Coding Schemes and Asymptotic Capacity for the Gaussian Broadcast and Interference Channels with Feedback

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

Abstract—A coding scheme is proposed for the memoryless Gaussian broadcast channel with correlated noises and feedback. For all noise correlations other than ± 1, the gap between the sum-rate the scheme achieves and the full-cooperation bound vanishes as the signal-to-noise ratio tends to infinity. When the correlation coefficient is -1, the gains afforded by feedback are unbounded and the prelog is doubled. When the correlation coefficient is $+1$ we demonstrate a dichotomy: If the noise variances are equal, then feedback is useless, and otherwise, feedback affords unbounded rate gains and doubles the prelog. The unbounded feedback gains, however, require perfect (noiseless) feedback. When the feedback links are noisy the feedback gains are bounded, unless the feedback noise decays to zero sufficiently fast with the signal-to-noise ratio.

Extensions to more receivers are also discussed as is the memoryless Gaussian interference channel with feedback.

Index Terms—broadcast channel, capacity, feedback, high SNR, interference channel, prelog.

I. INTRODUCTION

Among Shannon's most elegant results is that feedback cannot increase the capacity of a memoryless point-to-point channel. Feedback can, however, increase the capacity region of various memoryless multi-terminal networks such as the multiple-access channel (MAC) and the broadcast channel (BC). Exact expressions for the feedback capacities are known only for special networks, e.g., the memoryless Gaussian MAC [9], [27].

This paper considers the memoryless Gaussian BC with feedback. In the standard setting, the signals at the different receivers are corrupted by independent noises. For this setting, Ozarow and Leung [9], [10] showed that, indeed, feedback can enlarge the capacity region, though the exact capacity region with feedback remains to date unknown. A natural benchmark is the "full-cooperation bound," where the receivers

M. Gastpar is with the School of Computer and Communication Sciences at EPFL, Switzerland and with the EECS Department at University of California Berkeley {email: michael.gastpar@epfl.ch}. A. Lapidoth is with the Department of Information Technology and Electrical Engineering at ETH Zurich, Switzerland {email: lapidoth@isi.ee.ethz.ch}. Y. Steinberg is with the Department of Electrical Engineering at the Technion—Israel Institute of Technology {email: ysteinbe@ee.technion.ac.il}. M. Wigger is with the Communications and Electronics Department, at Telecom ParisTech, France {email: michele.wigger@telecom-paristech.fr}.

are allowed to cooperate. This turns the BC into a (single-input multiple-output) point-to-point channel whose capacity is well known and is not increased by feedback. Not surprisingly, even with feedback, the full-cooperation bound is generally not attainable.

In this paper we consider the case where the noises at the receivers are correlated, e.g., due to a common external interference. In the absence of feedback, such correlation does not impact the capacity region, because the latter depends only on the marginal channels. In the presence of feedback, however, the correlation is key.

Positively correlated noises were already considered by Ozarow and Leung. Willems and van der Meulen [33] extended Ozarow and Leung's scheme to negatively correlated noises. They also observed that when the two noises are of equal variances, the sum-rate achieved by the Ozarow-Leung scheme decreases as the correlation increases. (In the limiting case of fully correlated noises of equal variances, feedback does not increase capacity at all.) Willems and van der Meulen's observation inspired the current investigation: We wanted to see whether this observation is an artifact of the specific scheme they studied or whether it applies to the capacity region. And we wanted to see how crucial is the assumption that the noises are of equal variance (very much so!).

In this paper, we present a novel coding scheme and show that—in the high signal-to-noise ratio (SNR) limit—it achieves the full-cooperation bound for all noise correlations ρ_z satisfying $-1 < \rho_z < 1$. Consequently, the sum-rate capacity with noise-free feedback, $C_{\text{BC},\Sigma}$, satisfies

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

where P denotes the transmit power and $\sigma_1^2, \sigma_2^2 > 0$ the noise variances at the two receivers.

The case where $|\rho_z| = 1$ is special. As already mentioned, when $\rho_z = 1$ and the noises are of equal variance, feedback has no effect on capacity because in this case the BC is merely a point-to-point channel in disguise. But when $\rho_z = 1$ and the two noise variances differ, or when $\rho_z = -1$ the full-

cooperation bound is infinite for all SNR > 0, and it is thus useless. An alternative upper bound is the sum of the single-user capacities of the marginal channels to each receiver.

Prima Facie, it seems that this upper bound is completely out of reach because it ignores the tension between the users. Nevertheless, perhaps surprisingly, we show that in the high-SNR limit, this upper bound becomes achievable. More precisely, for noise correlation $\rho_z = -1$, as well as for noise correlation $\rho_z = 1$ provided that $\sigma_1^2 \neq \sigma_2^2$,

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

Without feedback, the sum-rate capacity is $\frac{1}{2} \log_2(1 + P/(\min\{\sigma_1^2, \sigma_2^2\}))$, so, for such noise correlations, feedback asymptotically doubles the sum-rate capacity in the high-SNR regime and the *prelog* becomes

$$
\overline{\lim_{P \to \infty}} \frac{C_{\mathrm{BC},\Sigma}}{\frac{1}{2} \log P} = 2 \tag{3}
$$
$$
\left(\text{when } \rho_z = -1, \text{ or when } \rho_z = 1 \text{ and } \sigma_1^2 \neq \sigma_2^2 \right).
$$

To put this result in context, it is important to note that, although feedback does provide capacity gains in many networks, these gains are typically modest and bounded in the SNR. By contrast, the present paper exhibits instances of networks where the capacity gains afforded by feedback are unbounded in the SNR. To the best of our knowledge, these are the first examples of such large feedback gains. Such examples were first reported in [2] for the two-user memoryless BC and for the symmetric two-user memoryless interference channel (IC) where the individual noise sequences corrupting the outputs at the two receivers are perfectly anti-correlated. In the meantime, other networks have been found where feedback affords unbounded capacity gains; see [3], [17] (based on the scheme proposed in [18]) for the two-user Gaussian IC when the noise sequences are independent. Multiplicative gains for the Gaussian IC with independent noises at moderate SNR were already reported in [15, Section VI-B].

There are several important ensuing questions concerning the special case of fully correlated noises. For example, we show that even if the correlation is not perfect but tends to one (or minus one) at least inversely proportionally to the SNR, we also obtain the same asymptotic capacity gain. Another question concerns the case where the feedback is noisy. We show that when the feedback links are corrupted by independent Gaussian noise sequences, then—irrespective of the positive feedback-noise variances and of the correlation of the forward noise-sequences—the prelog of the two-user

Gaussian 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 [14].

Finally, we consider the K-receivers memoryless Gaussian BC with $K > 2$, where no two of the Gaussian noise sequences corrupting the K received signals are of equal variance; none of the noise sequences is of zero variance; and the covariance matrix of the K noises is of rank 1. For this setup, our proposed coding scheme proves the achievability of a prelog of K. For a related recent result see [13].

The second network we consider is the two-user scalar memoryless Gaussian 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 if the noise sequences 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. Noise-free feedback thus approximately doubles the sum-rate capacity at high SNR and thus provides unbounded capacity 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 [15].)

Previously, a prelog of 2 was known to be achievable for the two-user scalar Gaussian IC only when the two transmitters (or the two receivers) could *fully* cooperate [16] in the sense that both transmitters could 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 Gaussian IC we do not consider noisy feedback. Rate-limited feedback for this setup has recently been studied in [19].

We conclude this section with some notation and a brief outline of the rest of the paper. Throughout the paper logarithms are base 2, and for convenience we define $-\log 0 = \infty$. We 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 is 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*.

The paper is organized as follows. In the following Section II we study the two-user Gaussian BC with noise-free or noisy feedback; in Section III we study the K-user Gaussian BC with noise-free feedback; and in Section IV the two-user Gaussian IC with noise-free feedback.

II. Two-User Broadcast Channel

A. Setup

We consider the real, scalar, memoryless Gaussian BC. Denoting the time-t transmitted symbol by $x_t \in \mathbb{R}$ and the time-t received symbols by $Y_{1,t}$ and $Y_{2,t}$,

$$
Y_{1,t} = x_t + Z_{1,t}, \tag{4a}
$$

$$
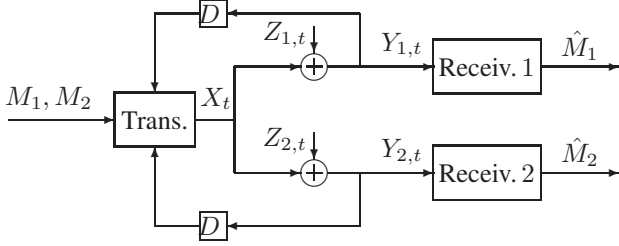Y_{2,t} = x_t + Z_{2,t}, \tag{4b}
$$

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

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

$$\mathsf{K}_z = \begin{pmatrix} \sigma_1^2 & \rho_z \sigma_1 \sigma_2 \\ \rho_z \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}. \tag{5}$$

We assume that both noise variances σ_1^2, σ_2^2 are strictly positive, and we denote their positive roots σ_1, σ_2.

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.

We depict the scenario with *noise-free feedback* in Figure 1 and with *noisy feedback* in Figure 2. In the former the transmitter learns the outputs $Y_{1,t-1}$ and $Y_{2,t-1}$ after sending X_{t-1}. It can thus choose its time-t channel input X_t as a function of both messages and all previous channel outputs:

$$X_t = f_{\mathrm{BC},t}^{(n)}\left(M_1, M_2, Y_1^{t-1}, Y_2^{t-1}\right), \quad t \in \{1, \ldots, n\}, \tag{6}$$

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

$$f_{\mathrm{BC},t}^{(n)} \colon \mathcal{M}_1 \times \mathcal{M}_2 \times \mathbb{R}^{t-1} \times \mathbb{R}^{t-1} \to \mathbb{R}. \tag{7}$$

In the scenario with *noisy feedback* the transmitter, after sending X_{t-1}, does not learn $Y_{1,t-1}$ and $Y_{2,t-1}$ but instead learns $V_{1,t-1}$ and $V_{2,t-1}$, which are noisy versions of $Y_{1,t-1}$ and $Y_{2,t-1}$:

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

where the sequence of pairs of feedback noises $\{(W_{1,t}, W_{2,t})\}$ is IID according to a zero-mean bivariate Gaussian distribution of diagonal[1] covariance matrix

$$\begin{pmatrix} \sigma_{W1}^2 & 0 \\ 0 & \sigma_{W2}^2 \end{pmatrix}, \quad \sigma_{W1}, \sigma_{W2} > 0. \tag{8}$$

The sequence $\{(W_{1,t}, W_{2,t})\}$ is assumed to be independent of the messages (M_1, M_2) and the noise sequences on the forward path $\{(Z_{1,t}, Z_{2,t})\}$. In this scenario the transmitter chooses its time-t channel input X_t as

$$X_t = f_{\mathrm{BCNoisy},t}^{(n)}\left(M_1, M_2, V_1^{t-1}, V_2^{t-1}\right), \quad t \in \{1, \ldots, n\}, \tag{9}$$

[1]We do not treat setups with correlated feedback noises or setups with feedback noises that are correlated with the forward noises.

where the encoding function is of the form

$$f_{\mathrm{BCNoisy},t}^{(n)} \colon \mathcal{M}_1 \times \mathcal{M}_2 \times \mathbb{R}^{t-1} \times \mathbb{R}^{t-1} \to \mathbb{R}. \tag{10}$$

In both scenarios, the channel inputs are subject to an expected average block-power constraint $P > 0$. Thus, we only allow encoding functions $\left\{f_{\mathrm{BC},t}^{(n)}\right\}_{t=1}^{n}$ or $\left\{f_{\mathrm{BCNoisy},t}^{(n)}\right\}_{t=1}^{n}$ for which

$$\frac{1}{n}\mathsf{E}\left[\sum_{t=1}^{n} X_t^2\right] \le P. \tag{11}$$

Receiver $k \in \{1, 2\}$ decodes its desired message M_k based on its observed channel output sequence Y_k^n. That is, it produces the estimate

$$\hat{M}_k = \phi_k^{(n)}(Y_k^n), \quad k \in \{1, 2\}, \tag{12}$$

using some decoding function

$$\phi_k^{(n)} \colon \mathbb{R}^n \to \{1, \ldots, \lfloor 2^{nR_k} \rfloor\}, \quad k \in \{1, 2\}. \tag{13}$$

For the scenario with noise-free feedback, 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 $\{f_{\mathrm{BC},t}^{(n)}\}_{t=1}^{n}$ satisfying the power constraint (11) and two decoding functions $\phi_1^{(n)}$ and $\phi_2^{(n)}$ such that

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

The closure of the set of all achievable rate pairs (R_1, R_2) is the *capacity region*. The supremum of the sum $R_1 + R_2$ over all achievable rate pairs (R_1, R_2) is the *sum-rate capacity*, which is denoted $C_{\mathrm{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)$.

For the scenario with noisy feedback, achievable rates, the capacity region, and the sum-rate capacity are defined analogously but using the encoding functions $\{f_{\mathrm{BCNoisy},t}^{(n)}\}_{t=1}^{n}$. The sum-rate capacity with noisy feedback is denoted by $C_{\mathrm{BCNoisy},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z, \sigma_{W1}^2, \sigma_{W2}^2)$.

