Linear-Feedback MAC-BC Duality for Correlated BC-Noises, and Iterative Coding

Selma Belhadj Amor, Michèle Wigger

Linear-Feedback MAC-BC Duality for Correlated BC-Noises, and Iterative Coding

Selma Belhadj Amor
Inria, Lyon, France
selma.belhadj-amor@inria.fr

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

***Abstract*—In this paper, we show that for the two-user Gaussian broadcast channel with correlated noises and perfect feedback the largest region that can be achieved by linear-feedback schemes equals the largest region that can be achieved over a dual multi-access channel when in this latter the channel inputs are subject to a "non-standard" sum-power constraint that depends on the BC-noise correlation. Combining this new duality result with Ozarow's MAC-scheme gives us an elegant achievable region for the Gaussian BC with correlated noises.**

We then present a constructive iterative coding scheme for the non-symmetric Gaussian BC with uncorrelated noises that is sum-rate optimal among all linear-feedback schemes. This coding scheme shows that the connection between the MAC and the BC optimal schemes is tighter than what is suggested by our duality result on achievable rates. In fact, it is linear-feedback sum-rate optimal to use Ozarow MAC-encoders and MAC-decoders—rearranged—to code over the BC.

I. Introduction

The *broadcast channel* (BC), which preluded to Cover [5], is a one-to-many channel that involves simultaneous communication of information from one single transmitter to multiple receivers. In this paper, we consider the two-user BC with perfect feedback, where at each time, upon receiving its dedicated channel output, each receiver is allowed to send back noiselessly its channel output to the transmitter via a feedback link. Unlike for point-to-point channels, feedback can increase capacity of BCs [6]. That this is also the case for BCs, was proved by Ozarow and Leung [14] in the case of two receivers, by Kramer when there are more than two receivers [12], and by Bhaskaran for two receivers and when there is only feedback from the stronger receiver. The main idea in their schemes is to send scaled versions of receiver-LMMSE estimation-errors, a strategy that is capacity-achieving for Gaussian point-to-point and multi-access channels [13], [15]. For the BC such an LMMSE-strategy is however not optimal and for some channel parameters it even fails to achieve the nofeedback capacity region. In fact, the Linear-Quadratic Gaussian (LQG) control-theory inspired codes by Elia [8], Wu *et al.* [18], and Ardestanizadeh *et al.* [1] can outperform Ozarow and Leung's and Kramer's schemes. For symmetric Gaussian two-user BCs these coding schemes achieve the largest known sum-rate with perfect feedback. In particular, they achieve the same sum-rate over the two-user symmetric BC with power P as Ozarow's and Kramer's [12], [13] schemes achieve over the two-user symmetric MAC with perfect feedback under individual power constraints $P_1 = P_2 = \frac{P}{2}$.

Gastpar *et al.* [10], [11] proposed a coding scheme for the two-user Gaussian BCs where the noises at the receivers are correlated. In the asymptotic regime where the allowed input power $P \to \infty$, the performance of their coding scheme approaches the sum-capacity, irrespective of the correlation between the noise sequences.

All these presented coding schemes are *linear-feedback coding schemes* where the feedback is used only linearly. Notice that linear-feedback coding schemes are optimal for the Gaussian point-to-point and multi-access channels and they achieve the largest known rates for other Gaussian networks. The only exception where linear-feedback schemes have shown to be strictly suboptimal is the Gaussian BC with only common message [19], which is not considered here.

We have recently shown a MAC-BC duality result [2], [3], [4] showing that with perfect feedback and when restricting to linear-feedback schemes the regions achieved over the two-user multi-antenna Gaussian memoryless MAC and BC coincide, if:

- the channel matrices of the MAC and the BC are transposes of each other;
- the same (sum-)power constraint P is imposed on the MAC and the BC inputs; and
- the noises at the BC-receivers are independent.

In the scalar case where the set of achievable rates over the Gaussian MAC using linear-feedback schemes is known—it equals Ozarow's perfect feedback capacity region under a sum-power constraint—our duality result readily established the set of achievable rates using linear-feedback schemes over the scalar Gaussian BC with uncorrelated noises. With feedback, the capacity region of the Gaussian BC depends on the noise correlation. In this paper, we prove a similar duality result for the scalar two-user Gaussian BC when the noises are correlated. As we shall show, such a BC is linear-feedback dual to (i.e., linear-feedback coding schemes achieve the same rates as over) a MAC with the same channel gains subject to a "non-standard" total power constraint which depends on the BC-noise correlation. This duality result combined with Ozarow's MAC-scheme for the Gaussian MAC with individual power constraints [13] allows us to readily obtain an elegant and simple achievable region for the Gaussian BC with correlated noises and perfect feedback. Previous achievable regions for the non-symmetric Gaussian BC with correlated noises [1],

[10] are more involved and difficult to compare to.

In the second part of this paper, we present a constructive, iterative linear-feedback coding scheme for the two-user scalar Gaussian BC with independent noises that achieves the linear-feedback sum-capacity. This result shows a much tighter connection between optimal codings for the scalar Gaussian MAC and BC with uncorrelated noises than what suggested by our duality result. In fact, Ozarow's sum-capacity achieving encoders and decoders for the MAC can be rearranged to form linear-feedback sum-rate optimal encoders and decoders for the BC. Specifically, the BC encoder should run the operations of the two MAC-encoders and send the sum of their outcomes over the BC, and each BC decoder should guess its desired message in exactly the same way that Ozarow's MAC decoder guessed this message. (A key observation here is that Ozarow's MAC decoder chooses to guess the two messages separately of each other.)

In the symmetric case, the linear-feedback sum-capacity of the BC with independent noises is achieved also by the Ardestanizadeh *et al.* scheme [1]. Our approach here is different because: a) it shows a much tighter connection between the optimal BC and MAC schemes than previously observed; and b) the analysis does not rely on control-theoretic arguments, but shows further dual relationships between quantities like power, error probability, and rate in the MAC and BC schemes.

Throughout the paper, the symbol E_d denotes the d-by-d *exchange matrix* which is 0 everywhere except on the counter-diagonal where it is 1 and for a given d_1-by-d_2 matrix A, the reverse-image of A is defined as $\bar{\mathsf{A}} \triangleq \mathsf{E}_{d_2}\mathsf{A}^{\mathsf{T}}\mathsf{E}_{d_1}$.

II. Gaussian Broadcast Channels with Feedback

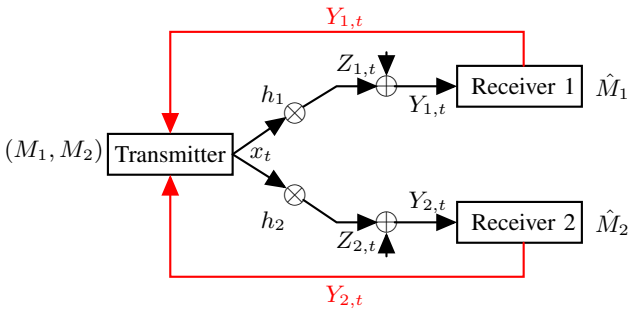

Fig. 1. Two-user scalar memoryless Gaussian BC with perfect feedback.

Consider the two-user scalar memoryless Gaussian BC with perfect feedback depicted in Figure 1. At each time $t \in \mathbb{N}$, if $x_t \in \mathbb{R}$ denotes the transmitter's channel input, Receiver $i \in \{1, 2\}$ observes the real channel output

$$Y_{i,t} = h_i x_t + Z_{i,t}, \tag{1}$$

where h_1 and h_2 are constant non-zero channel coefficients and the sequence of the noise-pairs $\{(Z_{1,t}, Z_{2,t})\}_{t=1}^n$ is drawn independently and identically distributed (i.i.d.) according to a centered Gaussian distribution of covariance matrix $\begin{pmatrix} 1 & \lambda \\ \lambda & 1 \end{pmatrix}$.

The transmitter wishes to convey a message M_1 to Receiver 1 and an independent message M_2 to Receiver 2. The messages are independent of the noise sequences $\{\mathbf{Z}_{1,t}\}_{t=1}^n$ and $\{\mathbf{Z}_{2,t}\}_{t=1}^n$ and 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 R_1 and R_2 denote the rates of transmission and n the blocklength.

