Feedback Can Double the Prelog of Some Memoryless Gaussian Networks

Michael Gastpar, Amos Lapidoth, Yossef Steinberg, and Michèle Wigger

***Abstract*—We exhibit two memoryless Gaussian networks where the capacity-gains afforded by feedback are unbounded in the signal-to-noise ratio (SNR). The networks are instances of the Gaussian broadcast channel and the two-user Gaussian interference channel. To demonstrate the capacity-gains we propose and analyze a novel feedback coding scheme.**

For the broadcast channel with two receivers it is shown that if the noise sequences at the two receivers are perfectly anticorrelated, then, at high SNR, feedback asymptotically doubles the sum-capacity. The same holds if the noise sequences are perfectly correlated provided that they are of unequal variances. This result extends to the multi-receiver broadcast channel: if the noise sequences are all different and have a rank-one covariance matrix, then, at high-SNR, feedback asymptotically multiplies the sum-capacity by the number of receivers. However, as we show, these multiplicative gains collapse when the feedback is noisy.

For the two-receiver Gaussian broadcast channel with noise-free feedback we also derive the high-SNR asymptotic sum-capacity. The expansion is exact in the sense that, as the SNR tends to infinity, the difference between the sum-capacity and our asymptotic expression tends to zero. If the noise sequences are perfectly anticorrelated or if they are perfectly correlated and of unequal variances, then the asymptotic expansion is as if the transmitter communicated to the two receivers over two parallel Gaussian channels. Otherwise, the asymptotic expansion is the same as if the receivers could cooperate.

For the two-user interference channel it is shown that if the noises experienced by the two receivers are perfectly correlated or perfectly anticorrelated, then for most channel-gains feedback doubles the high SNR sum-capacity.

I. Introduction

We present two memoryless Gaussian networks where the capacity gains afforded by noise-free feedback are unbounded in the signal-to-noise ratio (SNR). The networks are instances of the real, scalar, additive white Gaussian noise (AWGN) broadcast channel (BC) and the real, scalar, AWGN interference channel (IC).

For the two-user AWGN BC we prove that, when the AWGN sequences corrupting the outputs at the two receivers are *perfectly anti-correlated* (i.e, have correlation coefficient $\rho_z = -1$) or when they are *perfectly correlated* (i.e, $\rho_z = 1$) and of unequal noise variances ($\sigma_1^2 \neq \sigma_2^2$), at high SNR, noise-free feedback approximately doubles the sum-capacity from $\frac{1}{2}\log(1+\text{SNR})$ to approximately $2 \cdot \frac{1}{2}\log(1+\text{SNR})$. Thus, in these cases the noise-free feedback increases the sum-capacity by approximately $\frac{1}{2}\log(1+\text{SNR})$ which is unbounded in the SNR. The same holds also even when the two AWGN sequences are sufficiently close to perfectly anticorrelated (i.e., $\rho_z \approx -1+\frac{1}{\text{SNR}}$), or when they are sufficiently close to perfectly correlated (i.e., $\rho_z \approx 1-\frac{1}{\text{SNR}}$) and have different variances ($\sigma_1^2 \neq \sigma_2^2$).

In these cases the *prelog*[1] (the factor in the sum-capacity high-SNR expansion in front of $\frac{1}{2}\log(1+\text{SNR})$) is 2 and exceeds the number of transmit antennas. This is in sharp contrast to the case where the received signals are corrupted by *independent* AWGN sequences, a case in which the prelog (with or without feedback) cannot exceed the number of transmit (or receive) antennas [1].

To demonstrate these gains we propose and analyze a novel feedback coding scheme, similar to the schemes in [4], [5], and [6]. (The gains for anti-correlated AWGN sequences are also achieved by the Ozarow-Leung scheme in [8], [9] if the scheme's parameter are chosen carefully, see [10], [11].) Our proposed coding scheme also allows us to determine the high-SNR asymptotics of the sum-capacity of the two-receiver AWGN BC with noise-free feedback and arbitrary noise correlation. The expansion is exact in the sense that the difference between it and the sum-capacity $C_{\text{BC},\Sigma}$ vanishes as the SNR tends to infinity:

- When the two AWGN sequences are perfectly anticorrelated (i.e, have correlation coefficient $\rho_z = -1$) or when they are perfectly correlated and have unequal variances $\sigma_1^2 \neq \sigma_2^2$,

$$\lim_{P\to\infty}\left[C_{\text{BC},\Sigma} - \left(\frac{1}{2}\log\left(1+\frac{P}{\sigma_1^2}\right) + \frac{1}{2}\log\left(1+\frac{P}{\sigma_2^2}\right)\right)\right] = 0. \tag{1}$$

  Thus, the sum-capacity approaches the sum-capacity of two parallel AWGN channels where the transmitter can communicate with power P over each of them. In this case, noise-free feedback doubles the prelog and thus provides gains in the sum-capacity that are unbounded in the SNR.
- When the AWGN sequences are neither perfectly anticorrelated nor perfectly correlated,

$$\lim_{P\to\infty}\left[C_{\text{BC},\Sigma} - \frac{1}{2}\log\left(1+\frac{P(\sigma_1^2+\sigma_2^2-2\rho_z\sigma_1\sigma_2)}{\sigma_1^2\sigma_2^2(1-\rho_z^2)}\right)\right] = 0. \tag{2}$$

  and the sum-capacity approaches the sum-capacity of a setup where the two receivers can cooperate. Here feedback does not increase the prelog, and its benefit at high SNR is the additive term $\frac{1}{2}\log\left(\frac{\sigma_1^2+\sigma_2^2-2\rho_z\sigma_1\sigma_2}{\max\{\sigma_1^2,\sigma_2^2\}(1-\rho_z^2)}\right)$. This gain is bounded in the power, but it can be very

[1] The prelog is often also referred to as the *multiplexing gain* or *degrees of freedom*.

significant if ρ_z is close to -1 or if it is close to 1 with $\sigma_1^2 \neq \sigma_2^2$.

For the $K > 2$ user AWGN BC our proposed coding scheme allows us to prove that if the AWGN sequences corrupting the K received signals are all different but have covariance matrix of rank 1, then a prelog of K is achievable. For a related recent result see [12].

A naturally ensuing question is whether these gains in capacity can also be attained when the feedback is noisy. In this paper we concentrate on how feedback noise affects the reported gains in the prelog. We show that when the feedback links are corrupted by independent AWGN sequences, then—irrespective of the positive feedback-noise variances and of the correlation of the forward noise-sequences—the prelog of the two-user AWGN BC setup equals one (as in the absence of feedback). The proof of this result is based on a genie argument inspired by the work of Kim, Lapidoth, and Weissman [13].

The second network we consider is the two-user scalar AWGN IC with noise-free *one-sided* feedback where each of the two transmitters communicates with a different intended receiver, and each transmitter observes feedback from its corresponding receiver only. Our proposed coding scheme proves that when the two AWGN sequences experienced at the two receivers are perfectly anticorrelated or perfectly correlated, then for most channel gains noise-free feedback doubles the prelog from 1 to 2. Thus, noise-free feedback allows to approximately double the sum-capacity at high SNR and can provide unbounded gains. (When the interference channel is symmetric, the prelog 2 result can also be shown using a slight generalization (to account for the correlation between the noise sequences) of Kramer's memoryless LMMSE-scheme [14].)

Previously, a prelog of 2 was known to be achievable for the two-user scalar AWGN IC only when the two transmitters (or the two receivers) could *fully* cooperate [15] in the sense that both transmitters can compute their channel inputs as a function of both messages. Our result shows that *limited* cooperation through feedback can be sufficient.

For the two-user AWGN IC we do not consider noisy feedback. Rate-limited feedback for this setup has recently been studied in [18].

The fact that feedback can increase the capacity of memoryless networks without bounds was first reported by the authors in [2] for the two-user AWGN BC and for the symmetric two-user AWGN IC where the individual noise sequences corrupting the outputs at the two receivers are perfectly anticorrelated. The same conclusion was later also reached in [3], [16] (also based on the scheme proposed in [17]) for the two-user AWGN IC when the noise sequences are independent. Multiplicative gains for the AWGN IC with independent noises at moderate SNR were already reported in [14, Section VI-B].

We conclude this section with some notation and an outline of the rest of the paper. Throughout the paper the logarithm is taken with respect to the base 2 and for convenience we define $-\log 0 = \infty$. We will use the shorthand notation $\log^+(x)$ for $\max\{0, \log(x)\}$. Also, we denote by A^n and a^n the tuple of random variables $A_1, \ldots, A_n$ and their realizations $a_1, \ldots, a_n$, respectively. The set of real numbers will be denoted by $\mathbb{R}$, the set of positive real numbers by $\mathbb{R}^+$, and the set of positive integers by $\mathbb{Z}^+$. The abbreviation *IID* stands for *independent and identically distributed*.

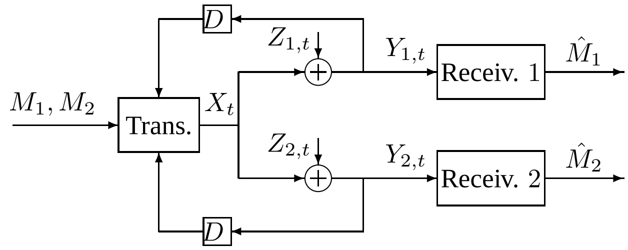

Fig. 1. The two-user AWGN BC with noise-free feedback.

The paper is organized as follows. In Sections II–V we present the channel models and the main results for the two-user AWGN BC with noise-free feedback, the two-user AWGN BC with noisy feedback, the $K \geq 2$ user AWGN BC with noise-free feedback, and the two-user AWGN IC with one-sided noise-free feedback. In Section VI we present a coding scheme for the two-user AWGN BC with noise-free feedback, and prove our results for this setup. In Section VII we prove our results for the AWGN BC with noisy feedback. In Sections VIII and IX we present coding schemes and prove our results for the K-user AWGN BC with noise-free feedback and the two-user AWGN IC with one-sided noise-free feedback, respectively.

II. Two-User Broadcast Channel with Noise-Free Feedback: Setup and Main Results

The real, scalar, AWGN BC with noise-free feedback is depicted in Figure 1. Denoting the time-t transmitted symbol by $x_t \in \mathbb{R}$ and the time-t received symbols by $Y_{1,t}$ and $Y_{2,t}$ respectively, we have

$$Y_{1,t} = x_t + Z_{1,t}, \tag{3a}$$

and

$$Y_{2,t} = x_t + Z_{2,t}, \tag{3b}$$

where the sequence of noise pairs $\{(Z_{1,t}, Z_{2,t})\}$ is drawn IID according to a centered Gaussian distribution of covariance matrix

$$\mathsf{K}_z = \begin{pmatrix} \sigma_1^2 & \rho_z\sigma_1\sigma_2 \\ \rho_z\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}. \tag{4}$$

We assume that both noise variances σ_1^2, σ_2^2 are positive (i.e., not equal to 0).

The transmitter wishes to send Message M_1 to Receiver 1 and an independent message M_2 to Receiver 2. The messages M_1 and M_2 are assumed to be uniformly distributed over the sets $\mathcal{M}_1 \triangleq \{1, \ldots, \lfloor 2^{nR_1} \rfloor\}$ and $\mathcal{M}_2 \triangleq \{1, \ldots, \lfloor 2^{nR_2} \rfloor\}$, where n denotes the blocklength and R_1 and R_2 the respective rates of transmission.

It is assumed that the transmitter has access to noise-free feedback from both receivers, i.e., that after sending x_{t-1} it learns both outputs $Y_{1,t-1}$ and $Y_{2,t-1}$. The transmitter can thus compute its time-t channel input as a function of both messages and all previous channel outputs:

$$X_t = f_{\text{BC},t}^{(n)}\left(M_1, M_2, Y_1^{t-1}, Y_2^{t-1}\right), \quad t \in \{1, \ldots, n\}, \tag{5}$$

where the encoding function $f_{\text{BC},t}^{(n)}$ is of the form

$$f_{\text{BC},t}^{(n)} \colon \mathcal{M}_1 \times \mathcal{M}_2 \times \mathbb{R}^{t-1} \times \mathbb{R}^{t-1} \to \mathbb{R}. \tag{6}$$

The channel inputs are subject to an expected average block-power constraint $P > 0$. Thus, in (5) we only allow for encoding functions $\left\{f_{\text{BC},t}^{(n)}\right\}_{t=1}^{n}$ such that

$$\frac{1}{n}\mathsf{E}\left[\sum_{t=1}^{n} X_t^2\right] \leq P. \tag{7}$$

After the n-th channel use each receiver decodes its intended message based on its observed channel output sequence. Receiver 1 produces the estimate

$$\hat{M}_1 \quad = \quad \phi_1^{(n)}(Y_1^n), \tag{8}$$

and Receiver 2 the estimate

$$\hat{M}_2 \quad = \quad \phi_2^{(n)}(Y_2^n), \tag{9}$$

where the decoding functions $\phi_1^{(n)}$ and $\phi_2^{(n)}$ are of the form

$$\phi_1^{(n)} \colon \mathbb{R}^n \to \{1, \ldots, \lfloor 2^{nR_1} \rfloor\}, \tag{10}$$

$$\phi_2^{(n)} \colon \mathbb{R}^n \to \{1, \ldots, \lfloor 2^{nR_2} \rfloor\}. \tag{11}$$

A rate pair (R_1, R_2) is said to be achievable if for every block-length n there exists a set of n encoding functions $\left\{f_{\text{BC},t}^{(n)}\right\}_{t=1}^{n}$ as in (6) satisfying the power constraint (7) and two decoding functions $\phi_1^{(n)}$ and $\phi_2^{(n)}$ as in (10) and (11) such that

$$\lim_{n\to\infty} \Pr\left[(M_1, M_2) \neq (\hat{M}_1, \hat{M}_2)\right] = 0.$$

The closure of the set of all achievable rate pairs (R_1, R_2) is called the *capacity region*. The supremum of the sum $R_1 + R_2$ over all achievable rate pairs (R_1, R_2) is called its *sum-capacity*, and we denote it by $C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)$. In this paper we are particularly interested in the *prelog*, which characterizes the logarithmic growth of the sum-capacity at high powers, and is defined as:

$$\varlimsup_{P\to\infty} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1+P)}. \tag{12}$$

The sum-capacity of our setup $C_{\text{BC},\Sigma}$ is unknown except for the *physically degraded* case where

$$\rho_z \in \left\{\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right\}. \tag{13}$$

In this case the capacity region is not increased with feedback [19], and thus, [20], [21]

$$C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z) = \frac{1}{2}\log\left(1 + \frac{P}{\min\{\sigma_1^2, \sigma_2^2\}}\right). \tag{14}$$

The AWGN BC with feedback is in particular physically degraded if $\rho_z = 1$ and $\sigma_1^2 = \sigma_2^2$, in which case both receivers observe exactly the same output sequence.