The *prelog*, characterizes the logarithmic growth of the sum-rate capacity at high SNR. In the scenario with noise-free feedback it is defined as

$$\varlimsup_{P \to \infty} \frac{C_{\mathrm{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1 + P)} \tag{14}$$

and in the scenario with noisy feedback as

$$\varlimsup_{P \to \infty} \frac{C_{\mathrm{BCNoisy},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z, \sigma_{W1}^2, \sigma_{W2}^2)}{\frac{1}{2}\log(1 + P)}. \tag{15}$$

B. Main Results

Our results depend on whether or not the channel is physically degraded. The Gaussian BC is physically degraded whenever

$$\rho_z = \frac{\sigma_1}{\sigma_2} \quad \text{or} \quad \rho_z = \frac{\sigma_2}{\sigma_1}. \tag{16}$$

For example, it is physically degraded when $\rho_z = 1$ and $\sigma_1 = \sigma_2$, in which case the receivers observe the same sequence.
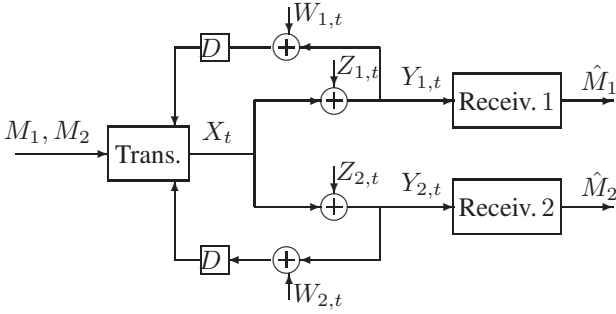
Fig. 2. The two-user Gaussian BC with noisy feedback.

When the Gaussian BC is physically degraded, feedback does not increase capacity [20], and thus [21], [22]

$$C_{\text{BC},\Sigma}(P,\sigma_1^2,\sigma_2^2,\rho_z) = C_{\text{BCNoisy},\Sigma}(P,\sigma_1^2,\sigma_2^2,\rho_z,\sigma_{W1}^2,\sigma_{W2}^2)$$
$$= \frac{1}{2}\log\left(1 + \frac{P}{\min\{\sigma_1^2,\sigma_2^2\}}\right), \quad (17)$$

irrespective of $\sigma_{W1}^2, \sigma_{W2}^2$ and ρ_z. If the channel is not physically degraded, then $C_{\text{BC},\Sigma}$ and $C_{\text{BCNoisy},\Sigma}$ are, in general, unknown, and bounds are called for.

We first present our results for noise-free feedback. For this scenario Theorem 1 reveals the high-SNR asymptotic sum-rate capacity. As we shall see, feedback strictly improves this asymptote whenever the BC is not physically degraded.

We shall express the asymptotic behavior using the function $C_{\text{Hi-SNR}}(P,\sigma_1^2,\sigma_2^2,\rho_z)$ whose definition depends on whether the BC is physically degraded and on whether ρ_z is strictly between -1 and $+1$:

Definition 1: Define $C_{\text{Hi-SNR}}(P,\sigma_1^2,\sigma_2^2,\rho_z)$ as follows.

- For channels that are physically degraded,

$$C_{\text{Hi-SNR}}(P,\sigma_1^2,\sigma_2^2,\rho_z) \triangleq \frac{1}{2}\log\left(1 + \frac{P}{\min\{\sigma_1^2,\sigma_2^2\}}\right). \quad (18)$$

- For channels that are not physically degraded and for which $\rho_z \in \{-1,1\}$,

$$C_{\text{Hi-SNR}}(P,\sigma_1^2,\sigma_2^2,\rho_z) \triangleq \frac{1}{2}\log\frac{P}{\sigma_1^2} + \frac{1}{2}\log\frac{P}{\sigma_2^2}. \quad (19)$$

- For channels that are not physically degraded and where the noises are only partially correlated, i.e., $\rho_z \in (-1,1)$ and $\rho_z \notin \left\{\frac{\sigma_1}{\sigma_2},\frac{\sigma_2}{\sigma_1}\right\}$,[2]

$$C_{\text{Hi-SNR}}(P,\sigma_1^2,\sigma_2^2,\rho_z)$$
$$\triangleq \frac{1}{2}\log\left(\frac{P(\sigma_1^2+\sigma_2^2-2\rho_z\sigma_1\sigma_2)}{\sigma_1^2\sigma_2^2(1-\rho_z^2)}\right). \quad (20)$$

Theorem 1: For all $\sigma_1,\sigma_2 > 0$ and $\rho_z \in [-1,1]$

$$\lim_{P\to\infty}\left(C_{\text{BC},\Sigma}(P,\sigma_1^2,\sigma_2^2,\rho_z) - C_{\text{Hi-SNR}}(P,\sigma_1^2,\sigma_2^2,\rho_z)\right) = 0. \quad (21)$$

Proof: When the channel is physically degraded, the result follows from (17).

[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 (18) and (20) coincide.

When $\rho_z \in \{-1,1\}$ and the BC is not physically degraded, the desired rates are achieved by a novel scheme that we describe in Section II-C ahead (see Corollary 9 in Section II-C5). The converse for this case follows by applying the cut-set bound with two cuts, one between the transmitter and each of the two receivers:

$$R_1 + R_2 \leq \max_{X\,:\,\mathsf{E}[X^2]\leq P}\{I(X;Y_1)+I(X;Y_2)\}$$
$$= \frac{1}{2}\log\left(1+\frac{P}{\sigma_1^2}\right) + \frac{1}{2}\log\left(1+\frac{P}{\sigma_2^2}\right), \quad (22)$$

where the equality follows because a Gaussian law maximizes the differential entropy under a variance constraint [23].

When $\rho_z \in (-1,1)$ and the channel is not physically degraded, the achievability is demonstrated using our scheme of Section II-C (see Corollary 8 in Section II-C5); the converse follows by applying the cut-set bound with a single cut separating the transmitter from the two receivers:

$$R_1 + R_2 \leq \max_{X\,:\,\mathsf{E}[X^2]\leq P} I(X;Y_1,Y_2)$$
$$\leq \frac{1}{2}\log\left(1+\frac{P(\sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2\rho_z)}{\sigma_1^2\sigma_2^2(1-\rho_z^2)}\right). \quad (23)$$
∎

In general, the previously proposed schemes in [9], [10], and [13] cannot achieve the high-SNR asymptotic sum-rate capacity $C_{\text{Hi-SNR}}$: The scheme in [9], [10], for example, achieves $C_{\text{Hi-SNR}}$ when $\rho_z \leq 0$ but not when $\rho_z > 0$. And the scheme in [13, Theorem 2] achieves $C_{\text{Hi-SNR}}$ when $\rho_z = 0$ and $\sigma_1 = \sigma_2$, but it does not apply when $\sigma_1 \neq \sigma_2$.

Note 1: If $|\rho_z| < 1$ and the feedback links are noise-free, then the high-SNR sum-rate capacity $C_{\text{Hi-SNR}}(P,\sigma_1^2,\sigma_2^2,\rho_z)$ is as though the two receivers could fully cooperate in their decoding. If $\rho_z \in \{-1,1\}$ and the channel is not physically degraded, then the high-SNR sum-rate capacity is as though there were a separate (non-interfering) link from the transmitter to each of the receivers and the transmitter could communicate with full power P over each of these links.

Note 2: Given $\sigma_2,\sigma_1 > 0$, define the *power offset*

$$\gamma\colon (-1,1) \to \mathbb{R}^+ \quad (24)$$
$$\rho_z \mapsto \frac{\sigma_1^2+\sigma_2^2-2\rho_z\sigma_1\sigma_2}{\sigma_1^2\sigma_2^2(1-\rho_z^2)} \quad (25)$$

so

$$C_{\text{Hi-SNR}}(P,\sigma_1^2,\sigma_2^2,\rho_z) = \frac{1}{2}\log\big(P\,\gamma(\rho_z)\big), \quad |\rho_z| < 1.$$

Notice that $\gamma(\rho_z) \to \infty$ as $\rho_z \to -1$. Also, if $\sigma_1 \neq \sigma_2$, then $\gamma(\rho_z) \to \infty$ as $\rho_z \to +1$. Moreover, $\gamma(\rho_z)$ is strictly decreasing for $\rho_z \in \left(-1,\min\left\{\frac{\sigma_1}{\sigma_2},\frac{\sigma_2}{\sigma_1}\right\}\right)$ and strictly increasing for $\rho_z \in \left(\min\left\{\frac{\sigma_1}{\sigma_2},\frac{\sigma_2}{\sigma_1}\right\},1\right)$. The power offset is thus minimal at $\rho_z = \min\left\{\frac{\sigma_1}{\sigma_2},\frac{\sigma_2}{\sigma_1}\right\}$ where $\gamma(\rho_z) = (\min\{\sigma_1^2,\sigma_2^2\})^{-1}$. Unless σ_1 and σ_2 are equal, the power off-set is not monotonic over the interval $(-1,1)$. Figure 3 shows the typical behavior of $\gamma(\rho_z)$.
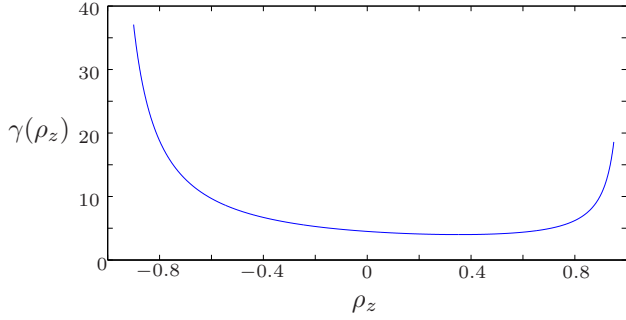
From Theorem 1 we obtain the following corollary.

Fig. 3. 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 $\left(-1, \sqrt{1/8}\right)$ and strictly increasing over $\left(\sqrt{1/8}, 1\right)$.
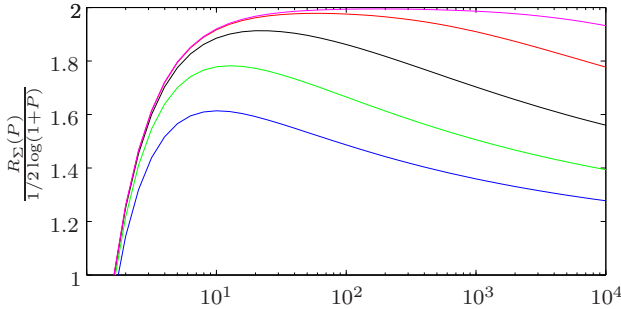

Fig. 4. The sum-rate $R_\Sigma(P)$ achieved by the scheme in Section II-C normalized by $\frac{1}{2}\log(1 + \frac{P}{\sigma_1^2})$ is plotted as a function of the power $P > 0$. The noise variances $\sigma_1^2 = \sigma_2^2 = 1$ are fixed, and the different curves correspond (in increasing order) to correlation coefficients $\rho_z = -0.85, -0.95, -0.99, -0.999 - 0.9999$.

Corollary 2: The prelog of the Gaussian BC with noise-free feedback is 2 if $\rho_z = -1$ or if $\rho_z = +1$ and $\sigma_1 \neq \sigma_2$; it is 1 otherwise:

$$\varlimsup_{P \to \infty} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1 + P)} \begin{cases} 2 & \rho_z = -1 \\ 2 & \rho_z = 1 \text{ and } \sigma_1^2 \neq \sigma_2^2 \\ 1 & \text{otherwise.} \end{cases} \quad (26)$$

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 $\{Y_{2,t}\}$. Indeed, for $\rho_z \in \{-1, 1\}$, the capacity regions with one-sided and two-sided noise-free feedback coincide: when $\rho_z \in \{-1, 1\}$ the transmitter can compute the output it does not observe from the input and the output it does observe.

Theorem 1 and Corollary 2 show that when $\rho_z \in \{-1, 1\}$ and the channel is not physically degraded, noise-free feedback approximately doubles the high-SNR sum-rate capacity. In Section II-C we present a coding scheme achieving these gains. (When $\rho_z = -1$ also the Ozarow-Leung scheme [9], [10] achieves such sum-rates [11], [12].) Figure 4 depicts the sum-rate achieved by the scheme of Section II-C as a function of the transmit power P, for various values of the correlation ρ_z. It shows that for large powers P (i.e., $P \geq 100$), noise-free feedback can nearly double the sum-rate 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 the power P and the correlation ρ_z that are required for noise-free feedback to roughly double the sum-rate capacity. 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 the noise variances $\sigma_1^2, \sigma_2^2 > 0$ be fixed. For every functional dependence $\rho_z(P)$ of the correlation coefficient on the power P, define

$$\zeta_{-1} \triangleq \varliminf_{P \to \infty} \frac{-\log(1 + \rho_z(P))}{\log(P)}, \quad (27)$$

$$\zeta_{+1} \triangleq \varliminf_{P \to \infty} \frac{-\log(1 - \rho_z(P))}{\log(P)}, \quad (28)$$

where $-\log 0 \triangleq \infty$. Notice that $\zeta_{-1} > 0$ only if $\varliminf_{P \to \infty} \rho_z(P) = -1$, and $\zeta_{+1} > 0$ only if $\varlimsup_{P \to \infty} \rho_z(P) = 1$.

