New Achievable Rates for the Gaussian Broadcast Channel with Feedback

Michael Gastpar
EPF Lausanne, Switzerland
michael.gastpar@epfl.ch

Amos Lapidoth
ETH Zürich, Switzerland
lapidoth@isi.ee.ethz.ch

Yossef Steinberg
Technion, Haifa, Israel
ysteinbe@ee.technion.ac.il

Michèle Wigger
Telecom ParisTech, France
wigger@telecom-paristech.fr

***Abstract*—A coding scheme for the two-receivers Gaussian broadcast channel (BC) with feedback is proposed. For some asymmetric settings it achieves new rate pairs. Moreover, it achieves prelog 2 when the noises at the two receivers are fully positively correlated and of unequal variances, thus allowing us to complete the characterization of the prelog of the two-receivers Gaussian BC with feedback. The new achievable rates also allow us to determine the asymptotic power offset when the noises at the two receivers are uncorrelated.**

I. Introduction and Channel Model

We consider the real, scalar, memoryless two-user Gaussian broadcast channel (BC) with perfect feedback. Achievable regions for this setup were proposed by Ozarow & Leung [9], Elia [4], Wu et al. [11], and Ardestanizadeh et al. [1]. The region in [9] is for the general setup, whereas the regions in [4], [1] are only for symmetric setups where the noises at the two receivers are of equal variances, and the regions in [4], [11] only for setups where these noises are uncorrelated. For such setups, the regions in [4], [11], [1] include the one in [9].

The results in [9], [4], [1] demonstrate that feedback can enlarge the capacity region of the two-user Gaussian BC. This gain can be even unbounded in the signal-to-noise ratio (SNR): in the absence of feedback the prelog is 1, whereas in some settings with feedback (e.g., when the noises at the two receivers are perfectly anticorrelated) it is 2 [5], [6].

Outer bounds on the capacity region were presented in [3], [9], [5],[6]. However, they do not coincide with the known achievable regions, except for the special case where the noises at the two receivers are identical (i.e., fully correlated noises of equal variances). In this case the capacity with feedback is the same as without [3]. For all other scenarios the capacity of the two-user Gaussian BC with feedback is still unknown.

Here, we present a new coding scheme which yields new achievable regions for some non-symmetric setups with unequal noise variances at the two receivers. Our achievable regions allow us to characterize the prelog of the two-user Gaussian BC with feedback and, when the noises at the two receivers are uncorrelated, also its asymptotic power offset.

II. Channel Model

Let $x_t \in \mathbb{R}$ denote the time-t channel input, and $Y_{1,t}, Y_{2,t} \in \mathbb{R}$ the time-t channel outputs observed at Receivers 1 and 2. Then,

$$Y_{k,t} \quad = \quad x_t + \alpha_k Z_{0,t} + Z_{k,t}, \qquad k \in \{1,2\}. \tag{1}$$

where α_1 and α_2 are real constants and $\{Z_{0,t}\}$, $\{Z_{1,t}\}$, and $\{Z_{2,t}\}$ are independent sequences of independent and identically distributed (IID) zero-mean Gaussians of variances 1, σ_1^2, and σ_2^2.

The total noise experienced at Receiver $k \in \{1,2\}$ is $\alpha_k^2 + \sigma_k^2$, which, by assumption, is strictly positive for both receivers. Thus, for $k \in \{1,2\}$

$$\alpha_k^2 > 0 \quad \text{or} \quad \sigma_k^2 > 0.$$

The correlation between the noises at the two receivers is

$$\rho_z \triangleq \frac{\alpha_1\alpha_2}{\sqrt{\alpha_1^2+\sigma_1^2}\sqrt{\alpha_2^2+\sigma_2^2}}. \tag{2}$$

Thus, when both σ_1^2 and σ_2^2 are zero, the correlation ρ_z is $\text{sign}(\alpha_1\alpha_2)$, i.e., either -1 or $+1$, and when both α_1 and α_2 are zero, ρ_z is zero.

For convenience, and without loss of generality, we will assume that α_1 and α_2 are either both zero or both nonzero.

The goal of the transmitter is to convey Message M_1 to Receiver 1 and Message M_2 to Receiver 2. The messages are independent and Message M_k, for $k \in \{1,2\}$, is uniformly distributed over the set $\mathcal{M}_k \triangleq \{1,\dots,\lfloor 2^{nR_k}\rfloor\}$, where n denotes the blocklength and R_k the rate of transmission.

