# MIMO MAC-BC Duality with Linear-Feedback Coding Schemes

Selma Belhadj Amor, Yossef Steinberg, and Michèle Wigger

***Abstract*—We show that the rate regions achieved by linear-feedback coding schemes over *dual* multi-antenna Gaussian multi-access channels (MAC) and broadcast channels (BC) with independent noises coincide. By dual here we mean:**

- **the channel matrices of the multi-access channel (MAC) and the broadcast channel (BC) are transposes of each other**
- **and the same total input-power constraint $P$ is imposed on both channels.**

**We also present multi-letter expressions for the linear-feedback capacity regions of the two channels, i.e., for the set of all rates that are achievable with linear-feedback coding schemes. We identify a sub-class of MAC and BC linear-feedback coding schemes that achieve the respective linear-feedback capacity regions, and within these subclasses we identify pairs of MAC and BC coding schemes that achieve the same rate regions.**

**In the two-user case, when the transmitters or the receiver are single-antenna, the capacity region for the Gaussian MAC is known [20], [15] and the capacity-achieving scheme is a linear-feedback coding scheme. With our results we can thus determine the linear-feedback capacity region of the two-user Gaussian BC when either transmitter or receivers are single-antenna and we can identify the corresponding linear-feedback capacity-achieving coding schemes. Our results show that the control-theory inspired linear-feedback coding scheme by Elia [11], by Wu *et al.* [30], and by Ardestanizadeh *et al.* [1] is sum-rate optimal among all linear-feedback coding schemes for the symmetric single-antenna Gaussian BC with equal channel gains. More generally, we show that the linear-feedback sum-capacity of the scalar Gaussian BC with independent noises is achieved using a simple rearrangement of Ozarow's MAC encodings and decodings.**

**In the $K \geq 3$-user case, Kramer [16] and Ardestanizadeh *et al.* [2] determined the linear-feedback sum-capacity for the symmetric single-antenna Gaussian MAC with equal channel gains. Using our duality result, in this paper we identify the linear-feedback sum-capacity for the $K \geq 3$-user single-antenna Gaussian BC with equal channel gains. It is equal to the sum-rate achieved by Ardestanizadeh *et al.*'s linear-feedback coding scheme [1].**

**Our results extend also to the setup where only a subset of the feedback links are present.**

This work was in part presented at the *2014 International Zurich Seminar on Communications*, in Zurich, Switzerland, and at the *2014 IEEE International Symposium on Information Theory*, Honolulu, Hawai'i, USA.

S. Belhadj Amor is with the Institut National de Recherche en Informatique et en Automatique (INRIA), Lyon, France. email: selma.belhadj-amor@inria.fr.

Y. Steinberg is with the Department of Electrical Engineering at the Technion—Israel Institute of Technology, Haifa, Israel. email: ysteinbe@ee.technion.ac.il.

M. Wigger is with the Department of Communications and Electronics, Telecom Paristech, Paris, France. email: wigger@telecom-paristech.fr.

The work of S. Belhadj Amor and M. Wigger has been supported by the city of Paris under the "Emergences" program. The work of Y. Steinberg has been supported by the Israel Science Foundation (grant no. 684/11).



***Index Terms*—Broadcast channel (BC), multiple-access channel (MAC), Gaussian noise, channel capacity, duality, linear-feedback coding schemes, perfect feedback, multiple-input multiple-output (MIMO) channels.**

## I. Introduction

Unlike for point-to-point channels, in multi-user networks feedback can enlarge capacity. For most multi-user networks the capacity region with feedback is however still unknown. Notable exceptions are the two-user memoryless single-input single-output (SISO) Gaussian multi-access channel (MAC) whose capacity region with feedback was determined by Ozarow [20], and the two-user single-input multi-output (SIMO) and multi-input single-output (MISO) memoryless Gaussian MAC, whose capacity regions were determined by Jafar and Goldsmith [15]. For more than two users or in the general multi-input multi-output (MIMO) case, the capacity region of the memoryless Gaussian MAC with feedback is still open. For $K > 2$ transmitters, Kramer [16] determined the sum-capacity of the SISO Gaussian MAC under equal power constraints $P$ at all the transmitters when $P$ is sufficiently large.

Ozarow's coding scheme [20], which achieves the capacity region of the two-user SISO Gaussian MAC with feedback, is a variation of the Schalkwijk-Kailath scheme for point-to-point channels. Each transmitter maps its message to a message point and sends this message point during one of the first two channel uses. In channel uses 3 and thereafter both transmitters send scaled versions of the linear minimum mean squared estimation (LMMSE) errors of their message points when observing all previous outputs. Ozarow showed that this scheme achieves the sum-capacity of the two-user SISO Gaussian MAC with perfect feedback. To achieve the entire capacity region, one of the two transmitters has to combine this scheme with a non-feedback scheme using rate-splitting. The described scheme falls into the class of *linear-feedback coding schemes* [2], where the transmitters can use the feedback signals only in a linear way. That means, a transmitter's channel input for a given channel use is a *linear* combination of the previously observed feedback signals and some information-carrying code symbols which only depend on the transmitter's message but not on the feedback.

As detailed in the following, linear schemes achieve capacity for many networks. For other networks they achieve the largest rate regions known to date, but it is unknown whether they achieve capacity. An important step to answer this question is to characterize the set of rates that are achievable with linear-feedback coding schemes. In fact, due to the simple appealing form of linear-feedback schemes, such a characterization is also of importance in its own right.

Jafar and Goldsmith's [15] capacity-achieving schemes for the two-user SIMO and MISO Gaussian MACs and Kramer's scheme for the $K$-user SISO Gaussian MAC are variations of Ozarow's scheme and also belong to the class of linear-feedback coding schemes. It has recently been shown [2] that under equal input-power constraints $P$ at all $K$ transmitters, irrespective of the values of $P$ and $K$, Kramer's scheme achieves the largest sum-rate among all linear-feedback coding schemes.

The capacity region of the memoryless Gaussian BC with perfect feedback is unknown even with only two receivers and in the SISO case. Achievable regions have been proposed by Ozarow & Leung [21], Elia [11], Kramer [16], Wu *et al.* [30], Ardestanizadeh *et al.* [1], Gastpar *et al.* [12], Wu & Wigger [32], Murin *et al.* [19], Shayevitz & Wigger [24], and Venkataramanan & Pradhan [25]. The schemes in [1], [11], [12], [16], [21], [30] are linear-feedback coding schemes and outperform the other schemes [24], [25], [32] when these latter are specialized to the SISO Gaussian BC.