Theorem 3 (Generalized Prelog with Noise-Free Feedback): The generalized prelog depends on whether or not the noise variances are equal. If $\sigma_1^2 = \sigma_2^2$, then the generalized prelog is

$$\varlimsup_{P \to \infty} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z(P))}{\frac{1}{2}\log(1 + P)} = \min\{1 + \zeta_{-1}, 2\} \quad (29)$$

and if $\sigma_1^2 \neq \sigma_2^2$, then the generalized prelog is

$$\varlimsup_{P \to \infty} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z(P))}{\frac{1}{2}\log(1 + P)}$$
$$= \min\left\{1 + \max\left\{\zeta_{-1}, \zeta_{+1}\right\}, 2\right\}. \quad (30)$$

Proof: See Appendix A. ∎

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

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

where

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

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

$$\varlimsup_{P \to \infty} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z(P))}{\frac{1}{2}\log(1 + P)} = 1 + \zeta. \quad (31)$$

The above results on the dramatic capacity gains afforded by feedback were predicated on the feedback being noise-free. Otherwise, as the next theorem shows, the gains are more moderate.

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

$$\varlimsup_{P \to \infty} \frac{C_{\text{BCNoisy},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z, \sigma_{W1}^2, \sigma_{W2}^2)}{\frac{1}{2}\log(1 + P)} = 1. \quad (32)$$

Proof: See Section II-D. ∎

Thus, if the feedback is noisy, then the prelog equals 1 also when the noise correlation ρ_z is ± 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 more slowly 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

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

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

Proof: See Appendix B. ∎

C. A Coding Scheme for Noise-Free Feedback

We present a "successive noise cancellation" coding strategy for the Gaussian BC with noise-free feedback. The scheme achieves the desired rates in Theorems 1 and 3; see Corollaries 7–9 and the proof of Theorem 3 in Appendix A.

Our scheme is similar in flavor to the schemes proposed by Cover and Pombra [4] for (non-white) Gaussian point-to-point channels with noise-free feedback, by Lapidoth & Wigger and Khisti & Lapidoth for the two-user Gaussian MAC with noisy feedback [5] or with intermittent feedback and side-information [6], and by Lapidoth, Steinberg, and Wigger [7] for the two-user Gaussian BC with one-sided noise-free feedback.

Before describing our scheme in Subsections II-C3 and II-C4 ahead, we first motivate it by sketching a simple scheme for the Gaussian point-to-point channel with noise-free feedback (Subsection II-C1) and a scheme for the two-user Gaussian BC with noise-free feedback when ρ_z is ± 1 (Subsection II-C2).

1) Motivation I: "Successive Noise Cancellation" scheme for the Gaussian point-to-point channel: Consider η transmissions over the standard memoryless Gaussian point-to-point channel

$$Y_t = x_t + Z_t, \tag{33}$$

where x_t and Y_t denote the time-t input and output, $\{Z_t\}$ is a sequence of IID zero-mean, variance $\sigma^2 > 0$ Gaussian random variables, and the inputs $\{x_t\}$ are subject to an expected average block-power constraint P. The transmitter is assumed to have access to noise-free feedback.

In the first channel use, the transmitted symbol is a unit-variance information-carrying symbol Ξ. In the subsequent $\eta - 1$ channel uses, the transmitted symbols are scaled versions of the preceding noise symbols (which are known thanks to the

feedback):

$$X_1 = \sqrt{P}\, \Xi \tag{34}$$

$$X_2 = \sqrt{\frac{P}{\sigma^2}}\, Z_1 \tag{35}$$

$$\vdots = \vdots$$

$$X_\eta = \sqrt{\frac{P}{\sigma^2}}\, Z_{\eta-1}. \tag{36}$$

Consider now a (suboptimal) receiver that replaces the η channel outputs $Y_1, \ldots, Y_\eta$ with the single random variable I, where

$$I = \sum_{\ell=1}^{\eta} \left(-\sqrt{\frac{P}{\sigma^2}}\right)^{\eta-\ell} Y_\ell \tag{37}$$

$$= \sqrt{P}\sqrt{\frac{P}{\sigma^2}}^{\eta-1} \Xi + Z_\eta. \tag{38}$$

From (38) we see that the noise samples $Z_1, Z_2, \ldots, Z_{\eta-1}$ have all been canceled out, and the only remaining noise sample is Z_η. The channel from Ξ to I is a Gaussian channel. Each use of this channel requires η transmissions on the original channel, so this scheme allows us to attain the rate

$$\frac{1}{2\eta} \log\left(1 + \left(\frac{P}{\sigma^2}\right)^\eta\right). \tag{39}$$

We obtain the largest achievable rate by choosing $\eta = 1$. However, when we extend the scheme to the Gaussian BC, we will be interested in the limit $\eta \to \infty$. In this limiting case, the described scheme achieves any rate $R > 0$ that satisfies

$$R < \frac{1}{2} \log^+\left(\frac{P}{\sigma^2}\right). \tag{40}$$

Thus, even though for $\eta \to \infty$ the described scheme does not achieve capacity, it nevertheless achieves prelog 1 and the gap to capacity vanishes as P tends to infinity.

The lesson from this example is that for high-SNR optimality it suffices to send "most of the time" past noise samples, rather than information symbols.

2) Motivation II: "Successive Noise Cancellation" scheme for the Gaussian BC when $\rho_z \in \{-1, 1\}$: Consider the two-receivers Gaussian BC with noise-free feedback and noise correlation $\rho_z \in \{-1, 1\}$. In this case,

$$\frac{Z_{1,t}}{\sigma_1} = \rho_z \frac{Z_{2,t}}{\sigma_2} \qquad \text{with probability 1}, \tag{41}$$

i.e., the noise samples at the two receivers are proportional. Consequently, using the successive noise cancellation scheme, the transmitter can simultaneously and asymptotically-optimally serve both receivers. To see how, let

the transmitted symbols be

$$X_1 = \gamma(\Xi_1 + \Xi_2) \tag{42}$$

$$X_2 = \gamma\left(\rho_z\sqrt{\frac{P}{\sigma_2^2}}\Xi_1 + \sqrt{\frac{P}{\sigma_1^2}}\Xi_2\right) + \sqrt{\frac{P}{\sigma_1^2}}Z_{1,1} \tag{43}$$

$$X_3 = \sqrt{\frac{P}{\sigma_1^2}}Z_{1,2} \tag{44}$$

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

$$X_\eta = \sqrt{\frac{P}{\sigma_1^2}}Z_{1,\eta-1} \tag{45}$$

where $\gamma > 0$ is a scaling factor that ensures that the sum of the powers of the first two inputs X_1 and X_2 does not exceed $2P$, and where now we have two information-carrying symbols: Ξ_1 is intended for Receiver 1 and Ξ_2 is intended for Receiver 2. We transmit the two information symbols Ξ_1 and Ξ_2 along the two signaling directions,

$$\mathbf{u}_1 = \begin{pmatrix} 1 & \rho_z\sqrt{P/\sigma_2^2} & 0 & \ldots & 0 \end{pmatrix}^\top \tag{46}$$

and

$$\mathbf{u}_2 = \begin{pmatrix} 1 & \sqrt{P/\sigma_1^2} & 0 & \ldots & 0 \end{pmatrix}^\top, \tag{47}$$

which are different whenever the BC is not physically degraded, i.e., whenever $\rho_z \notin \left\{\frac{\sigma_1}{\sigma_2}, \frac{\sigma_2}{\sigma_1}\right\}$.

Receiver 1 uses $Y_{1,1}, \ldots, Y_{1,\eta}$ to compute I_1, where

$$I_1 = \sum_{\ell=1}^{\eta} \left(-\sqrt{\frac{P}{\sigma_1^2}}\right)^{\eta-\ell} Y_{1,\ell}$$

$$= \gamma\left(-\sqrt{\frac{P}{\sigma_1^2}}\right)^{\eta-1}\left(1 - \rho_z\frac{\sigma_1}{\sigma_2}\right)\Xi_1 + Z_{1,\eta}, \tag{48}$$

and Receiver 2 uses $Y_{2,1}, \ldots, Y_{2,\eta}$ to compute I_2, where

$$I_2 = \sum_{\ell=1}^{\eta} \left(-\rho_z\sqrt{\frac{P}{\sigma_2^2}}\right)^{\eta-\ell} Y_{2,\ell}$$

$$= \gamma\left(-\sqrt{\frac{P}{\sigma_2^2}}\right)^{\eta-1}\left(1 - \rho_z\frac{\sigma_2}{\sigma_1}\right)\Xi_2 + Z_{2,\eta}. \tag{49}$$

In other words, each receiver projects its observed outputs onto a particular receive beam-forming vector:

$$\mathbf{v}_1 = \left(\left(-\sqrt{P/\sigma_1^2}\right)^{\eta-1} \quad \left(-\sqrt{P/\sigma_1^2}\right)^{\eta-2} \quad \ldots \quad 1\right)^\top \tag{50}$$

at Receiver 1 and

$$\mathbf{v}_2 = \left(\left(-\rho_z\sqrt{P/\sigma_2^2}\right)^{\eta-1} \quad \left(-\rho_z\sqrt{P/\sigma_2^2}\right)^{\eta-2} \quad \ldots \quad 1\right)^\top \tag{51}$$

at Receiver 2. These beam-forming vectors are different whenever the Gaussian BC is not physically degraded.

We see from (48) and (49) that the noise samples $Z_{1,1}, \ldots, Z_{1,\eta-1}$ and $Z_{2,1}, \ldots, Z_{2,\eta-1}$ are completely canceled and do not influence I_1 and I_2. Only the last noise samples $Z_{1,\eta}$ and $Z_{2,\eta}$ remain. Moreover, also the undesired information symbol is canceled out: Ξ_2 does not influence I_1, and Ξ_1 does not influence I_2. The channel from Ξ_1 to I_1 and the channel from Ξ_2 to I_2 are parallel Gaussian channels

(with dependent noises). We thus obtain from (48) and (49) the achievability of the rates

$$R_1 = \frac{1}{2\eta}\log\left(1 + \frac{\gamma^2}{P}\left(\frac{P}{\sigma_1^2}\right)^\eta\left(1 - \rho_z\frac{\sigma_1}{\sigma_2}\right)^2\right) \tag{52}$$

$$R_2 = \frac{1}{2\eta}\log\left(1 + \frac{\gamma^2}{P}\left(\frac{P}{\sigma_2^2}\right)^\eta\left(1 - \rho_z\frac{\sigma_2}{\sigma_1}\right)^2\right). \tag{53}$$

Notice that γ is nonzero and does not depend on η. Consequently, when the BC is not physically degraded, (16), and when η tends to infinity, the described scheme achieves all rate pairs (R_1, R_2) that satisfy (see also Corollary 9 ahead)

$$R_1 < \frac{1}{2}\log^+\left(\frac{P}{\sigma_1^2}\right) \tag{54}$$

$$R_2 < \frac{1}{2}\log^+\left(\frac{P}{\sigma_2^2}\right). \tag{55}$$

3) A General Scheme: The "Successive Noise Cancellation" schemes can be extended to the general Gaussian BC with noise-free feedback. The idea is to use the feedback to transform each block of η uses of the original scalar BC into a single use of a new MISO BC, which can be viewed as a BC with two transmit antennas and a single receive antenna for each receiver: the new BC's input is the vector $(\Xi_1, \Xi_2)^\top \in \mathbb{R}^2$, and its two scalar outputs are $I_1 \in \mathbb{R}$ at Receiver 1 and $I_2 \in \mathbb{R}$ at Receiver 2. We then code over this new BC ignoring the feedback. Scaling by η^{-1} any rate pair that is achievable on the new BC will yield a pair that is achievable on the original BC with feedback.

We next describe how to transform a block of η uses of the original scalar BC with feedback into a single use of the MISO BC. For simplicity, we restrict attention to the first block; the procedure for the subsequent blocks is analogous. The key parameters are the signaling vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, the coefficients according to which past noise symbols are retransmitted, which we collect into the strictly lower-triangular matrices $\mathsf{B}_1, \mathsf{B}_2$, and the receivers' beam-forming vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. (In the following Subsection II-C4, we present a specific choice for these parameters.)

We use the notation $\mathbf{X} \triangleq (X_1, \ldots, X_\eta)$, $\mathbf{Y}_1 \triangleq (Y_{1,1}, \ldots, Y_{1,\eta})$, and $\mathbf{Y}_2 \triangleq (Y_{2,1}, \ldots, Y_{2,\eta})$. The transmitter produces the η-length input vector $\mathbf{X}$ based on the information carrying symbols Ξ_1 and Ξ_2 and on the feedback signals it receives:[3]

$$\mathbf{X} = \Xi_1\mathbf{u}_1 + \Xi_2\mathbf{u}_2 + \mathsf{B}_1\mathbf{Z}_1 + \mathsf{B}_2\mathbf{Z}_2, \tag{56}$$

where $\mathbf{Z}_1 \triangleq (Z_{1,1}, \ldots, Z_{1,\eta})$ and $\mathbf{Z}_2 \triangleq (Z_{2,1}, \ldots, Z_{2,\eta})$, and where the assumption that B_1 and B_2 are lower-triangular guarantees that Ξ_1, Ξ_2, and the feedback signals suffice to compute $\mathbf{X}$.