The transmitter has access to noise-free feedback from both receivers. Thus, after sending x_{t-1} it learns both $Y_{1,t-1}$ and $Y_{2,t-1}$, and it can compute its next channel input

$$X_t = f_t^{(n)}\left(M_1, M_2, Y_{1,1},\dots,Y_{1,t-1},Y_{2,1},\dots,Y_{2,t-1}\right), \tag{3}$$

where

$$f_t^{(n)} \colon \mathcal{M}_1 \times \mathcal{M}_2 \times \mathbb{R}^{t-1} \times \mathbb{R}^{t-1} \to \mathbb{R}$$

is the time-t encoding function. We only allow encoding functions that produce inputs satisfying the expected average block-power constraint

$$\frac{1}{n}\mathsf{E}\left[\sum_{t=1}^{n} X_t^2\right] \le P. \tag{4}$$

Receiver k decodes its message M_k based on its observed sequence $Y_k^n = (Y_{k,1},\dots,Y_{k,n})$ by producing the guess

$$\hat{M}_k \quad = \quad \phi_k^{(n)}(Y_k^n) \tag{5}$$

for some decoding function $\phi_k^{(n)} \colon \mathbb{R}^n \to \{1,\dots,\lfloor 2^{nR_k}\rfloor\}$. We say that a rate pair (R_1,R_2) is achievable if for every blocklength n there exists a set of n encoding functions $\left\{f_t^{(n)}\right\}_{t=1}^{n}$ and two decoding functions $\phi_1^{(n)}$ and $\phi_2^{(n)}$ such that the probability of decoding error tends to 0 as the blocklength n tends to infinity, i.e., such that

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

A set of achievable rate pairs is an *achievable region*. The closure of the union of all achievable regions is the *capacity region*. The *sum-rate capacity* is the supremum of R_1+R_2 over all achievable rate pairs (R_1,R_2). We denote it by $C_\Sigma(\alpha_1,\alpha_2,\sigma_1^2,\sigma_2^2;P)$.

III. Main Results

Definition 1: Given a positive integer η, two strictly lower-triangular η-by-η matrices B_1 and B_2, two η-dimensional column vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, and two η-dimensional row-vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, define $\mathcal{R}(\eta,\mathsf{B}_1,\mathsf{B}_2,\mathbf{u}_1,\mathbf{u}_2,\mathbf{v}_1,\mathbf{v}_2)$ as the set of all nonnegative rate pairs (R_1,R_2) satisfying (6) on top of the next page.

Theorem 1: The following region is achievable:

$$\text{closure}\left(\bigcup_{\eta,\mathsf{B}_1,\mathsf{B}_2,\mathbf{u}_1,\mathbf{u}_2,\mathbf{v}_1,\mathbf{v}_2}\mathcal{R}(\eta,\mathsf{B}_1,\mathsf{B}_2,\mathbf{u}_1,\mathbf{u}_2,\mathbf{v}_1,\mathbf{v}_2)\right)\quad(7)$$

where the union is over all choices of parameters $(\eta,\mathsf{B}_1,\mathsf{B}_2,\mathbf{u}_1,\mathbf{u}_2,\mathbf{v}_1,\mathbf{v}_2)$ satisfying (8) on top of the next page, where $\mathrm{tr}(\cdot)$ denotes the trace-operator.

Remark 1: The set of achievable rates in [9], [4], and [1] are included in the achievable region in (7).

We have the following Corollaries 1–4 to Theorem 1.

Corollary 1 (Fully Correlated Noises): If $\sigma_1^2=\sigma_2^2=0$ and $\alpha_1\neq\alpha_2$, then all rate pairs (R_1,R_2) satisfying

$$0\le R_k\le\frac{1}{2}\log^+\left(\frac{P}{\alpha_k^2}\right),\quad k\in\{1,2\}\quad(9)$$

are achievable, where $\log^+(x)=\max\{0,\log(x)\}$.

Remark 2: For large powers P, some of the achievable rates in Corollary 1 are not contained in the achievable region proposed by Ozarow and Leung [9]. In particular, when $\sigma_1^2=\sigma_2^2=0$ and α_1 and α_2 have the same sign but $\alpha_1\neq\alpha_2$, then the achievability of prelog 2 (Theorem 2 ahead) can be proved with the achievable rates in Corollary 1 but not with the achievable rates in [9].

Combined with the results in [5], [6], Corollary 1 yields:

Theorem 2 (Prelog): The prelog of the two-user BC with feedback is:

$$\varlimsup_{P\to\infty}\frac{C_\Sigma(\alpha_1,\alpha_2,\sigma_1^2,\sigma_2^2;P)}{\frac{1}{2}\log(1+P)}=\begin{cases}2 & \text{if } \sigma_1^2=\sigma_2^2=0 \text{ and } \alpha_1\neq\alpha_2\\ 1 & \text{otherwise.}\end{cases}\quad(10)$$

Thus, when the noise correlation $\rho_z=1$ (i.e., $\sigma_1^2=\sigma_2^2=0$ and α_1 and α_2 have the same sign), the prelog is 2 whenever the total noise variances at the two receivers differ, and it is 1 when they are the same.

