\hat{Z}_2$  a zero-mean Gaussian random variable of variance  $\sigma_2^2 + \alpha^2 \sigma_{fb2}^2 - \sigma_1^2$  and independent of all other random variables. This is possible since  $\sigma_2^2 \geq \sigma_1^2$ .

Under this choice, (13a) holds since  $\hat{Y}_1 = X + Z_1$  under both  $P$  and  $P'$ , with  $Z_1 \perp\!\!\!\perp (Z_2, Z_{fb2})$ . The assumption (13b) also holds, since the pair  $(Z_a, Z_2 + Z_{fb2})$  has independent Gaussian entries and it has the same distribution as  $(Z_1 + \hat{Z}_2, Z_2 + Z_{fb2})$ . Since  $\hat{Y}_1 = \hat{Y}_2 = Y_c$ , the condition (13c) holds as well. Thus, we can employ Theorem 6 to conclude that feedback does not increase the capacity of the augmented BC, i.e.,

$$\mathcal{C}_{\text{aug}}^{\text{noisy-fb}} = \mathcal{C}_{\text{aug}}^{\text{wo-fb}}, \quad (20)$$

where  $\mathcal{C}_{\text{aug}}^{\text{wo-fb}}$  denotes the capacity region of the augmented BC without feedback. Using the well known expression for

the capacity region of a Gaussian vector BC [2],  $\mathcal{C}_{\text{aug}}^{\text{wo-fb}}$  is the collection of rate pairs  $(R_1, R_2)$  such that

$$R_1 \leq \frac{1}{2} \log \left( 1 + \frac{\theta P}{\sigma_{\text{comb}}^2} \right) \text{ and } R_2 \leq \frac{1}{2} \log \left( 1 + \frac{(1-\theta)P}{\theta P + \sigma_2^2} \right),$$

for some  $0 \leq \theta \leq 1$ , where  $\sigma_{\text{comb}}^2$  is defined in (15).

The desired converse proof is now established, based on (18) and (20).  $\blacksquare$

#### IV. OUTER BOUNDS FOR GBC WITH NOISY FEEDBACK (CONVERSE PARTS OF THEOREMS 1 AND 4)

Before we embark on proving the converse part in Theorem 1, notice that there is no obvious physical degradation in our GBC model with noisy feedback shown in Fig. 1. In addition, Theorem 6 cannot be applied as such. Interestingly, under the condition in (5), an outer bound to the capacity region can be constructed by equipping each receiver with suitable information about the noise processes in the model. That the proposed outer bound turns out to be achievable even without feedback clinches the deal, further underlining the novelty of the proposed bound. While some steps in our proof effectively employ the degradation between different variables, these are somewhat implicit in the manipulations. To make the exposition gradual, we consider a related problem first, that of feedback only from the strong receiver, where Theorem 6 turns out to be useful.

##### A. Noisy Feedback From the Strong Receiver Alone

With feedback only from the stronger receiver in a GBC, let us devise the converse for the second part of Corollary 2. The model is depicted in Fig. 4a. We have to show that for  $\sigma_{\text{fb}2}^2 = \infty$  and

$$\frac{\sigma_1^2}{\sigma_{\text{fb}1}^2} \leq \frac{\sigma_2^2}{\sigma_1^2} - 1, \quad (21)$$

feedback from the stronger receiver (receiver 1) does not enlarge the capacity region.

Let us first construct an augmented BC in which receiver 1 observes  $(Y_1^n + Z_{\text{fb}1}^n, Y_b^n)$ , where

$$Y_b^n := Y_1^n - \alpha Z_{\text{fb}1}^n = X^n + \underbrace{Z_1^n - \alpha Z_{\text{fb}1}^n}_{\triangleq Z_b^n}, \quad (22)$$

for

$$\alpha \triangleq \frac{\sigma_1^2}{\sigma_{\text{fb}1}^2}. \quad (23)$$

As before, the transmitter observes the feedback signal  $Y_{1i} + Z_{\text{fb}1i}$  after channel use  $i \in \{1, \dots, n\}$ . The capacity region  $\mathcal{C}_{\text{aug}}^{\text{noisy-fb}}$  of this augmented BC, which is depicted in Fig. 4b, includes the capacity region  $\mathcal{C}^{\text{noisy-fb}}$  of our original BC:

$$\mathcal{C}^{\text{noisy-fb}} \subseteq \mathcal{C}_{\text{aug}}^{\text{noisy-fb}}. \quad (24)$$

This is immediate since  $(1 + \alpha)Y_1 = \alpha(Y_1 + Z_{\text{fb}1}) + Y_b$ , and thus receiver 1 can compute  $Y_1^n$  from  $Y_1^n + Z_{\text{fb}1}^n$  and  $Y_b^n$ . We next argue that the augmented BC satisfies the assumptions in Theorem 6 and thus feedback does not enlarge its capacity region. To this end, let us specialize Theorem 6 by identifying

$\hat{Y}_1 = Y_1 + Z_{\text{fb}1}$ ,  $\hat{Y}_1 = Y_b$ ,  $\hat{Y}_2 = \emptyset$ , and  $\hat{Y}_2 = Y_2$ , and then choose  $P'_{\hat{Y}_2|\hat{Y}_1,\hat{Y}_1}$  such that under this new law

$$\hat{Y}_2 = \hat{Y}_1 + \hat{Z}_2, \quad (25)$$

with  $\hat{Z}_2$  a zero-mean Gaussian random variable of variance  $\sigma_2^2 - \alpha^2 \sigma_{\text{fb}1}^2 - \sigma_1^2 > 0$ , and independent of all other random variables. Clearly, (13a) and (13c) are immediate under this choice. The choice of parameter  $\alpha$  in (23) ensures that the condition (13b) is also met. To see this, notice that  $Z_2$  is independent of  $Z_1 + Z_{\text{fb}1}$  under the probability law  $P$ , whereas  $Z_b + \hat{Z}_2$  is independent of  $Z_1 + Z_{\text{fb}1}$  under  $P'$ . Thus, the noise sequences  $Z_2^n$  and  $Z_b^n + \hat{Z}_2^n$  are both independent of the feedback noise  $(Z_1^n + Z_{\text{fb}1}^n)$  in the respective models. Also, notice that the random variables  $Z_2$  and  $Z_b + \hat{Z}_2$  are identically distributed. Since all the required assumptions are met, we can employ Theorem 6 to conclude that feedback does not increase the capacity of the augmented BC, and thus

$$\mathcal{C}_{\text{aug}}^{\text{noisy-fb}} = \mathcal{C}^{\text{wo-fb}}. \quad (26)$$

The capacity region without feedback of the augmented BC is obtained from the well-known capacity region of a Gaussian vector BC [2]. Combining (24) and (26) establishes the desired converse.