Ozarow & Leung [21] presented the first linear-feedback coding scheme for the broadcast channel. It is inspired by Schalkwijk & Kailath's [23] coding scheme for the point-to-point channel and Ozarow's [20] coding scheme for the MAC: the transmitter sends a linear combination of the two receivers' LMMSE estimation errors about their desired message points. Kramer suggested to use LMMSE estimators with memory, as opposed to the memoryless estimators used in [21]. In some cases this modification leads indeed to improved achievable rates, see Murin *et al.* [19], but in other cases the set of achievable rates is decreased.

In a symmetric setup with equal noise variances and when the noises are uncorrelated, both LMMSE-based schemes (with and without memory) are outperformed by the linear-feedback coding schemes in [1], [11], and [30], which are designed based on control-theoretic considerations. These schemes achieve the same sum-rate over the symmetric SISO Gaussian BC under power constraint $P$ as Ozarow's scheme [20] achieves over the Gaussian MAC under a *sum-power* constraint $P$. There is thus a duality in terms of achievable sum-rate between the control-theoretic schemes for the BC [11], [30], and [1] and Ozarow's capacity-achieving scheme for the MAC. It is unknown whether the schemes in [1], [11], and [30], and achieve the sum-capacity with perfect feedback for symmetric BCs, and previous to this work, it was also unknown whether for the symmetric SISO Gaussian BC it is sum-rate optimal among all linear-feedback coding schemes. As detailed shortly, our results in this paper show that this is indeed the case.

Gastpar *et al.* [13] showed that in the asymptotic regime where the allowed input power $P \to \infty$ linear-feedback coding schemes achieve sum-capacity, irrespective of the correlation between the noise sequences at the two receivers [13]. For some setups where the noises at the two receivers are correlated, their scheme in [12] also provides the largest explicit achievable rates known to date.

As stated above, linear-feedback coding schemes are optimal for the Gaussian point-to-point and two-user Gaussian MAC multi-access channels and they achieve the largest known rates for other Gaussian networks. The only exception where linear-feedback coding schemes have shown to be strictly suboptimal is the Gaussian BC with only common message [31], which is not considered here.

Linear-feedback coding schemes are also interesting because of their connection to the scenario with active noisy feedback where the two receivers can communicate with the transmitter over a memoryless Gaussian MAC that does not interfere with the forward BC. Ben-Yishai and Shayevitz [3] recently proposed a clever reduction proving that any rate-pair that is achievable over the two-user Gaussian BC with perfect feedback using a linear-feedback coding scheme is achievable also over the two-user Gaussian BC with active noisy feedback whenever the forward-noise variance and the feedback-noise variances lie below certain thresholds.

Without feedback, the following duality relation is well known [26], [27], [28]: under the same sum input-power constraint the capacity regions of the MIMO Gaussian MAC and BC coincide when the channel matrices of the MAC and BC are transposes of each other. Such a pair of MAC and BC is called *dual*.

Our main contribution in this work is the following new duality result: with perfect feedback and when restricting to linear-feedback coding schemes, the set of all achievable rates, coincide for the MIMO Gaussian MAC and BC when the two channels are dual and when the same sum input-power constraint is imposed on their inputs. This result is particularly interesting in the two-user case and when either transmitter(s) or receiver(s) are single-antenna (SISO, MISO, and SIMO setups) because for these setups computable single-letter characterizations of the linear-feedback capacity regions of the Gaussian MAC are known. With our duality result, we thus immediately obtain single-letter characterizations of the linear-feedback capacity regions for the two-user SISO, SIMO, and MISO Gaussian BC. For more than $K \geq 3$ users the linear-feedback sum-capacity of the SISO Gaussian MAC is known when the channel gains are equal [2], [16]; with our results we thus obtain the linear-feedback sum-capacity of the SISO Gaussian BC when the channel gains are equal. Our results in particular show that the control-theory inspired linear-feedback coding schemes proposed by Elia [11], by Wu *et al.* [30], and by Ardestanizadeh *et al.* [1] are sum-rate optimal among all linear-feedback coding schemes for the symmetric SISO Gaussian BC with equal channel gains, irrespective of the number of receivers $K \geq 2$.

Our duality result extends also to a setup where only some of the feedback links are present.

Duality however does not extend in a straightforward manner to setups with passive noisy feedback where the receiver(s) observe the channel outputs corrupted by additive white Gaussian noise. Whereas for the two-user SISO Gaussian MAC the linear-feedback capacity region with noisy feedback is strictly larger than the no-feedback capacity region for all feedback-noise variances [18], this is not the case for the two-user SISO Gaussian BC: Pillai and Prabhakaran [22] showed recently that in asymmetric setups and when the feedback-noise variances exceed a certain threshold capacity is the same with or without noisy feedback. (Notice that when the receivers can actively code over the feedback links, then in asymmetric setups the capacity region with noisy feedback is always larger than without feedback [32].)

In this paper, we also introduce a class of linear-feedback coding schemes for the MIMO Gaussian MAC and BC that achieve the linear-feedback capacity regions. Within this class we can identify the pairs of coding schemes that achieve the same rate-regions over dual MACs and BCs. Since we know the optimal linear-feedback coding schemes for the two-user SISO, SIMO, and MISO Gaussian MAC [15], [20], we can identify the optimal linear-feedback coding schemes for the two-user SISO, MISO, and SIMO Gaussian BC. (For more than two users, the linear-feedback capacity of the MAC is unknown and our duality result does not provide a characterization of the linear-feedback capacity.)

For the two-user SISO case we can show an even tighter connection between the Gaussian MAC and BC under linear-feedback coding schemes than what is suggested by our duality results on achievable rates. In fact, the following simple rearrangement of the encoders and decoders in Ozarow's sum-capacity achieving MAC scheme yields a constructive sum-rate optimal BC scheme: 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 as Ozarow's MAC decoder guesses this message. (A key observation here is that Ozarow's MAC decoder chooses to guess the two messages separately of each other.)

The remainder of this paper is organized as follows. In Section II we explain the notations used in this paper and introduce some preliminaries. In Section III, we consider the two-user MIMO Gaussian BC with perfect feedback and in Section IV the two-user MIMO Gaussian MAC with perfect feedback: specifically, we describe the channel model, introduce the class of linear-feedback coding schemes, and summarize previous results. Section VI presents our main results on MAC-BC duality with linear-feedback schemes, the linear-feedback capacity-achieving schemes for MAC and BC. Section VII describes a constructive sum-rate optimal linear-feedback coding scheme for the two-user SISO Gaussian BC. In Sections VIII and IX, we explain how our results extend to setups with partial feedback and to arbitrary $K \geq 2$ users. Finally, Section X contains the major proofs.