The inputs to the original channel satisfy the average block-power constraint (11) whenever the information carrying symbols Ξ_1 and Ξ_2 are independent and satisfy $\mathsf{E}\left[\Xi_1^2\right], \mathsf{E}\left[\Xi_2^2\right] \leq 1$

---

[3]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.

and when

$$\|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + \text{tr}\left(\mathsf{B}_1\mathsf{B}_1^{\mathsf{T}}\right)\sigma_1^2 + \text{tr}\left(\mathsf{B}_2\mathsf{B}_2^{\mathsf{T}}\right)\sigma_2^2$$
$$+ 2\,\text{tr}\left(\mathsf{B}_1\mathsf{B}_2^{\mathsf{T}}\right)\rho_z\sigma_1\sigma_2 \le \eta P. \quad (57)$$

Receiver 1 observes

$$\mathbf{Y}_1 = \Xi_1\mathbf{u}_1 + \Xi_2\mathbf{u}_2 + (\mathsf{B}_1 + \mathsf{I})\,\mathbf{Z}_1 + \mathsf{B}_2\mathbf{Z}_2, \quad (58)$$

and computes

$$I_1 \triangleq \mathbf{v}_1^{\mathsf{T}}\mathbf{Y}_1. \quad (59)$$

Receiver 2 observes

$$\mathbf{Y}_2 = \Xi_1\mathbf{u}_1 + \Xi_2\mathbf{u}_2 + (\mathsf{B}_2 + \mathsf{I})\,\mathbf{Z}_2 + \mathsf{B}_1\mathbf{Z}_1, \quad (60)$$

and computes

$$I_2 \triangleq \mathbf{v}_2^{\mathsf{T}}\mathbf{Y}_2. \quad (61)$$

4) Choice of Parameters and Achievable Rates: Given η, we describe a choice of the parameters $\mathsf{B}_1, \mathsf{B}_2, \mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2$.

Choose $q > 0$ and $\delta \notin \{-1, 0\}$ to satisfy Equation (62) (shown at the top of the next page), and define

$$a_1 \triangleq q \quad (63)$$
$$a_2 \triangleq -\delta^2 q \quad (64)$$
$$b_1 \triangleq -\delta(1+\delta)q^2 \quad (65)$$
$$b_2 \triangleq -\delta^2(1+\delta)q^2. \quad (66)$$

Choose the $\eta \times \eta$ matrices B_1 and B_2 to be Toeplitz with non-zero entries only on the first and second diagonals below the main diagonal:

$$\mathsf{B}_k = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ a_k & 0 & 0 & \cdots & 0 & 0 & 0 \\ b_k & a_k & 0 & \cdots & 0 & 0 & 0 \\ 0 & b_k & a_k & 0 & \cdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & b_k & a_k & 0 & 0 \\ 0 & \cdots & \cdots & 0 & b_k & a_k & 0 \end{pmatrix}, \quad (67)$$

and choose the η-dimensional vectors

$$\mathbf{u}_1 = \sqrt{\frac{P}{2 + 2\frac{b_1^2}{a_1^2}}}\begin{pmatrix} 1 & \frac{b_1}{a_1} & 0 & \cdots & 0 \end{pmatrix}^{\mathsf{T}} \quad (68a)$$

$$\mathbf{u}_2 = \sqrt{\frac{P}{2 + 2\frac{b_2^2}{a_2^2}}}\begin{pmatrix} 1 & \frac{b_2}{a_2} & 0 & \cdots & 0 \end{pmatrix}^{\mathsf{T}} \quad (68b)$$

and

$$\mathbf{v}_1 = \left(\left(-\frac{b_2}{a_2}\right)^{\eta-1} \quad \left(-\frac{b_2}{a_2}\right)^{\eta-2} \quad \cdots \quad -\frac{b_2}{a_2} \quad 1\right)^{\mathsf{T}} \quad (69a)$$

$$\mathbf{v}_2 = \left(\left(-\frac{b_1}{a_1}\right)^{\eta-1} \quad \left(-\frac{b_1}{a_1}\right)^{\eta-2} \quad \cdots \quad -\frac{b_1}{a_1} \quad 1\right)^{\mathsf{T}}. \quad (69b)$$

By (62)–(68), this choice satisfies the power constraint (57). 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, the noise samples $Z_{1,1}, \ldots, Z_{1,\eta-2}$ and $Z_{2,1}, \ldots, Z_{2,\eta-2}$ are completely canceled out when forming the "new outputs" in (59) and only the noise samples $Z_{1,\eta-1}, Z_{1,\eta}, Z_{2,\eta-1}, Z_{2,\eta}$ remain. Moreover, the "interference" Ξ_2 is canceled in I_1 and the "interference" Ξ_1 is canceled in I_2. In fact,

$$I_1 = \sqrt{\frac{P}{2 + 2\frac{b_1^2}{a_1^2}}}\left(-\frac{b_2}{a_2}\right)^{\eta-1}\left(1 - \frac{a_2}{b_2}\frac{b_1}{a_1}\right)\Xi_{1,i}$$
$$+ \left(-\frac{b_2}{a_2} + a_1\right)Z_{1,\eta-1} + Z_{1,\eta} + a_2 Z_{2,\eta-1} \quad (70a)$$

and

$$I_2 = \sqrt{\frac{P}{2 + 2\frac{b_2^2}{a_2^2}}}\left(-\frac{b_1}{a_1}\right)^{\eta-1}\left(1 - \frac{a_1}{b_1}\frac{b_2}{a_2}\right)\Xi_{2,i}$$
$$+ \left(-\frac{b_1}{a_1} + a_2\right)Z_{2,\eta-1} + Z_{2,\eta} + a_1 Z_{1,\eta-1}. \quad (70b)$$

Over the original Gaussian BC with feedback we can achieve the scaled-by-η^{-1} capacity of the new MISO BC (70), and we thus have the following proposition.

Proposition 6: The noise-free feedback scheme of Section II-C3 with the choice of parameters presented here in Section II-C4, achieves all rate pairs (R_1, R_2) for which

$$R_1 \le \frac{1}{2\eta}\log\left(1 + \frac{\frac{P(1+\delta)^2}{2+2q^2\delta^2(1+\delta)^2}\left(q^2(1+\delta)^2\right)^{\eta-1}}{(q^2\delta^2 + 1)\sigma_1^2 + q^2\delta^4\sigma_2^2}\right) \quad (71a)$$

$$R_2 \le \frac{1}{2\eta}\log\left(1 + \frac{\frac{P(1+1/\delta^2)^2}{2+2q^2(1+\delta)^2}\left(q^2\delta^2(1+\delta)^2\right)^{\eta-1}}{(q^2\delta^2 + 1)\sigma_2^2 + q^2\sigma_1^2}\right) \quad (71b)$$

simultaneously hold for some real numbers $\delta \notin \{-1, 0\}$ and q such that (62) holds.

The choice of parameters in (63)–(69) that leads to Proposition 6 is, in general, sub-optimal; better choices can be found in [8]. However, whenever the BC is not physically degraded, the rates in Proposition 6 achieve the asymptotic high-SNR sum-rate capacity with noise-free feedback (Theorem 1); see Corollaries 8 and 9 ahead. Moreover, they also achieve the generalized prelog in Theorem 3. In fact, Corollary 7 ahead suffices to prove the achievability of Theorem 3, as is shown in Appendix A.

5) High-SNR Performance: By the following lemma, $\eta = 1$ maximizes the constraints in (71) for small powers P and $\eta \to \infty$ maximizes them for large powers P.

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 C. ∎

Letting $\eta \to \infty$, we obtain the following corollary to Proposition 6.

Corollary 7: All nonnegative rate-pairs (R_1, R_2) that satisfy

$$R_1 < \frac{1}{2}\log^+\left(q^2(1+\delta)^2\right) \quad (72a)$$

$$R_2 < \frac{1}{2}\log^+\left(q^2\delta^2(1+\delta)^2\right), \quad (72b)$$

$$q^2(\sigma_1^2 + \delta^4\sigma_2^2 - 2\delta^2\rho_z\sigma_1\sigma_2) + q^4(1+\delta)^2\delta^2(\sigma_1^2 + \delta^2\sigma_2^2 + 2\delta\rho_z\sigma_1\sigma_2) \le P. \qquad (62)$$

for some real numbers $\delta \notin \{-1, 0\}$ and q such that (62) holds, are achievable over the Gaussian BC with noise-free feedback.

From Corollary 7 with an appropriate choice of the parameters $\delta \notin \{-1, 0\}$ and q we further obtain:

Corollary 8: If $\rho_z \in (-1, 1)$, then for 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

$$R_1 + R_2 = \frac{1}{2}\log^+\left(\frac{(1-\epsilon)P(\sigma_1^2 + \sigma_2^2 - 2\rho_z\sigma_1\sigma_2)}{\sigma_1^2\sigma_2^2(1-\rho_z^2)}\right) \quad (73)$$

is achievable over the Gaussian BC with noise-free feedback when the allowed power P exceeds $P_0(\epsilon, \sigma_1^2, \sigma_2^2, \rho_z)$.

Proof: By choosing

$$\delta = \frac{\sigma_1}{\sigma_2} \cdot \frac{\sigma_1 - \rho_z\sigma_2}{\sigma_2 - \rho_z\sigma_1} \qquad (74a)$$

$$q = \left(\frac{(1-\epsilon)P}{\delta^2(1+\delta)^2(\sigma_1^2 + \delta^2\sigma_2^2 + 2\delta\rho_z\sigma_1\sigma_2)}\right)^{1/4}, \quad (74b)$$

in Corollary 7 and in power constraint (62). ∎

Corollary 9: If $\rho_z \in \{-1, 1\}$, 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

$$R_1 < \frac{1}{2}\log^+\left(\frac{P}{\sigma_1^2}\right) \qquad (75a)$$

$$R_2 < \frac{1}{2}\log^+\left(\frac{P}{\sigma_2^2}\right) \qquad (75b)$$

are achievable over the Gaussian BC with noise-free feedback.

Proof: Follows from Corollary 7 by choosing

$$\delta = -\rho_z\frac{\sigma_1}{\sigma_2} \qquad (76a)$$

$$q = \left(\frac{P}{\sigma_1^2\left(1 - \frac{\sigma_1}{\sigma_2}\rho_z\right)^2}\right)^{1/2} \qquad (76b)$$

and verifying the power constraint (62). ∎

Note 5: Specializing the rate constraints in (71) to the choice in (76) we conclude that when $\rho_z \in \{-1, 1\}$ all rate-pairs (R_1, R_2) that satisfy

$$R_1 \le \frac{1}{2\eta}\log\left(1 + \left(\frac{P}{\sigma_1^2}\right)^{\eta-1}\left(1 - \rho_z\frac{\sigma_1}{\sigma_2}\right)^2\frac{P/\sigma_1^2}{2 + 2P/\sigma_2^2}\right) \quad (77a)$$

$$R_2 \le \frac{1}{2\eta}\log\left(1 + \left(\frac{P}{\sigma_2^2}\right)^{\eta-1}\left(1 - \rho_z\frac{\sigma_2}{\sigma_1}\right)^2\frac{P/\sigma_2^2}{2 + 2P/\sigma_1^2}\right) \quad (77b)$$

for some positive integer η, are achievable. Consequently, for fixed η, when $\rho_z \in \{-1, 1\}$ and the BC is not physically degraded, our scheme achieves prelog $2\frac{\eta-1}{\eta}$. Thus, $\eta = 3$ suffices to increase the prelog compared to the non-feedback setup. Also, when $\eta \to \infty$ the scheme can achieve prelog 2.

D. Proof of Theorem 5 (Prelog with Noisy Feedback)

The interesting part is the converse, which we prove using a genie-argument inspired by [14]. It is based on the following three steps. 1.) We introduce a *genie-aided Gaussian BC* without feedback and show that its sum-rate capacity upper bounds the sum-rate capacity of the original Gaussian BC with noisy feedback. 2.) We introduce a *less noisy Gaussian BC* with neither genie-information nor feedback and show that its sum-rate capacity coincides with the sum-rate capacity of the genie-aided Gaussian BC. 3.) We show that the prelog of the less noisy Gaussian 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 next elaborate on these steps starting with the first. The genie-aided Gaussian BC is defined as the original Gaussian 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 compute the missing feedback outputs $V_{1,t}$ and $V_{2,t}$:

$$V_{1,t} = X_t + (Z_{1,t} + W_{1,t}), \qquad (78)$$

$$V_{2,t} = X_t + (Z_{2,t} + W_{2,t}). \qquad (79)$$

Consequently, the sum-rate capacity of the genie-aided Gaussian BC is at least as large as the sum-rate capacity of the original Gaussian BC with feedback.