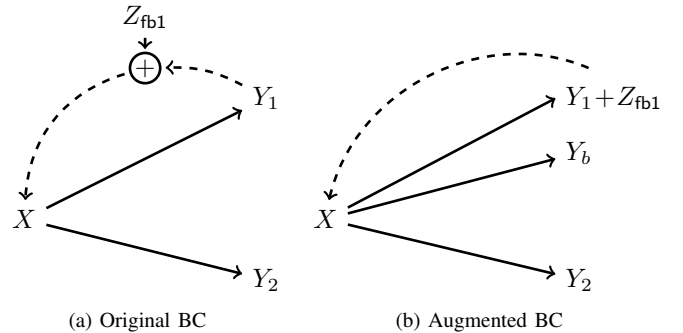


Fig. 4. Augmenting one-sided noisy feedback from the strong receiver

##### B. Noisy Feedback From Both Receivers (Converse for Theorem 1)

We already mentioned that the converse result for Theorem 1 can be obtained as a special case of Theorem 4. Nevertheless, we first provide a self-contained proof for the former as it is not only instructive, but also makes the proof of the latter appear more natural. While we successfully took recourse to identifying a physically degraded set up in the previous subsection, the arguments there do not immediately generalize when there is feedback from the weak receiver. Observe that feeding back the signal  $Y_2$  allows the transmissions to depend, in general, on  $Z_2 + Z_{\text{fb}2}$ . Thus, there is no obvious stochastic degradation between  $Y_2$  and any signal derived from the observed symbols at the strong receiver. Nevertheless, as we will show, the dependence between  $Z_2$  and the transmitted symbols can be decoupled by careful conditioning to enable the identification of a suitable degradation structure. While the condition in (5) permits such an identification and subsequent

single-letterization as we show below, it is surprising that the condition is tight from the capacity enlargement point of view.

By Fano's inequality [2], after ignoring the  $o(n)$  terms,

$$\begin{aligned}
nR_1 &= I(W_1; Y_1^n | W_2) \\
&\leq I(W_1; Y_1^n, Z_{fb1}^n, Z_2^n + Z_{fb2}^n | W_2) \\
&= \sum_{i=1}^n I(W_1; Y_{1i} | W_2, Y_1^{i-1}, Z_{fb1}^{i-1}, Z_2^{i-1} + Z_{fb2}^{i-1}) \\
&\quad + I(W_1; Z_{fb1i}, Z_{2i} + Z_{fb2i} | W_2, Z_{fb1}^{i-1}, Z_2^{i-1} + Z_{fb2}^{i-1}, Y_1^i) \\
&\stackrel{(a)}{=} \sum_{i=1}^n I(W_1; Y_{1i} | W_2, Y_1^{i-1}, Z_{fb1}^{i-1}, Z_2^{i-1} + Z_{fb2}^{i-1}) \\
&\stackrel{(b)}{=} \sum_{i=1}^n h(Y_{1i} | W_2, Y_1^{i-1}, Z_{fb1}^{i-1}, Z_2^{i-1} + Z_{fb2}^{i-1}) - h(Z_{1i}) \\
&\stackrel{(c)}{=} \sum_{i=1}^n h(Y_{1i} | U_i, X^{i-1} + Z_1^{i-1}, Z_{fb1}^{i-1}) - h(Z_{1i}) \\
&\stackrel{(d)}{=} \sum_{i=1}^n h(Y_{1i} | U_i, X^{i-1} + Z_1^{i-1} - \gamma Z_{fb1}^{i-1}) - h(Z_{1i}). \tag{27}
\end{aligned}$$

$$= \sum_{i=1}^n h(Y_{1i} | U_i, X^{i-1} + \tilde{Z}_1^{i-1}) - h(Z_{1i}). \tag{28}$$

Here (a) follows since  $(Z_{fb1i}, Z_{2i} + Z_{fb2i})$  is independent of  $(W_1, W_2, Z_{fb1}^{i-1}, Z_2^{i-1} + Z_{fb2}^{i-1}, Y_1^i)$ . In (b), we used the fact that  $(W_1, W_2, Y_1^{i-1} + Z_{fb1}^{i-1}, Z_2^{i-1} + Z_{fb2}^{i-1})$  determines the transmitted symbol  $X_i$ . Clearly, given  $X_i$ , the remaining uncertainty in  $Y_{1i}$  is only due to  $Z_{1i}$ , which is independent of  $(W_1, W_2, Y_1^{i-1}, Z_{fb1}^{i-1}, Z_2^{i-1} + Z_{fb2}^{i-1})$ . In step (c), we took  $U_i = (W_2, Z_2^{i-1} + Z_{fb2}^{i-1})$ , whereas (d) is true for any  $\gamma \in \mathbb{R}$ , since conditioning cannot increase entropy. Now, (28) follows by defining  $Z_1 - \gamma Z_{fb1}$  as  $\tilde{Z}_1$ .

For the weak receiver, again by applying Fano's inequality and ignoring the  $o(n)$  terms,

$$nR_2 \leq h(Y_2^n) - h(Y_2^n | W_2). \tag{29}$$

Let us now expand the second term as

$$\begin{aligned}
h(Y_2^n | W_2) &= \sum_{i=1}^n h(Y_{2i} | W_2, Y_2^{i-1}) \\
&\geq \sum_{i=1}^n h(Y_{2i} | W_2, X^{i-1} + Z_2^{i-1}, Z_2^{i-1} + Z_{fb2}^{i-1}) \\
&= \sum_{i=1}^n h(Y_{2i} | U_i, X^{i-1} + Z_2^{i-1} - \beta(Z_2^{i-1} + Z_{fb2}^{i-1})) \\
&= \sum_{i=1}^n h(Y_{2i} | U_i, X^{i-1} + \tilde{Z}_2^{i-1}), \tag{30}
\end{aligned}$$

where  $\tilde{Z}_2 \triangleq (1 - \beta)Z_2 - \beta Z_{fb2}$  for some  $\beta \in \mathbb{R}$ .