## II. Notation and Preliminaries

In the following, a random variable is denoted by an upper-case letter (e.g $X$, $Y$, $Z$) and its realization by a lower-case letter (e.g $x$, $y$, $z$). An $n$-dimensional random column-vector and its realization are denoted by boldface symbols (e.g. $\mathbf{X}$, $\mathbf{x}$). We use $\|\cdot\|$ to indicate the Euclidean norm and $\mathsf{E}[\cdot]$ for the expectation operator. The abbreviation i.i.d. stands for *independently and identically distributed.*

Sets are denoted by calligraphic letters (e.g., $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$) and $\mathcal{X} \times \mathcal{Y}$ denotes the Cartesian product of the sets $\mathcal{X}$ and $\mathcal{Y}$. The set of real numbers is denoted by $\mathbb{R}$ and its $d$-fold Cartesian product by $\mathbb{R}^d$. We use $\mathrm{cl}(X)$ to denote the convex closure of the set $\mathcal{X}$.

Throughout the paper, $\log(\cdot)$ refers to the binary logarithm-function.

To denote matrices we use the font $\mathsf{A}$. For the transpose of a matrix $\mathsf{A}$ we write $\mathsf{A}^\intercal$, for its determinant $|\mathsf{A}|$, and for its trace $\mathrm{tr}(\mathsf{A})$. For the Kronecker product of two matrices $\mathsf{A}$ and $\mathsf{B}$ we write $\mathsf{A} \otimes \mathsf{B}$. We use $\mathsf{I}_d$ to denote the $d$-by-$d$ identity matrix, where we drop the subscript whenever the dimensions are clear from the context. The symbol $\mathsf{E}_d$ denotes the $d$-by-$d$ *exchange matrix* which is 0 everywhere except on the counter-diagonal where it is 1. For example,

$$\mathsf{E}_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \tag{1}$$

Again, we drop the subscript whenever the dimensions are clear.

We will be interested in the mirror image of a given matrix along its counter-diagonal. We call this matrix operator the *reverse-image matrix operator* $\bar{\cdot}$. Formally:

**Definition 1.** *For a given $d_1$-by-$d_2$ matrix $\mathsf{A}$, the reverse-image of $\mathsf{A}$ is defined as*

$$\bar{\mathsf{A}} \triangleq \mathsf{E}_{d_2} \mathsf{A}^\intercal \mathsf{E}_{d_1}. \tag{2}$$

Simple algebraic manipulations give:

**Note 1.** *The reverse image matrix operator satisfies the following properties:*

1) *Applying the operator twice results in the identity operation: $\mathsf{A} = \bar{\bar{\mathsf{A}}}$.*
2) *For every matrix $\mathsf{A}$, $\mathrm{tr}(\mathsf{A}) = \mathrm{tr}(\bar{\mathsf{A}})$.*
3) *If $\mathsf{A}$ is a Toeplitz-matrix, then $\bar{\mathsf{A}} = \mathsf{A}$.*
4) *The operator commutes with the matrix inverse-operator, the product operator, and the tensor operator:*

$$(\bar{\mathsf{A}})^{-1} = \overline{(\mathsf{A}^{-1})} \tag{3}$$
$$\bar{\mathsf{A}}\bar{\mathsf{B}} = \overline{(\mathsf{B}\mathsf{A})} \tag{4}$$
$$\bar{\mathsf{A}} \otimes \bar{\mathsf{B}} = \overline{(\mathsf{A} \otimes \mathsf{B})}. \tag{5}$$

5) *The operator maps a strictly-lower block-triangular $\eta\kappa_1$-by-$\eta\kappa_2$ matrix of block sizes $\kappa_1 \times \kappa_2$ into a strictly-lower block-triangular $\eta\kappa_2$-by-$\eta\kappa_1$ matrix of block sizes $\kappa_2 \times \kappa_1$.*

## III. MIMO Gaussian BC with Feedback

### A. Setup

We consider the two-user memoryless MIMO Gaussian BC with perfect-output feedback depicted in Figure 1. The transmitter is equipped with $\kappa$ transmit-antennas and each Receiver $i$, for $i \in \{1,2\}$, is equipped with $\nu_i$ receive-antennas. At each time $t \in \mathbb{N}$, if $\mathbf{x}_t$ denotes the real vector-valued input symbol sent by the transmitter, Receiver $i \in \{1,2\}$ observes the real vector-valued channel output

$$\mathbf{Y}_{i,t} = \mathsf{H}_i \mathbf{x}_t + \mathbf{Z}_{i,t}, \tag{6}$$

where $\mathsf{H}_i$, for $i \in \{1,2\}$, is a deterministic real $\nu_i$-by-$\kappa$ channel matrix known to transmitter and receivers and $\{\mathbf{Z}_{1,t}\}_{t=1}^n$

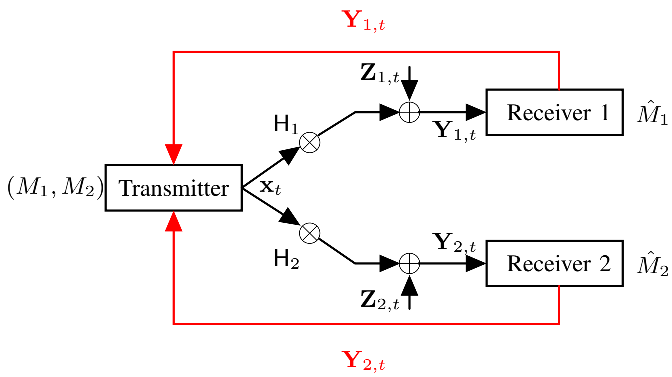


Fig. 1. Two-user memoryless MIMO Gaussian BC with perfect feedback.

and $\{\mathbf{Z}_{2,t}\}_{t=1}^n$ are independent sequences of i.i.d. centered Gaussian random vectors of identity covariance matrix.

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 $\mathcal{M}_1 \triangleq \{1,\dots,\lfloor 2^{nR_1}\rfloor\}$ and $\mathcal{M}_2 \triangleq \{1,\dots,\lfloor 2^{nR_2}\rfloor\}$, where $R_1$ and $R_2$ denote the rates of transmission and $n$ the blocklength.