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

$$Y'_{1,t} = x_t + Z'_{1,t}, \qquad (80)$$

$$Y'_{2,t} = x_t + Z'_{2,t}, \qquad (81)$$

where the reduced noise samples $Z'_{1,t}$, and $Z'_{2,t}$ are defined as

$$Z'_{1,t} \triangleq Z_{1,t} - \mathsf{E}[Z_{1,t}|(Z_{1,t} + W_{1,t}), (Z_{2,t} + W_{2,t})], \quad (82)$$

$$Z'_{2,t} \triangleq Z_{2,t} - \mathsf{E}[Z_{2,t}|(Z_{1,t} + W_{1,t}), (Z_{2,t} + W_{2,t})], \quad (83)$$

and are of variances

$$\mathsf{Var}(Z'_{1,t}) = \sigma_1^2 \frac{\sigma_{W1}^2\sigma_2^2(1-\rho_z^2) + \sigma_{W1}^2\sigma_{W2}^2}{(\sigma_1^2 + \sigma_{W1}^2)(\sigma_2^2 + \sigma_{W2}^2) - \sigma_1^2\sigma_2^2\rho_z^2}, \quad (84)$$

$$\mathsf{Var}(Z'_{2,t}) = \sigma_2^2 \frac{\sigma_{W2}^2\sigma_1^2(1-\rho_z^2) + \sigma_{W2}^2\sigma_{W1}^2}{(\sigma_1^2 + \sigma_{W1}^2)(\sigma_2^2 + \sigma_{W2}^2) - \sigma_1^2\sigma_2^2\rho_z^2}. \quad (85)$$

By the following two observations, the sum-rate capacity of this less noisy Gaussian BC coincides with the sum-rate capacity of the genie-aided Gaussian BC. The first is that the sum-rate capacity of the less noisy Gaussian 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. Indeed, by (82)–(83) 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 it thus plays only the role of common

randomness, which does not increase capacity. The second observation is that the sum-rate capacity of the genie-aided Gaussian BC coincides with the sum-rate capacity of the less noisy Gaussian BC, when in this latter case the transmitter and both receivers additionally know the genie-information $\{(Z_{1,t} + W_{1,t})\}$ and $\{(Z_{2,t} + W_{2,t})\}$. Indeed, 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 Gaussian BC is a classical Gaussian BC with neither feedback nor genie-information, and its sum-rate capacity is [21]

$$C_{\text{BCLessNoisy},\Sigma}\big(P, \sigma_1^2, \sigma_2^2, \rho_z, \sigma_{W1}^2, \sigma_{W2}^2\big)$$
$$= \frac{1}{2} \log \left(1 + \frac{P}{\min\left\{ \mathsf{Var}\big(Z'_{1,t}\big), \mathsf{Var}\big(Z'_{2,t}\big) \right\}} \right), \quad (86)$$

where the variances $\mathsf{Var}\big(Z'_{1,t}\big), \mathsf{Var}\big(Z'_{2,t}\big)$ are defined in (84) and (85). By (84)–(86) the prelog of the less noisy Gaussian 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.

III. K-USER BROADCAST CHANNEL

A. Setup and Results

We next extend the model of Section II by allowing the number of receivers K to exceed two. We assume noise-free feedback. For each $k \in \{1, \ldots, K\}$ we denote the message intended for Receiver k by M_k, and we assume that it is uniformly distributed over $\mathcal{M}_k \triangleq \{1, \ldots, \lfloor 2^{nR_k} \rfloor\}$ and that $M_1, \ldots, M_K$ are independent. The time-t symbol observed by Receiver k is

$$Y_{k,t} = x_t + Z_{k,t}, \qquad t \in \{1, \ldots, n\}, \quad (87)$$

where x_t is the time-t transmitted symbol, and $Z_{k,t}$ is the time-t noise sample at Receiver k. We assume that $\{(Z_{1,t}, \ldots, Z_{K,t})^\mathsf{T}\}_{t=1}^n$ is a sequence of IID centered Gaussian vectors of covariance matrix K_z and that this sequence is independent of the messages $(M_1, \ldots, M_K)$. We denote the variance of the noise at the k-th receiver by σ_k^2 and the standard deviation by σ_k. We assume that the standard deviations are all strictly positive

$$\sigma_k > 0, \qquad k \in \{1, \ldots, K\}. \quad (88)$$

Based on the messages and the feedback signals, the transmitter produces the time-t channel input

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

using encoding functions of the form

$$f_{\text{K-BC},t}^{(n)} \colon \mathcal{M}_1 \times \cdots \times \mathcal{M}_K \times \mathbb{R}^{K(t-1)} \to \mathbb{R} \quad (89)$$

that are constrained to produce channel inputs $X_1, \ldots, X_n$ satisfying the expected average block-power constraint (11).

Based on its received sequence $Y_{k,1}, \ldots, Y_{k,n}$, Receiver k forms the guess $\hat{M}_k$ of M_k. We say that an error occurred whenever at least one of the receivers errs, i.e., whenever

$$(M_1, \ldots, M_K) \neq (\hat{M}_1, \ldots, \hat{M}_K). \quad (90)$$

Achievable rate-tuples, the capacity region, the sum-rate capacity, and the prelog are defined as in the two-receiver case. We denote the sum-rate capacity by $C_{\text{K-BC},\Sigma}(P, \mathsf{K}_z)$.

Our main result for this model is the prelog when

$$\text{rank}(\mathsf{K}_z) = 1. \quad (91)$$

This case is the K-receivers analog of the two-receiver setting with noise correlation ± 1. In this case, the noise samples $Z_{1,t}, \ldots, Z_{K,t}$ are all multiples of each other, and we can rewrite the channel law as:

$$Y_{k,t} = X_t + \rho_{1,k} \sigma_k Z_{1,t}, \qquad t \in \{1, \ldots, n\}, \quad (92)$$

where $\sigma_k > 0$ (by (88)), and $\rho_{1,k}$ denotes the correlation coefficient between $Z_{1,t}$ and $Z_{k,t}$, which, by (91), is either -1 or 1:

$$\rho_{1,k} \in \{-1, +1\}, \qquad k \in \{1, \ldots, K\}. \quad (93)$$

Define

$$\alpha_k \triangleq \rho_{1,k} \sigma_k, \quad k \in \{1, \ldots, K\}, \quad (94)$$

and note that by (88) and (93)

$$\alpha_k \neq 0, \qquad k \in \{1, \ldots, K\}. \quad (95)$$

As we shall see, when K_z is of rank 1, the prelog depends on the number n_α of α_k's that are different:

$$n_\alpha = \text{cardinality of } \{\alpha_1, \ldots, \alpha_K\}.$$

Notice that n_α is also the number of noise samples $\{Z_{1,t}, \ldots, Z_{K,t}\}$ that are not exactly the same, but that differ by a constant factor not equal to 1. It is also equal to the number of different rows (or columns) in K_z.

Theorem 10: If all the noises are of positive variance (88), and if the covariance matrix K_z has rank 1, then the prelog is n_α:

$$\varlimsup_{P \to \infty} \frac{C_{\text{K-BC},\Sigma}(P, \mathsf{K}_z)}{\frac{1}{2} \log(1 + P)} = n_\alpha. \quad (96)$$

Proof: Since there are only n_α different channels, the prelog cannot exceed n_α. It thus only remains to prove achievability.

If $n_\alpha = K$, then a prelog n_α is achievable using the scheme presented in Section III-B ahead; see Proposition 12 at the end of that section. If $n_\alpha < K$, then we pick n_α receivers such that the corresponding α's are all different. We then apply the scheme in Section III-B to only these n_α receivers (and ignore the other receivers). ∎

Corollary 11: When K_z is of rank 1 and all its rows are different (i.e., $n_\alpha = K$), the prelog is K.

The achievability of prelog K, for $K \geq 3$, was proved in [13] for the *complex* memoryless Gaussian broadcast channel. In [13], however, it is assumed that the real and imaginary parts of the noise symbols are correlated. Consequently, Theorem 10 is not implied by [13].

Note 6: Theorem 10 remains valid also when the transmitter has feedback only from a single receiver. The proof is analogous to the proof of Note 3.

B. A Scheme for the case where $\text{rank}(\mathsf{K}_z) = 1$ *and* $n_\alpha = K$

We generalize the coding scheme of Section II-C to the case where there are $K \geq 2$ receivers. We focus on the case where K_z has rank 1 and

$$n_\alpha = K, \qquad (97)$$

i.e.,

$$\alpha_k \neq \alpha_{k'}, \quad \left(k, k' \in \{1, \ldots, K\}, \, k \neq k'\right). \qquad (98)$$

The idea is to exploit the feedback in order to transform each block of η uses of the original scalar BC into a single use of a new MISO BC with vector input $(\Xi_1, \ldots, \Xi_K)^\mathsf{T} \in \mathbb{R}^K$ and scalar outputs $I_1, \ldots, I_K$ at Receivers $1, \ldots, K$, and to then code over these blocks (ignoring the feedback). This allows us to achieve on the original scalar BC the scaled-by-η^{-1} capacity of the new MISO BC (without feedback).

We next describe how to transform the first block of η uses of the original BC into a single use of the new MISO BC; subsequent η-length blocks are transformed similarly. The key parameters are: 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 column-vectors $\mathbf{v}_1, \ldots, \mathbf{v}_K$. How to choose these parameters will be discussed later.

The transmitter produces the η-length vector of inputs

$$\mathbf{X} = \sum_{k=1}^{K} \Xi_k \mathbf{u}_k + \sum_{k=1}^{K} \mathsf{B}_k \alpha_k \mathbf{Z}_1 \qquad (99)$$

where $\mathbf{Z}_1 \triangleq (Z_{1,1}, \ldots, Z_{1,\eta})^\mathsf{T}$ are the first η samples of the noise experienced by Receiver 1, which can be computed strictly-causally by the transmitter thanks to the feedback. Using (92) and (94), we can express the channel outputs $\mathbf{Y}_k \triangleq (Y_{k,1}, \ldots, Y_{k,\eta})^\mathsf{T}$ observed by Receiver $k \in \{1, \ldots, K\}$ as

$$\mathbf{Y}_k = \sum_{k'=1}^{K} \Xi_k \mathbf{u}_k + \left(\sum_{k'=1}^{K} \mathsf{B}_{k'} \alpha_{k'} + \mathsf{I}\alpha_k\right) \mathbf{Z}_1. \qquad (100)$$

Based on these outputs, Receiver k computes its new scalar output

$$I_k \triangleq \mathbf{v}_k^\mathsf{T} \mathbf{Y}_k. \qquad (101)$$

The channel input sequence satisfies the average block-power constraint over the block of length η whenever the information carrying symbols $\Xi_1, \ldots, \Xi_K$ are independent; they satisfy $\mathsf{E}\left[\Xi_k^2\right] \leq 1$ for every $k \in \{1, \ldots, K\}$; and

$$\sum_{k=1}^{K} \|\mathbf{u}_k\|^2 + \text{tr}\left(\left(\sum_{k=1}^{K} \mathsf{B}_k \alpha_k\right)\left(\sum_{k=1}^{K} \mathsf{B}_k^\mathsf{T} \alpha_k\right)\right) \leq \eta P. \qquad (102)$$

For every integer $\eta \geq K$, we next present a choice of the parameters $\mathsf{B}_1, \ldots, \mathsf{B}_K, \mathbf{u}_1, \ldots, \mathbf{u}_K$, and $\mathbf{v}_1, \ldots, \mathbf{v}_K$ with the following properties:

(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) each inner product $\mathbf{v}_k^\mathsf{T} \mathbf{u}_k$ is proportional to $\left(\sqrt{P}/\alpha_k\right)^{\eta - K}$, where the proportionality factor is nonzero and does not depend on η or P; and

(iv) the power constraint (102) is satisfied for all $P \geq K$.

Properties (i) and (ii) guarantee that the new scalar output I_k formed at Receiver k has the form

$$I_k = \mathbf{v}_k^\mathsf{T} \mathbf{u}_k \Xi_k + \alpha_k Z_{1,\eta}, \qquad (103)$$

i.e., the first $\eta - 1$ noise symbols $Z_{1,1}, \ldots, Z_{1,\eta-1}$ and the interference symbols $\{\Xi_{k'}\}_{k' \neq k}$ are completely canceled out. As explained in more detail later, Property (iii) guarantees that, when $\eta \to \infty$, our scheme achieves prelog 1 to each Receiver $k \in \mathcal{K}$.

To describe our choice of the parameters, we need definitions (104) and (105) ahead. Define for each $k \in \{1, \ldots, K\}$ the K-dimensional column-vector

$$\boldsymbol{\alpha}_k = \begin{pmatrix} 1 & \alpha_k & \alpha_k^2 & \alpha_k^3 & \cdots & \alpha_k^K \end{pmatrix}^\mathsf{T}. \qquad (104)$$

Let $\hat{\mathbf{w}}_k$ be the projection of $\boldsymbol{\alpha}_k$ onto the linear subspace spanned by $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{k-1}, \boldsymbol{\alpha}_{k+1}, \ldots, \boldsymbol{\alpha}_K\}$, and define

$$\mathbf{w}_k \triangleq \frac{\boldsymbol{\alpha}_k - \hat{\mathbf{w}}_k}{\|\boldsymbol{\alpha}_k - \hat{\mathbf{w}}_k\|}. \qquad (105)$$

Note that the vectors $\{\boldsymbol{\alpha}_k\}$ and $\{\mathbf{w}_k\}$ do not depend on P or η.