Our main result in Theorem 1 is the asymptotic high-SNR sum-capacity of the AWGN BC with feedback. Whenever the BC is not physically degraded, the asymptotic high-SNR sum-capacity with feedback is strictly larger than without.

Definition 1: Let $C_{\text{HighSNR}}(P, \sigma_1^2, \sigma_2^2, \rho_z)$ be defined as follows.

- For physically degraded channels, i.e., $\rho_z \in \left\{\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right\}$,

$$C_{\text{HighSNR}}(P, \sigma_1^2, \sigma_2^2, \rho_z) \triangleq \frac{1}{2}\log\left(1 + \frac{P}{\min\{\sigma_1^2, \sigma_2^2\}}\right). \tag{15}$$

- For not physically-degraded channels with fully correlated noises, i.e., $|\rho_z| = 1$ and $\rho_z \notin \left\{\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right\}$,

$$C_{\text{HighSNR}}(P, \sigma_1^2, \sigma_2^2, \rho_z) \triangleq \frac{1}{2}\log\frac{P}{\sigma_1^2} + \frac{1}{2}\log\frac{P}{\sigma_2^2}. \tag{16}$$

- Finally, for not physically degraded channels with partially correlated noises, i.e., $\rho_z \in (-1, 1)$ and $\rho_z \notin \left\{\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right\}$,[2]

$$C_{\text{HighSNR}}(P, \sigma_1^2, \sigma_2^2, \rho_z) \triangleq \frac{1}{2}\log\left(\frac{P(\sigma_1^2 + \sigma_2^2 - 2\rho_z\sigma_1\sigma_2)}{\sigma_1^2\sigma_2^2(1-\rho_z^2)}\right). \tag{17}$$

Theorem 1 (Asymptotic High-SNR Sum-Capacity): For all $\sigma_1^2, \sigma_2^2 > 0$ and $\rho_z \in [-1, 1]$

$$\lim_{P\to\infty}\left(C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z) - C_{\text{HighSNR}}(P, \sigma_1^2, \sigma_2^2, \rho_z)\right) = 0. \tag{18}$$

Proof: When the channel is physically degraded, i.e., $\rho_z \in \left\{\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right\}$, the result follows from (14). For the other cases the result is proved in Section VI-C. ■

Note 1: If $|\rho_z| = 1$ and the channel is not physically degraded, the sum-capacity at high SNR is as if there was a separate, non-interfering link from the transmitter to each of the two receivers and the transmitter could communicate with power P over both these links. If $\rho_z \in (-1, 1)$, then the sum-capacity at high SNR is as if the two receivers could fully cooperate in their decoding.

Note 2: Given $\sigma_2^2, \sigma_1^2 > 0$, define $\gamma \colon (-1, 1) \to \mathbb{R}^+$ as

$$\gamma(\rho_z) \triangleq \frac{\sigma_1^2 + \sigma_2^2 - 2\rho_z\sigma_1\sigma_2}{\sigma_1^2\sigma_2^2(1-\rho_z^2)}. \tag{19}$$

By Theorem 1, $\gamma(\rho_z)$ describes the behavior of the sum-capacity in the asymptotic high-SNR regime when $\rho_z \in (-1, 1)$. Notice that $\gamma(\rho_z) \to \infty$ as $\rho_z \to -1$, and if $\sigma_1^2 \neq \sigma_2^2$, then also $\gamma(\rho_z) \to \infty$ as $\rho_z \to +1$. Moreover, $\gamma(\rho_z)$ is strictly decreasing over $\rho_z \in \left(-1, \min\left\{\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right\}\right)$ and strictly increasing over $\rho_z \in \left(\min\left\{\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right\}, 1\right)$. It thus takes its minimum at $\rho_z = \min\left\{\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right\}$ where $\gamma(\rho_z) = \frac{1}{\min\{\sigma_1^2, \sigma_2^2\}}$.

Figure 2 shows the typical behavior of $\gamma(\rho_z)$.

From Theorem 1 we obtain the following corollary.

[2]Notice that for physically degraded channels with partially correlated noises, i.e., for $\rho_z \in (-1, 1)$ and $\rho_z \in \left\{\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right\}$, the definitions in (15) and (17) coincide.

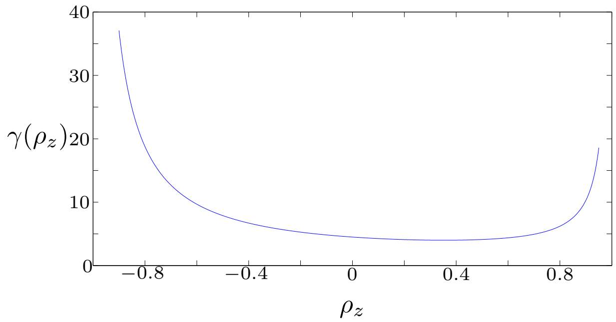

Fig. 2. The function $\gamma(\rho_z)$ is plotted over $\rho_z \in [-0.9, 0.95]$ for $\sigma_1^2 = 2$ and $\sigma_2^2 = 0.25$. The minimum is at $\sqrt{1/8} \approx 0.3536$, and the function is strictly decreasing over $\big(-1, \sqrt{1/8}\big)$ and strictly increasing over $\big(\sqrt{1/8}, 1\big)$.

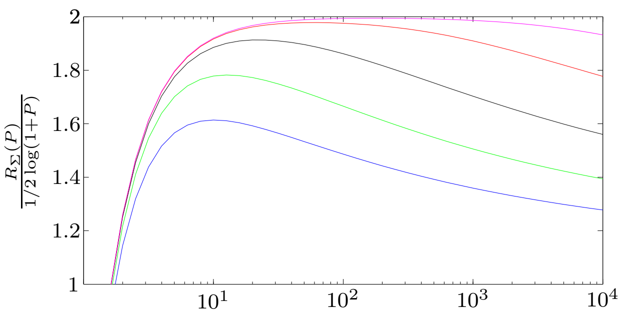

Fig. 3. The sum-rate $R_\Sigma(P)$ achieved by the scheme in Section VI-A normalized by $\frac{1}{2}\log(1+\frac{P}{\sigma_1^2})$ is plotted as a function of the power $P > 0$. The noise variances $\sigma_1^2 = \sigma_2^2 = 1$ are fixed, and the different curves correspond (in increasing order) to correlation coefficients $\rho_z = -0.85, -0.95, -0.99, -0.999 - 0.9999$.

Corollary 2 (Prelog): The prelog of the AWGN BC with noise-free feedback is 2 if $\rho_z = -1$ or if $\rho_z = +1$ and $\sigma_1^2 \neq \sigma_2^2$, otherwise it is 1.

Note 3: Our results for $\rho_z \in \{-1, 1\}$ remain valid when the transmitter has only one-sided noise-free feedback, i.e., when the transmitter for example only observes the outputs $\{Y_{1,t}\}$ but not $\{Y_{2,t}\}$. This holds because for $\rho_z \in \{-1, 1\}$ the capacity regions with one-sided and two-sided feedback coincide. In fact, in the setup with one-sided feedback the transmitter can locally compute the missing outputs $Y_{2,t}$ (or $Y_{1,t}$):

$$Y_{2,t} = \rho_z \frac{\sigma_2}{\sigma_1}(Y_{1,t} - X_t) + X_t, \quad t \in \{1, \dots, n\}.$$

Theorem 1 and Corollary 2 show that when $\rho_z \in \{-1, 1\}$ and the channel is not physically degraded, feedback approximately doubles the high-SNR sum-capacity. In Section VI-A we present a coding scheme with this desired performance. (When $\rho_z = -1$ also the Ozarow-Leung scheme [8], [9] achieves such sum-rates [10], [11].) In Figure 3 we have plotted the relationship between the sum-rate achieved by our scheme in Section VI-A and the transmitted power P for various values of the correlation ρ_z. It shows that for large powers P (i.e., $P \geq 100$), feedback can nearly double the capacity not only when the correlation ρ_z is exactly -1, but also when ρ_z is sufficiently close to -1, (i.e., when $\rho_z = -1+\epsilon$ for sufficiently small $\epsilon > 0$ depending on the power P). The same observation can be made when $\rho_z = 1 - \epsilon$ if $\sigma_1^2 \neq \sigma_2^2$.

Theorem 3 and Corollary 4 ahead explore the relationship between P and the required correlation ρ_z. Since the required correlation depends on the transmit power P, we make the dependence explicit and denote the correlation by $\rho_z(P)$. Theorem 3 and Corollary 4 thus characterize the *generalized prelog* where the channel parameters (here the noise correlation ρ_z) vary with the power P.

Let noise variances $\sigma_1^2, \sigma_2^2 > 0$ and for each power $P > 0$ a correlation coefficients $\rho_z(P)$ be given. Define

$$\zeta_{-1} \triangleq \overline{\lim_{P\to\infty}} \frac{-\log(1+\rho_z(P))}{\log(P)}, \tag{20}$$

$$\zeta_{+1} \triangleq \overline{\lim_{P\to\infty}} \frac{-\log(1-\rho_z(P))}{\log(P)}, \tag{21}$$

where recall that we defined $-\log 0 \triangleq \infty$. Notice that $\zeta_{-1} > 0$ only if $\underline{\lim}_{P\to\infty} \rho_z(P) = -1$, and $\zeta_{+1} > 0$ only if $\overline{\lim}_{P\to\infty} \rho_z(P) = 1$.

Theorem 3 (Generalized Prelog): The generalized prelog depends on whether the noise variances $\sigma_1^2, \sigma_2^2 > 0$ are equal or different. If $\sigma_1^2 = \sigma_2^2$, the generalized prelog is

$$\overline{\lim_{P\to\infty}} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z(P))}{\frac{1}{2}\log(1+P)} = \min\{1+\zeta_{-1}, 2\}. \tag{22}$$

and if $\sigma_1^2 \neq \sigma_2^2$, it is

$$\begin{aligned}\overline{\lim_{P\to\infty}} &\frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z(P))}{\frac{1}{2}\log(1+P)} \\ &= \min\Big\{1+\max\Big\{\zeta_{-1}, \zeta_{+1}\Big\}, 2\Big\}.\end{aligned} \tag{23}$$

Proof: See Section VI-C. ∎

Corollary 4 (Generalized Prelog): Let $\rho_z(P)$ be of the form

$$\rho_z(P) = \pm\left(1 - \frac{\epsilon(P)}{P^\zeta}\right), \qquad \nu \in \{1, 2\},\ \zeta \in [0, 1]$$

where

$$\lim_{P\to\infty} \frac{\log\big(\epsilon(P)\big)}{\log(P)} = 0.$$

Unless $\sigma_1^2 = \sigma_2^2$ and $\lim_{P\to\infty} \rho_z(P) = 1$,

$$\overline{\lim_{P\to\infty}} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z(P))}{\frac{1}{2}\log(1+P)} = 1+\zeta. \tag{24}$$

III. Two-User Broadcast Channel with Noisy Feedback: Setup and Results

In this section we study the AWGN BC with *noisy* feedback, which is depicted in Figure 4. The goal of the communication is the same as in the previous section. That is, the transmitter wishes to convey Message M_1 to Receiver 1 and Message M_2 to Receiver 2 by communicating over the AWGN BC described in (3). The transmitter has access to *noisy* feedback. Thus, instead of observing the channel outputs $Y_{1,t}$ and $Y_{2,t}$ as in the previous section, it observes the noisy feedback outputs

$$V_{1,t} = Y_{1,t} + W_{1,t},$$
$$V_{2,t} = Y_{2,t} + W_{2,t}.$$

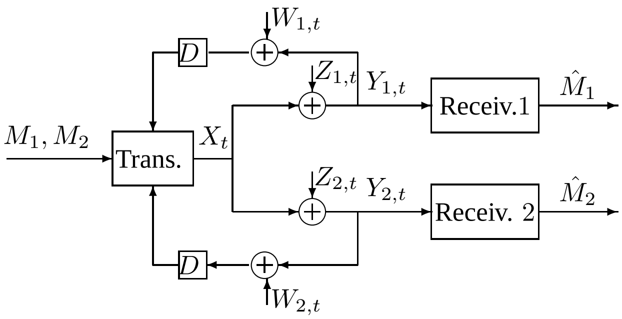

Fig. 4. The two-user AWGN BC with noisy feedback.

The feedback-noise sequences $\{(W_{1,t}, W_{2,t})\}$ are assumed to be independent of the messages (M_1, M_2) and of the noise sequences on the forward path $\{(Z_{1,t}, Z_{2,t})\}$ and IID according to a zero-mean bivariate Gaussian distribution of diagonal[3] covariance matrix $\begin{pmatrix} \sigma_{W1}^2 & 0 \\ 0 & \sigma_{W2}^2 \end{pmatrix}$. We shall assume throughout this section that the feedback noise variances are positive

$$\sigma_{W1}, \sigma_{W2} > 0. \tag{25}$$

In this setup the transmitter computes its channel inputs as

$$X_t = f_{\text{BCNoisy},t}^{(n)}\left(M_1, M_2, V_1^{t-1}, V_2^{t-1}\right), \quad t \in \{1, \ldots, n\}, \tag{26}$$

where the encoding function $f_{\text{BCNoisy},t}^{(n)}$ is of the form

$$f_{\text{BCNoisy},t}^{(n)} \colon \mathcal{M}_1 \times \mathcal{M}_2 \times \mathbb{R}^{t-1} \times \mathbb{R}^{t-1} \to \mathbb{R}. \tag{27}$$

The channel input sequence $(X_1, \ldots, X_n)$ is again subject to an expected average block-power constraint $P > 0$ as in (7).