We now choose  $\gamma$  and  $\beta$  in (28) and (30), so that  $(\tilde{Z}_1^n, \tilde{Z}_2^n)$  is independent of  $(Z_1^n + Z_{fb1}^n, Z_2^n + Z_{fb2}^n)$ , the latter being the information that feedback makes available to the encoder. Specifically, set

$$\gamma = \frac{\sigma_1^2}{\sigma_{fb1}^2} \quad \text{and} \quad \beta = \frac{\sigma_2^2}{\sigma_2^2 + \sigma_{fb2}^2}. \tag{31}$$

The required independence follows from observing that  $\mathbb{E}[(Z_1 + Z_{fb1})(Z_1 - \gamma Z_{fb1})] = \mathbb{E}[(Z_2 + Z_{fb2})((1 - \beta)Z_2 - \beta Z_{fb2})] = 0$ . Furthermore, since  $(\tilde{Z}_1^n, \tilde{Z}_2^n)$  is independent of all the information at the encoder, namely,  $(W_1, W_2, Z_1^n + Z_{fb1}^n, Z_2^n + Z_{fb2}^n)$ , it is also independent of  $X^n$ . Thus, we have  $(\tilde{Z}_1^n, \tilde{Z}_2^n)$  independent of  $(U^n, X^n)$ . Now consider (28); notice that the noise term  $Z_{1i}$  in  $Y_{1i} = X_i + Z_{1i}$  is independent of  $(Z_1^{i-1}, Z_{fb1}^{i-1})$  and hence of  $\tilde{Z}_1^{i-1}$ . Therefore, by the independence of  $\tilde{Z}_1^{i-1}$  and  $(U_i, X^i)$ , the value of (28) remains unchanged even if we replace  $\tilde{Z}_1^n$  by a noise process with the same marginal distribution, but independent of  $(Z_1^n, Z_{fb1}^n, Z_2^n, Z_{fb2}^n, W_1, W_2)$ . By a similar argument, without affecting (30) we may replace  $\tilde{Z}_2^n$  by a noise process with the same marginal, but independent of  $(Z_1^n, Z_{fb1}^n, Z_2^n, Z_{fb2}^n, W_1, W_2)$ . Hence, we will take  $(\tilde{Z}_1^n, \tilde{Z}_2^n)$  to be independent of  $(Z_1^n, Z_{fb1}^n, Z_2^n, Z_{fb2}^n, W_1, W_2)$  in further analyzing (28)–(30).

By the data processing theorem,

$$h(Y_{1i} | U_i, X^{i-1} + \tilde{Z}_1^{i-1}) \leq h(Y_{1i} | U_i, X^{i-1} + \tilde{Z}_2^{i-1}) \tag{32}$$

as long as

$$\text{Var}(\tilde{Z}_1) \leq \text{Var}(\tilde{Z}_2). \tag{33}$$

Notice that (33) is equivalent to

$$\frac{\sigma_2^2}{\sigma_1^2} \geq \left( \frac{\sigma_1^2}{\sigma_{fb1}^2} + 1 \right) \left( \frac{\sigma_2^2}{\sigma_{fb2}^2} + 1 \right). \tag{34}$$

Under the above condition, (28) yields

$$nR_1 \leq \sum_{i=1}^n h(Y_{1i} | U_i, V_i) - h(Z_{1i}), \tag{35}$$

where we defined  $V_i = X^{i-1} + \tilde{Z}_2^{i-1}$ . Using this definition, along with (29) and (30), we get

$$nR_2 \leq \sum_{i=1}^n h(Y_{2i}) - h(Y_{2i} | U_i, V_i). \tag{36}$$

Notice that  $(U_i, V_i) \rightarrow X_i \rightarrow (Y_{1i}, Y_{2i})$ , and we can now obtain single letter rate expressions using one auxiliary random variable  $\tilde{U} = (U, V)$ . The optimality of a Gaussian auxiliary variable can then be proved along the lines of [18], by incorporating feedback as in [19]. However, we proceed through a more standard route, by applying the following version of EPI (similar to [5]) to connect (35) and (36).

*Lemma 8:* For  $(\tilde{U}_i, Y_{1i}, Y_{2i})$  as defined above, we have

$$2^{\frac{2}{n} \sum_{i=1}^n h(Y_{2i} | \tilde{U}_i)} \geq 2^{\frac{2}{n} \sum_{i=1}^n h(Y_{1i} | \tilde{U}_i)} + 2\pi e(\sigma_2^2 - \sigma_1^2). \tag{37}$$

*Proof:* The proof is presented in Appendix A. ■

The remaining part of the proof is more routine. By well known results in information theory [17], one can write

$$\frac{n}{2} \log 2\pi e(P + \sigma_2^2) \geq \sum_{i=1}^n h(Y_{2i}) \geq \sum_{i=1}^n h(Y_{2i} | \tilde{U}_i) \geq h(Z_2^n). \tag{38}$$

Therefore, we can take

$$\sum_{i=1}^n h(Y_{2i} | \tilde{U}_i) = \frac{n}{2} \log(2\pi e(\sigma_2^2 + \theta P)), \tag{39}$$

for some  $\theta \in [0, 1]$ . Then, (37) will imply that

$$\sum_{i=1}^n h(Y_{1i}|\tilde{U}_i) \leq \frac{n}{2} \log(2\pi e(\sigma_1^2 + \theta P)). \quad (40)$$

From (35), (36), (38)–(40), we get for some  $\theta \in [0, 1]$ ,

$$R_2 \leq \frac{1}{2} \log \left( 1 + \frac{(1-\theta)P}{\theta P + \sigma_2^2} \right) \quad (41)$$

$$R_1 \leq \frac{1}{2} \log \left( 1 + \frac{\theta P}{\sigma_1^2} \right). \quad (42)$$

The bounds in (41)–(42) can be achieved by standard superposition coding, without using any feedback. This completes the proof of the converse part of Theorem 1. In a nutshell, identifying the respective noise components which are orthogonal to the encoder's view of the noise processes, is the key step. Then, the condition in (34) ensures that the weak receiver's conditioning  $(U_i, X^{i-1} + \tilde{Z}_2^{i-1})$  in (30) can be taken as a physically degraded version of the strong receiver's conditioning  $(U_i, X^{i-1} + \tilde{Z}_1^{i-1})$  in (27).