Corollary 2 (Noise Correlation in $(-1,0)$ or $(0,1)$): Let $\sigma_1^2,\sigma_2^2>0$ and $\alpha_1,\alpha_2\neq0$ be given and so that $\alpha_1\neq\alpha_2$. Also, let $p>0$ be defined through (11) on top of the next page. Then, the rate pair (R_1,R_2) is achievable if it satisfies

$$0\le R_k\le\frac{1}{2}\log^+\left(\frac{p^2}{\alpha_k^2}\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right)^2\right),\quad k\in\{1,2\}.\quad(12)$$

Remark 3: For certain channel parameters, the achievable region in Corollary 2 contains rate pairs that are not in Ozarow & Leung's achievable region [9]. In particular, when α_1 and α_2 have the same sign (i.e., the noise correlation is positive) but $\alpha_1\neq\alpha_2$, Theorem 3 ahead can be proved with the achievable rates in Corollary 2 but not with the ones in [9].

For the scope of the generalized prelog result in Theorem 3, we let the independent-noise variances σ_1^2 and σ_2^2 depend on the power P, and write $\sigma_1^2(P)$ and $\sigma_2^2(P)$.

Theorem 3 (Generalized Prelog): Let $\alpha_1\neq\alpha_2$ be fixed. Also, let $\{\sigma_1^2(P)\}_{\{P>0\}}$ and $\{\sigma_2^2(P)\}_{\{P>0\}}$ be given sequences of variances of the form

$$\sigma_k^2(P)=\frac{\epsilon_k(P)}{P^\zeta},\qquad\zeta\in[0,1],\quad k\in\{1,2\},\quad(13)$$

where

$$\lim_{P\to\infty}\frac{\log(\epsilon_k(P))}{\log(P)}=0.\quad(14)$$

Then,

$$\varlimsup_{P\to\infty}\frac{C_\Sigma(\alpha_1^2,\alpha_2^2,\sigma_1^2(P),\sigma_2^2(P);P)}{\frac{1}{2}\log\left(1+P\right)}=1+\zeta.\quad(15)$$

Proof: The achievability can be proved with the rates in Corollary 2. The converse is based on [6, Equation (60)]. ■

Corollary 3 (Independent Noises): Let $\alpha_1=\alpha_2=0$. The rate pair (R_1,R_2) is achievable if it satisfies

$$0\le R_1\le\frac{1}{2}\log^+\left(q^2(1+\delta)^2\right)\quad(16a)$$

$$0\le R_2\le\frac{1}{2}\log^+\left(q^2\delta^2(1+\delta)^2\right)\quad(16b)$$

for some $\delta\notin\{-1,0\}$ and $q>0$ satisfying

$$(q^2+q^4\delta^2(1+\delta)^2)\sigma_1^2+(q^2\delta^4+q^4\delta^4(1+\delta)^2)\sigma_2^2=P.\quad(17)$$

Choosing $\delta=\frac{\sigma_1^2}{\sigma_2^2}$ and $q>0$ as to satisfy (17), we obtain:

Corollary 4 (Independent Noises): Let $\sigma_1^2=\sigma_2^2=\sigma^2$. The equal rate pair (R,R) is achievable if it satisfies

$$0\le R\le\frac{1}{2}\log^+\left(-\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{2P}{\sigma^2}}\right).\quad(18)$$

For general $\sigma_1^2,\sigma_2^2\ge0$ and large powers P a sum-rate (R_1+R_2) is achievable if it satisfies

$$0\le R_1+R_2\le\frac{1}{2}\log^+\left(\frac{P(\sigma_1^2+\sigma_2^2)}{\sigma_1^2\sigma_2^2}\right)+o(P),$$

where $o(P)$ tends to 0 as $P\to\infty$.

Corollary 4 and the converse in [6, Equation (59)] yield:

$$R_1 \le \frac{1}{2}\left(1+\frac{(\mathbf{v}_1\mathbf{u}_1)^2}{(\mathbf{v}_1\mathbf{u}_2)^2+\|\mathbf{v}_1(\mathsf{B}_1\alpha_1+\mathsf{B}_2\alpha_2+\mathsf{I}\alpha_1)\|^2+\|\mathbf{v}_1(\mathsf{B}_1+\mathsf{I})\|^2\sigma_1^2+\|\mathbf{v}_1\mathsf{B}_2\|^2\sigma_2^2}\right) \tag{6a}$$

$$R_2 \le \frac{1}{2}\left(1+\frac{(\mathbf{v}_2\mathbf{u}_2)^2}{(\mathbf{v}_2\mathbf{u}_1)^2+\|\mathbf{v}_2(\mathsf{B}_1\alpha_1+\mathsf{B}_2\alpha_2+\mathsf{I}\alpha_2)\|^2+\|\mathbf{v}_2(\mathsf{B}_2+\mathsf{I})\|^2\sigma_2^2+\|\mathbf{v}_2\mathsf{B}_1\|^2\sigma_1^2}\right) \tag{6b}$$

$$\|\mathbf{u}_1\|^2+\|\mathbf{u}_2\|^2+\operatorname{tr}\left((\alpha_1\mathsf{B}_1+\alpha_2\mathsf{B}_2)(\alpha_1\mathsf{B}_1+\alpha_2\mathsf{B}_2)^\intercal\right)+\operatorname{tr}\left(\mathsf{B}_1\mathsf{B}_1^\intercal\right)\sigma_1^2+\operatorname{tr}\left(\mathsf{B}_2\mathsf{B}_2^\intercal\right)\sigma_2^2 \le \eta P \tag{8}$$