Decoding rules, achievable rates, capacity region, sum-capacity and prelog are defined as in the previous section. The sum-capacity for this setup with noisy feedback is denoted by $C_{\text{BCNoisy},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z, \sigma_{W1}^2, \sigma_{W2}^2)$.

Noisy feedback does not increase the prelog.

Theorem 5: Irrespective of the correlation $\rho_z \in [-1, 1]$, the prelog is one:

$$\overline{\lim_{P\to\infty}} \frac{C_{\text{BCNoisy},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z, \sigma_{W1}^2, \sigma_{W2}^2)}{\frac{1}{2}\log(1+P)} = 1. \tag{28}$$

Proof: See Section VI. ∎

Thus, with noisy feedback the prelog equals 1 also when the noise correlation $\rho_z \in \{-1, 1\}$. This result assumes that the feedback-noise variances $\sigma_{W1}^2, \sigma_{W2}^2 > 0$ are fixed. If instead they tend to 0 as the power $P \to \infty$, then for $\rho_z \in \{-1, 1\}$ the (generalized) prelog may be larger than 1, depending on the speed of convergence of the limits $\sigma_{W1}^2, \sigma_{W2}^2 \to 0$. The following note examines the generalized prelog when the feedback-noise variances tend to 0 slower than $P^{-\xi}$ for any $\xi > 0$.

Note 4: Theorem 5 remains valid if the feedback-noise variances $\sigma_{W1}^2, \sigma_{W2}^2$ tend to 0 as the power $P \to \infty$, if the convergence is slower than $P^{-\xi}$ for all $\xi > 0$. More precisely, if $\sigma_{W1}^2, \sigma_{W2}^2$ depend on P in a way that

$$\overline{\lim_{P\to\infty}} \frac{-\log\left(\sigma_{W\nu}^2\right)}{\log(P)} \leq 0, \qquad \nu \in \{1, 2\},$$

then the prelog is 1, irrespective of the noise correlation $\rho_z \in [-1, 1]$.

Proof: See Section VII-B. ∎

[3]For simplicity, we do not treat setups with correlated feedback noises or setups with feedback noises that are correlated with the forward noises.

IV. K-User Broadcast Channel with Noise-Free Feedback: Setup and Main Results

We generalize the setup in Section II to $K \geq 2$ users. Thus, here, the single transmitter wishes to communicate with K receivers. The goal of the communication is that each Receiver $k \in \{1, \ldots, K\}$ learns its desired Message M_k, where the messages $M_1, \ldots, M_K$ are independent, and each M_k is uniformly distributed over $\mathcal{M}_k \triangleq \{1, \ldots, \lfloor 2^{nR_k} \rfloor\}$.

The channel model is described as follows. Given that the transmitter sends the time-t input symbol x_t, Receiver k observes the time-t output symbol

$$Y_{k,t} = x_t + Z_{k,t}, \qquad t \in \{1, \ldots, n\}, \tag{29}$$

where $\{(Z_{1,t}, \ldots, Z_{K,t})\}_{t=1}^n$ is a sequence of IID centered Gaussian vectors of covariance matrix K_z and independent of the messages $(M_1, \ldots, M_K)$.

The transmitter has noise-free feedback from all channel outputs, and thus can compute its channel inputs as

$$X_t = f_{\text{K-BC},t}^{(n)}(M_1, \ldots, M_K, Y_1^{t-1}, \ldots, Y_K^{t-1})$$

for some encoding function $f_{\text{K-BC},t}^{(n)}$, such that the input sequence $(X_1, \ldots, X_n)$ satisfies the expected average block-power constraint in (7).

An error occurs in the communication whenever

$$(M_1, \ldots, M_K) \neq (\hat{M}_1, \ldots, \hat{M}_K). \tag{30}$$

Achievable rate-tuples, capacity region, and sum-capacity are defined in analogy to the setup with two users. The sum-capacity of this setup is denoted by $C_{\text{K-BC},\Sigma}(P, \mathsf{K}_z)$. We are interested in its prelog.

Theorem 6: If the covariance matrix K_z has rank 1, then the prelog is

$$\overline{\lim_{P\to\infty}} \frac{C_{\text{K-BC},\Sigma}(P, \mathsf{K}_z)}{\frac{1}{2}\log(1+P)} = n_\alpha. \tag{31}$$

where n_α denotes the number of rows in K_z that are different.

Proof: See Section VIII-B. ∎

Corollary 7: When K_z is of rank 1 and all its rows are different, the prelog is K.

The achievability of prelog $K \geq 3$ was first proved in [12] for the *complex* memoryless Gaussian broadcast channel. The prelog ≥ 3 result in [12] however requires that the real and imaginary parts of the noise symbols are correlated, and thus does not include our result in Theorem 6 as a special case.

Note 5: Theorem 6 remains valid also when the transmitter has feedback only from a single receiver. This can be seen by analogy to Note 3.

V. Two-User Interference Channel with Noise-Free Feedback: Setup and Main Results

In this section we consider the real scalar AWGN IC with noise-free feedback, which has two transmitters. Transmitter 1 wishes to send Message M_1 to Receiver 1, and Transmitter 2 Message M_2 to Receiver 2. Assuming that at time t Transmitter 1 sends the real symbol $x_{1,t}$ and Transmitter 2 sends the real symbol $x_{2,t}$, Receiver 1 observes

$$Y_{1,t} \quad = \quad a_{1,1}x_{1,t} + a_{1,2}x_{2,t} + Z_{1,t}, \tag{32a}$$

and Receiver 2 observes

$$Y_{2,t} \quad = \quad a_{2,1}x_{1,t} + a_{2,2}x_{2,t} + Z_{2,t}. \tag{32b}$$

The "channel gains"

$$a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2} \neq 0, \tag{33}$$

are non-zero real constants and the noise sequences $\{(Z_{1,t}, Z_{2,t})\}_{t=1}^n$ are defined as in Section II.

Each transmitter has access to noise-free feedback from its corresponding receiver. Thus, each of the two transmitters can compute its time-t channel input as

$$X_{k,t} \quad = \quad f_{\text{IC},k,t}^{(n)}\left(M_k, Y_k^{t-1}\right), \quad k \in \{1,2\},$$

for some encoding function $f_{\text{IC},k,t}^{(n)}$ of the form

$$f_{\text{IC},k,t}^{(n)}\colon \mathcal{M}_k \times \mathbb{R}^{t-1} \to \mathbb{R}, \quad \nu \in \{1,2\}.$$

The two channel input sequences are subject to the same average block-power constraint $P > 0$:

$$\frac{1}{n}\mathsf{E}\left[\sum_{t=1}^{n} X_{k,t}^2\right] \leq P, \quad k \in \{1,2\}. \tag{34}$$

Decoding rules, achievable rate pairs, capacity region, sum-capacity, and prelog are defined as for the AWGN BC. We denote the sum-capacity of the symmetric AWGN IC by $C_{\text{IC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)$.

Without feedback, the prelog of the AWGN IC equals 1; with noise-free feedback it can be 2, depending on the channel gains $a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2} \neq 0$ and on the noise parameters $\sigma_1^2, \sigma_2^2, \rho_z$.

Theorem 8: The prelog of the AWGN IC with noise-free feedback satisfies the following three statements.

- If $\rho_z \in (-1,1)$ or if $\rho_z \in \{-1,1\}$ and $\frac{a_{2,2}}{a_{1,2}} = \frac{a_{1,2}}{a_{1,1}} = \rho_z \frac{\sigma_2}{\sigma_1}$,
$$\overline{\lim_{P\to\infty}} \frac{C_{\text{IC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1+P)} = 1; \tag{35}$$
- if $\rho_z \in \{-1,1\}$ and moreover, both $\frac{a_{2,2}}{a_{1,2}}$ and $\frac{a_{1,2}}{a_{1,1}}$ are different from $\rho_z \frac{\sigma_2}{\sigma_1}$, then
$$\overline{\lim_{P\to\infty}} \frac{C_{\text{IC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1+P)} = 2; \tag{36}$$
- otherwise
$$1 \leq \overline{\lim_{P\to\infty}} \frac{C_{\text{IC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1+P)} \leq 2. \tag{37}$$

Proof: See Section IX-B. ■

VI. A Scheme and Proofs for the AWGN BC with Noise-Free Feedback

A. A coding scheme

We present a new coding scheme for the $K = 2$ user Gaussian BC with noise-free feedback, see also [7]. (The scheme is generalized to $K \geq 2$ users in Section VIII-A.) Our scheme is similar to the schemes proposed by Cover and Pombra [4] for (non-white) Gaussian point-to-point channels with noise-free feedback, by Lapidoth and Wigger [5] for the two-user AWGN multi-access channel with noisy feedback, and by Lapidoth, Steinberg, and Wigger [6] for the two-user AWGN BC with one-sided noise-free feedback. We also propose a choice of parameters for our scheme. Despite being suboptimal (see [7] for improved choices), this choice suffices to prove the achievability results in Theorems 1 and 3 when the channel is not physically degraded.

Our scheme has the following parameters: the positive integer number η; the η-by-η strictly lower-triangular matrices $\mathsf{B}_1, \mathsf{B}_2$; the η-dimensional column-vectors $\mathbf{u}_1, \mathbf{u}_2$; and the η-dimensional row-vectors $\mathbf{v}_1, \mathbf{v}_2$.

1) Code Construction: Let the block-length n of our scheme be a multiple of η, i.e., $n = \eta n'$ for some positive integer number n'.

We generate two codebooks $\{\mathcal{C}_k\}_{k=1}^2$, where $\mathcal{C}_k$ contains $\lfloor 2^{nR_k} \rfloor$ codewords of length n'. To generate the codebooks we draw each entry of each codeword randomly and independently according to a standard Gaussian distribution. Both codebooks are revealed to the transmitter, and each codebook $\mathcal{C}_k$ to the corresponding receiver k.

2) Encoding: For $k \in \{1,2\}$, let $\mathbf{\Xi}_k(M_k)$ denote the codeword in $\mathcal{C}_k$ that corresponds to Message M_k, and let $\Xi_{k,i}$ be its i-th symbol.

Also, for each $i \in \{1, \ldots, n'\}$ let $\mathbf{X}_i$ denote the η-length column vector obtained by stacking the η input symbols $X_{(i-1)\eta+1}, \ldots, X_{(i\eta)}$ on top of each other:

$$\mathbf{X}_i \triangleq (X_{(i-1)\eta+1}, \ldots, X_{(i\eta)})^{\mathsf{T}}, \tag{38}$$

and for each $k \in \{1,2\}$, let the η-length vectors $\mathbf{Y}_{k,i}$ and $\mathbf{Z}_{k,i}$ be defined similarly.

The encoding procedure is as follows. The transmitter picks from each codebook $\mathcal{C}_k$, for $k \in \{1,2\}$, the codeword that corresponds to the message M_k, and then sends in each subblock $i \in \{1, \ldots, n'\}$ a linear combination of the i-th symbols of these codewords and the subblock's past noise-symbols[4]

$$\mathbf{X}_i = \Xi_{1,i}\mathbf{u}_1 + \Xi_{2,i}\mathbf{u}_2 + \mathsf{B}_1\mathbf{Z}_{1,i} + \mathsf{B}_2\mathbf{Z}_{2,i}. \tag{39}$$

The strict lower-triangularity of the matrices B_1 and B_2 assures that the feedback is used only causally.

The inputs satisfy the average block-power constraint (7) whenever Inequality (40) on top of the next page is satisfied.

[4]The transmitter can compute all the past noise symbols because, through the feedback, it learns the past channel outputs and because it also knows the past channel inputs.

$$\|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + \mathrm{tr}\left(\mathsf{B}_1\mathsf{B}_1^\intercal\right)\sigma_1^2 + \mathrm{tr}\left(\mathsf{B}_2\mathsf{B}_2^\intercal\right)\sigma_2^2 + 2\,\mathrm{tr}\left(\mathsf{B}_1\mathsf{B}_2^\intercal\right)\rho_z\sigma_1\sigma_2 \le \eta P. \tag{40}$$

3) Decoding: In subblock $i \in \{1,\dots,n'\}$ Receiver 1 observes

$$\mathbf{Y}_{1,i} = \Xi_{1,i}\mathbf{u}_1 + \Xi_{2,i}\mathbf{u}_2 + (\mathsf{B}_1 + \mathsf{I})\,\mathbf{Z}_{1,i} + \mathsf{B}_2\mathbf{Z}_{2,i}, \tag{41a}$$

and Receiver 2 observes

$$\mathbf{Y}_{2,i} = \Xi_{1,i}\mathbf{u}_1 + \Xi_{2,i}\mathbf{u}_2 + (\mathsf{B}_2 + \mathsf{I})\,\mathbf{Z}_{2,i} + \mathsf{B}_1\mathbf{Z}_{1,i}. \tag{41b}$$

To decode its desired message M_k, each Receiver $k \in \{1,2\}$ forms the "new outputs"

$$I_{k,i} \triangleq \mathbf{v}_k\mathbf{Y}_{k,i}, \qquad i \in \{1,\dots,n'\}, \tag{42}$$

and optimally decodes M_k based on $(I_{k,1},\dots,I_{k,n'})$, e.g., using a maximum-likelihood decoder.

4) Choice of Parameters: We describe a choice of the parameters $\mathsf{B}_1, \mathsf{B}_2$, $\mathbf{u}_1, \mathbf{u}_2$, $\mathbf{v}_1, \mathbf{v}_2$ for every positive integer η, and then let $\eta \to \infty$.