### C. A General Outer Bound for GBC With Noisy Feedback (Proof of Theorem 4)

Let us now prove Theorem 4. Without loss of generality, we can assume that receiver 1 observes  $Y_1 = X + Z_a + Z_b$ , where  $Z_a$  and  $Z_b$  are two independent zero mean Gaussian random variables with respective variances  $\sigma_a^2$  and  $\sigma_b^2$  such that  $\sigma_a^2 + \sigma_b^2 = \sigma_1^2$ . In order to construct an outer bound, we equip receiver 1 with the observations  $Y_a = X_1 + Z_a$ , in addition to  $Y_1$ . For convenience, let us denote the tuple  $(Y_1, Y_a, Z_{fb1} + Z_b, Z_2 + Z_{fb2})$  by  $T$ . Similar to Section IV-B, we can bound the rates to user 1 as

$$\begin{aligned} nR_1 & \stackrel{(43)}{\leq} I(W_1; Y_1^n, Y_a^n | W_2) \\ & \leq I(W_1; Y_1^n, Y_a^n, Z_{fb1}^n + Z_b^n, Z_2^n + Z_{fb2}^n | W_2) \\ & \stackrel{(i)}{=} I(W_1; T^n | W_2) \\ & \stackrel{(ii)}{=} \sum_{i=1}^n I(W_1; Y_{ai} | W_2, T^{i-1}) \\ & \quad + I(W_1; Z_{fb1i} + Z_{bi} | Y_{ai}, W_2, T^{i-1}) \\ & \quad + I(W_1; Y_{1i} | Z_{fb1i} + Z_{bi}, Y_{ai}, W_2, T^{i-1}) \\ & \quad + I(W_1; Z_{2i} + Z_{fb2i} | Y_{1i}, Z_{fb1i} + Z_{bi}, Y_{ai}, W_2, T^{i-1}) \\ & \stackrel{(iii)}{=} \sum_{i=1}^n I(W_1; Y_{ai} | W_2, T^{i-1}) \\ & \quad + I(W_1; Y_{1i} | Z_{fb1i} + Z_{bi}, Y_{ai}, W_2, T^{i-1}) \\ & = \sum_{i=1}^n I(W_1; Y_{ai} | W_2, T^{i-1}) \\ & \quad + I(W_1; Z_{bi} | Z_{fb1i} + Z_{bi}, Y_{ai}, W_2, T^{i-1}) \quad (44) \\ & \stackrel{(iv)}{=} \sum_{i=1}^n I(W_1; Y_{ai} | W_2, T^{i-1}) \\ & \stackrel{(v)}{\leq} h(Y_{ai} | U_i, X^{i-1} + Z_a^{i-1} - \gamma(Z_{fb1}^{i-1} + Z_b^{i-1})) - h(Z_{ai}), \quad (45) \end{aligned}$$

where the justifications of (i)–(v) are as follows. In (i), we used  $T = (Y_1, Y_a, Z_{fb1} + Z_b, Z_2 + Z_{fb2})$ , (ii) is based on chain rule of mutual information, whereas (iii) uses the fact that both the second and fourth terms in the RHS of (ii) are zero. In (iv), we used the fact that  $(Z_{bi}) \rightarrow (Z_{bi} + Z_{fb1i}) \rightarrow (W_1, W_2, Y_a^i, Y_1^{i-1}, Z_{fb1}^{i-1} + Z_b^{i-1}, Z_2^{i-1} + Z_{fb2}^{i-1})$  forms a Markov chain. Finally, in (v), we took  $U_i$  as  $(W_2, Z_2^{i-1} + Z_{fb2}^{i-1})$ , and removed some conditioning from the first term in the RHS. Observe that (45) is the same as (27) with  $(Y_1, Z_{fb1})$  replaced by  $(Y_a, Z_{fb1} + Z_b)$ . The rate  $R_2$  to the weak receiver can be bounded exactly as in (30). Let us now choose an appropriate  $Y_a$  such that the bounds in (36) and (45) become easy to compute.

Suppose the given GBC has a set of noise variances such that the condition in (5) of Theorem 1 is violated, but  $\sigma_{fb2}^2 > 0$ . Let us choose  $Y_a = X + Z_a$  in (45) such that the condition in (5) is satisfied when  $(\sigma_1^2, \sigma_{fb1}^2)$  there is replaced by  $(\sigma_a^2, \sigma_b^2 + \sigma_{fb1}^2)$ , with  $\sigma_1^2 = \sigma_a^2 + \sigma_b^2$ . The expressions (30) and (45) can now be coupled as in Section IV-B to yield a region which is same as the capacity region of a GBC without feedback where receiver 1 observes  $Y_a$  and receiver 2 observes  $Y_2$ . We remind the reader that  $\gamma$  in (31) should be taken as  $\gamma = \sigma_a^2 / (\sigma_{fb1}^2 + \sigma_b^2)$  while using (45) instead of (27). Thus the no feedback capacity region of a GBC which observes  $Y_a$  and  $Y_2$  at the respective receivers will be an outer bound to the capacity region of the original BC with noisy feedback. The maximum value of the forward variance  $\sigma_a^2$  that also satisfies the condition in (5) is indeed the  $\sigma^2$  given in (9). It is evaluated as the solution to

$$\frac{\sigma^2}{\sigma^2} = \left( \frac{\sigma^2}{\sigma_1^2 + \sigma_{fb1}^2 - \sigma^2} + 1 \right) \left( \frac{\sigma^2}{\sigma_{fb2}^2} + 1 \right). \quad (46)$$

This completes the proof of Theorem 4.

## V. ZERO-FORCING ACHIEVABLE SCHEMES FOR THE GBC WITH NOISY FB (PROOF OF LEMMA 5 AND DIRECT PART TO THEOREM 1)

In order to complete the proof of Theorem 1, we now show that a rate pair outside  $\mathcal{C}^{\text{wo-fb}}$  is achievable if the condition in (5) is violated by the given tuple  $(\sigma_1^2, \sigma_2^2, \sigma_{fb1}^2, \sigma_{fb2}^2)$  of noise variances. While several feedback coding schemes are available in literature, the main difficulty is in having tractable rate expressions which can show the required enlargement. Rate regions incorporating noisy feedback are typically stated in terms of the intersections of several hyperplanes, and are thus difficult to express in suitable functional forms for comparison [9], [10]. Some simplifications are possible, for example, when  $\sigma_1^2 = \sigma_2^2$ , the rate region proposed in [10] is shown to achieve rate-pairs outside  $\mathcal{C}^{\text{wo-fb}}$ , by suitable substitution of auxiliary variables, and thereby simplifying the expressions. However extending this to find the set of noise variances for which an enlargement becomes possible seems difficult in general. Instead, we will use a simple linear feedback coding scheme to show the required enlargement. Lemma 5 will also be proved along the way, making a separate proof for it superfluous.