The transmitter observes causal noise-free channel output feedback from both receivers. Thus, the time-t channel input X_t can depend on all previous channel outputs Y_1^{t-1} and Y_2^{t-1} and messages M_1 and M_2:

$$X_t = \varphi_t^{(n)}(M_1, M_2, Y_1^{t-1}, Y_2^{t-1}), \quad t \in \{1, \ldots, n\}, \tag{2}$$

for some encoding function $\varphi_t^{(n)} \colon \mathcal{M}_1 \times \mathcal{M}_2 \times \mathbb{R}^{t-1} \times \mathbb{R}^{t-1} \to \mathbb{R}$. The channel inputs have to satisfy an *expected average block-power constraint*

$$\frac{1}{n}\sum_{t=1}^{n} \mathbf{E}[X_t^2] \le P, \tag{3}$$

where the expectation is over the messages and the realizations of the channel.

Each Receiver $i \in \{1, 2\}$ produces the guess $\hat{M}_i^{(n)} = \phi_i^{(n)}(Y_i^n)$ for some decoding function $\phi_i^{(n)} \colon \mathbb{R}^n \to \mathcal{M}_i$. The average error probability is

$$P_{\text{e,BC}}^{(n)} \triangleq \Pr\left\{(\hat{M}_1^{(n)} \neq M_1) \text{ or } (\hat{M}_2^{(n)} \neq M_2)\right\}. \tag{4}$$

A rate-pair (R_1, R_2) is achievable over the Gaussian BC with feedback and power constraint P, if there exists a sequence of encoding and decoding functions $\left\{\{\varphi_t^{(n)}\}_{t=1}^n, \phi_1^{(n)}, \phi_2^{(n)}\right\}_{n=1}^{\infty}$ satisfying the power constraint (3) and such that the average probability of error $P_{\text{e,BC}}^{(n)}$ tends to zero as n tends to infinity.

The closure of the union of all achievable regions is called *capacity region* and the supremum of the sum $R_1 + R_2$ over all achievable pairs (R_1, R_2) is called *sum-capacity*.

This Gaussian BC is physically degraded if and only if

$$\lambda = \frac{h_1}{h_2} \qquad \text{or} \qquad \lambda = \frac{h_2}{h_1}. \tag{5}$$

For instance, if $h_1 = h_2 = \lambda = 1$ the two receivers observe the same sequence.

In this paper, we focus on *linear-feedback schemes* where the transmitter's channel input is a linear combination of the previous feedback signals and an information-carrying vector that depends only on the messages (M_1, M_2). Specifically, we assume the channel input vector $\mathbf{X} \triangleq \left(X_1, \ldots, X_n\right)^{\mathsf{T}}$ can be written as:

$$\mathbf{X} = \mathbf{W} + \mathsf{A}_1\mathbf{Y}_1 + \mathsf{A}_2\mathbf{Y}_2, \tag{6}$$

where $\mathbf{Y}_i \triangleq \left(Y_{i,1}, \ldots, Y_{i,n}\right)^{\mathsf{T}}$ represents the channel output vector at Receiver i, A_1 and A_2 are n-by-n strictly lower-triangular matrices, and $\mathbf{W}$ is an n-dimensional information-carrying vector $\mathbf{W} = \xi^{(n)}(M_1, M_2)$. The strict lower-triangularity of A_1 and A_2 ensures that the feedback is used strictly causally. The mapping $\xi^{(n)} \colon \mathcal{M}_1 \times \mathcal{M}_2 \to \mathbb{R}^n$ and the decoding operations $\phi_1^{(n)}$ and $\phi_2^{(n)}$ can be arbitrary.

The set of all rate-pairs achieved by linear-feedback schemes is called *linear-feedback capacity region* and is

denoted by $\mathcal{C}_{\text{BC}}^{\text{linfb}}(h_1, h_2, \lambda; P)$. The largest sum-rate achieved by a linear-feedback scheme is called linear-feedback sum-capacity and is denoted by $\mathcal{C}_{\text{BC},\Sigma}^{\text{linfb}}(h_1, h_2, \lambda; P)$.

III. The Dual Gaussian Multi-Access Channel

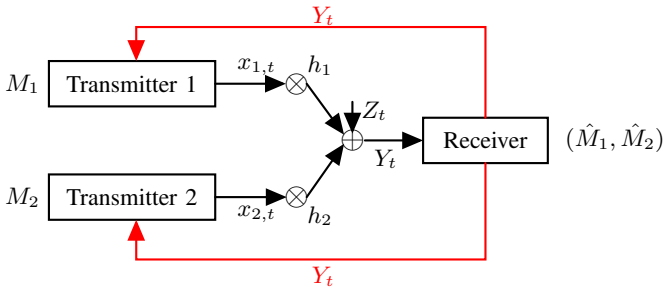

Fig. 2. Two-user Gaussian memoryless MAC with perfect feedback.

Consider the two-user memoryless scalar Gaussian MAC with perfect output-feedback in Figure 2. At each time $t \in \mathbb{N}$, if $x_{1,t}$ and $x_{2,t}$ denote the real symbols sent by Transmitters 1 and 2, the receiver observes the real channel output

$$Y_t = h_1 x_{1,t} + h_2 x_{2,t} + Z_t, \tag{7}$$

where h_1 and h_2 are constant nonzero channel coefficients and $\{Z_t\}$ is a sequence of i.i.d. zero-mean unit-variance Gaussian random variables.

The goal of communication is that Transmitters 1 and 2 convey the independent messages M_1 and M_2 to the common receiver, where the pair (M_1, M_2) is independent of the noise sequence $\{\mathbf{Z}_t\}$. (Recall that M_i is uniformly distributed over $\mathcal{M}_i = \{1, \ldots, \lfloor 2^{nR_i} \rfloor\}$)

The two transmitters observe perfect feedback from the channel outputs. Thus, the time-t channel input at Transmitter $i \in \{1, 2\}$ can depend on all previous channel outputs Y^{t-1} and its message M_i:

$$X_{i,t} = f_{i,t}^{(n)}(M_i, Y^{t-1}), \quad t \in \{1, \ldots, n\}, \tag{8}$$

for some encoding function $f_{i,t}^{(n)} \colon \mathcal{M}_i \times \mathbb{R}^{t-1} \to \mathbb{R}$.

The channel inputs $\{X_{1,t}\}_{t=1}^n$ and $\{X_{2,t}\}_{t=1}^n$ have to satisfy a sligthly non-standard total input-power constraint:

$$\mathsf{E}\big[\|\mathbf{X}_1\|^2\big] + \mathsf{E}\big[\|\mathbf{X}_2\|^2\big] + 2\lambda \mathsf{E}[\langle \mathbf{X}_1, \mathbf{X}_2 \rangle] \le nP \tag{9}$$

where $\mathbf{X}_i = (X_{i,1}, \ldots, X_{i,n})^\intercal$, for $i \in \{1, 2\}$, and where $\lambda \in [-1, 1]$ is a given parameter. Notice that for $\lambda = 0$ the power constraint (9) specializes to a standard sum-power constraint.

The receiver produces a guess $(\hat{M}_1^{(n)}, \hat{M}_2^{(n)}) = \Phi^{(n)}(Y^n)$ by means of a decoding function $\Phi^{(n)} \colon \mathbb{R}^n \to \mathcal{M}_1 \times \mathcal{M}_2$.

The average probability of error is

$$P_{\text{e,MAC}}^{(n)} \triangleq \Pr\big\{(\hat{M}_1^{(n)}, \hat{M}_2^{(n)}) \neq (M_1, M_2)\big\}. \tag{10}$$

Achievability, capacity region, and sum-capacity are defined analogously to the BC.

We focus on linear-feedback schemes where the channel inputs can be written as

$$\mathbf{X}_i = \mathbf{W}_i + \mathsf{C}_i \mathbf{Y}, \qquad i \in \{1, 2\}, \tag{11}$$

where $\mathbf{Y} \triangleq \big(Y_1, \ldots, Y_n\big)^\intercal$ is the channel output vector, C_1 and C_2 are n-by-n strictly lower-triangular matrices and $\mathbf{W}_i$ is an n-dimensional information-carrying vector $\mathbf{W}_i = \xi_i^{(n)}(M_i)$. The mapping $\xi_i^{(n)} \colon \mathcal{M}_i \to \mathbb{R}^n$ as well as the decoder mapping $\Phi^{(n)}$ can be arbitrary (also non-linear). The strict-lower-triangularity of the matrices C_1 and C_2 ensures that the feedback is used in a strictly causal way.