The transmitter observes causal, noise-free output feedback from both receivers. Thus, the time-$t$ channel input $\mathbf{X}_t$ can depend on both messages $M_1$ and $M_2$ and on all previous channel outputs $\mathbf{Y}_{1,1},\dots,\mathbf{Y}_{1,t-1}$ and $\mathbf{Y}_{2,1},\dots,\mathbf{Y}_{2,t-1}$:

$$\mathbf{X}_t = g_t^{(n)}(M_1, M_2, \mathbf{Y}_{1,1},\dots,\mathbf{Y}_{1,t-1}, \mathbf{Y}_{2,1},\dots,\mathbf{Y}_{2,t-1}), \quad t\in\{1,\dots,n\}, \tag{7}$$

for some encoding function of the form:

$$g_t^{(n)}: \mathcal{M}_1 \times \mathcal{M}_2 \times \mathbb{R}^{\nu_1(t-1)} \times \mathbb{R}^{\nu_2(t-1)} \to \mathbb{R}^{\kappa}. \tag{8}$$

We impose an *expected average block-power constraint*

$$\frac{1}{n}\sum_{t=1}^{n} \mathbf{E}[\|\mathbf{X}_t\|^2] \le P, \tag{9}$$

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

Each Receiver $i$ decodes its corresponding message $M_i$ by means of a decoding function $\phi_i^{(n)}$ of the form

$$\phi_i^{(n)}: \mathbb{R}^{\nu_i n} \to \mathcal{M}_i, \quad i\in\{1,2\}. \tag{10}$$

That means, based on the output sequence $\mathbf{Y}_{i,1},\dots,\mathbf{Y}_{i,n}$, Receiver $i$ produces the guess

$$\hat{M}_i^{(n)} = \phi_i^{(n)}(\mathbf{Y}_{i,1},\dots,\mathbf{Y}_{i,n}). \tag{11}$$

An error occurs in the communication if

$$(\hat{M}_1 \neq M_1) \text{ or } (\hat{M}_2 \neq M_2). \tag{12}$$

Thus, the average probability of error is

$$P_{\text{e,BC}}^{(n)} \triangleq \Pr\big[(\hat{M}_1 \neq M_1) \text{ or } (\hat{M}_2 \neq M_2)\big]. \tag{13}$$

A $(\lfloor 2^{nR_1}\rfloor, \lfloor 2^{nR_2}\rfloor, n)$ *MIMO BC feedback-code of power* $P$ is composed of a sequence of encoding functions $\{g_t^{(n)}\}_{t=1}^n$ as in (8) and satisfying (9) and of two decoding functions $\phi_1^{(n)}$ and $\phi_2^{(n)}$ as in (10).

We say that a rate-pair $(R_1, R_2)$ is achievable over the MIMO Gaussian BC with feedback under a power constraint $P$, if there exists a sequence of $\{(\lfloor 2^{nR_1}\rfloor, \lfloor 2^{nR_2}\rfloor, n)\}_{n=1}^{\infty}$ MIMO BC feedback-codes such that the average probability of error $P_{\text{e,BC}}^{(n)}$ tends to zero as the blocklength tends to infinity. The closure of the union of all achievable regions is called *capacity region*. We denote it by $\mathcal{C}_{\text{BC}}^{\text{fb}}(\mathsf{H}_1, \mathsf{H}_2, P)$. The supremum of the sum $R_1 + R_2$, where $(R_1, R_2)$ are in $\mathcal{C}_{\text{BC}}^{\text{fb}}(\mathsf{H}_1, \mathsf{H}_2, P)$ is called *sum-capacity* and is denoted $C_{\text{BC},\Sigma}^{\text{fb}}(\mathsf{H}_1, \mathsf{H}_2, P)$.

## *B. Linear-Feedback Coding Schemes for MIMO BC*

We restrict attention to *linear-feedback coding 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 $(M_1, M_2)$ (but not on the feedback). Specifically, we assume that the channel input vector has the form

$$\mathbf{X}_t = \mathbf{W}_t + \sum_{i=1}^{2}\sum_{\tau=1}^{t-1} \mathsf{A}_{i,\tau,t}\mathbf{Y}_{i,\tau}, \quad t\in\{1,\dots,n\}, \tag{14}$$

where $\mathbf{W}_t = \xi_t^{(n)}(M_1, M_2)$ and where $\{\mathsf{A}_{i,\tau,t}\}$ are arbitrary $\kappa$-by-$\nu_i$ matrices.

The mappings $\{\xi_t^{(n)}: \mathcal{M}_1 \times \mathcal{M}_2 \to \mathbb{R}^{\kappa}\}_{t=1}^n$ and the decoding operations $\phi_1^{(n)}$ and $\phi_2^{(n)}$ can be arbitrary.

Taking a linear combination of the information-carrying vector $\mathbf{W}_t$ and the past output vectors $\mathbf{Y}_{1,1},\dots,\mathbf{Y}_{1,t-1}$ and $\mathbf{Y}_{2,1},\dots,\mathbf{Y}_{2,t-1}$ is equivalent to taking a (different) linear combination of (a different information-carrying vector) $\tilde{\mathbf{W}}_t$ and the past noise vectors $\mathbf{Z}_{1,1},\dots,\mathbf{Z}_{1,t-1}$ and $\mathbf{Z}_{2,1},\dots,\mathbf{Z}_{2,t-1}$. Hence, we can equivalently write (14) as

$$\mathbf{X}_t = \tilde{\mathbf{W}}_t + \sum_{i=1}^{2}\sum_{\tau=1}^{t-1} \mathsf{B}_{i,\tau,t}\mathbf{Z}_{i,\tau}, \quad t\in\{1,\dots,n\}, \tag{15}$$

where $\tilde{\mathbf{W}}_t = \tilde{\xi}_t^{(n)}(M_1, M_2)$, for some arbitrary function $\tilde{\xi}_t^{(n)}: \mathcal{M}_1 \times \mathcal{M}_2 \to \mathbb{R}^{\kappa}$, and $\{\mathsf{B}_{i,\tau,t}\}$ are arbitrary $\kappa$-by-$\nu_i$ matrices.

The set of all rate-pairs achieved by linear-feedback coding schemes is called *linear-feedback capacity region* and is denoted $\mathcal{C}_{\text{BC}}^{\text{linfb}}(\mathsf{H}_1, \mathsf{H}_2; P)$. The largest sum-rate achieved by a linear-feedback coding scheme is called *linear-feedback sum-capacity* and is denoted $C_{\text{BC},\Sigma}^{\text{linfb}}(\mathsf{H}_1, \mathsf{H}_2; P)$.

## *C. Previous Results*

Without feedback, the capacity region of the MIMO Gaussian BC, $\mathcal{C}_{\text{BC}}^{\text{nofb}}(\mathsf{H}_1, \mathsf{H}_2; P)$ was determined by Weingarten, Steinberg, and Shamai [28]. Geng and Nair [14] have extended their result to also allow for an additional common message to be sent to the two receivers.

With feedback, the capacity region is unknown even in the scalar case. Achievable regions—based on linear-feedback coding schemes—have been proposed in [1], [5], [16], [11], [12], [21], [30]. Non-linear feedback schemes have been proposed in [24], [25], [32]. The best known achievable regions are due to linear-feedback schemes.