### A. A Simple Linear-Feedback Coding Scheme

Let us employ a linear feedback coding scheme where the noise realization from each receiver after an odd numbered transmission instant, perceived through the noisy feedback link, is linearly combined and sent along with new symbols in the very next instant. Thus the transmitted signal only depends on feedback during even numbered channel uses, and the feedback values need to be stored for just one instant at the receiver. The fresh symbols at each instant, which are linearly combined with feedback, are generated and conveyed to all parties, as in standard random coding arguments [2]. Specifically, we construct two independent Gaussian codebooks to convey the messages  $W_1$  and  $W_2$  to the respective users. Let  $U_1, \dots, U_n$  denote the codeword symbols to the first receiver and  $V_1, \dots, V_n$  denote the codeword symbols to the second receiver. In order to convey the symbols  $(u_i, v_i)$  chosen from the codebooks, the transmitter sends,

$$X_{2i-1} = u_i + v_i \quad (47)$$

$$X_{2i} = \sqrt{\alpha} \left[ u_i - v_i + \beta_1 (Z_{1,2i-1} + Z_{fb1,2i-1}) + \beta_2 (Z_{2,2i-1} + Z_{fb2,2i-1}) \right] \quad (48)$$

where  $\alpha, \beta_1$  and  $\beta_2$  are appropriate real valued parameters. Let  $\hat{Z}_{j,i}$  be the MMSE estimate of  $Z_{j,2i-1}$  given  $Z_{j,2i-1} + Z_{fbj,2i-1}$ , for  $j = 1, 2$ . Then,

$$\hat{Z}_{j,i} \triangleq \frac{\sigma_j^2}{\sigma_j^2 + \sigma_{fbj}^2} (Z_{j,2i-1} + Z_{fbj,2i-1}), \quad (49)$$

which has variance  $\hat{\sigma}_j^2 \triangleq \frac{\sigma_j^4}{\sigma_j^2 + \sigma_{fbj}^2}$ . Defining

$$\gamma_j \triangleq \beta_j \frac{\sigma_j^2}{\hat{\sigma}_j^2}, \quad j = 1, 2, \quad (50)$$

we can rewrite (48) as

$$X_{2i} = \sqrt{\alpha} (u_i - v_i + \gamma_1 \hat{Z}_{1,i} + \gamma_2 \hat{Z}_{2,i}). \quad (51)$$

To facilitate random coding arguments, let us choose the distribution  $p(u, v)$  according to  $(U, V) \sim \mathcal{N}(0, \text{diag}(\theta'P, (1 - \theta')P))$  for some  $\theta' \in [0, 1]$ . This choice of  $(U, V)$  ensures that the average power constraint is met over the transmissions  $X_{2i-1}$  at odd time instants. To ensure the same for  $X_{2i}$ , we choose

$$\frac{1}{\alpha} = 1 + \frac{\gamma_1^2 \hat{\sigma}_1^2}{P} + \frac{\gamma_2^2 \hat{\sigma}_2^2}{P}. \quad (52)$$

The following operations are performed at the two receivers for decoding the respective messages.

**Receiver 1:** In two consecutive instants, the observations are

$$Y_{1,2i-1} = u_i + v_i + Z_{1,2i-1} \quad (53)$$

$$Y_{1,2i} = \sqrt{\alpha} (u_i - v_i + \gamma_1 \hat{Z}_{1,i} + \gamma_2 \hat{Z}_{2,i}) + Z_{1,2i}. \quad (54)$$

A simple *zero forcing* is achieved by computing

$$\begin{aligned} S_{1,i} &\triangleq Y_{1,2i-1} + \frac{Y_{1,2i}}{\sqrt{\alpha}} \\ &= 2u_i + \gamma_1 \hat{Z}_{1,i} + \gamma_2 \hat{Z}_{2,i} + Z_{1,2i-1} + \frac{Z_{1,2i}}{\sqrt{\alpha}}. \end{aligned} \quad (55)$$

**Receiver 2:** Here also we do *zero forcing*. On observing

$$Y_{2,2i-1} = u_i + v_i + Z_{2,2i-1} \quad (56)$$

$$Y_{2,2i} = \sqrt{\alpha} (u_i - v_i + \gamma_1 \hat{Z}_{1,i} + \gamma_2 \hat{Z}_{2,i}) + Z_{2,2i} \quad (57)$$

in two consecutive instants, the receiver computes

$$\begin{aligned} S_{2,i} &\triangleq Y_{2,2i-1} - \frac{Y_{2,2i}}{\sqrt{\alpha}} \\ &= 2v_i - \gamma_1 \hat{Z}_{1,i} - \gamma_2 \hat{Z}_{2,i} + Z_{2,2i-1} - \frac{Z_{2,2i}}{\sqrt{\alpha}}. \end{aligned} \quad (58)$$

Each receiver  $j \in \{1, 2\}$  will attempt to decode its intended message  $W_j$  based on the symbols  $S_{j,1}, \dots, S_{j,n}$  where  $n$  is the codeword length in each codebook.

**Analysis of error probability:** Notice that the zero forcing performed above creates an equivalent point-to-point channel to each receiver, with no interference from the other user's symbols. However the transmissions take place over a block-length of  $2n$  instants now. Therefore, standard random coding arguments [17] imply that the rate pair  $(\tilde{R}_1, \tilde{R}_2)$  is achievable, where

$$\tilde{R}_1 = \frac{1}{4} \log \left( 1 + \frac{4\theta'P}{\sigma_1^2 \left(1 + \frac{1}{\alpha}\right) + (\gamma_1^2 + 2\gamma_1)\hat{\sigma}_1^2 + \gamma_2^2 \hat{\sigma}_2^2} \right) \quad (59)$$

$$\tilde{R}_2 = \frac{1}{4} \log \left( 1 + \frac{4(1 - \theta')P}{\sigma_2^2 \left(1 + \frac{1}{\alpha}\right) + (\gamma_2^2 - 2\gamma_2)\hat{\sigma}_2^2 + \gamma_1^2 \hat{\sigma}_1^2} \right), \quad (60)$$

with  $\alpha$  given by (52), and  $(\theta', \gamma_1, \gamma_2)$  being arbitrary real tuples satisfying  $\theta' \in [0, 1]$  and  $\gamma_1, \gamma_2 \in \mathbb{R}$ . We remark that the rate expressions above match the ones given in (11), thereby directly proving Lemma 5.

As discussed in Section II after the statement of Lemma 5, the expressions in (59) immediately show that when  $\sigma_1^2 = \sigma_2^2$ , for any power level  $P$ , rate-pairs outside the no-feedback capacity region are achievable. In the coming subsection, we show that the same applies also for unequal noise variances  $\sigma_1^2 < \sigma_2^2$  when the power level  $P$  is sufficiently small. This argument is then used in Section V-C to prove a capacity enlargement for arbitrary noise variances  $(\sigma_1^2, \sigma_2^2)$  and power level  $P$ , when (5) is violated.