Linear-feedback capacity for the MAC is defined similarly to the BC. We denote it by $\mathcal{C}_{\text{MAC,CorrPower}}^{\text{linfb}}(h_1, h_2; \lambda, P)$, where we use the subscript "CorrPower" to emphasize the non-standard power constraint in (9).

IV. MAC-BC Duality for BC with Correlated Noises

Theorem 1.

$$\mathcal{C}_{\text{BC}}^{\text{linfb}}(h_1, h_2, \lambda; P) = \mathcal{C}_{\text{MAC,CorrPower}}^{\text{linfb}}(h_1, h_2; \lambda, P). \tag{12}$$

Proof: A sketch of proof is provided in Appendix A. ■

Remark 1. *Let $|h_1| \ge |h_2|$ and assume that the BC in Section II is physically degraded, i.e., that $\lambda = \frac{h_2}{h_1}$. In this case, our MAC power-constraint in* (9) *is equivalent to:*

$$\begin{aligned} h_1^2 nP &\ge \mathsf{E}\big[\|h_1\mathbf{X}_1 + h_2\mathbf{X}_2\|^2\big] + (h_1^2 - h_2^2)\mathsf{E}\big[\|\mathbf{X}_2\|^2\big] \\ &\quad + 2(h_1^2\lambda - h_1 h_2)\mathsf{E}[\langle \mathbf{X}_1, \mathbf{X}_2 \rangle] \\ &= \mathsf{E}\big[\|h_1\mathbf{X}_1 + h_2\mathbf{X}_2\|^2\big] + (h_1^2 - h_2^2)\mathsf{E}\big[\|\mathbf{X}_2\|^2\big]. \end{aligned} \tag{13}$$

The term $\mathsf{E}\big[\|h_1\mathbf{X}_1 + h_2\mathbf{X}_2\|^2\big]$ corresponds to the received power. By our assumption $|h_1| \ge |h_2|$, the term $(h_1^2 - h_2^2)\mathsf{E}\big[\|\mathbf{X}_2\|^2\big]$ is nonnegative and thus, Inequality (13) *implies*

$$\mathsf{E}\big[\|h_1\mathbf{X}_1 + h_2\mathbf{X}_2\|^2\big] \le h_1^2 nP. \tag{14}$$

Even with feedback, the maximum sum-rate over the Gaussian MAC is upper bounded by $1/2\log(1 + (1/n)\mathsf{E}\big[\|h_1\mathbf{X}_1 + h_2\mathbf{X}_2\|^2\big])$, which cannot exceed $1/2\log(1 + h_1^2 P)$. Thus, in this case the sum-capacity of the MAC under consideration is the same with and without feedback [9]. Notice that for these parameters the same also applies for the dual BC: feedback does not increase capacity.

Ozarow's scheme [13] for the two-user Gaussian MAC with perfect feedback and *individual* power constraints P_1 and P_2 at both transmitters is easily adapted to our two-user Gaussian MAC in Section III with the non-standard power constraint in (9). It suffices that in Ozarow's scheme the individual powers $P_1 \ge 0$ and $P_2 \ge 0$ are chosen so that

$$P_1 + P_2 + 2\lambda\rho\sqrt{P_1 P_2} \le P, \tag{15a}$$

for some parameter

$$\rho \in [0, \rho^\star(h_1, h_2, P_1, P_2)] \tag{15b}$$

where $\rho^\star(h_1, h_2, P_1, P_2)$ denotes Ozarow's correlation coefficient, which is the unique positive solution to the following quartic equation in x:

$$\begin{aligned} &(1 + h_1^2 P_1(1 - x^2)) \cdot (1 + h_2^2 P_2(1 - x^2)) \\ &\qquad = 1 + h_1^2 P_1 + h_2^2 P_2 + 2|h_1||h_2| x\sqrt{P_1 P_2}. \end{aligned}$$

Combined with our duality result in Theorem 1 we then obtain:

Theorem 2. *All nonnegative rate-pairs (R_1, R_2) are achievable over the BC with correlated noises and output feedback, if they satisfy Ozarow's rate-constraints*

$$R_1 \le \frac{1}{2}\log(1+h_1^2P_1(1-\rho^2))$$
$$R_2 \le \frac{1}{2}\log(1+h_1^2P_2(1-\rho^2))$$
$$R_1+R_2 \le \frac{1}{2}\log(1+h_1^2P_1+h_2^2P_2+2|h_1||h_2|\rho\sqrt{P_1P_2})$$

for some P_1, P_2, and $\rho \in [0,1]$ satisfying (15).

V. Constructive Sum-Rate Optimal BC-Scheme for Uncorrelated Noises ($\lambda = 0$)

In this section, we assume that the noises at the BC-receivers are uncorrelated ($\lambda = 0$).

We describe a constructive coding scheme for the Gaussian BC that achieves the linear-feedback sum-capacity. Our duality results in a previous work [3] suggest to use the same feedback-matrices A_1 and A_2 as Ozarow used in his MAC perfect-feedback capacity-achieving coding scheme; they however do not give us an explicit construction on how to choose the codewords. Here, we show that on the two-user Gaussian BC it is linear-feedback sum-rate optimal to use exactly the same encoders and decoders as Ozarow used in his sum-rate optimal MAC scheme [13]. This is possible because in Ozarow's scheme, which is schematically depicted in Figure 3, the receiver decodes the two messages M_1 and M_2 separately. So, as we shall prove, it is linear-feedback sum-rate optimal for the two-user Gaussian BC that (see Figure 4):

- the BC-Transmitter implements both Ozarow's MAC-encoders 1 and 2, and then sends the sum of the symbols produced by these encoders;
- BC-Receiver 1 implements the part of Ozarow's MAC-decoder that decodes Message M_1; and
- BC-Receiver 2 implements the part of Ozarow's MAC-decoder that decodes Message M_2.

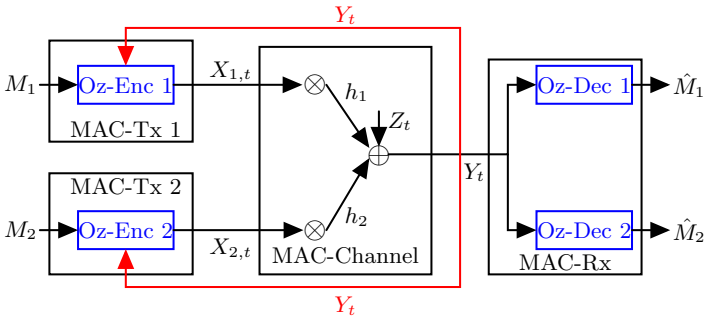

Fig. 3. Ozarow's MAC-scheme

We first revise the version of Ozarow's scheme in [7], where we allow the three terminals to share common randomness (that is independent of the messages). We then describe and analyze the equivalent BC scheme.