Choose $q > 0$ and $\delta \notin \{-1, 0\}$ to satisfy Equation (43) shown on top of the next page, and define

$$a_1 \triangleq q \tag{44}$$
$$a_2 \triangleq -\delta^2 q \tag{45}$$
$$b_1 \triangleq -\delta(1+\delta)q^2 \tag{46}$$
$$b_2 \triangleq -\delta^2(1+\delta)q^2. \tag{47}$$

For a given positive integer η we choose the $\eta\times\eta$ matrices B_1 and B_2 Toeplitz with non-zero entries only on the first and second diagonals below the main diagonal:

$$\mathsf{B}_\mathsf{k} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ a_k & 0 & 0 & \cdots & 0 & 0 & 0 \\ b_k & a_k & 0 & \cdots & 0 & 0 & 0 \\ 0 & b_k & a_k & 0 & \cdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & b_k & a_k & 0 & 0 \\ 0 & \cdots & \cdots & 0 & b_k & a_k & 0 \end{pmatrix}, \tag{48}$$

and we choose the η-dimensional vectors

$$\mathbf{u}_1 = \sqrt{\frac{P}{2+2\frac{b_1^2}{a_1^2}}}\begin{pmatrix}1 & \frac{b_1}{a_1} & 0 & \dots & 0\end{pmatrix}^\intercal \tag{49a}$$

$$\mathbf{u}_2 = \sqrt{\frac{P}{2+2\frac{b_2^2}{a_2^2}}}\begin{pmatrix}1 & \frac{b_2}{a_2} & 0 & \dots & 0\end{pmatrix}^\intercal \tag{49b}$$

$$\mathbf{v}_1 = \begin{pmatrix}\left(-\frac{b_2}{a_2}\right)^{\eta-1} & \left(-\frac{b_2}{a_2}\right)^{\eta-2} & \dots & -\frac{b_2}{a_2} & 1\end{pmatrix} \tag{49c}$$

$$\mathbf{v}_2 = \begin{pmatrix}\left(-\frac{b_1}{a_1}\right)^{\eta-1} & \left(-\frac{b_1}{a_1}\right)^{\eta-2} & \dots & -\frac{b_1}{a_1} & 1\end{pmatrix}. \tag{49d}$$

By (43)–(48), for every η this choice satisfies the power constraint (40). Moreover, the vector $\mathbf{v}_1$ is orthogonal to the first $\eta-2$ columns of the matrices $(\mathsf{B}_1+\mathsf{I})$ and B_2, and to the vector $\mathbf{u}_2$, but not to $\mathbf{u}_1$. Similarly, $\mathbf{v}_2$ is orthogonal to the first $\eta-2$ columns of the matrices B_1 and $(\mathsf{B}_2+\mathsf{I})$, and to the vector $\mathbf{u}_1$, but not to $\mathbf{u}_2$. Therefore,

$$I_{1,i} = \sqrt{\frac{P}{2+2\frac{b_1^2}{a_1^2}}}\left(-\frac{b_2}{a_2}\right)^{\eta-1}\left(1 - \frac{a_2}{b_2}\frac{b_1}{a_1}\right)\Xi_{1,i} + \left(-\frac{b_2}{a_2} + a_1\right)Z_{1,i\eta-1} + Z_{1,i\eta} + a_2 Z_{2,i\eta-1} \tag{50a}$$

$$I_{2,i} = \sqrt{\frac{P}{2+2\frac{b_2^2}{a_2^2}}}\left(-\frac{b_1}{a_1}\right)^{\eta-1}\left(1 - \frac{a_1}{b_1}\frac{b_2}{a_2}\right)\Xi_{2,i} + \left(-\frac{b_1}{a_1} + a_2\right)Z_{2,i\eta-1} + Z_{2,i\eta} + a_1 Z_{1,i\eta-1}. \tag{50b}$$

We see that with our choice of parameters the last $\eta-2$ channel uses in each subblock are used to iteratively cancel the previous noise symbols in this subblock. Therefore, the "new output symbols" $I_{1,i}$ or $I_{2,i}$ depend only on the noise symbols $Z_{1,i\eta-1}, Z_{1,i\eta}, Z_{2,i\eta-1}, Z_{2,i\eta}$ but not on the previous noise symbols. Also, Receiver 1's "new outputs" $I_{1,1},\dots,I_{1,n'}$ do not depend on the non-desired codeword $\Xi_{2,1},\dots,\Xi_{2,n'}$, and thus they can be viewed as the outputs of a point-to-point channel where the transmitter only sends the codeword $\Xi_{1,1},\dots,\Xi_{1,n'}$, but not $\Xi_{2,1},\dots,\Xi_{2,n'}$. Similarly for Receiver 2's "new outputs".

5) Achievable Rates: By (44)–(49d) and (50) and by standard arguments, our scheme achieves all rate pairs (R_1, R_2) that satisfy

$$R_1 \le \frac{1}{2\eta}\log\left(1 + \frac{\frac{P(1+\delta)^2}{2+2q^2\delta^2(1+\delta)^2}\left(q^2(1+\delta)^2\right)^{\eta-1}}{(q^2\delta^2+1)\sigma_1^2 + q^2\delta^4\sigma_2^2}\right) \tag{51a}$$

$$R_2 \le \frac{1}{2\eta}\log\left(1 + \frac{\frac{P(1+1/\delta^2)^2}{2+2q^2(1+\delta)^2}\left(q^2\delta^2(1+\delta)^2\right)^{\eta-1}}{(q^2\delta^2+1)\sigma_2^2 + q^2\sigma_1^2}\right). \tag{51b}$$

We have the following lemma.

Lemma 1: Let ξ, ζ be positive real numbers. If $1+\zeta \ge \xi$, then the mapping $\eta \in \mathbb{Z}^+ \mapsto \frac{1}{2\eta}\log(1+\xi^{\eta-1}\zeta)$, has its maximum at $\eta = 1$; otherwise it has its supremum at $\eta \to \infty$.

Proof: See Appendix A. ■

By Lemma 1, $\eta = 1$ maximizes the constraints in (51) for small powers P and $\eta \to \infty$ maximizes them for large powers P. Letting $\eta \to \infty$, we conclude the following.

Proposition 9: All nonnegative rate-pairs (R_1, R_2) that satisfy

$$R_1 \le \frac{1}{2}\log^+\left(q^2(1+\delta)^2\right) \tag{52a}$$

$$R_2 \le \frac{1}{2}\log^+\left(q^2\delta^2(1+\delta)^2\right), \tag{52b}$$

for some $\delta \notin \{-1, 0\}$ and q such that (43) holds, are achievable.

$$q^2(\sigma_1^2+\delta^4\sigma_2^2-2\delta^2\rho_z\sigma_1\sigma_2)+q^4(1+\delta)^2\delta^2(\sigma_1^2+\delta^2\sigma_2^2+2\delta\rho_z\sigma_1\sigma_2)\leq P. \tag{43}$$

When $\rho_z(P)\in(-1,1)$, the choice

$$\delta=\frac{\sigma_1}{\sigma_2}\cdot\frac{\sigma_1-\rho_z\sigma_2}{\sigma_2-\rho_z\sigma_1} \tag{53a}$$

$$q^2=\sqrt{\frac{(1-\epsilon)P}{\delta^2(1+\delta)^2(\sigma_1^2+\delta^2\sigma_2^2+2\delta\rho_z\sigma_1\sigma_2)}} \tag{53b}$$

satisfies the power constraint (43) for all $\epsilon>0$ and all powers P exceeding some threshold $P_0(\epsilon,\sigma_1^2,\sigma_2^2,\rho_z)$. Plugging this choice into the achievable rates in Proposition 9 results in the following corollary.

Corollary 10: For $\rho_z\in(-1,1)$ and every $\epsilon\in(0,1)$ there exists a positive real number $P_0(\epsilon,\sigma_1^2,\sigma_2^2,\rho_z)$ such that the sum-rate

$$R_1+R_2=\frac{1}{2}\log^+\left(\frac{(1-\epsilon)P(\sigma_1^2+\sigma_2^2-2\rho_z\sigma_1\sigma_2)}{\sigma_1^2\sigma_2^2(1-\rho_z^2)}\right) \tag{54}$$

is achievable for all powers $P>P_0(\epsilon,\sigma_1^2,\sigma_2^2,\rho_z)$.

When $|\rho_z|=1$, the choice

$$\delta=-\rho_z\frac{\sigma_1}{\sigma_2} \tag{55a}$$

$$q=\sqrt{\frac{P}{\sigma_1^2\left(1-\frac{\sigma_1}{\sigma_2}\rho_z\right)^2}} \tag{55b}$$

satisfies (43), and the matrices B_1 and B_2 defined through (44)–(48) and (55) satisfy

$$\mathsf{B}_1\sigma_1+\mathsf{B}_2\rho_z\sigma_2=\mathsf{B}, \tag{56}$$

where B is a Toeplitz matrix with all zero entries except on the first diagonal below the main diagonal where the entries equal $\sqrt{P}$:

$$\mathsf{B}=\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0\\ \sqrt{P} & 0 & 0 & \cdots & 0 & 0\\ 0 & \sqrt{P} & 0 & \cdots & 0 & 0\\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots\\ 0 & \cdots & 0 & \sqrt{P} & 0 & 0\\ 0 & \cdots & 0 & 0 & \sqrt{P} & 0 \end{pmatrix}. \tag{57}$$

Since $\rho_z\in\{-1,1\}$ this implies that—similarly to the scheme for the AWGN point-to-point channel with feedback sketched in Note 7 ahead—in each channel use the transmitter only sends the previous noise symbol, but not the ones before. This is unlike our scheme for $\rho_z\in(-1,1)$ where the transmitter sends the *two* previous noise symbols.

Specializing the rate constraints in (51) to the choice in (55) results in

$$R_1\leq\frac{1}{2\eta}\log\left(1+\left(\frac{P}{\sigma_1^2}\right)^{\eta-1}\left(1-\rho_z\frac{\sigma_1}{\sigma_2}\right)^2\frac{P/\sigma_1^2}{2+2P/\sigma_2^2}\right) \tag{58a}$$

$$R_2\leq\frac{1}{2\eta}\log\left(1+\left(\frac{P}{\sigma_2^2}\right)^{\eta-1}\left(1-\rho_z\frac{\sigma_2}{\sigma_1}\right)^2\frac{P/\sigma_2^2}{2+2P/\sigma_1^2}\right). \tag{58b}$$

Letting $\eta\to\infty$ we thus obtain the following corollary for $|\rho_z|=1$.

Corollary 11: If $|\rho_z|=1$, and the channel is not physically degraded, i.e., $\rho_z\notin\left\{\frac{\sigma_1}{\sigma_2},\frac{\sigma_2}{\sigma_1}\right\}$, then all nonnegative rate-pairs (R_1,R_2) that satisfy

$$R_1\leq\frac{1}{2}\log^+\left(\frac{P}{\sigma_1^2}\right) \tag{59a}$$

$$R_2\leq\frac{1}{2}\log^+\left(\frac{P}{\sigma_2^2}\right) \tag{59b}$$

are achievable.

Note 6: For $|\rho_z|=1$ and when the channel is not physically degraded our scheme achieves prelog 2 when $\eta\to\infty$. For finite η our scheme achieves prelog $2\frac{\eta-1}{\eta}$, see (58). Thus, $\eta=3$ suffices to increase the prelog compared to the non-feedback setup.

Note 7: To gain intuition about the scheme and our choice of parameters, it is instructive to consider the following scheme for the AWGN point-to-point channel with feedback. We consider an AWGN channel with noise variance σ^2 and power constraint P. The scheme is obtained by specializing our BC-scheme to $K=1$. We describe our choice of parameters through an explicit description of the encoding and decoding strategies.

Let $\Xi_1,\ldots,\Xi_{n'}$ denote the symbols of the single codeword chosen by the transmitter. In the first channel use of each subblock $i\in\{1,\ldots,n'\}$ the transmitter sends the codeword symbol Ξ_i scaled by $\sqrt{P}$, and in all other channel uses of the subblock it sends the previous noise symbol scaled by $\sqrt{\frac{P}{\sigma^2}}$:

$$X_{(i-1)\eta+1}=\sqrt{P}\Xi_i \tag{60}$$

$$X_{(i-1)\eta+2}=\sqrt{\frac{P}{\sigma^2}}Z_{(i-1)\eta+1} \tag{61}$$

$$\vdots \quad = \quad \vdots$$

$$X_{i\eta}=\sqrt{\frac{P}{\sigma^2}}Z_{i\eta-1} \tag{62}$$

where Z_t denotes the Gaussian noise symbol corrupting the output at time t. The receiver uses the last $\eta-1$ channel outputs of each subblock to iteratively cancel the noise corrupting the first output in the subblock

$$I_i\triangleq\sum_{\ell=1}^{\eta}\sqrt{\frac{P}{\sigma^2}}^{\eta-\ell}Y_{(i-1)\eta+\ell}=\sqrt{P}\sqrt{\frac{P}{\sigma^2}}^{\eta-1}\Xi_i+Z_{i\eta}, \tag{63}$$

and then decodes its desired message M by applying a maximum-likelihood decoder to the resulting sequence $I_1,\ldots,I_{n'}$. Letting $\eta\to\infty$, the described scheme achieves all rates satisfying

$$0\leq R\leq\frac{1}{2}\log^+\left(\frac{P}{\sigma^2}\right), \tag{64}$$

and is thus optimal in the asymptotic high-SNR regime.

Notice that for finite η the performance of the scheme could be improved, if instead of as in (63) the symbol I_i is chosen as the linear minimum-mean square estimate (LMMSE) of Ξ_i given $\{Y_{(i-1)\eta+\ell}\}_{\ell=1}^{\eta}$. However, in the limit $\eta \to \infty$ both choices achieve the same rates.

Note 8: The simple scheme in Note 7 illustrates that with feedback one can achieve a rate $R = \frac{1}{2}\log^+\left(\frac{P}{\sigma^2}\right)$ over an AWGN point-to-point channel by sending most of the time—in fact, in all channel uses except for the first channel use of each subblock—the previous noise-symbol scaled to satisfy the power constraint P.