### B. Improving on $C^{\text{wo-fb}}$ at Low Power Levels

Recall our assumption that  $\sigma_1^2 \leq \sigma_2^2$ , and notice that the Pareto optimal rate-pairs on the boundary of  $C^{\text{wo-fb}}$  given in (3) can be indexed by the continuous parameter  $\theta \in [0, 1]$ . We fix a suitable  $\theta \in [0, 1]$  and show that for small enough  $P$  the corresponding Pareto optimal rate pair in (3) is dominated by the achievable rate pair  $(\tilde{R}_1, \tilde{R}_2)$  in (59) for some appropriate choice of  $(\theta', \gamma_1, \gamma_2)$ .

Consider an arbitrary quadruple  $(\theta, \theta', \gamma_1, \gamma_2)$  and define

$$\zeta \triangleq \gamma_1^2 \hat{\sigma}_1^2 + \gamma_2^2 \hat{\sigma}_2^2. \quad (61)$$

By (59), the pair  $(\tilde{R}_1, \tilde{R}_2)$  (for the parameters  $\theta', \gamma_1, \gamma_2$ ) dominates the pair  $(R_1, R_2)$  (for  $\theta$ ) if

$$1 + \frac{4\theta'P}{\sigma_1^2(1 + \frac{1}{\alpha}) + 2\gamma_1\hat{\sigma}_1^2 + \zeta} \geq \left(1 + \frac{\theta P}{\sigma_1^2}\right)^2 \quad (62)$$

$$1 + \frac{4(1 - \theta')P}{\sigma_2^2(1 + \frac{1}{\alpha}) - 2\gamma_2\hat{\sigma}_2^2 + \zeta} > \left(1 + \frac{(1 - \theta)P}{\theta P + \sigma_2^2}\right)^2,$$

or equivalently,

$$\frac{4\theta'}{\sigma_1^2(1 + \frac{1}{\alpha}) + 2\gamma_1\hat{\sigma}_1^2 + \zeta} \geq \frac{2\theta}{\sigma_1^2} + \frac{\theta^2 P}{\sigma_1^4} \quad (63)$$

$$\frac{4(1 - \theta')}{\sigma_2^2(1 + \frac{1}{\alpha}) - 2\gamma_2\hat{\sigma}_2^2 + \zeta} > \frac{2(1 - \theta)}{\theta P + \sigma_2^2} + \frac{(1 - \theta)^2 P}{(\theta P + \sigma_2^2)^2}.$$

Let us change the variables from  $(\gamma_1, \gamma_2, \theta')$  to  $(a_1, a_2, \mu)$  by defining

$$a_j \triangleq (-1)^j \frac{\gamma_j \hat{\sigma}_j}{P}, \quad j = 1, 2 \quad (64)$$

$$\mu \triangleq \frac{1}{P} \left( \frac{\theta'}{\theta} - 1 \right). \quad (65)$$

While there are no restrictions on the parameters  $(a_1, a_2)$ , the parameter  $\mu$  needs to lie in the interval  $[-\frac{1}{P}, \frac{1}{P}(\frac{1}{\theta} - 1)]$  so that  $0 \leq \theta' \leq 1$ . Notice that the described interval for possible  $\mu$  can be made to include any desired real value by choosing  $P$  sufficiently small.

Using the transformations (64)–(65) on (63), we get

$$\frac{4(1 + \mu P)}{(a_1^2 + a_2^2)(P^2 + P\sigma_1^2) + 2\sigma_1^2 - 2a_1\hat{\sigma}_1 P} \geq \frac{2}{\sigma_1^2} + \frac{\theta P}{\sigma_1^4} \quad (66)$$

$$\frac{4(1 - \theta(1 + \mu P))}{(a_1^2 + a_2^2)(P^2 + P\sigma_2^2) + 2\sigma_2^2 - 2a_2\hat{\sigma}_2 P} > \frac{2(1 - \theta)}{\theta P + \sigma_2^2} + \frac{(1 - \theta)^2 P}{(\theta P + \sigma_2^2)^2}. \quad (67)$$

Clearly, the above expressions hold with equality at  $P = 0$ . Therefore, if the pair of derivatives on the LHS dominates the corresponding RHS derivatives as  $P \rightarrow 0^+$ , this will show that the required capacity enlargement is possible at low enough powers. Differentiating with respect to  $P$  and setting  $P = 0$ , we get the conditions

$$2\mu\sigma_1^2 \geq \theta + (a_1^2 + a_2^2)\sigma_1^2 - 2a_1\hat{\sigma}_1 \quad (68)$$

$$(1 - \theta)(2a_2\hat{\sigma}_2 - (a_1^2 + a_2^2)\sigma_2^2) > (1 - \theta)(1 - 3\theta) + 2\theta\mu\sigma_2^2. \quad (69)$$

We remind the reader that (68)–(69) will imply (66)–(67) only at sufficiently low powers. Additional arguments will be provided later in Section V-C for extensions to more general average power levels. The first of the above two equations can be guaranteed by choosing

$$\mu \triangleq \frac{\theta}{2\sigma_1^2} + \frac{a_1^2 + a_2^2}{2} - \frac{a_1\hat{\sigma}_1}{\sigma_1^2}, \quad (70)$$

whereas the second equation, on substitution of (70), will yield

$$a_1^2 + a_2^2 < \frac{2a_1\theta}{\sqrt{\sigma_1^2 + \sigma_{fb1}^2}} + \frac{2a_2(1 - \theta)}{\sqrt{\sigma_2^2 + \sigma_{fb2}^2}} - \frac{\theta^2}{\sigma_1^2} - \frac{(1 - \theta)(1 - 3\theta)}{\sigma_2^2}. \quad (71)$$

Notice that above choice of  $\mu$  does not depend on  $P$  and thus lies in the desired interval for all sufficiently small values of  $P > 0$ . Since  $a_1, a_2$  are free parameters, we can choose them as

$$a_1 = \frac{\theta}{\sqrt{\sigma_1^2 + \sigma_{fb1}^2}}$$

$$a_2 = \frac{1 - \theta}{\sqrt{\sigma_2^2 + \sigma_{fb2}^2}}.$$

By substituting this into (71), we need to verify

$$\frac{\theta^2}{\sigma_1^2} + \frac{(1 - \theta)(1 - 3\theta)}{\sigma_2^2} - \frac{\theta^2}{\sigma_1^2 + \sigma_{fb1}^2} - \frac{(1 - \theta)^2}{\sigma_2^2 + \sigma_{fb2}^2} < 0. \quad (72)$$

for some value of  $\theta \in [0, 1]$ . Defining  $x = \frac{\theta}{1 - \theta}$ , one can equivalently check if  $g(x) < 0$  for some value of  $x > 0$ , where

$$g(x) \triangleq \frac{x\sigma_{fb1}^2}{\sigma_1^2(\sigma_1^2 + \sigma_{fb1}^2)} + \frac{\sigma_{fb2}^2}{x\sigma_2^2(\sigma_2^2 + \sigma_{fb2}^2)} - \frac{2}{\sigma_2^2}.$$