$$p^2 = \frac{-\left(\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right)^2+\frac{\sigma_1^2}{\alpha_1^4}+\frac{\sigma_2^2}{\alpha_2^4}\right)+\sqrt{\left(\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right)^2+\frac{\sigma_1^2}{\alpha_1^4}+\frac{\sigma_2^2}{\alpha_2^4}\right)^2+4P\frac{1}{\alpha_1^2\alpha_2^2}\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right)^2\left(\frac{\sigma_1^2}{\alpha_1^2}+\frac{\sigma_2^2}{\alpha_2^2}\right)}}{2\frac{1}{\alpha_1^2\alpha_2^2}\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right)^2\left(\frac{\sigma_1^2}{\alpha_1^2}+\frac{\sigma_2^2}{\alpha_2^2}\right)} \tag{11}$$

Theorem 4 (Asymptotic Power Offset): When $\alpha_1 = \alpha_2 = 0$, i.e., when the noises at the two receivers are independent,

$$\varlimsup_{P\to\infty}\left(C_\Sigma(0,0,\sigma_1^2,\sigma_2^2;P)-\frac{1}{2}\log\left(\frac{P(\sigma_1^2+\sigma_2^2)}{\sigma_1^2\sigma_2^2}\right)\right)=0.$$

Thus, when the two noises are uncorrelated, then in the asymptotic high-power regime the sum-rate capacity of the two-user Gaussian BC with feedback approaches the sum-rate capacity of a single-user setup (with or without feedback) where the two receivers can decode their messages jointly.

For the symmetric setup $\sigma_1^2 = \sigma_2^2$ the achievability result in Theorem 4 has previously been reported in [1].

IV. A Block-Scheme

We first describe a general coding scheme (Section IV-A) achieving the rates in Theorem 1. Then we propose choices of parameters for this scheme (Sections IV-B and IV-C) which establish Corollaries 1–9.

A. General Scheme

Our scheme is similar to the schemes in [2], [7], [8]. It has the following parameters: the positive integer η; the η-by-η strictly lower-triangular matrices $\mathsf{B}_1, \mathsf{B}_2$; the η-dimensional column-vectors $\mathbf{u}_1, \mathbf{u}_2$; and the η-dimensional row-vectors $\mathbf{v}_1, \mathbf{v}_2$.

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

Independently generate the two codebooks $\{\mathcal{C}_k\}_{k=1}^2$, each containing $\lfloor 2^{nR_k}\rfloor$ codewords of length n', by randomly drawing all entries of all codewords IID according to a standard Gaussian distribution. The codebooks are revealed to the transmitter and the receivers.

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

Also, for $i \in \{1,\dots,n'\}$, let $\mathbf{X}_i$ be the η-length column-vector $\mathbf{X}_i \triangleq (X_{(i-1)\eta+1},\dots,X_{(i\eta)})^\intercal$, and let $\mathbf{Z}_{0,i}$ and for each $k \in \{1,2\}$ also $\mathbf{Y}_{k,i}$ and $\mathbf{Z}_{k,i}$ be defined similarly.

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

$$\mathbf{X}_i = \Xi_{1,i}\mathbf{u}_1+\Xi_{2,i}\mathbf{u}_2+\mathsf{B}_1(\alpha_1\mathbf{Z}_{0,i}+\mathbf{Z}_{1,i})+\mathsf{B}_2(\alpha_2\mathbf{Z}_{0,i}+\mathbf{Z}_{2,i}).$$

The transmitter can compute the past noise symbols because it knows the past inputs and, through the feedback, also the past outputs. Also, the strict lower-triangularity of the matrices B_1 and B_2 assures that only past noise symbols are sent.

The inputs satisfy the average block-power constraint (4) whenever Inequality (8) is satisfied.

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

$$\begin{aligned}\mathbf{Y}_{1,i} = {} & \Xi_{1,i}\mathbf{u}_1+\Xi_{2,i}\mathbf{u}_2+(\mathsf{B}_1\alpha_1+\mathsf{B}_2\alpha_2+\mathsf{I}\alpha_1)\mathbf{Z}_{0,i}\\ & +(\mathsf{B}_1+\mathsf{I})\mathbf{Z}_{1,i}+\mathsf{B}_2\mathbf{Z}_{2,i},\end{aligned} \tag{19a}$$

and Receiver 2 observes

$$\begin{aligned}\mathbf{Y}_{2,i} = {} & \Xi_{1,i}\mathbf{u}_1+\Xi_{2,i}\mathbf{u}_2+(\mathsf{B}_1\alpha_1+\mathsf{B}_2\alpha_2+\mathsf{I}\alpha_2)\mathbf{Z}_{0,i}\\ & +(\mathsf{B}_2+\mathsf{I})\mathbf{Z}_{2,i}+\mathsf{B}_1\mathbf{Z}_{1,i}.\end{aligned} \tag{19b}$$

Each Receiver $k \in \{1,2\}$ forms

$$I_{k,i} \triangleq \mathbf{v}_k\mathbf{Y}_{k,i}, \qquad i \in \{1,\dots,n'\}, \tag{20}$$

and decodes its desired message M_k by applying a maximum-likelihood decoder to the sequence $I_{k,1},\dots,I_{k,n'}$.