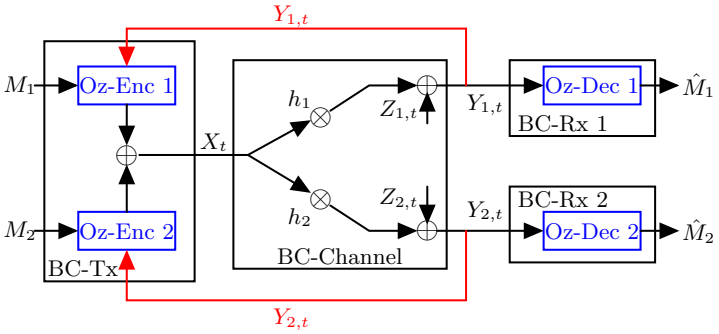

Fig. 4. New BC-scheme

A. Ozarow's Sum-Capacity Achieving Scheme for the Scalar Gaussian MAC with Feedback

Before transmission starts, each message M_i, for $i \in \{1,2\}$, is mapped into the real-valued message point

$$\Theta_i(M_i) \triangleq -(M_i-1)\Delta_i + \sqrt{P_i}, \tag{16}$$

where

$$\Delta_i \triangleq \frac{2\sqrt{P_i}}{\lfloor 2^{nR_i}\rfloor}. \tag{17}$$

The first two channel uses are part of an initialization procedure. To simplify notation, we assume that the initialization takes place at times $t=-2$ and $t=-1$. The channel inputs during the initialization procedure are,

$$t=-2: \qquad X_{1,-2}=0, \qquad X_{2,-2}=\Theta_2(M_2), \tag{18a}$$
$$t=-1: \qquad X_{1,-1}=\Theta_1(M_1), \qquad X_{2,-1}=0. \tag{18b}$$

Define for each $i \in \{1,2\}$,

$$\Xi_i \triangleq \sqrt{1-\rho^\star}Z_{-i} + \sqrt{\rho^\star}Z_0, \tag{19}$$

where we write $\rho^\star$ as a short-hand notation for $\rho^\star(h_1,h_2;P_1,P_2)$ and where Z_0 denotes a standard Gaussian random variable that acts as common randomness known at all terminals and independent of the messages M_1 and M_2.

Through the feedback, Transmitter 1 learns Z_{-1} and Transmitter 2 learns Z_{-2}. After the initialization phase, each Transmitter $i \in \{1,2\}$ can thus compute Ξ_i.

The idea is that during the remaining channel uses $0,\ldots,n-1$, Transmitter i describes Ξ_i to the receiver. Since

$$\sqrt{1-\rho^\star}Y_{-i} + \sqrt{\rho^\star}Z_0 = \sqrt{1-\rho^\star}h_i\Theta_i(M_i) + \Xi_i, \tag{20}$$

for any estimate of Ξ_i, Receiver i immediately obtains also an estimate of $\Theta_i(M_i)$. To describe Ξ_i to the receiver, Transmitter i produces the following inputs:

$$X_{i,0} = \sqrt{P_i}\Xi_i, \tag{21a}$$
$$X_{i,t} = \gamma_i\left(X_{i,t-1} - \delta_i Y_{t-1}\right), \quad t=1,\ldots,n-1, \tag{21b}$$

where

$$\gamma_1 \triangleq \sqrt{1+h_1^2P_1(1-(\rho^\star)^2)}, \tag{22a}$$
$$\gamma_2 \triangleq -\sqrt{1+h_2^2P_2(1-(\rho^\star)^2)}. \tag{22b}$$

and where

$$\delta_1 \triangleq \frac{h_1 P_1 + \rho^\star h_2 \sqrt{P_1 P_2}}{h_1^2 P_1 + h_2^2 P_2 + 2\rho^\star h_1 h_2 \sqrt{P_1 P_2} + 1}, \tag{23a}$$

$$\delta_2 \triangleq \frac{h_2 P_2 + \rho^\star h_1 \sqrt{P_1 P_2}}{h_1^2 P_1 + h_2^2 P_2 + 2\rho^\star h_1 h_2 \sqrt{P_1 P_2} + 1}. \tag{23b}$$

Notice that δ_i is so that $(X_{i,t} - \delta_i Y_t)$ equals the LMMSE estimation error when estimating $X_{i,t}$ based on Y_t, for all $t = 0, \ldots, n-1$.

After reception of output symbols $Y_{-2}, \ldots, Y_{n-1}$, for each $i \in \{1, 2\}$, the receiver calculates the LMMSE-estimate $\hat{\Xi}_i^{(n-1)}$ of Ξ_i based on $Y_0, \ldots, Y_{n-1}$:

$$\hat{\Xi}_i^{(n-1)} \triangleq \delta_i (\sqrt{P_i})^{-1} \sum_{\tau=0}^{n-1} \gamma_i^{1-\tau} Y_\tau, \tag{24}$$

and forms its estimate of message point Θ_i:

$$\begin{aligned}\hat{\Theta}_i &\triangleq \frac{1}{h_i} \left(Y_{-i} + \frac{\sqrt{\rho^\star}}{\sqrt{1-\rho^\star}} Y_0 - \frac{1}{\sqrt{1-\rho^\star}} \hat{\Xi}_i^{(n-1)} \right) \\ &= \Theta_i(M_i) + \frac{1}{h_i \sqrt{1-\rho^\star}} \left(\Xi_i - \hat{\Xi}_i^{(n-1)} \right). \end{aligned} \tag{25}$$

It then decodes Message M_i using nearest-neighbor decoding based on its guess of message point $\hat{\Theta}_i$:

$$\hat{M}_i = \underset{m_i \in \{1, \ldots, \lfloor 2^{nR_i} \rfloor\}}{\operatorname{argmin}} \left| \Theta_i(m_i) - \hat{\Theta}_i \right|. \tag{26}$$

The single MAC-receiver thus decodes the two messages M_1 and M_2 completely independently. We make use of this observation to adapt the scheme to the BC:

B. Scheme for the Scalar Gaussian BC

In our scheme the single BC-transmitter mimics the two distributed MAC-transmitters of Ozarow's scheme. For each $t \in \{-2, \ldots, n-1\}$ the single transmitter calculates the two MAC-inputs $\{X_{1,t}\}_{t=-2}^{n-1}$ and $\{X_{2,t}\}_{t=-2}^{n-1}$ as in Subsection (V-A), but where the MAC-output Y_t and the MAC-noise Z_t need to be replaced by one of the BC-outputs $Y_{1,t}$ or $Y_{2,t}$ and BC-noises $Z_{1,t}$ or $Z_{2,t}$, respectively. The BC transmitter then sends the sum

$$X_t = X_{1,t} + X_{2,t}, \qquad t = 0, \ldots, n-1. \tag{27}$$

After reception of the outputs $Y_{i,-i}, Y_{i,0}, \ldots, Y_{i,n-1}$, each Receiver $i \in \{1, 2\}$ calculates the estimate $\hat{\Xi}_i^{(n-1)}$ of Ξ_i based on $Y_{i,0}, \ldots, Y_{i,n-1}$ as

$$\hat{\Xi}_i^{(n-1)} \triangleq \delta_i (\sqrt{P_i})^{-1} \sum_{\tau=0}^{n-1} \gamma_i^{1-\tau} Y_{i,t}, \quad i \in \{1, 2\}, \tag{28}$$

and forms

$$\begin{aligned}\hat{\Theta}_i &\triangleq \frac{1}{h_i} \left(Y_{i,-i} + \frac{\sqrt{\rho^\star}}{\sqrt{1-\rho^\star}} Z_0 - \frac{1}{\sqrt{1-\rho^\star}} \hat{\Xi}_i^{(n-1)} \right) \\ &= \Theta_i(M_i) + \frac{1}{h_i \sqrt{1-\rho^\star}} \left(\Xi_i - \hat{\Xi}_i^{(n-1)} \right). \end{aligned} \tag{29}$$

Finally, Receiver i performs the nearest-neighbor decoding in (26).

In the following Section VI we show that in the limit as $n \to \infty$ this linear-feedback coding scheme for the BC has average block-power tending to P[1], and that it achieves the same rates as Ozarow's MAC scheme, i.e., all rate-pairs (R_1, R_2) satisfying

$$0 \le R_1 < \frac{1}{2} \log \left(1 + h_1^2 P_1 (1 - \rho^{\star 2}) \right), \tag{30a}$$

$$0 \le R_2 < \frac{1}{2} \log \left(1 + h_2^2 P_2 (1 - \rho^{\star 2}) \right). \tag{30b}$$

VI. Proof of Performance of the BC Scheme in Section V

Before analyzing our BC scheme we recall parts of the analysis of Ozarow's MAC scheme from [7], [13]. These parts will help us analyze the BC scheme.