In fact, the function  $g(x)$  is minimized for  $x > 0$  by

$$x^* = \sqrt{\left(\frac{1 + \frac{\sigma_1^2}{\sigma_{fb1}^2}}{1 + \frac{\sigma_2^2}{\sigma_{fb2}^2}}\right) \left(\frac{\sigma_1^2}{\sigma_2^2}\right)}. \quad (73)$$

Now the condition for  $g(x^*) < 0$  can be seen to be equivalent to

$$\frac{\sigma_2^2}{\sigma_1^2} < \left(\frac{\sigma_1^2}{\sigma_{fb1}^2} + 1\right) \left(\frac{\sigma_2^2}{\sigma_{fb2}^2} + 1\right), \quad (74)$$

which is the complement of condition (5) given in Theorem 1.

Putting it all together, we have shown a capacity enlargement at sufficiently small powers, when at least one of the passive feedback links is not too noisy, as implied by (74).

### C. Improving on $C^{\text{wo-fb}}$ for General Power Levels

We now show that any capacity enlargement at low power levels naturally extends to a capacity enlargement at arbitrary transmit powers. This can be shown by message splitting and bootstrapping the low power achievable scheme. Consider a rate-pair in which user 1 demands a small enough positive rate  $R_1$ . Let us split the message  $W_2$  for user 2 into 2 sub-messages  $W_{2,1}$  and  $W_{2,2}$  of respective rates  $R_{2,1}$  and  $R_{2,2}$ . Consider a small positive  $\epsilon$ , and appropriate parameters  $\theta' \in [0, 1]$  and  $\gamma_1, \gamma_2 \in \mathbb{R}$ . The message  $W_{2,2}$  is conveyed to receiver 2 using a simple point-to-point scheme of power  $P - \epsilon$ . We use the variables  $U \sim \mathcal{N}(0, \theta'\epsilon)$  and  $V_1 \sim \mathcal{N}(0, (1 - \theta')\epsilon)$  to denote the code symbols employed in the above linear feedback code construction to encode  $(W_1, W_{2,1})$ , and the variable  $V_2 \sim \mathcal{N}(0, P - \epsilon)$  denotes the symbol employed in the point-to-point code encoding  $W_{2,2}$ . All codebooks are generated independently, and we employ standard random

coding arguments to find the error probability. Over two successive channel uses, the sender transmits

$$\begin{aligned} X_{2i-1} &= u_i + v_{1,i} + v_{2,2i-1} \\ X_{2i} &= \sqrt{\alpha}(u_i - v_{1,i} + \gamma_1 \hat{Z}_{1,2i-1} + \gamma_2 \hat{Z}_{2,2i-1}) + v_{2,2i}. \end{aligned} \quad (75)$$

Similar to (51) and (52), the parameters  $\alpha, \gamma_1$ , and  $\gamma_2$  are chosen so that the sum  $\sqrt{\alpha}(u_i - v_{1,i} + \gamma_1 \hat{Z}_{1,2i-1} + \gamma_2 \hat{Z}_{2,2i-1})$  satisfies the average power constraint of  $\epsilon$ .

Receiver 2 infers its desired messages via the following steps:

- It decodes message  $W_{2,2}$  by treating everything as noise. Clearly the transmissions in successive instants are independent, and we can treat the rest of the transmissions as memoryless Gaussian noise sequences while decoding the  $V_2$  codewords.
- After inferring  $W_{2,2}$ , it subtracts the  $V_2$  codeword to obtain a *more clean* BC with equivalent average transmit power  $\epsilon$ , and it applies the *zero forcing* decoding scheme as in (58).

Receiver 1 decodes in a similar way: it first decodes messages  $W_{2,2}$  (even though this message is not intended for it), subtracts the  $V_2$  codeword, and applies the zero forcing decoding scheme described in (55).

We show that for appropriate choices of  $\epsilon, \theta', \gamma_1, \gamma_2$ , the proposed scheme achieves a rate-point outside the no-feedback capacity region  $\mathcal{C}^{\text{wo-fb}}$ . We first notice that  $R_1$  and  $R_{2,1}$  can be chosen as the rates  $\tilde{R}_1$  and  $\tilde{R}_2$  in (59) and (60), however with the average power  $P$  replaced by  $\epsilon$ . We have shown in the previous subsection that for sufficiently small  $\epsilon > 0$  there exist choices of  $\theta', \gamma_1, \gamma_2$  and  $\theta$  such that

$$R_1 = \frac{1}{2} \log_2 \left( 1 + \frac{\theta \epsilon}{\sigma_1^2} \right), \quad (76)$$

$$R_{2,1} > \frac{1}{2} \log_2 \left( 1 + \frac{(1-\theta)\epsilon}{\theta \epsilon + \sigma_2^2} \right). \quad (77)$$

We continue with such a sufficiently small value of  $\epsilon$ . Since Message  $W_{2,2}$  can be transmitted at a rate

$$R_{2,2} = \frac{1}{2} \log_2 \left( 1 + \frac{P - \epsilon}{\epsilon + \sigma_2^2} \right), \quad (78)$$

by introducing  $\theta'' \triangleq \theta \frac{\epsilon}{P}$ , we conclude that the rate pair

$$R_1 = \frac{1}{2} \log_2 \left( 1 + \frac{\theta'' P}{\sigma_1^2} \right), \quad (79)$$

$$\begin{aligned} R_2 &= R_{2,1} + R_{2,2} \\ &> \frac{1}{2} \log_2 \left( 1 + \frac{(1-\theta)\epsilon}{\theta \epsilon + \sigma_2^2} \right) + \frac{1}{2} \log_2 \left( 1 + \frac{P - \epsilon}{\epsilon + \sigma_2^2} \right) \\ &= \frac{1}{2} \log_2 \left( 1 + \frac{(1-\theta'')P}{\theta'' P + \sigma_2^2} \right). \end{aligned} \quad (80)$$

is achievable. Observing that this rate-pair lies outside  $\mathcal{C}^{\text{wo-fb}}$  is sufficient to conclude the proof. In short, we used the enlargement for small values of power, and allotted all the remaining power for transmissions to the weak receiver. This allowed us to strictly improve the no-feedback capacity region at high power levels as well.