The analysis of the scheme is standard and omitted.

B. Simple Choice of Parameters

We present a simple choice of parameters (in particular, simple matrices $\mathsf{B}_1, \mathsf{B}_2$). The choice is sub-optimal (see Subsection IV-C) but suffices to prove the results in Section III.

Given a positive integer η we choose $\mathsf{B}_1, \mathsf{B}_2$, $\mathbf{u}_1, \mathbf{u}_2$, and $\mathbf{v}_1, \mathbf{v}_2$ in a way that

1) $\mathbf{v}_1$ is orthogonal to the first $\eta-1$ columns of $(\mathsf{B}_1\alpha_1+\mathsf{B}_2\alpha_2+\mathsf{I}\alpha_1)$, and if $\sigma_1^2, \sigma_2^2 > 0$, also to the first $\eta-2$ columns of the matrices $(\mathsf{B}_1+\mathsf{I})$ and B_2;
2) $\mathbf{v}_2$ is orthogonal to the first $\eta-1$ columns of $(\mathsf{B}_1\alpha_1+\mathsf{B}_2\alpha_2+\mathsf{I}\alpha_2)$, and if $\sigma_1^2, \sigma_2^2 > 0$, also to the first $\eta-2$ columns of the matrices $(\mathsf{B}_2+\mathsf{I})$ and B_1;

3) $\mathbf{u}_1$ is orthogonal to $\mathbf{v}_2$ but not to $\mathbf{v}_1$;

4) $\mathbf{u}_2$ is orthogonal to $\mathbf{v}_1$ but not to $\mathbf{v}_2$;

By 3) and 4), the symbols $I_{1,1},\ldots,I_{1,n'}$ can be viewed as the outputs of a point-to-point channel where the transmitter sends the codeword $\Xi_{1,1},\ldots,\Xi_{1,n'}$ only. Moreover, by 1) each $I_{1,i}$ depends only on the last two noises of subblock i, i.e., on $\{Z_{k,i\eta-1},Z_{k,i\eta}\}_{k=0}^2$; all previous noise symbols are cancelled, see (23). Analogous observations hold for $I_{2,1},\ldots,I_{2,n'}$.

Remark 4: When $\sigma_1^2,\sigma_2^2>0$, Conditions 1)–4) can only be satisfied if the matrices $\mathsf{B}_1,\mathsf{B}_2$ have non-zero entries on at least two diagonals (unless $\mathsf{B}_1=\mathsf{B}_2=0$). When $\sigma_1^2=\sigma_2^2=0$ non-zero entries on one diagonal suffice to satisfy 1)–4).

We choose B_k Toeplitz with non-zero entries only on the first and second diagonals below the main diagonal

$$\mathsf{B}_k=\begin{pmatrix}0&0&\cdots&0&0&0\\ a_k&0&\cdots&0&0&0\\ b_k&a_k&0&\cdots&0&0\\ 0&b_k&a_k&0&\cdots&0\\ \vdots&\ddots&\ddots&\ddots&\ddots&\vdots\\ 0&0&0&b_k&a_k&0\end{pmatrix}, \tag{21}$$

and we choose

$$\mathbf{u}_1=\sqrt{P/(2+2\gamma_2^2)}\cdot\begin{pmatrix}1&-\gamma_2&0&\ldots&0\end{pmatrix}^{\mathsf{T}} \tag{22a}$$
$$\mathbf{u}_2=\sqrt{P/(2+2\gamma_1^2)}\cdot\begin{pmatrix}1&-\gamma_1&0&\ldots&0\end{pmatrix}^{\mathsf{T}} \tag{22b}$$
$$\mathbf{v}_1=\begin{pmatrix}\gamma_1^{\eta-1}&\gamma_1^{\eta-2}&\ldots&\gamma_1&1\end{pmatrix} \tag{22c}$$
$$\mathbf{v}_2=\begin{pmatrix}\gamma_2^{\eta-1}&\gamma_2^{\eta-2}&\ldots&\gamma_2&1\end{pmatrix} \tag{22d}$$

for some real values $a_1,a_2,b_1,b_2,\gamma_1,\gamma_2$ which will be defined shortly. Notice that by (19), (20), and (21)–(22) we have

$$\begin{aligned}I_{1,i}=&\sqrt{\frac{P}{(2+2\gamma_2^2)}}\gamma_1^{\eta-1}\left(1-\frac{\gamma_2}{\gamma_1}\right)\Xi_{1,i}+\alpha_1Z_{0,i\eta}\\&+(\gamma_1+a_1)Z_{1,i\eta-1}+Z_{1,i\eta}+a_2Z_{2,i\eta-1}\end{aligned} \tag{23}$$
$$\begin{aligned}I_{2,i}=&\sqrt{\frac{P}{(2+2\gamma_1^2)}}\gamma_2^{\eta-1}\left(1-\frac{\gamma_1}{\gamma_2}\right)\Xi_{2,i}+\alpha_2Z_{0,i\eta}\\&+(\gamma_2+a_2)Z_{2,i\eta-1}+Z_{2,i\eta}+a_1Z_{1,i\eta-1}.\end{aligned} \tag{24}$$