For every $\eta \geq K$, choose the matrices $\mathsf{B}_1, \ldots, \mathsf{B}_K$ so that

$$\sum_{k'=1}^{K} \mathsf{B}_{k'} \alpha_{k'} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0 \\ \sqrt{P} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \sqrt{P} & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \sqrt{P} & 0 & 0 \\ 0 & \cdots & 0 & 0 & \sqrt{P} & 0 \end{pmatrix}, \qquad (106)$$

and choose for $k \in \{1, \ldots, K\}$:

$$\mathbf{u}_k = \left(\begin{pmatrix} \frac{w_{k,1}}{\sqrt{P}^{K-1}} & \frac{w_{k,2}}{\sqrt{P}^{K-2}} & \cdots & \frac{w_{k,K-1}}{\sqrt{P}} & w_{k,K} & 0 & \cdots & 0 \end{pmatrix}\right)^\mathsf{T} \qquad (107)$$

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

$$\mathbf{v}_k = \left(\begin{pmatrix} \left(-\frac{\sqrt{P}}{\alpha_k}\right)^{\eta-1} & \left(-\frac{\sqrt{P}}{\alpha_k}\right)^{\eta-2} & \cdots & -\frac{\sqrt{P}}{\alpha_k} & 1 \end{pmatrix}\right)^\mathsf{T}. \qquad (108)$$

We next verify that this parameter choice satisfies Properties (i)–(iv). Property (i) follows from (106) and (108). By (95) and (98) the vectors $\{\boldsymbol{\alpha}_k\}_{k=1}^{K}$ are linearly independent. Therefore, by (105) and by the definition of the vectors $\{\hat{\mathbf{w}}_k\}$, each vector $\mathbf{w}_k$ is orthogonal to $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{k-1}, \boldsymbol{\alpha}_{k+1}, \ldots, \boldsymbol{\alpha}_K\}$ but not to $\boldsymbol{\alpha}_k$:

$$\boldsymbol{\alpha}_{k'}^\mathsf{T} \mathbf{w}_k = 0, \qquad k' \in \{1, \ldots, k-1, k+1, \ldots, K\}, \quad (109a)$$

and

$$\boldsymbol{\alpha}_k^\mathsf{T} \mathbf{w}_k \neq 0. \qquad (109b)$$

Property (ii) follows now by (109b) and because by (107) and (108):

$$\mathbf{v}_{k'}^\mathsf{T} \mathbf{u}_k = \frac{\sqrt{P}^{\eta-K}}{\alpha_k^{\eta-1}} \boldsymbol{\alpha}_{k'}^\mathsf{T} \mathbf{w}_k, \qquad k, k' \in \{1, \ldots, K\} \quad (110)$$

Property (iii) follows by (109) and (110), and because $\boldsymbol{\alpha}_k$ and $\mathbf{w}_k$ do not depend on η or P. Finally, the last Property (iv)

(the power constraint for $P > K$) follows by combining (106) with (107) and because, by definition (105), $\|\mathbf{w}_k\|^2 = 1$.

We conclude that the new output I_k at Receiver k is of the form in (103). Therefore, by (110), when $P > K$ our scheme with the described choice of parameters achieves all rate tuples $(R_1, \ldots, R_K)$ that satisfy

$$R_k \leq \frac{1}{2\eta} \log \left(1 + \frac{P^{\eta-K}(\boldsymbol{\alpha}_k^\mathsf{T}\mathbf{w}_k)^2}{\alpha_k^{2\eta}} \right), \quad k \in \{1, \ldots, K\}, \tag{111}$$

for some $\eta \geq K$. Since $\boldsymbol{\alpha}_k^\mathsf{T}\mathbf{w}_k$ is not zero (109b) and does not depend on η, by letting η tend to infinity, we obtain from (111):

Proposition 12: If $\mathrm{rank}(\mathsf{K}_z) = 1$ and $n_\alpha = K$ and if the power constraint $P > K$, then with noise-free feedback all nonnegative rate-tuples $(R_1, \ldots, R_K)$ satisfying

$$R_k < \frac{1}{2} \log^+ \left(\frac{P}{\alpha_k^2} \right), \quad k \in \{1, \ldots, K\}, \tag{112}$$

are achievable over the K-user Gaussian BC.

Thus, for each Receiver k, we can achieve prelog 1; and therefore we achieve prelog K for the sum-rate.

IV. TWO-USER INTERFERENCE CHANNEL
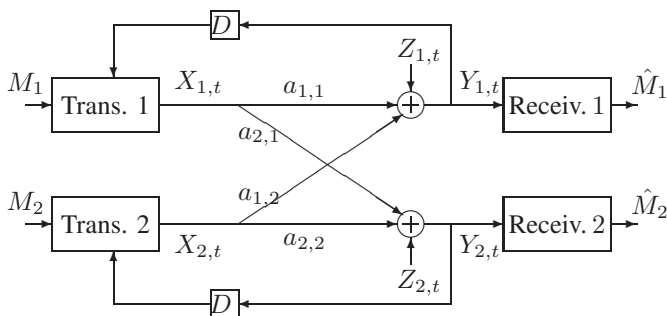
A. Setup and Results

Fig. 5.   The two-user Gaussian IC with one-sided noise-free feedback.

In this section we study the real scalar memoryless Gaussian IC with noise-free feedback, which is depicted in Figure 5. This network has two transmitters and two receivers: Transmitter 1 wishes to send Message M_1 to Receiver 1, and Transmitter 2 wishes to send 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

$$Y_{1,t} = a_{1,1}x_{1,t} + a_{1,2}x_{2,t} + Z_{1,t}, \tag{113a}$$

and Receiver 2 observes

$$Y_{2,t} = a_{2,1}x_{1,t} + a_{2,2}x_{2,t} + Z_{2,t}. \tag{113b}$$

The channel gains are non-zero real constants

$$a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2} \neq 0, \tag{114}$$

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

Each transmitter has access to noise-free feedback from its intended receiver. Thus, each transmitter can choose its time-t channel input as

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

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

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

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

$$\frac{1}{n}\mathsf{E}\left[\sum_{t=1}^{n} X_{k,t}^2\right] \leq P, \quad k \in \{1, 2\}. \tag{115}$$

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

Without feedback, the prelog of the Gaussian 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 13: The prelog of the Gaussian IC with noise-free feedback satisfies the following three statements.

- If $|\rho_z| < 1$ or if $|\rho_z| = 1$ and $\frac{a_{2,2}}{a_{1,2}} = \frac{a_{1,2}}{a_{1,1}} = \rho_z\frac{\sigma_2}{\sigma_1}$, then

$$\varliminf_{P \to \infty} \frac{C_{\mathrm{IC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1 + P)} = 1; \tag{116}$$

- if $|\rho_z| = 1$ and neither $\frac{a_{2,2}}{a_{1,2}}$ nor $\frac{a_{1,2}}{a_{1,1}}$ equals $\rho_z\frac{\sigma_2}{\sigma_1}$, then

$$\varliminf_{P \to \infty} \frac{C_{\mathrm{IC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1 + P)} = 2; \tag{117}$$

- otherwise

$$1 \leq \varliminf_{P \to \infty} \frac{C_{\mathrm{IC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1 + P)} \leq 2. \tag{118}$$

Proof: See Section IV-C. ∎

B. A Scheme

We present a scheme similar to the scheme for the BC in Section II-C. Thus, the idea is to transform each block of η channel uses of the original IC into a single use of a new IC with inputs $\Xi_1 \in \mathbb{R}$ at Transmitter 1 and $\Xi_2 \in \mathbb{R}$ at Transmitter 2 and with outputs I_1 at Receiver 1 and I_2 at Receiver 2, and to then code over this new IC. In this way we can achieve on the original IC the scaling-by-η^{-1} of any rate pair that is achievable on the new IC.

We describe how to transform the first block of η uses of the original IC into a single use of the new IC; subsequent blocks are transformed similarly. The key parameters are: 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 $\mathbf{v}_1, \mathbf{v}_2$.

Denoting by

$$\mathbf{X}_k \triangleq (X_{k,1}, \ldots, X_{k,\eta})^\mathsf{T}, \quad k \in \{1, 2\}, \tag{119}$$

the η-length vector of symbols that Transmitter k sends in this first block, we choose

$$\mathbf{X}_1 = \mathbf{u}_1\Xi_1 + \mathsf{B}_1(a_{1,2}\mathbf{X}_2 + \mathbf{Z}_1) \tag{120a}$$

$$\mathbf{X}_2 = \mathbf{u}_2\Xi_2 + \mathsf{B}_2(a_{2,1}\mathbf{X}_1 + \mathbf{Z}_2). \tag{120b}$$

Receiver k observes the corresponding η-length vector of outputs $\mathbf{Y}_k$ and computes the new output

$$I_k = \mathbf{v}_k^\top\mathbf{Y}_k. \tag{121}$$

In the following, we present a choice of parameters for the case where $\rho_z \in \{-1, 1\}$. In this case,

$$Z_{2,t} = \rho_z\frac{\sigma_2}{\sigma_1}Z_{1,t}, \tag{122}$$

and we can rewrite the channel outputs at Receiver 2 as

$$Y_{2,t} = a_{2,1}X_{1,t} + a_{2,2}X_{2,t} + \rho_z\frac{\sigma_2}{\sigma_1}Z_{1,t}, \tag{123a}$$

We choose B_2 the all-zero matrix and

$$\mathsf{B}_1 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ \frac{\sqrt{P}}{\sigma_1} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \frac{\sqrt{P}}{\sigma_1} & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \frac{\sqrt{P}}{\sigma_1} & 0 & 0 \\ 0 & 0 & \cdots & 0 & \frac{\sqrt{P}}{\sigma_1} & 0 \end{pmatrix}. \tag{124}$$

By (120a) and (123), with this choice

$$\mathbf{Y}_1 = a_{1,1}\mathbf{u}_1\Xi_1 + a_{1,2}(a_{1,1}\mathsf{B}_1 + \mathsf{I})\mathbf{u}_2\Xi_2$$
$$+ (a_{1,1}\mathsf{B}_1 + \mathsf{I})\mathbf{Z}_1$$

and

$$\mathbf{Y}_2 = a_{2,1}\mathbf{u}_1\Xi_1 + (a_{2,1}a_{1,2}\mathsf{B}_1 + a_{2,2}\mathsf{I})\mathbf{u}_2\Xi_2$$
$$+ \left(a_{2,1}\mathsf{B}_1 + \rho_z\frac{\sigma_2}{\sigma_1}\mathsf{I}\right)\mathbf{Z}_1.$$

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:

$$\mathbf{u}_1 = \sqrt{\frac{P/2}{1 + a_{2,1}^2 P/\sigma_2^2}} \cdot \begin{pmatrix} 1 & \frac{a_{2,1}\sqrt{P}}{\rho_z\sigma_2} & 0 \ldots & 0 \end{pmatrix}^\top$$

$$\mathbf{u}_2 = \sqrt{\frac{\sigma_1^2}{2a_{1,2}^2}} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}^\top$$

$$\mathbf{v}_1 = \left(\left(\frac{-a_{1,1}\sqrt{P}}{\sigma_1}\right)^{\eta-1} \quad \left(\frac{-a_{1,1}\sqrt{P}}{\sigma_1}\right)^{\eta-2} \quad \cdots \quad \frac{-a_{1,1}\sqrt{P}}{\sigma_1} \quad 1\right)$$

$$\mathbf{v}_2 = \left(\left(\frac{-a_{2,1}\sqrt{P}}{\rho_z\sigma_2}\right)^{\eta-1} \quad \left(\frac{-a_{2,1}\sqrt{P}}{\rho_z\sigma_2}\right)^{\eta-2} \quad \cdots \quad \frac{-a_{2,1}\sqrt{P}}{\rho_z\sigma_2} \quad 1\right).$$

If the information symbols Ξ_1 and Ξ_2 are independent and satisfy $\mathsf{E}[\Xi_1^2], \mathsf{E}[\Xi_2^2] \leq 1$, then the chosen $\mathbf{u}_1, \mathbf{u}_2, \mathsf{B}_1, \mathsf{B}_2$ result in a scheme satisfying the blocklength-η average power constraint at Transmitter 1 for any positive integer η and also at Transmitter 2 for $\eta > \frac{\sigma_1^2}{2P}$.

We obtain:

$$I_1 \triangleq \mathbf{v}_1^\top\mathbf{Y}_1$$
$$= \sqrt{\frac{P\sigma_2^2}{\sigma_2^2 + Pa_{2,1}^2}}\left(-\frac{\sqrt{P}a_{1,1}}{\sigma_1}\right)^{\eta-1}\left(1 - \frac{a_{2,1}\sigma_1}{a_{1,1}\rho_z\sigma_2}\right)a_{1,1}\Xi_1$$
$$+ Z_{1,\eta} \tag{125a}$$

and

$$I_2 \triangleq \mathbf{v}_2^\top\mathbf{Y}_2$$
$$= \left(-\frac{\sqrt{P}a_{2,1}}{\rho_z\sigma_2}\right)^{\eta-1}\left(a_{2,2} - a_{1,2}\frac{\rho_z\sigma_2}{\sigma_1}\right)\Xi_2 + \rho_z\frac{\sigma_2}{\sigma_1}Z_{1,\eta}. \tag{125b}$$

Thus, the noise symbols $Z_{1,1}, \ldots, Z_{1,\eta-1}$ are completely canceled out when forming the "new outputs" in (125a) and (125b) and only $Z_{1,\eta}$ remains. Moreover, the "interference symbol" Ξ_2 is canceled out in I_1 and the "interference symbol" Ξ_1 is canceled out in I_2.