## VI. CONCLUSION

We have analyzed the effect of passive noisy feedback in enlarging the capacity region of a Gaussian broadcast channel. Interestingly, too much noise in both the feedback links does not lead to any enlargement in the asymmetric user case, where the two receivers have different noise variances. This is in sharp contrast to the case of identical noise variances at the receivers, where it is known that any noise of finite variance in the feedback link is beneficial for capacity enlargement. While we have characterized the regime of noise variances for which no enlargement occurs, our achievable scheme shows an enlargement outside this proposed regime, thus making the characterization sharp. While showing an enlargement, however small, was sufficient for our purposes here, there is still some way to go for finding the actual capacity region with enlargement. More sophisticated coding schemes than the ones proposed here may be required to achieve that.

The genie aided converse constructions, and associated ideas, seem to be beneficial in analyzing other multi-user systems with feedback. This is currently under consideration.

## APPENDIX A PROOF OF LEMMA 8

*Proof:* The proof is very similar to that in [5]. For the sake of completeness, here we repeat the arguments, which proceed by induction on  $n$ . Notice that since  $Z_{1i}$  and  $Z_{2i}$  are both Gaussian with  $\text{Var}(Z_{1i}) \leq \text{Var}(Z_{2i})$  and independent of  $(\tilde{U}_i, X_i)$ , in (37) we can take  $Y_{2i} = Y_{1i} + \tilde{Z}_i$ , where  $\tilde{Z}_i \sim \mathcal{N}(0, \sigma_2^2 - \sigma_1^2)$  is independent of  $(X_i, Z_{1i}, \tilde{U}_i)$ . Recall that  $\tilde{U}_i \rightarrow X_i \rightarrow (Y_{1i}, Y_{2i})$ . For  $n = 1$ , the inequality follows from entropy power inequality [2, pg. 22] since

$$h(Y_{21}|\tilde{U}_1) = h(Y_{11} + \tilde{Z}_1|\tilde{U}_1).$$

Let us now assume that (37) is true for  $n = m - 1$ . By the conditional EPI [2],

$$2^{2h(Y_{2m}|\tilde{U}_m)} \geq 2^{2h(Y_{1m}|\tilde{U}_m)} + 2^{2h(\tilde{Z}_m)}.$$

i.e.,

$$2h(Y_{2m}|\tilde{U}_m) \geq \log \left( 2^{2h(Y_{1m}|\tilde{U}_m)} + 2\pi e(\sigma_2^2 - \sigma_1^2) \right).$$

Therefore,

$$\begin{aligned} &\frac{2}{m} \sum_{i=1}^m h(Y_{2i}|\tilde{U}_i) \\ &= \frac{m-1}{m} \frac{2}{m-1} \sum_{i=1}^{m-1} h(Y_{2i}|\tilde{U}_i) + \frac{2}{m} h(Y_{2m}|\tilde{U}_m) \\ &\stackrel{(a)}{\geq} \frac{m-1}{m} \log \left( 2^{\frac{2}{m-1} \sum_{i=1}^{m-1} h(Y_{1i}|\tilde{U}_i)} + 2\pi e(\sigma_2^2 - \sigma_1^2) \right) \\ &\quad + \frac{1}{m} \log \left( 2^{2h(Y_{1m}|\tilde{U}_m)} + 2\pi e(\sigma_2^2 - \sigma_1^2) \right) \\ &\stackrel{(b)}{\geq} \log \left( 2^{\frac{2}{m} \sum_{i=1}^m h(Y_{1i}|\tilde{U}_i)} + 2\pi e(\sigma_2^2 - \sigma_1^2) \right). \end{aligned}$$

Here (a) follows from the induction hypothesis and the EPI above, and (b) follows from convexity of  $\log(2^u + v)$  in  $u$  for  $v \geq 0$ . ■

## REFERENCES

- [1] C. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, 1956.
- [2] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.
- [3] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Problems Inform. Transmission*, vol. 10, no. 3, pp. 185–193, 1974.
- [4] A. El Gamal, "The feedback capacity of degraded broadcast channels (corresp.)," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 379–381, 1978.
- [5] —, "The capacity of the physically degraded Gaussian broadcast channel with feedback (corresp.)," *IEEE Transactions on Information Theory*, vol. 27, no. 4, pp. 508–511, 1981.
- [6] G. Dueck, "Partial feedback for two-way and broadcast channels," *Information and Control*, vol. 46, no. 1, pp. 1–15, 1980.
- [7] 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, 1984.
- [8] S. R. Bhashkar, "Gaussian broadcast channel with feedback," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5252–5257, 2008.
- [9] O. Shayevitz and M. Wigger, "On the capacity of the discrete memoryless broadcast channel with feedback," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1329–1345, 2012.
- [10] R. Venkataramanan and S. S. Pradhan, "An achievable rate region for the broadcast channel with feedback," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6175–6191, 2013.
- [11] Z. Ahmad, Z. Chance, D. J. Love, and C. Wang, "Concatenated coding using linear schemes for Gaussian broadcast channels with noisy channel output feedback," *IEEE Transactions on Communications*, vol. 63, no. 11, pp. 4576–4590, 2015.
- [12] Y. Wu and M. Wigger, "Coding schemes with rate-limited feedback that improve over the no feedback capacity for a large class of broadcast channels," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 2009–2033, 2016.
- [13] A. Lapidoth and M. Wigger, "On the AWGN MAC with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5432–5476, 2010.
- [14] S. B. Amor, Y. Steinberg, and M. Wigger, "MAC-BC duality with linear-feedback schemes," in *2014 IEEE International Symposium on Information Theory*. IEEE, 2014, pp. 1737–1741.
- [15] S. R. B. Pillai and V. M. Prabhakaran, "On the noisy feedback capacity of Gaussian broadcast channels," in *2015 IEEE Information Theory Workshop (ITW)*. IEEE, 2015, pp. 1–5.
- [16] A. Narayan, S. R. B. Pillai, V. M. Prabhakaran, and M. Wigger, "When does partial noisy feedback enlarge the capacity of a Gaussian broadcast channel?" in *2020 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2020.
- [17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 1991.
- [18] Y. Geng and C. Nair, "The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2087–2104, 2014.
- [19] V. Ramachandran and S. R. B. Pillai, "Feedback-capacity of degraded Gaussian vector BC using directed information and concave envelopes," in *2017 Twenty-third National Conference on Communications India (NCC)*, 2017, pp. 1–6.