Now, in our scheme for the AWGN BC with noise correlation $|\rho_z| = 1$ we apply the same idea with the only difference that we wish to convey two different codewords (one for each receiver) and that we send their corresponding symbols over the first two channel uses of each subblock. Sending the codesymbols during two channel uses allows us to choose their signaling directions ($\mathbf{u}_1$ and $\mathbf{u}_2$, respectively) orthogonal to the beamforming direction at their non-intended receiver ($\mathbf{v}_2$ and $\mathbf{v}_1$, respectively). When the channel is not physically degraded, the beamforming directions at the two receivers are different, and each receiver observes essentially the same channel outputs as if the transmitter was using the point-to-point scheme in Note 7. Therefore, in this case, our scheme achieves the rate pair $R_1 = \frac{1}{2}\log^+\left(\frac{P}{\sigma_1^2}\right)$, $R_2 = \frac{1}{2}\log^+\left(\frac{P}{\sigma_2^2}\right)$ (see Corollary 11).

The special case $|\rho_z| = 1$ illustrates very nicely that the intuition put forth in [23], [9], [24], [25] applies also to our scheme: feedback is helpful for a memoryless BC because it allows the transmitter to identify and transmit information that at the same time is useful for both receivers, and therefore the BC can be used in a more efficient way.

B. Proof of Theorem 1

For physically degraded channels where $\rho_z \in \left\{\frac{\sigma_2}{\sigma_1}, \frac{\sigma_1}{\sigma_2}\right\}$ the result is proved in [19] and [20].

When $|\rho_z| = 1$ and $\rho_z \notin \left\{\frac{\sigma_2}{\sigma_1}, \frac{\sigma_1}{\sigma_2}\right\}$, the achievability follows from Corollary 11, and the converse by applying the cutset bound with two individual cuts between the transmitter and each of the two receivers:

$$\begin{aligned} R_1 + R_2 &\le \max_{X:\, \mathsf{E}[X^2]\le P} \{I(X;Y_1) + I(X;Y_2)\} \\ &= \frac{1}{2}\log\left(1 + \frac{P}{\sigma_1^2}\right) + \frac{1}{2}\log\left(1 + \frac{P}{\sigma_2^2}\right), \end{aligned} \quad (65)$$

where the equality follows because a Gaussian law maximizes the differential entropy under a variance constraint [22].

When $\rho_z \in (-1,1)$ and $\rho_z \notin \left\{\frac{\sigma_2}{\sigma_1}, \frac{\sigma_1}{\sigma_2}\right\}$, the achievability follows from Corollary 10 and the converse by applying the cutset bound with a single cut between the transmitter and both receivers:

$$\begin{aligned} R_1 + R_2 &\le \max_{X:\, \mathsf{E}[X^2]\le P} I(X;Y_1,Y_2) \\ &\le \frac{1}{2}\log\left(1 + \frac{P(\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_z)}{\sigma_1^2\sigma_2^2(1-\rho_z^2)}\right). \end{aligned} \quad (66)$$

C. Proof of Theorem 3

Recall that here $\rho_z(P)$ depends on the power P.

We first prove the converse to (22) where $\sigma_1^2 = \sigma_2^2$. In this case, Upper bound (66) specializes to

$$\begin{aligned} &C_{\text{BC},\Sigma}(P,\sigma_1^2,\sigma_1^2,\rho_z(P)) \\ &\quad \le \frac{1}{2}\log\left(1 + \frac{P}{\frac{\sigma_1^2}{2}(1+\rho_z(P))}\right), \quad \rho_z(P) \in (-1,1). \end{aligned} \quad (67)$$

In view of (14) and since we defined $-\log(0) = \infty$, Upper bound (67) holds also for $\rho_z(P) \in \{-1,1\}$, and thus for all $\rho_z(P) \in [-1,1]$. Therefore, by the definition of ζ_{-1} in (20),

$$\varlimsup_{P\to\infty} \frac{C_{\text{BC},\Sigma}(P,\sigma_1^2,\sigma_1^2,\rho_z(P))}{\frac{1}{2}\log(1+P)} \le 1 + \zeta_{-1}. \quad (68)$$

On the other hand, by (65), irrespective of $\{\rho_z(P)\}_{\{P>0\}}$,

$$\varlimsup_{P\to\infty} \frac{C_{\text{BC},\Sigma}(P,\sigma_1^2,\sigma_1^2,\rho_z(P))}{\frac{1}{2}\log(1+P)} \le 2. \quad (69)$$

Combining (68) and (69) establishes the converse to (22).

We now prove the converse to (23) where $\sigma_1^2 \neq \sigma_2^2$. By Upper bound (66) and since $\sigma_1^2+\sigma_2^2-2\rho_z(P)\sigma_1\sigma_2 < 2(\sigma_1^2+\sigma_2^2)$ and $1-\rho_z^2(P) \ge 1-|\rho_z(P)|$ for all $\rho_z(P) \in (-1,1)$,

$$\begin{aligned} &C_{\text{BC},\Sigma}(P,\sigma_1^2,\sigma_2^2,\rho_z(P)) \\ &\le \frac{1}{2}\log\left(1 + \frac{P}{\frac{\sigma_1^2\sigma_2^2}{2(\sigma_1^2+\sigma_2^2)}(1-|\rho_z(P)|)}\right), \quad \rho_z(P) \in (-1,1). \end{aligned} \quad (70)$$

Now, since we defined $-\log(0) = \infty$, Upper bound (70) holds also for $\rho_z(P) \in \{-1,1\}$, and hence for all $\rho_z(P) \in [-1,1]$. Moreover, since by the definitions of ζ_{-1} and ζ_{+1} in (20) and (21),

$$\varlimsup_{P\to\infty} \frac{-\log(1-|\rho_z(P)|)}{\log(P)} = \max\{\zeta_{-1},\zeta_{+1}\}, \quad (71)$$

we conclude that

$$\varlimsup_{P\to\infty} \frac{C_{\text{BC},\Sigma}(P,\sigma_1^2,\sigma_2^2,\rho_z(P))}{\frac{1}{2}\log(1+P)} \le 1 + \max\{\zeta_{-1},\zeta_{+1}\}. \quad (72)$$

Combining (72) with (69) establishes the converse to (23).

We next prove that for arbitrary σ_1^2, σ_2^2:

$$\varlimsup_{P\to\infty} \frac{C_{\text{BC},\Sigma}(P,\sigma_1^2,\sigma_2^2,\rho_z(P))}{\frac{1}{2}\log(1+P)} \ge \min\{1+\zeta_{-1}, 2\}. \quad (73)$$

Since a generalized prelog of 1 is achievable even without feedback [20], [21] the interesting case is $\zeta_{-1} > 0$. In the following, assume that $\zeta_{-1} > 0$, which implies the existence of an increasing unbounded sequence $\{P_\ell\}_{\ell=1}^{\infty}$ such that

$$\lim_{\ell\to\infty} \frac{-\log(1+\rho_z(P_\ell))}{\frac{1}{2}\log(P_\ell)} = \zeta_{-1} > 0, \quad (74)$$

and in particular

$$\lim_{\ell\to\infty} \rho_z(P_\ell) = -1. \quad (75)$$

For each ℓ we choose parameters q_ℓ and δ_ℓ and show that Inequality (73) follows from Proposition 9 specialized to these parameters. Let $\epsilon \in (0,1)$ be a small positive number, and define the limit (not necessarily finite)

$$\kappa \triangleq \overline{\lim_{\ell\to\infty}} P_\ell\big(\sigma_1^2+\delta_\ell^2\sigma_2^2+2\delta_\ell\rho_z(P_\ell)\sigma_1\sigma_2\big). \tag{76}$$

We choose

$$\delta_\ell = \begin{cases} \frac{\sigma_1}{\sigma_2}\cdot\frac{\sigma_1-\rho_z(P_\ell)\sigma_2}{\sigma_2-\rho_z(P_\ell)\sigma_1} & \text{if } \rho_z(P_\ell)\in(-1,1)\\ -\rho_z(P_\ell)\frac{\sigma_1}{\sigma_2} & \text{if } \rho_z(P_\ell)\in\{-1,1\},\end{cases} \tag{77}$$

and depending on κ, we choose $q_\ell > 0$ as follows.

- If $\kappa=\infty$, we choose $q_\ell>0$ such that

$$q_\ell^2 = \sqrt{\frac{(1-\epsilon)P_\ell}{\delta_\ell^2(1+\delta_\ell)^2\left(\sigma_1^2+\sigma_2^2\delta_\ell^2+2\rho_z(P_\ell)\delta_\ell\sigma_1\sigma_2\right)}}; \tag{78}$$

- if $\kappa\in[0,\infty)$, we choose $q_\ell>0$ such that

$$q_\ell^2=\beta(1-\epsilon)P_\ell \tag{79}$$

where $\beta>0$ is a solution to

$$\sigma_1^2\left(1+\frac{\sigma_1}{\sigma_2}\right)^2\beta\left(1+\beta\frac{\kappa}{\sigma_2^2}\right)=1. \tag{80}$$

Notice that for any $\epsilon>0$ there exists a positive integer $\ell_0(\epsilon,\sigma_1^2,\sigma_2^2,\kappa)$ such that our choice (δ_ℓ,q_ℓ) satisfies the power constraint (43) for all $\ell>\ell_0(\epsilon,\sigma_1^2,\sigma_2^2,\kappa)$.

Moreover, if $\kappa\in[0,\infty)$, then specializing the rates in Proposition 9 to the choices in (77) and (79) proves that

$$\overline{\lim_{\ell\to\infty}}\frac{C_{\text{BC},\Sigma}(P_\ell,\sigma_1^2,\sigma_2^2,\rho_z(P_\ell))}{\frac{1}{2}\log(1+P_\ell)}\geq 2. \tag{81}$$

If $\kappa=\infty$, then for all sufficiently large ℓ the correlation coefficient $\rho_z(P_\ell)\in(-1,1)$, and by (77) the choice in (78) evaluates to

$$q_\ell^2=\sqrt{\frac{(\sigma_2-\sigma_1\rho_z(P_\ell))^2}{\delta_\ell^2(1+\delta_\ell)^2\sigma_1^2\left(\sigma_2^2+\sigma_1^2-2\rho_z(P_\ell)\sigma_1\sigma_2\right)(1+|\rho_z(P_\ell)|)}}\cdot\sqrt{\frac{(1-\epsilon)P_\ell}{1-|\rho_z(P)|}} \tag{82}$$

Notice that by (75)

$$\lim_{\ell\to\infty}\frac{\log\left(\frac{(\sigma_2-\sigma_1\rho_z(P_\ell))^2}{\delta_\ell^2(1+\delta_\ell)^2\sigma_1^2\left(\sigma_2^2+\sigma_1^2-2\rho_z(P_\ell)\sigma_1\sigma_2\right)(1+|\rho_z(P_\ell)|)}\right)}{\frac{1}{2}\log(P)}=0, \tag{83}$$

and therefore, specializing the rates in Proposition 9 to the choice in (77) and (82) proves that

$$\overline{\lim_{\ell\to\infty}}\frac{C_{\text{BC},\Sigma}(P_\ell,\sigma_1^2,\sigma_2^2,\rho_z(P_\ell))}{\frac{1}{2}\log(1+P_\ell)}\geq 1+\zeta_{-1}. \tag{84}$$

Combining (84) with (81) establishes (73).

In a similar way we can also prove that when $\sigma_1^2\neq\sigma_2^2$

$$\overline{\lim_{P\to\infty}}\frac{C_{\text{BC},\Sigma}(P,\sigma_1^2,\sigma_2^2,\rho_z(P))}{\frac{1}{2}\log(1+P)}\geq\min\{1+\zeta_{+1},2\}. \tag{85}$$

To this end, it suffices if in the proof to (73) we replace ζ_{-1} by ζ_{+1}, (74) by

$$\lim_{\ell\to\infty}\frac{-\log(1-\rho_z(P_\ell))}{\frac{1}{2}\log(P)}=\zeta_{+1}>0, \tag{86}$$

(75) by $\lim_{\ell\to\infty}\rho_z(P)=1$, and (80) by

$$\sigma_1^2\left(1-\frac{\sigma_1}{\sigma_2}\right)^2\beta\left(1+\beta\frac{\kappa}{\sigma_2^2}\right)=1. \tag{87}$$

The assumption of non-equal noise variances $\sigma_1^2\neq\sigma_2^2$ is needed here to conclude that (83) holds and that (87) has a finite solution for β.

Combining finally (85) with (73) establishes the achievability of (22) and (23) and concludes the proof.

VII. Proofs for the AWGN BC with Noisy Feedback

A. Proof of Theorem 5

The interesting part is the converse result, which we prove using a genie-argument similar to [13].

The following three steps establish the desired converse result. 1.) We introduce a *genie-aided AWGN BC* without feedback and show that its sum-capacity upper bounds the sum-capacity of the original AWGN BC with noisy feedback. 2.) We introduce a *less noisy AWGN BC* with neither genie-information nor feedback and show that its sum-capacity coincides with the sum-capacity of the genie-aided AWGN BC. 3.) We show that the prelog of the less noisy AWGN BC equals 1, irrespective of the noise variances $\sigma_1^2,\sigma_2^2,\sigma_{W1}^2,\sigma_{W2}^2>0$ and the correlation coefficient $\rho_z\in[-1,1]$.

We elaborate on these three steps starting with the first one. The genie-aided AWGN BC is defined as the original AWGN BC *without feedback*, but with a genie that prior to transmission reveals the sequences $\{(Z_{1,t}+W_{1,t})\}_{t=1}^n$ and $\{(Z_{2,t}+W_{2,t})\}_{t=1}^n$ to the transmitter and both receivers. Notice that with this genie information, after each channel use t, the transmitter can locally compute the missing feedback outputs $V_{1,t}$ and $V_{2,t}$:

$$V_{1,t}=X_t+(Z_{1,t}+W_{1,t}), \tag{88}$$
$$V_{2,t}=X_t+(Z_{2,t}+W_{2,t}). \tag{89}$$

We can thus conclude that the sum-capacity of the genie-aided AWGN BC is at least as large as the sum-capacity of the original AWGN BC with feedback.