We now present our choice of $a_1,a_2,b_1,b_2,\gamma_1,\gamma_2$ depending on the channel parameters $\alpha_1,\alpha_2,\sigma_1^2,\sigma_2^2$.

1) $\sigma_1^2=\sigma_2^2=0$ (Fully Correlated Noises): Choose

$$b_1=b_2=0 \tag{25a}$$
$$a_1\alpha_1+a_2\alpha_2=\sqrt{P} \tag{25b}$$
$$\gamma_k=-\frac{\sqrt{P}}{\alpha_k},\qquad k\in\{1,2\}. \tag{25c}$$

Specializing (21)–(22) to our choice (25) results in parameters for our scheme that satisfy Conditions 1)–4) and constraint (8). Among all such choices of $a_1,a_2,b_1,b_2,\gamma_1,\gamma_2$ that also satisfy (25a), the one in (25) maximizes the rate constraints in (6).

Here, B_1 and B_2 have non-zero entries only on the first lower-diagonal (see Remark 4). Thus, by (22), starting from the third channel use in each subblock the transmitter simply sends a scaled version of the previous noise symbols.

Specializing Theorem 1 to (21)–(22d) and (25) yields:

Corollary 5: If $\sigma_1^2=\sigma_2^2=0$ our scheme achieves all rate pairs (R_1,R_2) that satisfy

$$0\le R_1\le\frac{1}{2\eta}\log\left(1+\left(\frac{P}{\alpha_1^2}\right)^{\eta-1}\left(1-\frac{\alpha_1}{\alpha_2}\right)^2\frac{P/\alpha_1^2}{2+2P/\alpha_2^2}\right),$$
$$0\le R_2\le\frac{1}{2\eta}\log\left(1+\left(\frac{P}{\alpha_2^2}\right)^{\eta-1}\left(1-\frac{\alpha_2}{\alpha_1}\right)^2\frac{P/\alpha_2^2}{2+2P/\alpha_1^2}\right).$$

Lemma 1: Let ξ,ζ be positive integers. If $1+\zeta\ge\xi$, then the function $f\colon\mathbb{Z}^+\to\mathbb{R}_0^+$, $f(\eta)=\frac{1}{2\eta}\log(1+\xi^{\eta-1}\zeta)$, has a maximum at $\eta=1$; otherwise it has a supremum at $\eta\to\infty$. Thus, for small P, $\eta=1$ is optimal in (26) (i.e., the feedback should not be used at all), and for large P, $\eta\to\infty$ is optimal. Letting $\eta\to\infty$, we obtain Corollary 1 in Section III.

Remark 5: Corollary 5 shows that when $\sigma_1^2=\sigma_2^2=0$ and $\alpha_1\neq\alpha_2$, our scheme achieves prelog $2\frac{\eta-1}{\eta}$ for every finite η. Thus, $\eta=3$ suffices to achieve a prelog larger than 1.

2) $\alpha_1,\alpha_2\neq0$ and $\sigma_1^2,\sigma_2^2>0$ (Noise Correlation in $(-1,0)$ or $(0,1)$): Let $p>0$ be defined through (11), and choose

$$a_k=\frac{(-1)^kp}{\alpha_k^2},\quad k\in\{1,2\}, \tag{27a}$$
$$b_k=-\frac{(-1)^k}{\alpha_k}\cdot\frac{p^2}{\alpha_1\alpha_2}\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right),\quad k\in\{1,2\}, \tag{27b}$$
$$\gamma_k=\frac{p}{\alpha_k}\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right),\quad k\in\{1,2\}. \tag{27c}$$

Specializing (21)–(22d) to the choice (27) results in parameters for our scheme that satisfy Conditions 1)–4) and the power constraint (8). Among all such choices of $a_1,a_2,b_1,b_2,\gamma_1,\gamma_2$, the one in (27) maximizes the constraints in (6) as $\eta\to\infty$.

Specializing Theorem 1 to (21)–(22d) and (27) results in:

Corollary 6: When $\alpha_1,\alpha_2\neq0$ and $\sigma_1^2,\sigma_2^2>0$, a rate pair (R_1,R_2) is achievable if it satisfies (28) on the next page.