## IV. MIMO Gaussian MAC with Feedback

### A. Setup

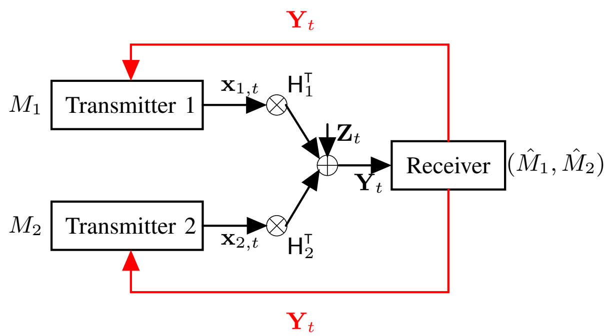


Fig. 2. Two-user memoryless MIMO Gaussian MAC with perfect feedback.

We consider the two-user memoryless MIMO Gaussian MAC with perfect output-feedback in Figure 2. Each Transmitter $i$, for $i \in \{1,2\}$, is equipped with $\nu_i$ transmit-antennas and the receiver is equipped with $\kappa$ receive-antennas. At each time $t \in \mathbb{N}$, if $\mathbf{x}_{1,t}$ and $\mathbf{x}_{2,t}$ denote the vector signals sent by Transmitters 1 and 2, the receiver observes the real vector-valued channel output

$$\mathbf{Y}_t = \mathsf{H}_1^\intercal \mathbf{x}_{1,t} + \mathsf{H}_2^\intercal \mathbf{x}_{2,t} + \mathbf{Z}_t, \tag{16}$$

where $\mathsf{H}_i$, for $i \in \{1,2\}$, is a deterministic real $\nu_i$-by-$\kappa$ channel matrix known to transmitters and receiver and $\{\mathbf{Z}_t\}$ is a sequence of independent and identically distributed $\kappa$-dimensional centered Gaussian random vectors of identity covariance matrix.

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, Transmitter $i$'s, $i \in \{1,2\}$, channel input at time $t$, $\mathbf{X}_{i,t}$, can depend on its message $M_i$ and the prior output vectors $\mathbf{Y}_1, \ldots, \mathbf{Y}_{t-1}$ :

$$\mathbf{X}_{i,t} = \varphi_{i,t}^{(n)}(M_i, \mathbf{Y}_1, \ldots, \mathbf{Y}_{t-1}), \quad t \in \{1, \ldots, n\}, \tag{17}$$

for some encoding functions of the form:

$$\varphi_{i,t}^{(n)} : \mathcal{M}_i \times \mathbb{R}^{\kappa(t-1)} \to \mathbb{R}. \tag{18}$$

The channel input sequences $\{\mathbf{X}_{1,t}\}_{t=1}^n$ and $\{\mathbf{X}_{2,t}\}_{t=1}^n$ have to satisfy a *total expected average block-power constraint* $P$:

$$\frac{1}{n} \sum_{t=1}^{n} \left( \mathbf{E}[\|\mathbf{X}_{1,t}\|^2] + \mathbf{E}[\|\mathbf{X}_{2,t}\|^2] \right) \leq P, \tag{19}$$

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

The receiver decodes the messages $(M_1, M_2)$ by means of a decoding function $\phi^{(n)}$ of the form

$$\phi^{(n)} : \mathbb{R}^{\kappa n} \to \mathcal{M}_1 \times \mathcal{M}_2. \tag{20}$$

This means, based on the output sequence $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$, the receiver produces its guess

$$(\hat{M}_1, \hat{M}_2) = \phi^{(n)}(\mathbf{Y}_1, \ldots, \mathbf{Y}_n). \tag{21}$$

An error occurs in the communication if

$$(\hat{M}_1, \hat{M}_2) \neq (M_1, M_2), \tag{22}$$

and thus the average probability of error is defined as

$$P_{\text{e,MAC}}^{(n)} \triangleq \Pr\left[(\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)\right]. \tag{23}$$

A $(\lfloor 2^{nR_1} \rfloor, \lfloor 2^{nR_2} \rfloor, n)$ *MIMO MAC feedback-code of sum-power* $P$ is a triple

$$\left( \{\varphi_{1,t}^{(n)}\}_{t=1}^n, \{\varphi_{2,t}^{(n)}\}_{t=1}^n, \Phi^{(n)} \right)$$

where $\{\varphi_{1,t}^{(n)}\}_{t=1}^n$ and $\{\varphi_{2,t}^{(n)}\}_{t=1}^n$ are of the form (18) and satisfy (19) and $\phi^{(n)}$ is as in (20).

We say that a rate-pair $(R_1, R_2)$ is achievable over the Gaussian MIMO MAC with feedback under a sum-power constraint $P$, if there exists a sequence of $\{(\lfloor 2^{nR_1} \rfloor, \lfloor 2^{nR_2} \rfloor, n)\}_{n=1}^\infty$ MIMO MAC feedback-codes such that the average probability of a decoding error $P_{\text{e,MAC}}^{(n)}$ tends to zero as the block-length $n$ tends to infinity. The closure of the union of all achievable regions is called *capacity region*. We denote it by $\mathcal{C}_{\text{MAC}}^{\text{fb}}(\mathsf{H}_1^\intercal, \mathsf{H}_2^\intercal, P)$. The supremum of the sum $R_1 + R_2$ over all pairs $(R_1, R_2)$ in $\mathcal{C}_{\text{MAC}}^{\text{fb}}(\mathsf{H}_1^\intercal, \mathsf{H}_2^\intercal, P)$ is called *sum-capacity* and is denoted by $C_{\text{MAC},\Sigma}^{\text{fb}}(\mathsf{H}_1^\intercal, \mathsf{H}_2^\intercal, P)$.

### B. Linear-Feedback Coding Schemes for MIMO MAC

In the present paper, we focus on the class of *linear-feedback* coding schemes where the channel inputs at Transmitter $i$, for $i \in \{1,2\}$, are given by *linear* combinations of the previous feedback signals and an information-carrying vector that only depends on the message $M_i$ (but not on the feedback).

Specifically, we assume that the channel input vectors have the form

$$\mathbf{X}_{i,t} = \mathbf{W}_{i,t} + \sum_{\tau=1}^{t-1} \mathsf{C}_{i,\tau,t} \mathbf{Y}_\tau, \quad i \in \{1,2\}, \quad t \in \{1, \ldots, n\}, \tag{24}$$

where $\mathbf{W}_{i,t}$ is an information-carrying vector

$$\mathbf{W}_{i,t} = \xi_{i,t}^{(n)}(M_i), \tag{25}$$

and $\{\mathsf{C}_{i,\tau,t}\}$ are arbitrary $\nu_i$-by-$\kappa$ matrices.

The mappings $\{\xi_{i,t}^{(n)} : \mathcal{M}_i \to \mathbb{R}^{\nu_i n}\}$ as well as the decoder mapping $\phi^{(n)}$ can be arbitrary (also non-linear).