A. Vector-Description of Ozarow's MAC Scheme & Analysis

We first rewrite inputs and outputs of Ozarow's MAC-scheme (Subsection V-A) after the initialization phase using vector notation. We collect the inputs and outputs of channel uses $0, \ldots, n-1$ in the n-dimensional vectors

$$\mathbf{X}_1 \triangleq (X_{1,0}, \ldots, X_{1,n-1})^\mathsf{T} \tag{31}$$

$$\mathbf{X}_2 \triangleq (X_{2,0}, \ldots, X_{2,n-1})^\mathsf{T} \tag{32}$$

$$\mathbf{Y} \triangleq (Y_0, \ldots, Y_{n-1})^\mathsf{T}. \tag{33}$$

Recall the definitions of γ_i and δ_i in (22) and (23), and define the matrix

$$\mathsf{C}_i = -\delta_i \begin{pmatrix} 0 & 0 & \ldots & 0 \\ \gamma_i & 0 & & \\ \gamma_i^2 & \gamma_i & 0 & \\ \vdots & & & \\ \gamma_i^{n-1} & \ldots & \gamma_i & 0 \end{pmatrix}, \tag{34}$$

as well as

$$\mathsf{G} \triangleq (\mathsf{I} - h_1 \mathsf{C}_1 - h_2 \mathsf{C}_2)^{-1}, \tag{35}$$

$$\mathsf{D}_i = \mathsf{C}_i \mathsf{G}. \tag{36}$$

Define also the vectors

$$\mathbf{u}_i \triangleq \sqrt{P_i} \left(1, \gamma_i, \ldots, \gamma_i^{n-1}\right)^\mathsf{T}, \quad \text{and} \tag{37}$$

$$\mathbf{v}_i \triangleq \delta_i (\sqrt{P_i})^{-1} (1, \gamma_i^{-1}, \ldots, \gamma_i^{-n+1})^\mathsf{T}. \tag{38}$$

With these definitions, we can write the channel inputs as

$$\mathbf{X}_i = \mathbf{u}_i \Xi_i + \mathsf{C}_i \mathbf{Y}, \quad i \in \{1, 2\}, \tag{39}$$

and the channel outputs as

$$\begin{aligned}\mathbf{Y} &= h_1 \mathbf{X}_1 + h_2 \mathbf{X}_2 + \mathbf{Z} \\ &= h_1 \mathbf{u}_1 \Xi_1 + h_2 \mathbf{u}_2 \Xi_2 + (h_1 \mathsf{C}_1 + h_2 \mathsf{C}_2) \mathbf{Y} + \mathbf{Z} \end{aligned} \tag{40}$$

[1]To change our BC scheme to a scheme that satisfies the average block-power constraint P for all sufficiently large n, it suffices to scale the first inputs appropriately. As can be verified at hand of the proof steps in Section VI such a scaling does not change the set of achievable rates.

Subtracting $(h_1\mathsf{C}_1+h_2\mathsf{C}_2)\mathbf{Y}$ from both sides of (40) and then multiplying both sides from the left by the matrix G in (35), we obtain

$$\mathbf{Y} = \mathsf{G}\big(h_1\mathbf{u}_1\Xi_1 + h_2\mathbf{u}_2\Xi_2 + \mathbf{Z}\big). \tag{41}$$

The LMMSE estimate $\hat{\Xi}_i^{(n-1)}$ in (24) can be expressed as

$$\hat{\Xi}_i^{(n-1)} = \mathbf{v}_i^{\mathsf{T}}\mathbf{Y}. \tag{42}$$

Define now the two symmetric matrices Q_1 and Q_2 as the unique positive square roots of the square matrices

$$\mathsf{M}_1 = (\mathsf{I}+h_1\mathsf{D}_1)^{\mathsf{T}}(\mathsf{I}+h_1\mathsf{D}_1) + (h_1\mathsf{D}_2)^{\mathsf{T}}(h_1\mathsf{D}_2), \tag{43a}$$
$$\mathsf{M}_2 = (h_2\mathsf{D}_1)^{\mathsf{T}}(h_2\mathsf{D}_1) + (\mathsf{I}+h_2\mathsf{D}_2)^{\mathsf{T}}(\mathsf{I}+h_2\mathsf{D}_2). \tag{43b}$$

Define also the vectors α_1, β_1, α_2, and β_2 as:

$$\alpha_i = (\mathsf{I}+h_i\mathsf{D}_i)\mathbf{u}_i, \quad i\in\{1,2\} \tag{44a}$$
$$\beta_1 = h_1\mathsf{D}_2\mathbf{u}_1 \tag{44b}$$
$$\beta_2 = h_2\mathsf{D}_1\mathbf{u}_2. \tag{44c}$$

We can rewrite the channel input vectors in (39) as

$$\mathbf{X}_1 = \alpha_1\Xi_1 + \beta_2\Xi_2 + \mathsf{D}_1\mathbf{Z} \tag{45a}$$
$$\mathbf{X}_2 = \alpha_2\Xi_2 + \beta_1\Xi_1 + \mathsf{D}_2\mathbf{Z}. \tag{45b}$$

Note that $\|\alpha_i\|^2 + \|\beta_i\|^2 = \|\mathsf{Q}_i\mathbf{u}_i\|^2$, for $i\in\{1,2\}$.

We can show that the total power the two transmitters use after the initialization phase is

$$\pi \triangleq \|\mathsf{Q}_1\mathbf{u}_1\|^2 + \|\mathsf{Q}_2\mathbf{u}_2\|^2 + \Omega + \mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^{\mathsf{T}}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^{\mathsf{T}}\right), \tag{46}$$

where $\Omega \triangleq 2\rho^\star(\alpha_1^{\mathsf{T}}\beta_2 + \alpha_2^{\mathsf{T}}\beta_1)$ and $\rho^\star\in(0,1)$ denotes the correlation factor between Ξ_1 and Ξ_2 and is constant. By construction of the scheme, $\pi = nP$, see [7], [13].

Using Cauchy-Schwarz inequality, we can show that

$$\Omega \geq -\rho^\star(\|\alpha_1\|^2 + \|\beta_2\|^2 + \|\alpha_2\|^2 + \|\beta_1\|^2). \tag{47}$$

Hence, we have

$$\begin{aligned} nP \geq (1-\rho^\star)\left(\|\mathsf{Q}_1\mathbf{u}_1\|^2 + \|\mathsf{Q}_2\mathbf{u}_2\|^2\right) \\ +\mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^{\mathsf{T}}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^{\mathsf{T}}\right). \end{aligned} \tag{48}$$

Thus, by the positivity of matrices D_1 and D_2 and because $\|\mathsf{Q}_1\mathbf{u}_1\|^2 + \|\mathsf{Q}_2\mathbf{u}_2\|^2 > 0$ and $1-\rho^\star > 0$,

$$\mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^{\mathsf{T}}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^{\mathsf{T}}\right) < nP, \tag{49a}$$
$$\|\mathsf{Q}_1\mathbf{u}_1\|^2 + \|\mathsf{Q}_2\mathbf{u}_2\|^2 < \frac{nP}{1-\rho^\star}. \tag{49b}$$

From the analysis in [7], [13], we also know that the receiver's error variances about Ξ_1 and Ξ_2 satisfy

$$\mathsf{Var}\Big(\Xi_i - \hat{\Xi}_i^{(n-1)}\Big) = \gamma_i^{-2n}, \qquad i\in\{1,2\}. \tag{50}$$

By (40) and (42)

$$\hat{\Xi}_i^{(n-1)} = h_1\mathbf{v}_i^{\mathsf{T}}\mathsf{G}\mathbf{u}_1\Xi_1 + h_2\mathbf{v}_i^{\mathsf{T}}\mathsf{G}\mathbf{u}_2\Xi_2 + \mathbf{v}_i^{\mathsf{T}}\mathsf{G}\mathbf{Z}. \tag{51}$$

Using the independence of the vector $\mathbf{Z}$ with the message points Ξ_1 and Ξ_2, for $i\in\{1,2\}$,

$$\begin{aligned} &\mathsf{Var}\Big(\Xi_i - \hat{\Xi}_i^{(n-1)}\Big) \\ &= \mathsf{Var}(\Xi_i - h_1\mathbf{v}_i^{\mathsf{T}}\mathsf{G}\mathbf{u}_1\Xi_1 - h_2\mathbf{v}_i^{\mathsf{T}}\mathsf{G}\mathbf{u}_2\Xi_2) + \|\mathbf{v}_i^{\mathsf{T}}\mathsf{G}\|^2, \end{aligned} \tag{52}$$

which combines with (50) to

$$\mathsf{Var}(\Xi_i - h_1\mathbf{v}_i^{\mathsf{T}}\mathsf{G}\mathbf{u}_1\Xi_1 - h_2\mathbf{v}_i^{\mathsf{T}}\mathsf{G}\mathbf{u}_2\Xi_2) \leq \gamma_i^{-2n}, \tag{53a}$$
$$\|\mathbf{v}_i^{\mathsf{T}}\mathsf{G}\|^2 \leq \gamma_i^{-2n}. \tag{53b}$$

Inequalities (49) and (53) will be key to analyze our BC scheme.