By Lemma 1, $\eta=1$ is optimal for small p^2 (and thus small P, see (11)), and $\eta\to\infty$ is optimal for large p^2 (i.e., large P). Letting $\eta\to\infty$, one obtains Corollary 2 in Section III.

3) $\alpha_1=\alpha_2=0$ (Independent Noises): Let $q>0$ and $\delta\notin\{-1,0\}$ satisfy (17). Choose

$$a_k=(-\delta^2)^{k-1}q,\qquad k\in\{1,2\}, \tag{29a}$$
$$b_k=-q^2\delta^k(1+\delta),\qquad k\in\{1,2\}, \tag{29b}$$
$$\gamma_k=-q(-\delta)^{k-1}(1+\delta),\qquad k\in\{1,2\}. \tag{29c}$$

Specializing (21)–(22d) to the choice (29) results in parameters that satisfy Conditions 1)–4) and (8). Among all such choices, the one in (29) maximizes the constraints in (6) as $\eta\to\infty$.

Corollary 7: For $\alpha_1=\alpha_2=0$ our scheme achieves all rate pairs (R_1,R_2) satisfying (30) on top of the next page.

Again, $\eta=1$ is optimal for small P and $\eta\to\infty$ for large P. Letting $\eta\to\infty$ we obtain Corollary 3 in Section III.

C. Improved Choice of Parameters

We improve (generalize) our choices of parameters using an idea from [10]. For $k\in\{1,2\}$ and some real numbers a_1, a_2, b_1, b_2, c_1, c_2, γ_1, γ_2, we choose the vectors $\mathbf{u}_k,\mathbf{v}_k$ as in

$$0 \le R_1 \le \frac{1}{2\eta} \log\left(1 + \frac{\left(\frac{p^2}{\alpha_1^2}\right)^{\eta-1}\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right)^{2\eta}\alpha_1^2}{\left(\alpha_1^2 + \frac{p^2}{\alpha_1^2\alpha_2^2}\sigma_1^2 + \sigma_1^2 + \frac{p^2}{\alpha_2^4}\sigma_2^2\right)} \cdot \frac{P}{\left(2 + 2\frac{p^2}{\alpha_2^2}\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right)^2\right)}\right) \tag{28a}$$

$$0 \le R_1 \le \frac{1}{2\eta} \log\left(1 + \frac{\left(\frac{p^2}{\alpha_2^2}\right)^{\eta-1}\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right)^{2\eta}\alpha_2^2}{\left(\alpha_2^2 + \frac{p^2}{\alpha_1^2\alpha_2^2}\sigma_2^2 + \sigma_2^2 + \frac{p^2}{\alpha_1^4}\sigma_1^2\right)} \cdot \frac{P}{\left(2 + 2\frac{p^2}{\alpha_1^2}\left(\frac{1}{\alpha_1}-\frac{1}{\alpha_2}\right)^2\right)}\right) \tag{28b}$$

$$0 \le R_1 \le \frac{1}{2\eta} \log\left(1 + \frac{\left(q^2(1+\delta)^2\right)^{\eta-1}(1+\delta)^2}{(\delta^2 q^2\sigma_1^2 + \sigma_1^2 + \delta^4 q^2 \sigma_2^2)} \cdot \frac{P}{(2 + 2q^2\delta^2(1+\delta)^2)}\right) \tag{30a}$$

$$0 \le R_2 \le \frac{1}{2\eta} \log\left(1 + \frac{\left(q^2\delta^2(1+\delta)^2\right)^{\eta-1}\left(1+\frac{1}{\delta}\right)^2}{(\delta^2 q^2\sigma_2^2 + \sigma_2^2 + q^2 \sigma_1^2)} \cdot \frac{P}{(2 + 2q^2(1+\delta)^2)}\right) \tag{30b}$$

---

(22) and the matrix B_k Toeplitz with the following entries. For $j \in \{1, \ldots, \lceil \frac{\eta-1}{2} \rceil\}$ the entries on the $2j$-th diagonal below the main diagonal of B_k equal $a_k c_k^{i-1}$, and for $j \in \{1, \ldots, \lfloor \frac{\eta-1}{2} \rfloor\}$ the entries on its $2j+1$-th diagonal below the main diagonal equal $b_k c_k^{j-1}$. The idea is to choose $a_1, a_2, b_1, b_2, c_1, c_2, \gamma_1, \gamma_2$ in a way that:

1) For every fixed positive integer j, the contribution of $Z_{k,(i-1)\eta+j}$ on $I_{1,i}$ and $I_{2,i}$ vanishes exponentially as $\eta \to \infty$, for $k \in \{0,1,2\}$. This way the total noise-variances of $I_{1,i}$ and $I_{2,i}$ are bounded in η, and do not influence the achievable rates as $\eta \to \infty$.
2) The vector $\mathbf{v}_1$ is orthogonal to $\mathbf{u}_2$ but not to $\mathbf{u}_1$.
3) The vector $\mathbf{v}_2$ is orthogonal to $\mathbf{u}_1$ but not to $\mathbf{u}_2$.