The set of all rate-pairs achieved by linear-feedback coding schemes is called *linear-feedback capacity region* and is denoted $\mathcal{C}_{\text{MAC}}^{\text{linfb}}(\mathsf{H}_1^\intercal, \mathsf{H}_2^\intercal; P)$. The largest sum-rate achieved by a linear-feedback coding scheme is called *linear-feedback sum-capacity* and is denoted $C_{\text{MAC},\Sigma}^{\text{linfb}}(\mathsf{H}_1^\intercal, \mathsf{H}_2^\intercal; P)$.

**Remark 1.** *For any channel matrices $\mathsf{H}_1^\intercal$ and $\mathsf{H}_2^\intercal$ and power constraint $P > 0$:*

$$\mathcal{C}_{\text{MAC}}^{\text{linfb}}(\mathsf{H}_1^\intercal, \mathsf{H}_2^\intercal; P) = \mathcal{C}_{\text{MAC}}^{\text{linfb}}(\bar{\mathsf{H}}_1, \bar{\mathsf{H}}_2; P). \tag{26}$$

To see this, note that if each transmitter of a MIMO MAC with channel matrices $(\mathsf{H}_1^\intercal, \mathsf{H}_2^\intercal)$ multiplies its input vectors by $\mathsf{E}$ (from the left) before sending the result over the MAC and if the receiver and the transmitters multiply their observed output vectors and feedback vectors by $\mathsf{E}$ (from the left) before using them, then the MIMO MAC is transformed into a MIMO MAC with channel matrices $(\bar{\mathsf{H}}_1, \bar{\mathsf{H}}_2)$. And in the same way the MIMO MAC with channel matrices $(\bar{\mathsf{H}}_1, \bar{\mathsf{H}}_2)$ can be transformed into a MIMO MAC with channel matrices $(\mathsf{H}_1^\intercal, \mathsf{H}_2^\intercal)$.

### C. Previous Results

Without feedback, the capacity region of the Gaussian MIMO MAC under a sum-power constraint $P$, $\mathcal{C}_{\text{MAC}}^{\text{nofb}}(\mathsf{H}_1^\intercal, \mathsf{H}_2^\intercal; P)$ is readily obtained from the results in [7].

With perfect feedback, the capacity region of the MIMO Gaussian MAC under sum-power constraint $P$ is known only in few special cases. An example is the scalar case $\nu_1 = \nu_2 = \kappa = 1$, which we also call single-input single-output (SISO) setup. In this setup, the channel matrices $(\mathsf{H}_1^\intercal, \mathsf{H}_2^\intercal)$ reduce to the scalar coefficients $(h_1, h_2)$. Ozarow [20] determined the capacity region of the two-user scalar Gaussian MAC with perfect feedback under individual power constraints $P_1$ and $P_2$ on the two transmitters' input sequences. It is given by

$$\mathcal{R}_{\text{Oz}}(h_1, h_2; P_1, P_2) = \bigcup_{\rho \in [0,1]} \mathcal{R}_{\text{Oz}}^{\rho}(h_1, h_2; P_1, P_2) \quad (27)$$

where for each $\rho \in [0,1]$, $\mathcal{R}_{\text{Oz}}^{\rho}(h_1, h_2; P_1, P_2)$ denotes the set of all nonnegative rate-pairs $(R_1, R_2)$ that satisfy

$$R_1 \leq \frac{1}{2}\log\left(1 + h_1^2 P_1 (1-\rho^2)\right), \quad (28a)$$

$$R_2 \leq \frac{1}{2}\log\left(1 + h_2^2 P_2 (1-\rho^2)\right), \quad (28b)$$

$$R_1 + R_2 \leq \frac{1}{2}\log\left(1 + h_1^2 P_1 + h_2^2 P_2 + 2\sqrt{h_1^2 h_2^2 P_1 P_2}\rho\right). \quad (28c)$$

From Ozarow's result, we can directly deduce the capacity region of the scalar Gaussian MAC with perfect feedback under a sum-power constraint:

$$\mathcal{C}_{\text{MAC,SISO}}^{\text{fb}}(h_1, h_2; P) = \bigcup_{\substack{P_1, P_2 \geq 0: \\ P_1 + P_2 = P}} \mathcal{R}_{\text{Oz}}(h_1, h_2; P_1, P_2). \quad (29)$$

Thus the capacity region $\mathcal{C}_{\text{MAC,SISO}}^{\text{fb}}$ is achieved by applying Ozarow's scheme with different power splits between the two transmitters. The sum-capacity $C_{\text{MAC,SISO},\Sigma}^{\text{fb}}(h_1, h_2; P)$ is

$$\begin{aligned} C_{\text{MAC,SISO},\Sigma}^{\text{fb}}&(h_1, h_2; P) \\ = \sup_{\substack{P_1, P_2 \geq 0: \\ P_1 + P_2 = P}} & \frac{1}{2}\log\Big(1 + h_1^2 P_1 + h_2^2 P_2 \\ & + 2\sqrt{h_1^2 h_2^2 P_1 P_2} \cdot \rho^\star(h_1, h_2; P_1, P_2)\Big) \end{aligned} \quad (30)$$

where $\rho^\star(h_1, h_2; P_1, P_2)$ is the unique solution in $(0,1)$ to the following quartic equation in $\rho$

$$\begin{aligned} 1 + h_1^2 P_1 + h_2^2 P_2 + 2\sqrt{h_1^2 h_2^2 P_1 P_2}\rho \\ = \left(1 + h_1^2 P_1 (1-\rho^2)\right)\left(1 + h_2^2 P_2 (1-\rho^2)\right). \end{aligned} \quad (31)$$

In Appendix A-A, we show that in a symmetric setup where $h_1 = h_2 = h$,

$$\begin{aligned} C_{\text{MAC,SISO},\Sigma}^{\text{fb}}&(h, h; P) \\ &= \frac{1}{2}\log\left(1 + h^2 P (1 + \rho^\star(h, h; P/2, P/2))\right). \end{aligned} \quad (32)$$

Ozarow's coding scheme is a *linear feedback coding scheme* since it combines a Schalkwijk-Kailath [23] type scheme at both transmitters with a non-feedback scheme at one of the two transmitters. Specifically, one transmitter sends scaled versions of the actual errors when performing linear minimum mean squared error estimation (LMMSE) of its message point (which depend only on the message) based on the previous feedback signals. The other transmitter sends the sum of the symbols of a non-feedback coding scheme and the scaled errors about its message point based on the previous feedback signals. Since any non-feedback coding scheme is a linear-feedback coding scheme and also the errors are by definition linear in the feedback, the overall Ozarow's coding scheme is also a linear-feedback coding scheme.