B. Vector-Description of our BC scheme & Analysis

We now rewrite our BC scheme of Subsection V-B in vector notation. To this end, collect the inputs, the outputs, and the noise symbols after the initialization phase in the vectors

$$\mathbf{X} \triangleq (X_0,\ldots,X_{n-1})^{\mathsf{T}}, \tag{54}$$
$$\mathbf{Y}_i \triangleq (Y_{1,0},\ldots,Y_{i,n-1})^{\mathsf{T}}, \qquad i\in\{1,2\}, \tag{55}$$
$$\mathbf{Z}_i \triangleq (Z_{1,0},\ldots,Z_{i,n-1})^{\mathsf{T}}, \qquad i\in\{1,2\}. \tag{56}$$

Also, let

$$\mathsf{A}_i = \bar{\mathsf{C}}_i \overset{(a)}{=} \mathsf{C}_i, \text{ and} \tag{57}$$
$$\mathsf{B}_i = \bar{\mathsf{D}}_i \overset{(b)}{=} \mathsf{D}_i, \quad i\in\{1,2\} \tag{58}$$

where (a) and (b) hold because the matrices $(\mathsf{C}_1,\mathsf{C}_2)$ and $(\mathsf{D}_1,\mathsf{D}_2)$ are Toeplitz.

The inputs of our BC scheme after the initialization phase satisfy

$$\mathbf{X} = \mathbf{u}_1\Xi_1 + \mathbf{u}_2\Xi_2 + \mathsf{A}_1\mathbf{Y}_1 + \mathsf{A}_2\mathbf{Y}_2, \tag{59}$$
$$= \mathbf{u}_1\Xi_1 + \mathbf{u}_2\Xi_2 + (h_1\mathsf{A}_1 + h_2\mathsf{A}_2)\mathbf{X} + \mathsf{A}_1\mathbf{Z}_1 + \mathsf{A}_2\mathbf{Z}_2, \tag{60}$$

Subtracting $(h_1\mathsf{A}_1+h_2\mathsf{A}_2)\mathbf{X}$ from both sides of (60) and then multiplying both sides from the left by the matrix $(\mathsf{I}-h_1\mathsf{A}_1-h_2\mathsf{A}_2)^{-1}$, by equivalence (57) and by the definition of the matrix G in (35), we obtain

$$\mathbf{X} = \mathsf{G}\mathbf{u}_1\Xi_1 + \mathsf{G}\mathbf{u}_2\Xi_2 + \mathsf{B}_1\mathbf{Z}_1 + \mathsf{B}_2\mathbf{Z}_2. \tag{61}$$

The channel outputs are thus

$$\mathbf{Y}_1 = h_1\left(\mathsf{G}\mathbf{u}_1\Xi_1 + \mathsf{G}\mathbf{u}_2\Xi_2\right) + (\mathsf{I}+h_1\mathsf{B}_1)\mathbf{Z}_1 + h_1\mathsf{B}_2\mathbf{Z}_2, \tag{62a}$$
$$\mathbf{Y}_2 = h_2\left(\mathsf{G}\mathbf{u}_1\Xi_1 + \mathsf{G}\mathbf{u}_2\Xi_2\right) + (\mathsf{I}+h_2\mathsf{B}_2)\mathbf{Z}_2 + h_2\mathsf{B}_1\mathbf{Z}_1. \tag{62b}$$

Receiver i estimates Ξ_i as

$$\hat{\Xi}_i^{(n-1)} \triangleq \mathbf{v}_i^{\mathsf{T}}\mathbf{Y}_i, \quad i\in\{1,2\}. \tag{63}$$

The following will be key in our analysis: since

$$\mathbf{u}_i = \frac{P_i}{\delta_i}\gamma_i^{n-1}\mathsf{E}\mathbf{v}_i, \tag{64}$$

for any n-by-n matrix M, and for $i\in\{1,2\}$,

$$\begin{aligned} \|\mathbf{v}_i^{\mathsf{T}}\mathsf{M}\|^2 &= \|\mathsf{M}^{\mathsf{T}}\mathbf{v}_i\|^2 = \|\mathsf{M}^{\mathsf{T}}\mathsf{E}\mathbf{u}_i\|^2\frac{\delta_i^2}{P_i^2}\gamma_i^{2(-n+1)} \\ &= \|\bar{\mathsf{M}}\mathbf{u}_i\|^2\frac{\delta_i^2}{P_i^2}\gamma_i^{2(-n+1)}. \end{aligned} \tag{65}$$

In particular, when the matrix M is Toeplitz,

$$\|\mathbf{v}_i^{\mathsf{T}}\mathsf{M}\|^2 = \|\mathsf{M}\mathbf{u}_i\|^2\frac{\delta_i^2}{P_i^2}\gamma_i^{2(-n+1)}. \tag{66}$$

We first analyze the blockpower used by our BC scheme. By (61), in channel uses $0,\ldots,n-1$ the transmitter uses power

$$\|\mathsf{G}\mathbf{u}_1\|^2+\|\mathsf{G}\mathbf{u}_2\|^2+2\rho^\star \mathrm{tr}\left(\mathsf{G}\mathbf{u}_1\mathbf{u}_2^\mathsf{T}\mathsf{G}^\mathsf{T}\right)+\mathrm{tr}\left(\mathsf{B}_1\mathsf{B}_1^\mathsf{T}\right)+\mathrm{tr}\left(\mathsf{B}_2\mathsf{B}_2^\mathsf{T}\right),$$

which can be upper-bounded using Cauchy-Schwarz and because $0<\rho^\star<1$ by

$$2\|\mathsf{G}\mathbf{u}_1\|^2+2\|\mathsf{G}\mathbf{u}_2\|^2+\mathrm{tr}\left(\mathsf{B}_1\mathsf{B}_1^\mathsf{T}\right)+\mathrm{tr}\left(\mathsf{B}_2\mathsf{B}_2^\mathsf{T}\right). \tag{67}$$

By (66) and since G is Toeplitz, (67) equals

$$\begin{aligned}&2\|\mathbf{v}_1^\mathsf{T}\mathsf{G}\|^2\frac{P_1^2}{\delta_1^2}\gamma_1^{2(n-1)}+2\|\mathbf{v}_2^\mathsf{T}\mathsf{G}\|^2\frac{P_2^2}{\delta_2^2}\gamma_2^{2(n-1)}\\&+\mathrm{tr}\left(\mathsf{B}_1\mathsf{B}_1^\mathsf{T}\right)+\mathrm{tr}\left(\mathsf{B}_2\mathsf{B}_2^\mathsf{T}\right)\\&\qquad\le\frac{2P_1^2}{\delta_1^2\gamma_1^2}+\frac{2P_2^2}{\delta_2^2\gamma_2^2}+\mathrm{tr}\left(\mathsf{B}_1\mathsf{B}_1^\mathsf{T}\right)+\mathrm{tr}\left(\mathsf{B}_2\mathsf{B}_2^\mathsf{T}\right),\end{aligned} \tag{68}$$

where the inequality follows by Upper bound (53b).

By (58), we have

$$\mathrm{tr}\left(\mathsf{B}_i\mathsf{B}_i^\mathsf{T}\right)=\mathrm{tr}\left(\mathsf{D}_i\mathsf{D}_i^\mathsf{T}\right),\quad i\in\{1,2\}. \tag{69}$$

Hence, by (49a)

$$\mathrm{tr}\left(\mathsf{B}_1\mathsf{B}_1^\mathsf{T}\right)+\mathrm{tr}\left(\mathsf{B}_2\mathsf{B}_2^\mathsf{T}\right)<nP. \tag{70}$$

Since the term $\frac{P_1^2}{\delta_1^2\gamma_1^2}+\frac{P_2^2}{\delta_2^2\gamma_2^2}$ is bounded and does not grow with n, and since the power used during the initialization phase is bounded as well, our BC scheme satisfies the average block-power constraint asymptotically as $n\to\infty$.