We next elaborate on the second step. The less noisy AWGN BC is described by the channel law

$$Y'_{1,t}=x_t+Z'_{1,t}, \tag{90}$$
$$Y'_{2,t}=x_t+Z'_{2,t}, \tag{91}$$

where the noise sequences are defined as

$$Z'_{1,t}\triangleq Z_{1,t}-\mathsf{E}[Z_{1,t}|(Z_{1,t}+W_{1,t}),(Z_{2,t}+W_{2,t})], \tag{92}$$
$$Z'_{2,t}\triangleq Z_{2,t}-\mathsf{E}[Z_{2,t}|(Z_{1,t}+W_{1,t}),(Z_{2,t}+W_{2,t})], \tag{93}$$

and are of variances

$$\mathsf{Var}(Z'_{1,t})=\sigma_1^2\frac{\sigma_{W1}^2\sigma_2^2(1-\rho_z^2)+\sigma_{W1}^2\sigma_{W2}^2}{(\sigma_1^2+\sigma_{W1}^2)(\sigma_2^2+\sigma_{W2}^2)-\sigma_1^2\sigma_2^2\rho_z^2}, \tag{94}$$

$$\mathsf{Var}(Z'_{2,t})=\sigma_2^2\frac{\sigma_{W2}^2\sigma_1^2(1-\rho_z^2)+\sigma_{W2}^2\sigma_{W1}^2}{(\sigma_1^2+\sigma_{W1}^2)(\sigma_2^2+\sigma_{W2}^2)-\sigma_1^2\sigma_2^2\rho_z^2}. \tag{95}$$

By the following two observations, the sum-capacity of this less noisy AWGN BC coincides with the sum-capacity of the genie-aided AWGN BC. Firstly, the sum-capacity of the less noisy AWGN BC remains unchanged if prior to transmission a genie reveals the sequences $\{Z_{1,t}+W_{1,t}\}$ and $\{Z_{2,t}+W_{2,t}\}$ to the transmitter and both receivers. This follows because by Definitions (92) and (93) and the Gaussianity of all involved sequences the genie-information $\{Z_{1,t}+W_{1,t}\}$ and $\{Z_{2,t}+W_{2,t}\}$ is independent of the reduced noise sequences $\{Z'_{1,t}, Z'_{2,t}\}$, and thus plays only the role of common randomness which does not increase capacity. Secondly, the sum-capacity of the genie-aided AWGN BC coincides with the sum-capacity of the less noisy AWGN BC, when in this latter case the transmitter and both receivers additionally know the genie-information $\{(Z_{1,t}+W_{1,t})\}$ and $\{(Z_{2,t}+W_{2,t})\}$. This holds because knowing the genie-information $\{(Z_{1,t}+W_{1,t})\}_{t=1}^n$ and $\{(Z_{2,t}+W_{2,t})\}_{t=1}^n$ the outputs $Y'_{1,t}$ and $Y'_{2,t}$ can be transformed into the outputs $Y_{1,t}$ and $Y_{2,t}$, and vice versa.

We finally elaborate on the third step. The less noisy AWGN BC is a classical AWGN BC with neither feedback nor genie-information, and its sum-capacity is [20]

$$C_{\text{BCLessNoisy},\Sigma}\left(P,\sigma_1^2,\sigma_2^2,\rho_z,\sigma_{W1}^2,\sigma_{W2}^2\right) = \frac{1}{2}\log\left(1+\frac{P}{\min\left\{\mathsf{Var}\left(Z'_{1,t}\right),\mathsf{Var}\left(Z'_{2,t}\right)\right\}}\right), \quad (96)$$

where the variances $\mathsf{Var}\left(Z'_{1,t}\right), \mathsf{Var}\left(Z'_{2,t}\right)$ are defined in (94) and (95). By (94)–(96) the prelog of the less noisy AWGN BC equals 1, irrespective of the noise variances $\sigma_1^2,\sigma_2^2,\sigma_{W1}^2,\sigma_{W2}^2>0$ and the noise correlation $\rho_z\in[-1,1]$. This concludes the third step, and thus our proof.

B. Proof of Note 4

Let $\sigma_1^2,\sigma_2^2>0$ and $\rho_z\in[-1,1]$ be fixed and for every power $P>0$ let the feedback-noise variances $\sigma_{W1}^2(P)$ and $\sigma_{W2}^2(P)$ be given, where

$$\varlimsup_{P\to\infty}\frac{-\log(\sigma_{W\nu}^2(P))}{\log P}\le 0, \qquad \nu\in\{1,2\}. \quad (97)$$

Since for each power P a prelog 1 is achievable even without feedback we have to prove

$$\varlimsup_{P\to\infty}\frac{C_{\text{BCNoisy},\Sigma}\left(P,\sigma_1^2,\sigma_2^2,\rho_z,\sigma_{W1}^2(P),\sigma_{W2}^2(P)\right)}{\frac{1}{2}\log(1+P)}\le 1. \quad (98)$$

For $\rho_z\in(-1,1)$ Inequality (98) follows immediately from Corollary 2, because with noisy feedback the prelog cannot be larger than with noise-free feedback. (The transmitter can always add the feedback noise itself.)

For $\rho_z\in\{-1,1\}$ the proof of (98) is similar to the proof in Section VII-A. In fact, following the same steps as before, we can conclude that for each $P>0$

$$C_{\text{BCNoisy},\Sigma}(P,\sigma_1^2,\sigma_2^2,\rho_z,\sigma_{W1}^2(P),\sigma_{W2}^2(P)) \le \frac{1}{2}\log\left(1+\frac{P}{\min\left\{\mathsf{Var}\left(Z'_{1,t}\right),\mathsf{Var}\left(Z'_{2,t}\right)\right\}}\right), \quad (99)$$

where here, because $|\rho_z|=1$, the variances in (94) and (95) simplify to

$$\mathsf{Var}\left(Z'_{1,t}\right)=\frac{\sigma_1^2\sigma_{W1}^2(P)\sigma_{W2}^2(P)}{\sigma_1^2\sigma_{W2}^2(P)+\sigma_2^2\sigma_{W1}^2(P)+\sigma_{W1}^2(P)\sigma_{W2}^2(P)}, \quad (100)$$

$$\mathsf{Var}\left(Z'_{2,t}\right)=\frac{\sigma_2^2\sigma_{W1}^2(P)\sigma_{W2}^2(P)}{\sigma_1^2\sigma_{W2}^2(P)+\sigma_2^2\sigma_{W1}^2(P)+\sigma_{W1}^2(P)\sigma_{W2}^2(P)}. \quad (101)$$

The desired inequality (98) for $\rho_z\in\{-1,1\}$ follows now simply by combining (97) with (99)–(101).

VIII. Scheme and Proof for the K-user AWGN BC with Noise-Free Feedback

A. A Scheme for the special case $\text{rank}(\mathsf{K}_z)=1$

We generalize the coding scheme in the previous section to $K\ge 2$ users. Here, we focus on the special case where the noise covariance matrix K_z has rank 1, which suffices to prove the achievability of Theorem 6. The scheme is however easily extended to general covariance matrices.

To make the description of the scheme more intuitive, we rewrite the channel law in (29) in an equivalent form. Since K_z is of rank 1, (29) is equivalent to

$$Y_{k,t}=X_t+\alpha_k Z_{0,t}, \qquad t\in\{1,\ldots,n\}, \quad (102)$$

where $\{Z_{0,t}\}$ is an IID zero-mean Gaussian sequence of variance 1, and where

$$\alpha_k=\text{sign}(\rho_{1,k})\sigma_k, \qquad k\in\{1,\ldots,K\};$$

here $\rho_{1,k}$ denotes the correlation coefficient (either -1 or 1) between $Z_{1,t}$ and $Z_{k,t}$ and σ_k is the positive root of σ_k^2.

We now describe the scheme. It has the following parameters: the positive integer number η; the η-by-η strictly lower-triangular matrices $\mathsf{B}_1,\ldots,\mathsf{B}_K$; the η-dimensional column-vectors $\mathbf{u}_1,\ldots,\mathbf{u}_K$; and the η-dimensional row-vectors $\mathbf{v}_1,\ldots,\mathbf{v}_K$.

1) Code Construction, Encoding, Decoding: Code construction, encoding, and decoding procedures are as described in Section VI-A, with the only difference that now the parameter k runs from $1,\ldots,K$.

Thus, in each subblock $i\in\{1,\ldots,n'\}$ the transmitter sends the channel inputs

$$\mathbf{X}_i=\sum_{k=1}^K \Xi_{k,i}\mathbf{u}_k+\sum_{k=1}^K \mathsf{B}_k\alpha_k\mathbf{Z}_{0,i}, \quad (103)$$

where $\mathbf{Z}_{0,i}\triangleq(Z_{0,(i-1)\eta+1},\ldots,Z_{0,i\eta})^{\mathsf{T}}$, and Receiver $k\in\{1,\ldots,K\}$ observes the channel outputs

$$\mathbf{Y}_{k,i}=\sum_{k'=1}^K \Xi_{k,i}\mathbf{u}_k+\left(\sum_{k'=1}^K \mathsf{B}_{k'}\alpha_{k'}+\mathsf{I}\alpha_k\right)\mathbf{Z}_{0,i}, \quad (104)$$

where $\mathbf{Y}_{k,i}\triangleq(Y_{k,(i-1)\eta+1},\ldots,Y_{k,i\eta})^{\mathsf{T}}$. The channel input sequence satisfies the average block-power constraint (7) whenever Inequality (105) shown on top of the next page is satisfied.

$$\sum_{k=1}^{K}\|\mathbf{u}_k\|^2+\operatorname{tr}\left(\left(\sum_{k=1}^{K}\mathsf{B}_k\alpha_k\right)\left(\sum_{k=1}^{K}\mathsf{B}_k^{\mathsf{T}}\alpha_k\right)\right)\leq \eta P \tag{105}$$

2) Choice of Parameters: For every positive integer η we present a relatively simple choice of parameters such that

(i) each vector $\mathbf{v}_k$ is orthogonal to the first $\eta-1$ columns of the matrix $\left(\sum_{k'=1}^{K}\mathsf{B}_{k'}\alpha_{k'}+\mathsf{I}\alpha_k\right)$;
(ii) each vector $\mathbf{u}_k$ is orthogonal to the vectors $\mathbf{v}_1,\ldots,\mathbf{v}_{k-1},\mathbf{v}_{k+1},\ldots,\mathbf{v}_K$ but not to $\mathbf{v}_k$;
(iii) the inner product $\mathbf{v}_k\mathbf{u}_k$ equals $\sqrt{P}^{\eta-K}$ times a constant that does not depend on P,

and such that the power constraint (105) is satisfied for all sufficiently large η and all powers $P\geq 1$.[5]

Notice that with such a choice, by (i) and (ii), the "new outputs" formed at Receiver k are of the form

$$I_{k,i}=\mathbf{v}_k\mathbf{u}_k\Xi_{k,i}+\alpha_k Z_{0,i\eta}, \tag{106}$$

and thus correspond to the outputs of a point-to-point channel where the transmitter only sends the desired codeword $(\Xi_{k,1},\ldots,\Xi_{k,n'})$. By (iii) this point-to-point channel has prelog $\frac{\eta-K}{\eta}$, and thus prelog 1 when $\eta\to\infty$.

To present our choice, we need the following two definitions. Define for each $k\in\{1,\ldots,K\}$ the K-dimensional column-vector

$$\boldsymbol{\alpha}_k=\begin{pmatrix}1 & \alpha_k & \alpha_k^2 & \alpha_k^3 & \ldots & \alpha_k^K\end{pmatrix}^{\mathsf{T}}. \tag{107}$$

Let now $\hat{\mathbf{w}}_k$ be the projection of the vector $\boldsymbol{\alpha}_k$ on the span of $\{\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_{k-1},\boldsymbol{\alpha}_{k+1},\boldsymbol{\alpha}_K\}$, and define the K-dimensional row-vector

$$\mathbf{w}_k\triangleq\left(\frac{\boldsymbol{\alpha}_k-\hat{\mathbf{w}}_k}{\|\boldsymbol{\alpha}_k-\hat{\mathbf{w}}_k\|}\right)^{\mathsf{T}}. \tag{108}$$

If $\alpha_1,\ldots,\alpha_K$ are all different and not equal to 0, then the vectors $\{\boldsymbol{\alpha}_k\}_{k=1}^K$ are linearly independent and for each $k\in\{1,\ldots,K\}$ the vector $\mathbf{w}_k\neq\mathbf{0}$ is orthogonal to $\{\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_{k-1},\boldsymbol{\alpha}_{k+1},\ldots,\boldsymbol{\alpha}_K\}$ but not to $\boldsymbol{\alpha}_k$.

Now, for given positive integer η, choose the matrices $\mathsf{B}_1,\ldots,\mathsf{B}_K$ such that

$$\sum_{k=1}^{K}\mathsf{B}_k\alpha_k=\mathsf{B}, \tag{109}$$

where B is defined in (57), and choose for $k\in\{1,\ldots,K\}$

$$\mathbf{u}_k=\begin{pmatrix}\frac{w_{k,1}}{\sqrt{P}^{K-1}} & \frac{w_{k,2}}{\sqrt{P}^{K-2}} & \ldots & \frac{w_{k,K-1}}{\sqrt{P}} & w_{k,K} & 0 & \ldots & 0\end{pmatrix}^{\mathsf{T}} \tag{110}$$

$$\mathbf{v}_k=\begin{pmatrix}\left(-\frac{\sqrt{P}}{\alpha_k}\right)^{\eta-1} & \left(-\frac{\sqrt{P}}{\alpha_k}\right)^{\eta-2} & \ldots & -\frac{\sqrt{P}}{\alpha_k} & 1\end{pmatrix}^{\mathsf{T}} \tag{111}$$

where $w_{k,j}$ denotes the j-th entry of the vector $\mathbf{w}_k$.

This choice has the desired properties (i)–(iii), and in particular it satisfies (iii) because

$$\mathbf{v}_k\mathbf{u}_k=\frac{\sqrt{P}^{\eta-K}}{\alpha_k^{\eta-1}}\boldsymbol{\alpha}_k\mathbf{w}_k, \tag{112}$$

where $\mathbf{w}_k\boldsymbol{\alpha}_k>0$ is a constant that neither depends on P nor on η.