For the special cases of fully correlated noises and of uncorrelated noises we present choices for $a_1, a_2, b_1, b_2, c_1, c_2, \gamma_1, \gamma_2$ satisfying Conditions 1)–3) and the power constraint (8).

1) $\sigma_1^2 = \sigma_2^2 = 0$: Choose $r \notin \{0, -\alpha_1, -\alpha_2\}$ and

$$a_1\alpha_1 + a_2\alpha_2 = -r\sqrt{\frac{P}{r^2+P}} \tag{31a}$$

$$b_1\alpha_1 + b_2\alpha_2 = -r\frac{P}{r^2+P} \tag{31b}$$

$$c_1 = c_2 = \frac{P}{r^2+P} \tag{31c}$$

$$\gamma_k = \sqrt{\frac{P}{r^2+P}}\frac{(\alpha_k + r)}{\alpha_k}, \qquad k \in \{1,2\}. \tag{31d}$$

Corollary 8: Let $\sigma_1^2 = \sigma_2^2 = 0$ and $\alpha_1 \neq \alpha_2$. With the choice in (31) and $\eta \to \infty$ our scheme achieves all rate pairs (R_1, R_2) satisfying

$$0 \le R_k \le \frac{1}{2}\log^+\left(\frac{P}{\alpha_k^2} \cdot \frac{(\alpha_k + r)^2}{(r^2+P)}\right), \quad r \notin \{0, -\alpha_1, -\alpha_2\}.$$

Here, $r = \frac{P}{\alpha_2}$ maximizes R_1 and $r = \frac{P}{\alpha_2}$ maximizes R_2.

2) $\alpha_1 = \alpha_2 = 0$: Choose $a_1, a_2, b_1, b_2 \neq 0$ and $c_1, c_2 \in (-1, 1)$ so that

$$\frac{b_2^2}{a_2^2} - a_1\frac{b_2}{a_2} + b_1 = c_1 \tag{32a}$$

$$\frac{b_1^2}{a_1^2} - a_2\frac{b_1}{a_1} + b_2 = c_2 \tag{32b}$$

$$\frac{a_1^2 + b_1^2}{1 - c_1^2}\sigma_1^2 + \frac{a_2^2 + b_2^2}{1 - c_2^2}\sigma_2^2 = P, \tag{32c}$$

and let

$$\gamma_1 = -\frac{b_2}{a_2} \quad \text{and} \quad \gamma_2 = -\frac{b_1}{a_1}. \tag{32d}$$

Corollary 9: Let $\alpha_1 = \alpha_2 = 0$. With the choice in (32) and $\eta \to \infty$ our scheme achieves all rate pairs (R_1, R_2) satisfying

$$0 \le R_k \le \frac{1}{2}\log\left(\frac{b_k^2}{a_k^2}\right), \qquad k \in \{1,2\}, \tag{33}$$

for $a_1, a_2, b_1, b_2 \neq 0$ satisfying (32) for some $c_1, c_2 \in (-1, 1)$.

REFERENCES

[1] E. Ardestanizadeh, P. Minero, and M. Franceschetti, "LQG control approach to Gaussian broadcast channels with feedback." submitted to *IEEE Trans on IT.*, Feb 2011.

[2] T. M. Cover and S. Pombra, "Gaussian feedback capacity," *IEEE Trans. on Inf. Th.* vol. 35, pp. 37–43, Jan. 1989.

[3] A. El Gamal, The feedback capacity of degraded broadcast channels, *IEEE Trans. on Inf. Th.*, vol. 24, no. 3, pp. 379-381, 1978.

[4] N. Elia, "When Bode meets Shannon: Control oriented feedback communication schemes," *IEEE Trans. on Autom. Control*, vol 47, pp. 1477–1488, 2004.

[5] M. Gastpar and M. Wigger, "The pre-log of Gaussian broadcast with feedback can be two," in *Proc. of ISIT 2008*, July 2008.

[6] M. Gastpar, A. Lapidoth, and M. Wigger, "When feedback doubles the prelog of Gaussian networks," submitted to *IEEE Trans. on Inf. Th.*, avaiable at `http://arxiv.org/abs/1003.6082`.

[7] A. Lapidoth and M. Wigger, "On the Gaussian MAC with imperfect feedback," *IEEE Trans. on Inf. Th.*, vol. 56, pp. 5432–5477, Nov. 2010.

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

[9] L. H. Ozarow and C. S. K. Leung, "An achievable region and outer bound for the Gaussian broadcast channel with feedback." *IEEE Trans. on Inf. Th.*, vol. 30, pp. 667–671, July 1984.

[10] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback—Part I: No bandwidth constraint," *IEEE Trans. on Inf. Th.L*, vol. 12, pp. 172–182, April 1966.

[11] W. Wu, S. Vishwanath, and A. Arapostathis, "Gaussian Interference Networks with Feedback: Duality, Sum capacity and Dynamic Team Problem," in Proc. of *43rd Allerton Conference, Sep. 2005.*