We now analyze the variance of Receiver i's estimation error $\hat{\Xi}_i^{(n-1)}-\Xi_i$. By (62) and (63), we have

$$\begin{aligned}\hat{\Xi}_1^{(n-1)}&=h_1\mathbf{v}_1^\mathsf{T}\mathsf{G}\mathbf{u}_1\Xi_1+h_1\mathbf{v}_1^\mathsf{T}\mathsf{G}\mathbf{u}_2\Xi_2\\&\quad+\mathbf{v}_1^\mathsf{T}\big((\mathsf{I}+h_1\mathsf{B}_1)\mathbf{Z}_1+h_1\mathsf{B}_2\mathbf{Z}_2\big),\end{aligned} \tag{71}$$

$$\begin{aligned}\hat{\Xi}_2^{(n-1)}&=h_2\mathbf{v}_2^\mathsf{T}\mathsf{G}\mathbf{u}_1\Xi_1+h_2\mathbf{v}_2^\mathsf{T}\mathsf{G}\mathbf{u}_2\Xi_2\\&\quad+\mathbf{v}_2^\mathsf{T}\big((\mathsf{I}+h_2\mathsf{B}_2)\mathbf{Z}_2+h_2\mathsf{B}_1\mathbf{Z}_1\big).\end{aligned} \tag{72}$$

Define

$$\mathsf{S}_i=\mathsf{E}\mathsf{Q}_i\mathsf{E},\quad i\in\{1,2\}. \tag{73}$$

It can be shown that the product $\mathsf{S}_1\mathsf{S}_1$ equals the covariance matrix of the noise term $\big((\mathsf{I}+h_1\mathsf{B}_1)\mathbf{Z}_1+h_1\mathsf{B}_2\mathbf{Z}_2\big)$, and the product $\mathsf{S}_2\mathsf{S}_2$ equals the covariance matrix of the noise term $\big((\mathsf{I}+h_2\mathsf{B}_2)\mathbf{Z}_2+h_2\mathsf{B}_1\mathbf{Z}_1\big)$. Thus, by the independence of the symbols Ξ_1 and Ξ_2 with the noise vectors $\mathbf{Z}_1$ and $\mathbf{Z}_2$,

$$\begin{aligned}&\mathsf{Var}\Big(\Xi_i-\hat{\Xi}_i^{(n-1)}\Big)\\&=\mathsf{Var}(\Xi_i-h_1\mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{u}_1\Xi_1-h_2\mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{u}_2\Xi_2)+\|\mathbf{v}_i^\mathsf{T}\mathsf{S}_i\|^2\\&\le\gamma_i^{-2n}+\|\bar{\mathsf{S}}_i\mathsf{u}_i\|^2\frac{\delta_i^2}{P_i^2}\gamma_i^{-2(n-1)}\\&\le\gamma_i^{-2n}+\|\mathsf{Q}_i\mathsf{u}_i\|^2\frac{\delta_i^2}{P_i^2}\gamma_i^{-2(n-1)}\\&\le\gamma_i^{-2n}+nP\frac{\delta_i^2}{P_i^2}\gamma_i^{-2(n-1)}\\&=\left(1+nP\frac{\delta_i^2\gamma_i^2}{P_i^2}\right)\gamma_i^{-2n},\end{aligned} \tag{74}$$

where in the first inequality we used (53a) and (65); in the second inequality we used that since Q_i (and S_i) is symmetric, $\mathsf{Q}_i=\mathsf{Q}_i^\mathsf{T}$, and thus by (73): $\bar{\mathsf{S}}_i=\mathsf{Q}_i$; and in the last inequality we used (49b).

By (29) we conclude that Receiver i's estimation error on message point $\Theta_i(M_i)$ is of variance

$$\mathsf{Var}\Big(\Theta_i(M_i)-\hat{\Theta}_i\Big)\le\frac{1}{h_i^2(1-\rho^\star)}\left(1+nP\frac{\delta_i^2\gamma_i^2}{P_i^2}\right)\gamma_i^{-2n}.$$

The nearest-neighbor decoding rule (26) produces the correct estimate whenever $\left|\Theta_i(M_i)-\hat{\Theta}_i\right|<\Delta_i/2$, where Δ_i is defined in (17). Since by (29), the estimation error $\Theta_i(M_i)-\hat{\Theta}_i$ is zero-mean Gaussian, this happens with probability $1-2\mathcal{Q}\left(\frac{\Delta_i/2}{\sqrt{\mathsf{Var}\left(\Theta_i(M_i)-\hat{\Theta}_i\right)}}\right)$ and thus Receiver i's probability of error is bounded by

$$\Pr\Big[\hat{M}_i\neq M_i\Big]\le 2\mathcal{Q}\left(\frac{\sqrt{P_i}}{2\lfloor 2^{nR_i}\rfloor}\left(1+nP\frac{\delta_i^2\gamma_i^2}{P_i^2}\right)^{-1}\gamma_i^{2n}\right).$$

This probability tends to 0 as $n\to\infty$ whenever

$$R_i<\frac{1}{2}\log(\gamma_i^2),\quad i\in\{1,2\}, \tag{75}$$

which is equivalent to (30) and concludes the proof.

Appendix A
Sketch of Proof of Theorem 1

We first introduce a class of multi-letter block-feedback schemes for MAC and BC as in [4] that achieve the linear-feedback capacity regions. Within this class, we identify pairs of schemes that achieve the same rate-regions over dual channels.

Each BC scheme in our capacity-achieving class is parametrized by a positive integer η and two η-by-η matrices A_1 and A_2, and each MAC scheme by a positive integer η' and two η'-by-η' matrices C_1 and C_2.

The idea of the schemes is to divide the blocklength n into subblocks of length η and to apply an inner linear-feedback code parametrized by $(\mathsf{A}_1,\mathsf{A}_2)$ (or by $(\mathsf{C}_1,\mathsf{C}_2)$) to transform each subblock of η channel uses of the scalar channel into a single use of a new super MIMO channel. An optimal MIMO nofeedback code is then used to communicate over the resulting super BC (or MAC) without using the feedback.

Now, choose for the BC scheme and the MAC scheme the same parameter η and feedback matrices that satisfy

$$\mathsf{A}_i=\bar{\mathsf{C}}_i,\quad i\in\{1,2\}. \tag{76}$$

We shall show, using nofeedback MIMO MAC-BC duality [16], [17], that the resulting super MIMO MAC and BC have the same capacity regions. By taking the union over all possible parameters, this will conclude the proof of the theorem.

Consider first the BC and define the matrices

$$\mathsf{B}_i\triangleq(\mathsf{I}-h_1\mathsf{A}_1-h_2\mathsf{A}_2)^{-1}\mathsf{A}_i,\quad i\in\{1,2\}. \tag{77}$$

The channel inputs in a subblock can be written as

$$\mathbf{X} = (\mathsf{I} - h_1\mathsf{A}_1 - h_2\mathsf{A}_2)\mathbf{U} + \mathsf{A}_1\mathbf{Y}_1 + \mathsf{A}_2\mathbf{Y}_2 \quad (78)$$
$$= \mathbf{U} + \mathsf{B}_1\mathbf{Z}_1 + \mathsf{B}_2\mathbf{Z}_2 \quad (79)$$

for some code vector $\mathbf{U}$ that is average power constrained to

$$\eta P - \mathrm{tr}\left(\mathsf{B}_1\mathsf{B}_1^{\mathsf{T}}\right) - \mathrm{tr}\left(\mathsf{B}_2\mathsf{B}_2^{\mathsf{T}}\right) - 2\lambda \mathrm{tr}\left(\mathsf{B}_1\mathsf{B}_2^{\mathsf{T}}\right). \quad (80)$$

Note that the precoding factor $(\mathsf{I}-h_1\mathsf{A}_1-h_2\mathsf{A}_2)$ in (78) is used only to simplify calculations. This definition remains general.