Thus, in the SISO case,

$$\mathcal{C}_{\text{MAC,SISO}}^{\text{fb}}(h_1, h_2; P) = \mathcal{C}_{\text{MAC,SISO}}^{\text{linfb}}(h_1, h_2; P), \quad (33)$$

and

$$C_{\text{MAC,SISO},\Sigma}^{\text{fb}}(h_1, h_2; P) = C_{\text{MAC,SISO},\Sigma}^{\text{linfb}}(h_1, h_2; P). \quad (34)$$

Jafar & Goldsmith [15] derived the capacity region with perfect feedback under individual power constraints in the multi-input single-output (MISO) case ($\nu_1, \nu_2$ arbitrary and $\kappa = 1$) and in the single-input multi-output (SIMO) case ($\nu_1 = \nu_2 = 1$ and $\kappa$ arbitrary). In both cases the capacity is achieved by a variation of Ozarow's scheme. Based on these results we immediately obtain the linear-feedback capacity region under a total sum-power constraint.

In the MISO case, the channel matrices $\mathsf{H}_1^\intercal$ and $\mathsf{H}_2^\intercal$ reduce to the $1 \times \nu_1$ and $1 \times \nu_2$ vectors $\mathbf{h}_1^\intercal$ and $\mathbf{h}_2^\intercal$ and the channel output can be written as

$$Y_t = \mathbf{h}_1^\intercal \mathbf{x}_{1,t} + \mathbf{h}_2^\intercal \mathbf{x}_{2,t} + Z_t. \quad (35)$$

The linear-feedback capacity region is given by

$$\begin{aligned} \mathcal{C}_{\text{MAC,MISO}}^{\text{linfb}}(\mathbf{h}_1^\intercal, \mathbf{h}_2^\intercal; P) &= \mathcal{C}_{\text{MAC,MISO}}^{\text{fb}}(\mathbf{h}_1^\intercal, \mathbf{h}_2^\intercal; P) \\ &= \mathcal{C}_{\text{MAC,SISO}}^{\text{fb}}(\|\mathbf{h}_1\|, \|\mathbf{h}_2\|; P), \end{aligned} \quad (36)$$

where notice that the last expression involves the *SISO* capacity region $\mathcal{C}_{\text{MAC,SISO}}^{\text{fb}}(\|\mathbf{h}_1\|, \|\mathbf{h}_2\|; P)$.

In the SIMO case, the channel matrices reduce to the $\kappa \times 1$ vectors $\mathbf{h}_1^\intercal, \mathbf{h}_2^\intercal$ and the channel output vector can be written as

$$\mathbf{Y}_t = \mathbf{h}_1 x_{1,t} + \mathbf{h}_2 x_{2,t} + \mathbf{Z}_t. \quad (37)$$

The linear-feedback capacity region is given by:

$$\begin{aligned} \mathcal{C}_{\text{MAC,SIMO}}^{\text{linfb}}&(\mathbf{h}_1^\intercal, \mathbf{h}_2^\intercal; P) \\ &= \mathcal{C}_{\text{MAC,SIMO}}^{\text{fb}}(\mathbf{h}_1^\intercal, \mathbf{h}_2^\intercal; P) \end{aligned} \quad (38)$$

$$=\bigcup_{\substack{P_1,P_2\geq 0:\\ P_1+P_2=P}} \mathrm{cl}\left(\bigcup_{\rho\in[0,1]} \mathcal{R}^{\rho}_{\text{SIMO}}(\mathbf{h}_1^{\mathsf{T}},\mathbf{h}_2^{\mathsf{T}};P_1,P_2)\right) \tag{39}$$

where for each $\rho \in [0,1]$, $\mathcal{R}^{\rho}_{\text{SIMO}}(\mathbf{h}_1^{\mathsf{T}},\mathbf{h}_2^{\mathsf{T}};P_1,P_2)$ denotes the set of all nonnegative rate-pairs $(R_1,R_2)$ that satisfy

$$R_1 \leq \frac{1}{2}\log\left(1+\|\mathbf{h}_1\|^2P_1(1-\rho^2)\right), \tag{40a}$$

$$R_2 \leq \frac{1}{2}\log\left(1+\|\mathbf{h}_2\|^2P_2(1-\rho^2)\right), \tag{40b}$$

$$\begin{aligned}R_1+R_2 \leq \frac{1}{2}\log\big(&1+\|\mathbf{h}_1\|^2P_1+\|\mathbf{h}_2\|^2P_2\\ &+2\rho\beta\sqrt{\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2P_1P_2}\\ &+\|\mathbf{h}_1\|^2\|\mathbf{h}_2\|^2P_1P_2(1-\rho^2)(1-\beta^2)\big),\end{aligned} \tag{40c}$$

where $\beta \triangleq \frac{\mathbf{h}_1^{\mathsf{T}}\mathbf{h}_2}{\|\mathbf{h}_1\|\|\mathbf{h}_2\|}$ may be interpreted as the cosine of the angle between the channel vectors. If $\beta=1$ the channel reduces to a SISO MAC with coefficients $(\|\mathbf{h}_1\|,\|\mathbf{h}_2\|)$.

## V. Previous Results: MAC-BC Duality without Feedback

Combining the results in [26] and [28], the following duality on the capacity regions of the Gaussian MIMO MAC and the Gaussian MIMO BC is obtained:

**Theorem 1** (From [26] and [28])**.**

$$\mathcal{C}^{\text{nofb}}_{\text{BC}}(\mathsf{H}_1,\mathsf{H}_2;P)=\mathcal{C}^{\text{nofb}}_{\text{MAC}}(\mathsf{H}_1^{\mathsf{T}},\mathsf{H}_2^{\mathsf{T}};P). \tag{41}$$

More specifically, Vishwanath, Jindal, and Goldsmith [26] showed that the following two "dual" schemes for the Gaussian MIMO MAC and the Gaussian MIMO BC achieve the same (rectangular) rate-region. Taking the union over all the choices of parameters, we obtain the capacity regions of Gaussian MIMO MAC and the Gaussian MIMO BC.