3) Achievable Rates: Combining (106)–(112), we conclude the following. Our scheme achieves all rate tuples $(R_1,\ldots,R_K)$ that simultaneously satisfy

$$R_k\leq\frac{1}{2\eta}\log\left(1+\frac{P^{\eta-K}(\boldsymbol{\alpha}_k\mathbf{w}_k)^2}{\alpha_k^{2\eta}}\right),\quad k\in\{1,\ldots,K\}. \tag{113}$$

Letting $\eta\to\infty$ we obtain the following.

Proposition 12: If $\alpha_1,\ldots,\alpha_K$ are all different, and $P>1$ the rate-tuple $(R_1,\ldots,R_K)$ satisfying

$$R_k\leq\frac{1}{2}\log^+\left(\frac{P}{\alpha_k^2}\right),\quad k\in\{1,\ldots,K\} \tag{114}$$

is achievable.

B. Proof of Theorem 6

By the definition of n_α there are only n_α different channels and therefore the prelog cannot exceed n_α.

If $n_\alpha=K$, then $\alpha_1,\ldots,\alpha_K$ in (102) are all different, and the achievability of prelog n_α follows directly from Proposition 12. If $n_\alpha<K$, then we pick n_α receivers such that the corresponding rows of K_z are all different. We then apply the scheme in the previous section to only these n_α receivers (and ignore the other receivers). By Proposition 12 this scheme achieves the desired prelog n_α to the selected receivers.

IX. Scheme and Proofs for the AWGN IC with One-Sided Feedback

A. A Scheme

We present a block-scheme similar to the scheme for the two-user broadcast channel in Section VI-A. Our scheme takes the same parameters as the scheme in Section VI-A: a positive integer number η, two strictly lower-triangular η-by-η matrices B_1 and B_2, two η-dimensional column-vectors $\mathbf{u}_1,\mathbf{u}_2$, and two η-dimensional row-vectors $\mathbf{v}_1,\mathbf{v}_2$.

1) Code construction, Encoding, Decoding: The code construction and the decoding procedures are the same as for the two-user broadcast scheme in Section VI-A. The encoding procedure differs because here there are two transmitters.

As before, for $k\in\{1,2\}$ let $\boldsymbol{\Xi}_k(M_k)$ denote the codeword in $\mathcal{C}_k$ corresponding to Message M_k, and let $\Xi_{k,i}$ denote its i-th symbol. Let for each $k\in\{1,2\}$ the η-length vectors $\mathbf{Y}_{k,i}$ and $\mathbf{Z}_{k,i}$ be defined as before, and similarly, define for $k\in\{1,2\}$

$$\mathbf{X}_{k,i}\triangleq\left(X_{k,(i-1)\eta+1},\ldots,X_{k,(i\eta)}\right)^{\mathsf{T}}. \tag{115}$$

Transmitter $k\in\{1,2\}$ performs the following encoding procedure. It picks codeword $\boldsymbol{\Xi}_k(M_k)$ and in each subblock

[5] A choice of parameters that satisfies the average block-power constraint for all powers $P>0$ and exhibits the same asymptotic performance as $\eta\to\infty$ is obtained by appropriately scaling the vectors $\{\mathbf{u}_k\}$.

$i \in \{1, \ldots, n'\}$ it sends a linear combination of the i-th symbol of this codeword and the past feedback signals observed in this subblock:

$$\mathbf{X}_{1,i} = \mathbf{u}_1 \Xi_{1,i} + \mathsf{B}_1(a_{1,2}\mathbf{X}_{2,i} + \mathbf{Z}_{1,i}) \quad (116a)$$
$$\mathbf{X}_{2,i} = \mathbf{u}_2 \Xi_{2,i} + \mathsf{B}_2(a_{2,1}\mathbf{X}_{1,i} + \mathbf{Z}_{2,i}). \quad (116b)$$

2) Choice of Parameters: We only present a choice of parameters for the case where $\rho_z \in \{-1, 1\}$. In this case $Z_{2,t} = \rho_z \frac{\sigma_2}{\sigma_1} Z_{1,t}$, and we can rewrite the channel outputs at Receiver 2 as

$$Y_{2,t} = a_{2,1}X_{1,t} + a_{2,2}X_{2,t} + \rho_z \frac{\sigma_2}{\sigma_1} Z_{1,t}, \quad (117a)$$

We fix a positive integer η, and choose B_2 the all-zero matrix and

$$\mathsf{B}_1 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ \frac{\sqrt{P}}{\sigma_1} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \frac{\sqrt{P}}{\sigma_1} & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \ldots & 0 & \frac{\sqrt{P}}{\sigma_1} & 0 & 0 \\ 0 & 0 & \ldots & 0 & \frac{\sqrt{P}}{\sigma_1} & 0 \end{pmatrix}. \quad (118)$$

With this choice, by (116a) and (117),

$$\mathbf{Y}_{1,i} = a_{1,1}\mathbf{u}_1\Xi_{1,i} + a_{1,2}(a_{1,1}\mathsf{B}_1 + \mathsf{I})\mathbf{u}_2\Xi_{2,i} + (a_{1,1}\mathsf{B}_1 + \mathsf{I})\mathbf{Z}_{1,i}$$

and

$$\mathbf{Y}_{2,i} = a_{2,1}\mathbf{u}_1\Xi_{1,i} + (a_{2,1}a_{1,2}\mathsf{B}_1 + a_{2,2}\mathsf{I})\mathbf{u}_2\Xi_{2,i} + \left(a_{2,1}\mathsf{B}_1 + \rho_z\frac{\sigma_2}{\sigma_1}\mathsf{I}\right)\mathbf{Z}_{1,i}.$$

We now choose the vector $\mathbf{v}_1$ to be orthogonal to the first $(\eta-1)$ columns of the matrix $(a_{1,1}\mathsf{B}_1+\mathsf{I})$; $\mathbf{v}_2$ to be orthogonal to the first $(\eta-1)$ columns of the matrix $\left(a_{2,1}\mathsf{B}_1+\rho_z\frac{\sigma_2}{\sigma_1}\mathsf{I}\right)$; the vector $\mathbf{u}_1$ to be orthogonal to $\mathbf{v}_2$ but not to $\mathbf{v}_1$, and the vector $\mathbf{u}_2$ simply not orthogonal to $(a_{2,1}a_{1,2}\mathsf{B}_1 + a_{2,2}\mathsf{I})\mathbf{v}_2$. Such a choice is:

$$\mathbf{u}_1 = \sqrt{\frac{P\sigma_2^2}{\sigma_2^2 + Pa_{2,1}^2}} \begin{pmatrix} 1 & \frac{a_{2,1}\sqrt{P}}{\rho_z\sigma_2} & 0\ldots & 0 \end{pmatrix}^{\mathsf{T}}$$
$$\mathbf{u}_2 = \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix}^{\mathsf{T}}$$
$$\mathbf{v}_1 = \begin{pmatrix} \left(\frac{-a_{1,1}\sqrt{P}}{\sigma_1}\right)^{\eta-1} & \left(\frac{-a_{1,1}\sqrt{P}}{\sigma_1}\right)^{\eta-2} & \ldots & \frac{-a_{1,1}\sqrt{P}}{\sigma_1} & 1 \end{pmatrix}$$
$$\mathbf{v}_2 = \begin{pmatrix} \left(\frac{-a_{2,1}\sqrt{P}}{\rho_z\sigma_2}\right)^{\eta-1} & \left(\frac{-a_{2,1}\sqrt{P}}{\rho_z\sigma_2}\right)^{\eta-2} & \ldots & \frac{-a_{2,1}\sqrt{P}}{\rho_z\sigma_2} & 1 \end{pmatrix}.$$

The chosen $\mathbf{u}_1, \mathbf{u}_2, \mathsf{B}_1, \mathsf{B}_2$ satisfy the average power constraints (34) at Transmitters 1 and 2 whenever P and η are sufficiently large. Moreover,

$$I_{1,i} = \sqrt{\frac{P\sigma_2^2}{\sigma_2^2 + Pa_{2,1}^2}} \left(-\frac{\sqrt{P}a_{1,1}}{\sigma_1}\right)^{\eta-1} \left(1 - \frac{a_{2,1}\sigma_1}{a_{1,1}\rho_z\sigma_2}\right) a_{1,1}\Xi_{1,i} + Z_{1,i\eta}$$

and

$$I_{2,i} = \left(-\frac{\sqrt{P}a_{2,1}}{\rho_z\sigma_2}\right)^{\eta-1} \left(a_{2,2} - a_{1,2}\frac{\rho_z\sigma_2}{\sigma_1}\right)\Xi_{2,i} + \rho_z\frac{\sigma_2}{\sigma_1}Z_{1,i\eta}.$$

Again, each receiver uses its last $\eta - 2$ outputs of each subblock to iteratively cancel the previous noise symbols in this subblock. Moreover, the "new outputs" $I_{1,1}, \ldots, I_{1,n'}$ do not depend on $\Xi_{2,1}, \ldots, \Xi_{2,n'}$ but only on $\Xi_{1,1}, \ldots, \Xi_{1,n'}$, and similarly, the "new outputs" $I_{2,1}, \ldots, I_{2,n'}$ do not depend on $\Xi_{1,1}, \ldots, \Xi_{1,n'}$ but only on $\Xi_{2,1}, \ldots, \Xi_{2,n'}$.

3) Achievable Rates: For every $\eta \in \mathbb{Z}^+$ our scheme achieves all nonnegative rate pairs that satisfy

$$R_1 \leq \frac{1}{2\eta}\log\left(1 + \frac{P^\eta a_{1,1}^{2\eta}}{\sigma_1^{2\eta}}\frac{\sigma_2^2}{(\sigma_2^2 + Pa_{2,1}^2)}\left(1 - \frac{a_{2,1}\sigma_1}{a_{1,1}\rho_z\sigma_2}\right)^2\right)$$
$$R_2 \leq \frac{1}{2\eta}\log\left(1 + \frac{P^{\eta-1}a_{2,1}^{2\eta-2}}{\sigma_2^{2\eta}}\left(a_{2,2} - \frac{a_{1,2}\rho_z\sigma_2}{\sigma_1}\right)^2\right).$$

Taking the limit $\eta \to \infty$ leads to the following.

Proposition 13: When $\frac{a_{2,1}}{\alpha_{1,1}}$ and $\frac{a_{2,2}}{a_{1,2}}$ are both different from $\rho_z\frac{\sigma_2}{\sigma_1}$, then all rate-pairs satisfying

$$R_1 \leq \frac{1}{2}\log^+\left(\frac{a_{1,1}^2 P}{\sigma_1^2}\right) \quad (119)$$
$$R_2 \leq \frac{1}{2}\log^+\left(\frac{a_{2,1}^2 P}{\sigma_2^2}\right) \quad (120)$$

are achievable.

Note 9: Exchanging the roles of the two transmitters we obtain: when $\frac{a_{2,1}}{\alpha_{1,1}}$ and $\frac{a_{2,2}}{a_{1,2}}$ are both different from $\frac{\rho_z\sigma_2}{\sigma_1}$, then all rate-pairs satisfying

$$R_1 \leq \frac{1}{2}\log^+\left(\frac{a_{1,2}^2 P}{\sigma_1^2}\right) \quad (121)$$
$$R_2 \leq \frac{1}{2}\log^+\left(\frac{a_{2,2}^2 P}{\sigma_1^2}\right) \quad (122)$$

are achievable.

Note 10: With the proposed choice of parameters our scheme achieves prelog 2 when $\frac{a_{2,1}}{\alpha_{1,1}}$ and $\frac{a_{2,2}}{a_{1,2}}$ are both different from $\rho_z\frac{\sigma_2}{\sigma_1}$ and when $\eta \to \infty$. For fixed η the scheme achieves a prelog of $2\frac{\eta-1}{\eta}$. Thus, choosing $\eta = 3$ suffices to achieve a prelog larger than 1.

Note 11: For a symmetric setup where $a_{1,1} = a_{2,2}$ and $a_{1,2} = a_{2,1}$ the achievability of (36) can also be shown using a slight generalization of Kramer's memoryless LMMSE scheme [14], see [11].

B. *Proof of Theorem 8*

Relation (37) follows from the following more general relation. Irrespective of the channel parameters the prelog satisfies

$$1 \leq \varlimsup_{P\to\infty} \frac{C_{\text{IC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1+P)} \leq 2. \quad (123)$$

The lower bound in (123) can be achieved, e.g., by silencing Transmitter 1 and letting Transmitter 2 communicate its Message M_2 to Receiver 2 over the resulting interference-free AWGN channel $Y_{2,t} = a_{2,2}X_{2,t} + Z_{2,t}$ at a rate $R_2 = \frac{1}{2}\log\left(1 + \frac{a_{2,2}^2 P}{\sigma_2^2}\right)$. The upper bound can be derived using the cutset bound and the entropy maximizing property of the Gaussian distribution under a covariance matrix constraint. In fact, applying two cuts between both transmitters and each of the two receivers yields the following upper bounds

$$R_k \leq \frac{1}{2}\log\left(1 + \frac{(|a_{k,1}| + |a_{k,2}|)^2 P}{\sigma_k^2}\right), \quad k \in \{1, 2\},$$

which establish the converse result in (123).

We next prove Equation (35). The achievability follows directly from the general relation (123). When $\rho_z \in \{-1, 1\}$ and $\frac{a_{2,1}}{a_{1,1}} = \frac{a_{2,2}}{a_{1,2}} = \text{sign}(\rho_z)\frac{\sigma_2}{\sigma_1}$, the converse follows because in this case

$$Y_{1,t} = \text{sign}(\rho_z)\frac{\sigma_1}{\sigma_2} Y_{2,t}, \qquad t \in \{1, \ldots, n\}, \tag{124}$$

and thus, each receiver can reconstruct the other receiver's outputs. Consequently, the feedback capacity of our AWGN IC coincides with the feedback capacity of the AWGN MAC from both transmitters to one of the two receivers, and its prelog is 1 [26].