The corresponding channel outputs are

$$\mathbf{Y}_1 = h_1\mathbf{U} + (\mathsf{I} + h_1\mathsf{B}_1)\mathbf{Z}_1 + h_1\mathsf{B}_2\mathbf{Z}_2, \quad (81a)$$
$$\mathbf{Y}_2 = h_2\mathbf{U} + (\mathsf{I} + h_2\mathsf{B}_2)\mathbf{Z}_2 + h_2\mathsf{B}_1\mathbf{Z}_1. \quad (81b)$$

Let S_1 and S_2 denote the unique positive square-roots of the the positive definite noise covariance matrices

$$\mathsf{N}_1 \triangleq (\mathsf{I} + h_1\mathsf{B}_1)(\mathsf{I} + h_1\mathsf{B}_1)^{\mathsf{T}} + h_1\mathsf{B}_2(h_1\mathsf{B}_2)^{\mathsf{T}} + \lambda(\mathsf{I} + h_1\mathsf{B}_1)(h_1\mathsf{B}_2)^{\mathsf{T}} + \lambda(h_1\mathsf{B}_2)(\mathsf{I} + h_1\mathsf{B}_1)^{\mathsf{T}}, \quad (82a)$$
$$\mathsf{N}_2 \triangleq (\mathsf{I} + h_2\mathsf{B}_2)(\mathsf{I} + h_2\mathsf{B}_2)^{\mathsf{T}} + h_2\mathsf{B}_1(h_2\mathsf{B}_1)^{\mathsf{T}} + \lambda(\mathsf{I} + h_2\mathsf{B}_2)(h_2\mathsf{B}_1)^{\mathsf{T}} + \lambda(h_2\mathsf{B}_1)(\mathsf{I} + h_2\mathsf{B}_2)^{\mathsf{T}}. \quad (82b)$$

Define $\tilde{\mathbf{U}} = \mathsf{E}\mathbf{U}$ and notice that $\|\tilde{\mathbf{U}}\|^2 = \|\mathbf{U}\|^2$. Since ES_1 and ES_2 are invertible, the following MIMO BC has the same capacity region as the MIMO BC in (81):

$$\mathbf{Y}_i' = h_i\mathsf{E}\mathsf{S}_i^{-1}\mathsf{E}\tilde{\mathbf{U}} + \tilde{\mathbf{Z}}_i, \quad i \in \{1, 2\}, \quad (83)$$

where $\tilde{\mathbf{Z}}_1$ and $\tilde{\mathbf{Z}}_2$ denote independent centered Gaussian vectors of identity covariance matrix and $\tilde{\mathbf{U}}$ and $\mathbf{U}$ are constrained to (80).

Now we turn to the MAC and choose feedback matrices that satisfy (76). Define the matrices

$$\mathsf{D}_i \triangleq \mathsf{C}_i(\mathsf{I}_\eta - h_1\mathsf{C}_1 - h_2\mathsf{C}_2)^{-1}, \quad (84)$$

and let Q_1 and Q_2 denote the unique positive square roots of the positive definite matrices

$$\mathsf{M}_1 \triangleq (\mathsf{I} + h_1\mathsf{D}_1)^{\mathsf{T}}(\mathsf{I} + h_1\mathsf{D}_1) + (h_1\mathsf{D}_2)^{\mathsf{T}}(h_1\mathsf{D}_2) + \lambda(\mathsf{I} + h_1\mathsf{D}_1)^{\mathsf{T}}h_1\mathsf{D}_2 + \lambda h_1\mathsf{D}_2^{\mathsf{T}}(\mathsf{I} + h_1\mathsf{D}_1), \quad (85a)$$
$$\mathsf{M}_2 \triangleq (h_2\mathsf{D}_1)^{\mathsf{T}}(h_2\mathsf{D}_1) + (\mathsf{I} + h_2\mathsf{D}_2)^{\mathsf{T}}(\mathsf{I} + h_2\mathsf{D}_2) + \lambda(\mathsf{I} + h_2\mathsf{D}_2)^{\mathsf{T}}h_2\mathsf{D}_1 + \lambda(h_2\mathsf{D}_1)^{\mathsf{T}}(\mathsf{I} + h_2\mathsf{D}_2). \quad (85b)$$

The channel inputs in a subblock are given by

$$\mathbf{X}_i = \mathsf{Q}_i^{-1}\mathbf{U}_i + \mathsf{C}_i\mathbf{Y}, \quad i \in \{1, 2\}, \quad (86)$$

for some code-vectors $\mathbf{U}_1$ and $\mathbf{U}_2$ that are average sum-power constrained to

$$\eta P - \mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^{\mathsf{T}}\right) - \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^{\mathsf{T}}\right) - 2\lambda \mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_2^{\mathsf{T}}\right). \quad (87)$$

The corresponding channel output vector can be written as

$$\mathbf{Y} = (\mathsf{I} + h_1\mathsf{D}_1 + h_2\mathsf{D}_2)(h_1\mathsf{Q}_1^{-1}\mathbf{U}_1 + h_2\mathsf{Q}_2^{-1}\mathbf{U}_2 + \mathbf{Z}). \quad (88)$$

Since $(\mathsf{I} + h_1\mathsf{D}_1 + h_2\mathsf{D}_2)$ is invertible, under the same power constraint, the MIMO MAC described by (88) has the same capacity region as

$$\tilde{\mathbf{Y}} = h_1\mathsf{Q}_1^{-1}\mathbf{U}_1 + h_2\mathsf{Q}_2^{-1}\mathbf{U}_2 + \mathbf{Z}. \quad (89)$$

In view of (76), which is equivalent to $\mathsf{B}_i = \bar{\mathsf{D}}_i$, after some algebraic manipulations we can show that $\mathsf{N}_i = \mathsf{E}\mathsf{M}_i\mathsf{E}$. Consequently, we have $\mathsf{S}_i = \mathsf{E}\mathsf{Q}_i\mathsf{E}$. We can also show that the power constrains (80) and (87) coincide. Hence, the MIMO BC and MAC channels in (83) and (89) are dual (channel matrices are transposes of each other) and are subject to the same total power constraint. Using nofeedback MIMO MAC-BC duality in [16], [17] we show that they have same nofeedback capacity regions.

REFERENCES

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

[2] S. Belhadj Amor, Y. Steinberg, and M. Wigger, "Duality with linear-feedback schemes for the scalar Gaussian MAC and BC," in *Proc. of 2014 International Zurich Seminar on Communications*, pp. 25–28.

[3] ——, "MAC-BC duality with linear-feedback schemes," in *Proc. of IEEE ISIT*, 2014, pp. 1737–1741.

[4] ——, "MIMO MAC-BC duality with linear-feedback coding schemes," submitted to IEEE Trans. on Inf. Theory, March 2014. [Online]. Available: http://arxiv.org/abs/1404.2584

[5] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, 1972.

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

[7] A. El Gamal and Y. H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[8] N. Elia, "When Bode meets Shannon: control-oriented feedback communication schemes," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1477–1488, 2004.

[9] M. Gastpar, "Gaussian multiple-access channels under received-power constraints," in *Proc. of IEEE ITW*, 2004, pp. 452–457.

[10] M. Gastpar, A. Lapidoth, Y. Steinberg, and M. Wigger, "New achievable rates for the Gaussian broadcast channel with feedback," in *Proc. of ISWCS*, 2011, pp. 579–583.

[11] ——, "Coding schemes and asymptotic capacity for the Gaussian broadcast and interference channels with feedback," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 54–71, 2014.

[12] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1423–1438, 2002.

[13] L. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 623–629, 1984.

[14] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 667–671, 1984.

[15] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–I: No bandwidth constraint," *IEEE Trans. Inf. Theory*, vol. 12, no. 2, pp. 172–182, 1966.

[16] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.

[17] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, 2003.

[18] W. Wu, S. Vishwanath, and A. Arapostathis, "Gaussian interference networks with feedback: duality, sum capacity, and dynamic team problems," in *Proc. 43th Ann. Allerton Conf.*, 2005, pp. 926–935.

[19] Y. Wu, P. Minero, and M. Wigger, "Insufficiency of linear-feedback schemes in gaussian broadcast channels with common message," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4553–4566, Aug 2014.