By (125), we conclude that our scheme achieves all non-negative rate pairs that satisfy

$$R_1 < \frac{1}{2\eta}\log\left(1 + \frac{P^\eta a_{1,1}^{2\eta}}{\sigma_1^{2\eta}}\frac{\sigma_2^2}{(\sigma_2^2 + Pa_{2,1}^2)}\left(1 - \frac{a_{2,1}\sigma_1}{a_{1,1}\rho_z\sigma_2}\right)^2\right) \tag{126a}$$

$$R_2 < \frac{1}{2\eta}\log\left(1 + \frac{P^{\eta-1}a_{2,1}^{2\eta-2}}{\sigma_2^{2\eta}}\left(a_{2,2} - \frac{a_{1,2}\rho_z\sigma_2}{\sigma_1}\right)^2\right) \tag{126b}$$

for all $\eta \geq \frac{\sigma_1^2}{2P}$.

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

Proposition 14: If $\rho_z \in \{-1, 1\}$ and neither $\frac{a_{2,1}}{\alpha_{1,1}}$ nor $\frac{a_{2,2}}{a_{1,2}}$ equals $\rho_z\frac{\sigma_2}{\sigma_1}$, then all rate-pairs satisfying

$$R_1 < \frac{1}{2}\log^+\left(\frac{a_{1,1}^2 P}{\sigma_1^2}\right) \tag{127}$$

$$R_2 < \frac{1}{2}\log^+\left(\frac{a_{2,1}^2 P}{\sigma_2^2}\right) \tag{128}$$

are achievable over the Gaussian IC with noise-free feedback.

Note 7: With the proposed choice of parameters our scheme achieves prelog 2 when $\rho_z \in \{-1, 1\}$ and $\frac{a_{2,1}}{\alpha_{1,1}}$ 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 $2\frac{\eta-1}{\eta}$; see the rate constraints in (126). Thus, choosing $\eta = 3$ suffices to achieve a prelog larger than 1.

Note 8: Exchanging the roles of the two transmitters we obtain: when $\rho_z \in \{-1, 1\}$ and $\frac{a_{2,1}}{\alpha_{1,1}}$ and $\frac{a_{2,2}}{a_{1,2}}$ are both different from $\frac{\rho_z\sigma_2}{\sigma_1}$, then all rate-pairs satisfying

$$R_1 \leq \frac{1}{2}\log^+\left(\frac{a_{1,2}^2 P}{\sigma_2^2}\right) \tag{129}$$

$$R_2 \leq \frac{1}{2}\log^+\left(\frac{a_{2,2}^2 P}{\sigma_1^2}\right) \tag{130}$$

are achievable over the Gaussian IC with noise-free feedback.

Note 9: For a symmetric setup where $a_{1,1} = a_{2,2}$ and $a_{1,2} = a_{2,1}$ the achievability of (117) can also be shown

using a slight generalization of Kramer's memoryless LMMSE scheme [15], see [12].

C. Proof of Theorem 13

Relation (118) follows from the following more general result: Irrespective of the channel parameters,

$$1 \le \varlimsup_{P \to \infty} \frac{C_{\text{IC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z)}{\frac{1}{2}\log(1+P)} \le 2. \qquad (131)$$

The lower bound in (131) can be achieved by silencing Transmitter 1 and letting Transmitter 2 communicate its Message M_2 to Receiver 2 over the resulting interference-free Gaussian channel $Y_{2,t} = a_{2,2}X_{2,t} + Z_{2,t}$ at 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 cut-set 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

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

which establish the converse result in (131).

We next prove (116). The achievability follows from (131). When $\rho_z \in \{-1, 1\}$ and $\frac{a_{2,1}}{a_{1,1}} = \frac{a_{2,2}}{a_{1,2}} = \rho_z \frac{\sigma_2}{\sigma_1}$, the converse holds because in this case

$$Y_{1,t} = \rho_z \frac{\sigma_1}{\sigma_2} Y_{2,t} \quad \text{with probability 1} \qquad (132)$$

and thus, each receiver can reconstruct the other receiver's outputs. Consequently, the feedback capacity of our Gaussian IC coincides with the feedback capacity of the Gaussian MAC from the two transmitters to one of the two receivers, and its prelog is 1 [27].

To prove the converse to (116) when $\rho_z \in (-1, 1)$ we use a genie-argument and a generalized Sato-MAC bound [28], similar to the upper bounds in [29, Section V-B], [30], [31], [32]. Our proof consists of the following three steps. In the first step we let a genie reveal the symbols

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

to Receiver 1 before the transmission begins. This obviously can only increase the sum-rate capacity of our channel. We refer to the resulting setup as the *genie-aided IC*.

In the second step, we apply Sato's MAC-bound argument [28] to this genie-aided IC.[4] That is, 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 requiring that the sole remaining Receiver 1 decode 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

---

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

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.

Given encoding/decoding functions for the genie-aided IC, we 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 sole 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

$$\hat{X}_{1,t} = f_{\text{IC},1,t}^{(n)}(\hat{M}_1, Y_1^{t-1}), \qquad t \in \{1, \dots, n\}, \qquad (133)$$

and

$$\hat{Y}_2^n = \frac{a_{2,2}}{a_{1,2}}(Y_1^n - a_{1,1}\hat{X}_1^n) + a_{2,1}\hat{X}_1^n + U^n, \qquad (134)$$

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 if the MAC-receiver (and thus also IC-Receiver 1) decodes M_1 correctly, then $\hat{X}_1^n = X_1^n$, and $\hat{Y}_2^n = Y_2^n$, and the MAC's guess of M_2 is identical to that of the IC's. Consequently, whenever the IC-Receivers 1 and 2 decode their intended messages M_1 and M_2 correctly, so does the sole MAC-receiver, and the probability of error over the MAC cannot therefore 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 (116). Before elaborating on this third step, we recall that in the genie-aided MAC the channel law is

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

the two transmitters observe the generalized feedback signals $\{Y_{1,t}\}$ and $\{Y_{2,t}\}$; and before the transmission begins 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. For every blocklength n we then have:

$$R_1 + R_2$$
$$\le \frac{1}{n}I(M_1, M_2; Y_1^n, U^n) + \frac{\epsilon(n)}{n}$$
$$= \frac{1}{n}I(M_1, M_2; Y_1^n | U^n) + \frac{\epsilon(n)}{n}$$
$$= \frac{1}{n}\sum_{t=1}^n \left(h(Y_{1,t}|Y_1^{t-1}, U^n)\right.$$
$$\left. - h(Y_{1,t}|Y_1^{t-1}, M_1, M_2, U^n)\right) + \frac{\epsilon(n)}{n}$$
$$\le \frac{1}{n}\sum_{t=1}^n \left(h(Y_{1,t}|U_t)\right.$$

$$-h(Y_{1,t}|Y_1^{t-1}, M_1, M_2, Y_2^{t-1}, U^n)\Big) + \frac{\epsilon(n)}{n}$$

$$= \frac{1}{n}\sum_{t=1}^{n}\Big(h(Y_{1,t}|U_t) - h(Y_{1,t}|X_{1,t}, X_{2,t}, U_t)\Big)$$

$$= \frac{1}{n}\sum_{t=1}^{n} I(Y_{1,t}; X_{1,t}, X_{2,t}|U_t)$$

$$\leq \frac{1}{2}\log\left(1 + \frac{(|a_{1,1}| + |a_{1,2}|)^2 P}{\mathsf{Var}\left(Z_{1,t}|Z_{2,t} - \frac{a_{2,2}}{a_{1,2}}Z_{1,t}\right)}\right) \qquad (135)$$

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 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 (134); 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 $\mathsf{Var}\left(Z_{1,t}|Z_{2,t} - \frac{a_{2,2}}{a_{1,2}}Z_{1,t}\right)$ does not depend on P and is strictly positive whenever $\rho_z \in (-1,1)$, by (135) we conclude that (116) holds also when $\rho_z \in (-1,1)$.

The converse to (117) follows from the general Relation (131), and its achievability from Proposition 14 in Section IV-B.

ACKNOWLEDGMENT

APPENDIX A
PROOF OF THEOREM 3

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

The following two upper bounds are obtained from the cutset bound and the fact that a Gaussian law maximizes the differential entropy under a variance constraint [23]. With two individual cuts between the transmitter and each of the two receivers we obtain

$$R_1 + R_2 \leq \max_{X : \mathsf{E}[X^2] \leq P}\{I(X; Y_1) + I(X; Y_2)\}$$

$$= \frac{1}{2}\log\left(1 + \frac{P}{\sigma_1^2}\right) + \frac{1}{2}\log\left(1 + \frac{P}{\sigma_2^2}\right), (136)$$

and with a single cut between the transmitter and both receivers

$$R_1 + R_2 \leq \max_{X : \mathsf{E}[X^2] \leq P} I(X; Y_1, Y_2)$$

$$\leq \frac{1}{2}\log\left(1 + \frac{P(\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho_z)}{\sigma_1^2\sigma_2^2(1 - \rho_z^2)}\right). \quad (137)$$

We first prove the converse to (29) where $\sigma_1^2 = \sigma_2^2$. In this case, Upper bound (137) specializes to

$$C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_1^2, \rho_z(P))$$

$$\leq \frac{1}{2}\log\left(1 + \frac{P}{\frac{\sigma_1^2}{2}(1 + \rho_z(P))}\right), \quad \rho_z(P) \in (-1,1). (138)$$

In view of (17), and since we assume $\sigma_1^2 = \sigma_2^2$ and we defined $-\log(0) = \infty$, Upper bound (138) 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 (27),

$$\varlimsup_{P \to \infty} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_1^2, \rho_z(P))}{\frac{1}{2}\log(1 + P)} \leq 1 + \zeta_{-1}. \qquad (139)$$

On the other hand, by (136), irrespective of $\{\rho_z(P)\}_{\{P>0\}}$,

$$\varlimsup_{P \to \infty} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_1^2, \rho_z(P))}{\frac{1}{2}\log(1 + P)} \leq 2. \qquad (140)$$

Combining (139) and (140) establishes the converse to (29).

We now prove the converse to (30) where $\sigma_1^2 \neq \sigma_2^2$. Using the facts that

$$\sigma_1^2 + \sigma_2^2 - 2\rho_z(P)\sigma_1\sigma_2 < 2(\sigma_1^2 + \sigma_2^2) \qquad (141)$$

and

$$1 - \rho_z^2(P) \geq 1 - |\rho_z(P)|, \qquad (142)$$

we can further upper bound the right-hand side of (137) to obtain:

$$C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z(P))$$

$$\leq \frac{1}{2}\log\left(1 + \frac{P}{\frac{\sigma_1^2\sigma_2^2}{2(\sigma_1^2+\sigma_2^2)}(1 - |\rho_z(P)|)}\right), \quad \rho_z(P) \in (-1,1).$$

$$(143)$$

Now, since we defined $-\log(0) = \infty$, Upper bound (143) 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 (27) and (28),

$$\varlimsup_{P \to \infty} \frac{-\log(1 - |\rho_z(P)|)}{\log(P)} = \max\{\zeta_{-1}, \zeta_{+1}\}, \qquad (144)$$

we conclude that

$$\varlimsup_{P \to \infty} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z(P))}{\frac{1}{2}\log(1 + P)} \leq 1 + \max\{\zeta_{-1}, \zeta_{+1}\}. (145)$$

Combining (145) with (140) establishes the converse to (30).

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

$$\varlimsup_{P \to \infty} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z(P))}{\frac{1}{2}\log(1 + P)} \geq \min\{1 + \zeta_{-1}, 2\}. \quad (146)$$

Since a generalized prelog of 1 is achievable even without feedback [21], [22] 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

$$\lim_{\ell \to \infty} \frac{-\log(1 + \rho_z(P_\ell))}{\frac{1}{2}\log(P_\ell)} = \zeta_{-1} > 0, \qquad (147)$$

and in particular

$$\lim_{\ell \to \infty} \rho_z(P_\ell) = -1. \qquad (148)$$

For each ℓ we choose parameters q_ℓ and δ_ℓ and show that Inequality (146) follows from Corollary 7 specialized to these parameters. We choose

$$\delta_\ell = \begin{cases} \frac{\sigma_1}{\sigma_2} \cdot \frac{\sigma_1 - \rho_z(P_\ell)\sigma_2}{\sigma_2 - \rho_z(P_\ell)\sigma_1} & \text{if } \rho_z(P_\ell) \in (-1, 1) \\ -\rho_z(P_\ell)\frac{\sigma_1}{\sigma_2} & \text{if } \rho_z(P_\ell) \in \{-1, 1\}, \end{cases} \quad (149)$$

and define the limit (not necessarily finite)

$$\kappa \triangleq \overline{\lim_{\ell \to \infty}} \, P_\ell \big(\sigma_1^2 + \delta_\ell^2 \sigma_2^2 + 2\delta_\ell \rho_z(P_\ell)\sigma_1\sigma_2\big). \quad (150)$$

Depending on κ, we choose $q_\ell > 0$ as follows. Let $\epsilon \in (0, 1)$ be a small positive number.