To prove the converse to (35) when $\rho_z \in (-1, 1)$ we use a genie-argument and a generalized Sato-MAC bound [27], similar to the upper bounds in [28, Section V-B], [29], [30], [31]. Our proof consists of the following three steps. In the first step we let a genie prior to transmission reveal the symbols

$$U^n = Z_2^n - \frac{a_{2,2}}{a_{1,2}} Z_1^n$$

to Receiver 1. This obviously can only increase the sum-capacity of our channel. The resulting setup will be called the *genie-aided IC*.

In the second step, we apply Sato's MAC-bound argument [27] to this genie-aided IC.[6] That means, we define an appropriate genie-aided MAC and show that the capacity of the genie-aided IC is contained in the capacity of this genie-aided MAC. The genie-aided MAC is obtained from the genie-aided IC by eliminating Receiver 2 and instead requiring that the only remaining receiver 1 decodes both messages M_1 and M_2. The desired inclusion of the capacities is proved by showing that for any encoding and decoding strategies for the genie-aided IC it is possible to find encoding/decoding strategies for the genie-aided MAC such that the probability of error over the MAC is no larger than over the IC.

We fix encoding/decoding functions for the genie-aided IC, and choose the encoding/decoding functions for the genie-aided MAC as follows. The MAC transmitters apply the same encoding functions as the IC transmitters. The only MAC-receiver decodes the pair (M_1, M_2) as follows: 1.) It applies IC-Receiver 1's decoding rule to decode Message M_1. 2.) It computes

$$\hat{X}_{1,t} = f_{\text{IC},1,t}^{(n)}(\hat{M}_1, Y_1^{t-1}), \qquad t \in \{1, \ldots, n\}, \tag{125}$$

and

$$\hat{Y}_2^n = \frac{a_{2,2}}{a_{1,2}}\big(Y_1^n - a_{1,1}\hat{X}_1^n\big) + a_{2,1}\hat{X}_1^n + U^n, \tag{126}$$

where $\hat{M}_1$ denotes the decoded message in 1.) and $f_{\text{IC},1,t}^{(n)}$ denotes IC-Transmitter 1's encoding function. 3.) It finally applies IC-Receiver 2's decoding rule to decode Message M_2 based on the sequence $\hat{Y}_2^n$.

Notice that whenever the MAC-receiver (and thus also IC-Receiver 1) decode M_1 correctly, then $\hat{X}_1^n = X_1^n$, and $\hat{Y}_2^n = Y_2^n$. Therefore, whenever the IC-Receivers 1 and 2 decode their intended messages M_1 and M_2 correctly, then so does the only MAC-receiver, and the probability of error over the MAC cannot exceed the probability of error over the IC. This concludes the second step.

In the third step we show that the genie-aided MAC has prelog no larger than 1. Combined with the previous two steps this yields the desired converse to (35). Before elaborating on this third step, we recall that in the considered genie-aided MAC the channel law is

$$Y_{1,t} = a_{1,1}X_{1,t} + a_{1,2}X_{2,t} + Z_{1,t}, \quad t \in \{1, \ldots, n\};$$

the two transmitters observe the generalized feedback signals $\{Y_{1,t}\}$ and $\{Y_{2,t}\}$; and before the transmission starts, the receiver learns the genie-information U^n.

We now prove that the prelog of this genie-aided MAC is upper bounded by 1. To this end we fix an arbitrary sequence of blocklength-n, rates-(R_1, R_2) coding schemes for the considered MAC such that the probability of error $\epsilon(n)$ tends to zero as n tends to infinity. Then, for every blocklength n we have:

$$\begin{aligned}
&R_1 + R_2 \\
&\leq \frac{1}{n} I(M_1, M_2; Y_1^n, U^n) + \frac{\epsilon(n)}{n} \\
&= \frac{1}{n} I(M_1, M_2; Y_1^n | U^n) + \frac{\epsilon(n)}{n} \\
&= \frac{1}{n}\sum_{t=1}^{n}\Big(h(Y_{1,t}|Y_1^{t-1}, U^n) \\
&\qquad\qquad - h(Y_{1,t}|Y_1^{t-1}, M_1, M_2, U^n)\Big) + \frac{\epsilon(n)}{n} \\
&\leq \frac{1}{n}\sum_{t=1}^{n}\Big(h(Y_{1,t}|U_t) \\
&\qquad\qquad - h(Y_{1,t}|Y_1^{t-1}, M_1, M_2, Y_2^{t-1}, U^n)\Big) + \frac{\epsilon(n)}{n} \\
&= \frac{1}{n}\sum_{t=1}^{n}\Big(h(Y_{1,t}|U_t) - h(Y_{1,t}|X_{1,t}, X_{2,t}, U_t)\Big) \\
&= \frac{1}{n}\sum_{t=1}^{n} I(Y_{1,t}; X_{1,t}, X_{2,t}|U_t) \\
&\leq \frac{1}{2}\log\left(1 + \frac{(|a_{1,1}| + |a_{1,2}|)^2 P}{\mathsf{Var}\Big(Z_{1,t}|Z_{2,t} - \frac{a_{2,2}}{a_{1,2}}Z_{1,t}\Big)}\right)
\end{aligned} \tag{127}$$

where the first inequality follows by Fano's inequality; the first equality follows by the independence of the genie-information U^n and the messages M_1 and M_2; the third equality by

[6] Unlike in Sato's setup, here both transmitters have feedback from their corresponding receivers. However, as we shall see, also in our setup with one-sided feedback we can use the same arguments.

noting that the vector Y_2^{t-1} can be computed as a function of M_1, Y_1^{t-1}, and U^{t-1}, see (126); the fourth equality follows because the input $X_{1,t}$ is a function of the Message M_1 and the feedback outputs Y_1^{t-1}, and similarly $X_{2,t}$ is a function of M_2 and Y_2^{t-1}, and because of the Markov relation

$$(M_1, M_2, Y_1^{t-1}, Y_2^{t-1}, U^{t-1}, U_{t+1}^n) - (X_{1,t}, X_{2,t}, U_t) - Y_{1,t};$$

and the last inequality follows because the Gaussian distribution maximizes differential entropy under a covariance constraint.

Since $\mathsf{Var}\Big(Z_{1,t}|Z_{2,t} - \frac{a_{2,2}}{a_{1,2}} Z_{1,t}\Big)$ does not depend on P and is strictly positive whenever $\rho_z \in (-1,1)$, by (127) we conclude that (35) holds also when $\rho_z \in (-1,1)$.

The converse to (36) follows from the general Relation (123), and its achievability from Proposition 13 (Section IX-A).

Acknowledgment

The authors thank Prof. Frans M. J. Willems, TU Eindhoven, for pointing them to [32], which inspired the investigation leading to this work. They also thank the Associate Editor and the anonymous reviewers for their valuable comments.

Appendix A
Proof of Lemma 1

Let $\eta \in \mathbb{Z}^+$. If $1+\zeta \geq \xi$, then

$$\begin{aligned}
\frac{1}{2\eta}\log(1+\xi^{\eta-1}\zeta) &\leq \frac{1}{2\eta}\log(1+(1+\zeta)^{\eta-1}\zeta) \\
&= \frac{1}{2\eta}\log\Big((1+\zeta)^\eta - \underbrace{((1+\zeta)^{\eta-1}-1)}_{\geq 0}\Big) \\
&\leq \frac{1}{2\eta}\log\big((1+\zeta)^\eta\big) \\
&= \frac{1}{2}\log(1+\zeta).
\end{aligned}$$

Otherwise, if $1+\zeta < \xi$, then

$$\begin{aligned}
\frac{1}{2\eta}\log(1+\xi^{\eta-1}\zeta) &\leq \frac{1}{2\eta}\log(1+\xi^{\eta-1}(\xi-1)) \\
&= \frac{1}{2\eta}\log\Big(\xi^\eta - \underbrace{(\xi^{\eta-1}-1)}_{\geq 0}\Big) \\
&\leq \frac{1}{2\eta}\log\big(\xi^\eta\big) \\
&= \frac{1}{2}\log(\xi).
\end{aligned}$$

References

[1] P. Algoet and J. Cioffi, “The capacity of a channel with Gaussian noise and intersymbol interference,” in *Proc. of ISIT 91*, (Budapest, Hungary), p. 16, June 1991.

[2] M. Gastpar and M. A. Wigger, “The pre-log of Gaussian broadcast with feedback can be two,” in *Proc. 2008 ITA Workshop*, (San Diego, USA), Jan. 27 - Feb. 1 2008.

[3] C. Suh and D. Tse, “Symmetric feedback capacity of the Gaussian interference channel to within one bit,” in *Proc. of ISIT 09*, (Seoul, Korea), pp. 1609–1613, June/July 2009.

[4] T. M. Cover and S. Pombra, “Gaussian feedback capacity,” *IEEE Transactions on Information Theory*, vol. 35, pp. 37–43, Jan. 1989.

[5] A. Lapidoth and M. Wigger, “On the Gaussian MAC with imperfect feedback,” *IEEE Transactions on Information Theory*, vol. 56, pp. 5432–5477, Nov. 2010.

[6] A. Lapidoth, Y. Steinberg, and M. Wigger, “Gaussian broadcast channel with partial feedback,” in *2010 IEEE 26-th Convention of Electrical and Electronics Engineers in Israel*, (Eilath, Israel), p. 641, Nov 2010.

[7] M. Gastpar, A. Lapidoth, Y. Steinberg, and M. Wigger, “New achievable rates for the Gaussian broadcast channel with feedback,” in *Proc. of ISWCS’11*, (Aachen, Germany), November 2011.

[8] L. H. Ozarow, *Coding and Capacity for Additive White Gaussian Noise Multi-user Channels with Feedback*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1979.

[9] L. H. Ozarow and C. S. K. Leung., “An achievable region and outer bound for the Gaussian broadcast channel with feedback,” *IEEE Transactions on Information Theory*, vol. IT–30, pp. 667–671, July 1984.

[10] M. Wigger and M. Gastpar, “The pre-log of Gaussian broadcast with feedback can be two,” in *Proceedings of ISIT 2008*, (Toronto, Canada), pp. 545–546, July 2008.

[11] M. Gastpar, A. Lapidoth, and M. Wigger, “When feedback doubles the prelog in AWGN networks.” Posted on arxiv: http://arxiv.org/pdf/1003.6082, March 2010.

[12] E. Ardestanizadeh, P. Minero, and M. Franceschetti, “LQG control approach to Gaussian broadcast channels with feedback.” submitted to *IEEE Transactions on Information Theory*, Feb 2011.

[13] Y.-H. Kim, A. Lapidoth, and T. Weissman, “Bounds on the error exponent of the AWGN channel with AWGN corrupted feedback,” in *Proc. IEEE 24th Convention of Electrical and Electronics Engineers in Israel*, (Eilat, Israel), pp. 184–188, Nov 2006.

[14] G. Kramer, “Feedback strategies for white Gaussian interference networks,” *IEEE Transactions on Information Theory*, vol. 48, pp. 1423–1437, June 2002.

[15] N. Devroye and M. Sharif, “The multiplexing gain of MIMO X-channels wiht partial transmit side information,” in *Proc. of ISIT 2007*, (Nice, France), June 2007.

[16] C. Suh and D. Tse, “Feedback capacity of the Gaussian interference channel to within 2 bits,” *IEEE Transactions on Information Theory*, vol. 57, pp. 2667–2685, May 2011.

[17] D. Tuninetti, “On interference channels with generalized feedback,” in *ISIT 2007*, (Nice, France), pp. 2861–2865, June 2007.

[18] A. Vahid, C. Suh, and A. S. Avestimehr, “Interference channels with rate-limited feedback.” submitted to *IEEE Transactions on Information Theory*, March 2011.

[19] A. E. Gamal, “The feedback capacity of degraded broadcast channels,” *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 379–381, 1978.

[20] P. Bergmans, “A simple converse for broadcast channels with additive white Gaussian noise,” *IEEE Transactions on Information Theory*, vol. IT–20, pp. 279–280, Mar 1974.

[21] T. M. Cover, “Broadcast channels,” *IEEE Transactions on Information Theory*, vol. 18, pp. 2–14, Jan. 1972.

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

[23] G. Dueck, “Partial feedback for two-way and broadcast channels,” *Problems of Information and Control*, vol. 46, no. 1, pp. 1–15, 1980.

[24] O. Shayevitz and M. Wigger, “An achievable region for the discrete memoryless broadcast channel with feedback,” *in Proc. ISIT 2010*, Austin, Texas, June 13–18, 2010.

[25] O. Shayevitz and M. Wigger, “On the Capacity of the Discrete Memoryless Broadcast Channel with Feedback,” submitted to *IEEE Transactions on Information Theory*, December 2010.

[26] L. H. Ozarow, “The capacity of the white Gaussian multiple access channel with feedback,” *IEEE Transactions on Information Theory*, vol. IT–30, pp. 623–629, July 1984.

[27] H. Sato, “The capacity of the Gaussian interference channel under strong interference,” *IEEE Transactions on Information Theory*, vol. IT–27, pp. 786–788, November 1981.

[28] T. S. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” *IEEE Transactions on Information Theory*, vol. IT–27, pp. 49–60, January 1981.

[29] G. Kramer, “Outer bounds on the capacity of Gaussian interference channels,” *IEEE Transactions on Information Theory*, vol. 50, pp. 581–586, March 2004.

[30] A. Lapidoth, S. S. (Shitz), and M. A. Wigger, “On cognitive interference networks,” in *IEEE Information Theory Workshop*, (Lake Tahoe, USA), pp. 325–330, September 2007.

[31] A. Lapidoth, N. Levy, S. S. (Shitz), and M. A. Wigger, “A cognitive interference network with clustered decoding,” in *Proc. of ISIT 09*, (Seoul, Korea), pp. 198–202, June/July 2009.

[32] F. M. J. Willems and E. C. van der Meulen, “Een verbetering en veralgemening van het transmissiegebied van Ozarow voor het Gaussische broadcast kanaal met feedback,” in *Tweede Symposium over Informatietheorie in de Benelux*, (Zoetermeer, The Netherlands), pp. 129–138, May 1981. In Dutch.