- If $\kappa = \infty$, we choose

$$q_\ell = \left(\frac{(1 - \epsilon)P_\ell}{\delta_\ell^2(1 + \delta_\ell)^2 \left(\sigma_1^2 + \sigma_2^2 \delta_\ell^2 + 2\rho_z(P_\ell)\delta_\ell\sigma_1\sigma_2\right)}\right)^{1/4}; \quad (151)$$

- if $\kappa \in [0, \infty)$, we choose

$$q_\ell = \big(\beta(1 - \epsilon)P_\ell\big)^{1/2} \qquad (152)$$

where $\beta > 0$ is a solution to

$$\sigma_1^2 \left(1 + \frac{\sigma_1}{\sigma_2}\right)^2 \beta \left(1 + \beta \frac{\kappa}{\sigma_2^2}\right) = 1. \qquad (153)$$

Notice that for every $\epsilon > 0$ 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 (62) for all $\ell > \ell_0(\epsilon, \sigma_1^2, \sigma_2^2, \kappa)$.

Moreover, if $\kappa \in [0, \infty)$, then specializing the rates in Corollary 7 to the choices in (149) and (152) proves that

$$\overline{\lim_{\ell \to \infty}} \frac{C_{\text{BC},\Sigma}(P_\ell, \sigma_1^2, \sigma_2^2, \rho_z(P_\ell))}{\frac{1}{2}\log(1 + P_\ell)} \geq 2. \qquad (154)$$

If $\kappa = \infty$, then for all sufficiently large ℓ the correlation coefficient $\rho_z(P_\ell) \in (-1, 1)$, and by (149) the choice in (151) evaluates to

$$q_\ell = \left(\frac{(\sigma_2 - \sigma_1\rho_z(P_\ell))^2}{\delta_\ell^2(1 + \delta_\ell)^2\sigma_1^2 \left(\sigma_2^2 + \sigma_1^2 - 2\rho_z(P_\ell)\sigma_1\sigma_2\right)}\right)^{1/4}$$
$$\cdot \left(\frac{(1 - \epsilon)P_\ell}{1 - |\rho_z(P_\ell)|^2}\right)^{1/4} \qquad (155)$$

Notice that by (148) and (149)

$$\lim_{\ell \to \infty} \frac{\log\left(\frac{(\sigma_2 - \sigma_1\rho_z(P_\ell))^2}{\delta_\ell^2(1+\delta_\ell)^2\sigma_1^2\left(\sigma_2^2+\sigma_1^2-2\rho_z(P_\ell)\sigma_1\sigma_2\right)(1+|\rho_z(P_\ell)|)}\right)}{\frac{1}{2}\log(P)} = 0, \qquad (156)$$

and therefore, specializing the rates in Proposition 7 to the choice in (149) and (155) proves that

$$\overline{\lim_{\ell \to \infty}} \frac{C_{\text{BC},\Sigma}(P_\ell, \sigma_1^2, \sigma_2^2, \rho_z(P_\ell))}{\frac{1}{2}\log(1 + P_\ell)} \geq 1 + \zeta_{-1}. \qquad (157)$$

Combining (157) with (154) establishes (146).

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

$$\overline{\lim_{P \to \infty}} \frac{C_{\text{BC},\Sigma}(P, \sigma_1^2, \sigma_2^2, \rho_z(P))}{\frac{1}{2}\log(1 + P)} \geq \min\{1 + \zeta_{+1}, 2\}. \quad (158)$$

To this end, it suffices that in the proof to (146) we replace ζ_{-1} by ζ_{+1}; (147) by

$$\lim_{\ell \to \infty} \frac{-\log(1 - \rho_z(P_\ell))}{\frac{1}{2}\log(P)} = \zeta_{+1} > 0; \qquad (159)$$

(148) by

$$\lim_{\ell \to \infty} \rho_z(P) = 1; \qquad (160)$$

and (153) by

$$\sigma_1^2 \left(1 - \frac{\sigma_1}{\sigma_2}\right)^2 \beta \left(1 + \beta \frac{\kappa}{\sigma_2^2}\right) = 1. \qquad (161)$$

The assumption of non-equal noise variances $\sigma_1^2 \neq \sigma_2^2$ is needed here to conclude that (156) holds and that (161) has a finite solution for β.

Combining finally (158) with (146) establishes the achievability of (29) and (30) and concludes the proof.

APPENDIX 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

$$\overline{\lim_{P \to \infty}} \frac{-\log(\sigma_{W\nu}^2(P))}{\log P} \leq 0, \qquad \nu \in \{1, 2\}. \quad (162)$$

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

$$\overline{\lim_{P \to \infty}} \frac{C_{\text{BCNoisy},\Sigma}\big(P, \sigma_1^2, \sigma_2^2, \rho_z, \sigma_{W1}^2(P), \sigma_{W2}^2(P)\big)}{\frac{1}{2}\log(1 + P)} \leq 1. \quad (163)$$

For $\rho_z \in (-1, 1)$ Inequality (163) 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 (163) is similar to the proof in Section II-D. In fact, following the same steps as before, we can conclude that for each $P > 0$

$$C_{\text{BCNoisy},\Sigma}\big(P, \sigma_1^2, \sigma_2^2, \rho_z, \sigma_{W1}^2(P), \sigma_{W2}^2(P)\big)$$
$$\leq \frac{1}{2}\log\left(1 + \frac{P}{\min\left\{\text{Var}\big(Z'_{1,t}\big), \text{Var}\big(Z'_{2,t}\big)\right\}}\right), \quad (164)$$

where here, because $|\rho_z| = 1$, the variances in (84) and (85) simplify to

$$\text{Var}\big(Z'_{1,t}\big) = \frac{\sigma_1^2\sigma_{W1}^2(P)\sigma_{W2}^2(P)}{\sigma_1^2\sigma_{W2}^2(P) + \sigma_2^2\sigma_{W1}^2(P) + \sigma_{W1}^2(P)\sigma_{W2}^2(P)}, \quad (165)$$

$$\text{Var}\big(Z'_{2,t}\big) = \frac{\sigma_2^2\sigma_{W1}^2(P)\sigma_{W2}^2(P)}{\sigma_1^2\sigma_{W2}^2(P) + \sigma_2^2\sigma_{W1}^2(P) + \sigma_{W1}^2(P)\sigma_{W2}^2(P)}. \quad (166)$$

The desired inequality (163) for $\rho_z \in \{-1, 1\}$ follows now simply by combining (162) with (164)–(166).

Appendix C
Proof of Lemma 1

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

$$\frac{1}{2\eta} \log(1 + \xi^{\eta-1}\zeta) \leq \frac{1}{2\eta} \log(1 + (1+\zeta)^{\eta-1}\zeta)$$

$$= \frac{1}{2\eta} \log\left((1+\zeta)^{\eta} - \underbrace{\left((1+\zeta)^{\eta-1} - 1\right)}_{\geq 0}\right)$$

$$\leq \frac{1}{2\eta} \log\left((1+\zeta)^{\eta}\right)$$

$$= \frac{1}{2} \log(1+\zeta).$$

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

$$\frac{1}{2\eta} \log(1 + \xi^{\eta-1}\zeta) \leq \frac{1}{2\eta} \log(1 + \xi^{\eta-1}(\xi - 1))$$

$$= \frac{1}{2\eta} \log\left(\xi^{\eta} - \underbrace{\left(\xi^{\eta-1} - 1\right)}_{\geq 0}\right)$$

$$\leq \frac{1}{2\eta} \log\left(\xi^{\eta}\right)$$

$$= \frac{1}{2} \log(\xi).$$

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./Feb. 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. Khisti and A. Lapidoth, "Multiple access channels with intermittent feedback and side information," in *Proc. ISIT 2013*, (Istanbul, Turkey), July 2013.

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

[8] 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), Nov. 2011.

[9] 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.

[10] 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. 30, pp. 667–671, July 1984.

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

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

[13] E. Ardestanizadeh, P. Minero, and M. Franceschetti, "LQG control approach to Gaussian broadcast channels with feedback." *IEEE Transactions on Information Theory*, vol. 58 , no. 8, pp. 5267–5278, Aug. 2012.

[14] 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.

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

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

[17] C. Suh and D. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2667–2685, May 2011.

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

[19] A. Vahid, C. Suh, and A. S. Avestimehr, "Interference channels with rate-limited feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2788–2812, May 2012.

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

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

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

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

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

[25] O. Shayevitz and M. Wigger, "An achievable region for the discrete memoryless broadcast channel with feedback," in *Proc. ISIT 2010*, (Austin, Texas), June 2010.

[26] 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, Mar. 2013.

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

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

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

[30] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 581–586, Mar. 2004.

[31] A. Lapidoth, S. Shamai (Shitz), and M. A. Wigger, "On cognitive interference networks," in *IEEE Information Theory Workshop*, (Lake Tahoe, USA), pp. 325–330, Sep. 2007.

[32] A. Lapidoth, N. Levy, S. Shamai (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.

[33] F. M. J. Willems and E. C. van der Meulen, "Een verbetering en ver-algemening van het transmissiegebied van Ozarow voor het Gaussische broadcast kanaal met feedback," in *Tweede Symposium over Informati-etheorie in de Benelux*, (Zoetermeer, The Netherlands), pp. 129–138, May 1981. In Dutch.

Michael Gastpar received the Dipl. El.-Ing. degree from ETH Zürich, in 1997, the M.S. degree from the University of Illinois at Urbana-Champaign, Urbana, IL, in 1999, and the Doctorat ès Science degree from Ecole Poly-technique Fédérale (EPFL), Lausanne, Switzerland, in 2002, all in electrical engineering. He was also a student in engineering and philosophy at the Universities of Edinburgh and Lausanne.

He is a Professor in the School of Computer and Communication Sciences, Ecole Polytechnique Fédérale (EPFL), Lausanne, Switzerland. He was an Assistant (2003-2008) and tenured Associate Professor (2008-2011) with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, where he still holds a faculty position. He also holds a faculty position at Delft University of Technology, The Netherlands, and was a Researcher with the Mathematics of Communications Department, Bell Labs, Lucent Technologies, Murray Hill, NJ. His research interests are in network information theory and related coding and signal processing techniques, with applications to sensor networks and neuroscience.

Dr. Gastpar won the 2002 EPFL Best Thesis Award, an NSF CAREER Award in 2004, an Okawa Foundation Research Grant in 2008, and an ERC Starting Grant in 2010. He is the co-recipient of the 2013 Communications Society & Information Theory Society Joint Paper Award. He was an Information Theory Society Distinguished Lecturer (20092011), an Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY (2008-2011), and he has served as Technical Program Committee Co-Chair for the 2010 International Symposium on Information Theory, Austin, TX.

Amos Lapidoth (S'89, M'95, SM'00, F'04) received the B.A. degree in mathematics (*summa cum laude*, 1986), the B.Sc. degree in electrical engineering (*summa cum laude*, 1986), and the M.Sc. degree in electrical engineering (1990) all from the Technion—Israel Institute of Technology. He received the Ph.D. degree in electrical engineering from Stanford University in 1995.

In the years 1995–1999 he was an Assistant and Associate Professor at the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, and was the KDD Career Development Associate Professor in Communications and Technology. He is now Professor of Information Theory at ETH Zurich in Switzerland. He is the author of the book *A Foundation in Digital Communication*, published by Cambridge University Press in 2009. His research interests are in digital communications and information theory.

Dr. Lapidoth served in the years 2003–2004 and 2009 as Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.

Yossef Steinberg (M96SM09, F'11) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering in 1983, 1986, and 1990, respectively, all from Tel-Aviv University, Tel-Aviv, Israel. He was a Lady Davis Fellow in the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel, and held visiting appointments in the Department of Electrical Engineering at Princeton University, Princeton, NJ, and at the C I Center, George Mason University, Fairfax, VA. From 1995 to 1999, he was with the Department of Electrical Engineering, Ben Gurion University, Beer-Sheva, Israel. In 1999, he joined the Department of Electrical Engineering at the Technion. Dr. Steinberg served in the years 2004–2007 as Associate Editor for Shannon Theory, and currently serves as Associate Editor at large, for the IEEE TRANSACTIONS ON INFORMATION THEORY. Dr. Steinberg's research interests are in Digital Communications, Information Theory, and Estimation. He won the 2007 Information Theory Society Paper Award, jointly with Hanan Weingarten and Shlomo Shamai.

Michèle Wigger (S'05-M'09) received the M.Sc. degree in electrical engineering (with distinction) and the Ph.D. degree in electrical engineering both from ETH Zurich in 2003 and 2008, respectively. In 2009 she was a postdoctoral researcher at the ITA center at the University of California, San Diego. Since December 2009 she is an Assistant Professor at Telecom ParisTech, in Paris, France. Her research interests are in information and communications theory.