In the MAC-scheme, the two transmitters use vector-valued Gaussian codebooks of covariance matrices $\Xi_1$ and $\Xi_2$ and the only receiver performs successive decoding and stripping, first decoding Message $M_1$ followed by Message $M_2$. In the BC-scheme, the only transmitter encodes $M_2$ using a Gaussian codebook of covariance matrix $\Psi_2$ and it encodes $M_1$ using a vector-valued dirty-paper code of covariance matrix $\Psi_1$ that treats the codeword produced for $M_2$ as interference. The encoder then sends the sum of the two produced sequences over the channel. Receiver 1 decodes $M_1$ using a point-to-point decoding and Receiver 2 decodes $M_2$ using dirty-paper decoding. Vishwanath, Jindal, and Goldsmith [26] showed that when $\Psi_1$ and $\Psi_2$ are chosen as a function of $\Xi_1$ and $\Xi_2$ as

$$\Psi_1=(\mathsf{I}+\mathsf{H}_2^{\mathsf{T}}\Xi_2\mathsf{H}_2)^{-\frac{1}{2}}(\mathsf{F}_1\mathsf{G}_1^{\mathsf{T}}\Xi_1\mathsf{G}_1\mathsf{F}_1^{\mathsf{T}})(\mathsf{I}+\mathsf{H}_2^{\mathsf{T}}\Xi_2\mathsf{H}_2)^{-\frac{1}{2}}, \tag{42a}$$

$$\Psi_2=\mathsf{F}_2\mathsf{G}_2^{\mathsf{T}}(\mathsf{I}+\mathsf{H}_2^{\mathsf{T}}\Psi_1\mathsf{H}_2)^{\frac{1}{2}}\,\Xi_2\,(\mathsf{I}+\mathsf{H}_2^{\mathsf{T}}\Psi_1\mathsf{H}_2)^{\frac{1}{2}}\,\mathsf{G}_2\mathsf{F}_2^{\mathsf{T}}, \tag{42b}$$

then the two schemes achieve the same regions and the sum-power used in the BC scheme cannot exceed the sum-power of the MAC scheme. In (42), the matrices $\mathsf{F}_i$ and $\mathsf{G}_i$, $i\in\{1,2\}$, are defined through the singular value decompositions (SVD)

$$(\mathsf{I}+\mathsf{H}_2^{\mathsf{T}}\Xi_2\mathsf{H}_2)^{-\frac{1}{2}}\,\mathsf{H}_1^{\mathsf{T}}=\mathsf{F}_1\Lambda_1\mathsf{G}_1^{\mathsf{T}}, \tag{43a}$$

$$\mathsf{H}_2^{\mathsf{T}}(\mathsf{I}+\mathsf{H}_2^{\mathsf{T}}\Psi_1\mathsf{H}_2)^{-\frac{1}{2}}=\mathsf{F}_2\Lambda_2\mathsf{G}_2^{\mathsf{T}}, \tag{43b}$$

where $\Lambda_i$ denotes a square and diagonal matrix.

They also showed that when $\Xi_1$ and $\Xi_2$ are chosen depending on $\Psi_1$ and $\Psi_2$ as:

$$\Xi_2=(\mathsf{I}+\mathsf{H}_2\Psi_1\mathsf{H}_2^{\mathsf{T}})^{-\frac{1}{2}}\,\mathsf{U}_2\mathsf{V}_2^{\mathsf{T}}\Psi_2\mathsf{V}_2\mathsf{U}_2^{\mathsf{T}}\,(\mathsf{I}+\mathsf{H}_2\Psi_1\mathsf{H}_2^{\mathsf{T}})^{-\frac{1}{2}}, \tag{44a}$$

$$\Xi_1=\mathsf{U}_1\mathsf{V}_1^{\mathsf{T}}(\mathsf{I}+\mathsf{H}_2^{\mathsf{T}}\Xi_2\mathsf{H}_2)^{\frac{1}{2}}\,\Psi_1\,(\mathsf{I}+\mathsf{H}_2^{\mathsf{T}}\Xi_2\mathsf{H}_2)^{\frac{1}{2}}\,\mathsf{V}_1\mathsf{U}_1^{\mathsf{T}}, \tag{44b}$$

then the two schemes achieve the same regions and the sum-power used in the MAC scheme cannot exceed the sum-power of the BC scheme. In (44), for $i\in\{1,2\}$, the matrices $\mathsf{U}_i$ and $\mathsf{V}_i$ are defined through the SVD decompositions

$$\mathsf{H}_1(\mathsf{I}+\mathsf{H}_2^{\mathsf{T}}\Xi_2\mathsf{H}_2)^{\frac{1}{2}}=\mathsf{U}_1\Delta_1\mathsf{V}_1^{\mathsf{T}}, \tag{45a}$$

$$(\mathsf{I}+\mathsf{H}_2\Psi_1\mathsf{H}_2^{\mathsf{T}})^{-\frac{1}{2}}\,\mathsf{H}_2=\mathsf{U}_2\Delta_2\mathsf{V}_2^{\mathsf{T}}, \tag{45b}$$

where $\Delta_i$ denotes a square and diagonal matrix.

## VI. Main Results

In this section, we state our main results for the two-user case. Partial extensions to more than two users and to different setups are given in Sections VIII and IX.

### A. MAC-BC Duality with Linear-Feedback Coding Schemes

**Theorem 2.**

$$\mathcal{C}^{\text{linfb}}_{\text{BC}}(\mathsf{H}_1,\mathsf{H}_2;P)=\mathcal{C}^{\text{linfb}}_{\text{MAC}}(\mathsf{H}_1^{\mathsf{T}},\mathsf{H}_2^{\mathsf{T}};P). \tag{46}$$

*Proof.* Follows by Propositions 1, 2, and 3 ahead, by point 2 of Note 1, and because the capacity regions of the MACs with channel matrices $\mathsf{H}_1^{\mathsf{T}}$ and $\mathsf{H}_2^{\mathsf{T}}$ and $\bar{\mathsf{H}}_1$ and $\bar{\mathsf{H}}_2$ coincide, see Remark 1. □

In the scalar case, Theorem 2 combined with (33) and (34) specialize to:

**Corollary 1.**

$$\mathcal{C}^{\text{linfb}}_{\text{BC,SISO}}(h_1,h_2;P)=\mathcal{C}^{\text{linfb}}_{\text{MAC,SISO}}(h_1,h_2;P) \tag{47}$$

$$=\mathcal{C}^{\text{fb}}_{\text{MAC,SISO}}(h_1,h_2;P) \tag{48}$$

*and*

$$C^{\text{linfb}}_{\text{BC,SISO},\Sigma}(h_1,h_2;P)=C^{\text{linfb}}_{\text{MAC,SISO},\Sigma}(h_1,h_2;P) \tag{49}$$

$$=C^{\text{fb}}_{\text{MAC,SISO},\Sigma}(h_1,h_2;P). \tag{50}$$

Figures 3 and 4 compare the linear-feedback capacity region for the SISO Gaussian BC to the non-feedback capacity region [4], [8], to Ozarow & Leung's achievable region [21], and to Ozarow & Leung [21] physically degraded outer bound [21]. This outer bound is given by $\mathcal{C}_1^{\text{enh}}\cap\mathcal{C}_2^{\text{enh}}$ where $\mathcal{C}_i^{\text{enh}}$, $i\in\{1,2\}$, is the capacity region of the enhanced physically degraded BC channel where Receiver $i$ obtains both channel outputs $Y_1$ and $Y_2$, and thus is unchanged with and without feedback.

**Remark 2.** *Our proposed linear-feedback capacity region $\mathcal{C}^{\text{linfb}}_{\text{BC,MIMO}}$ for the MIMO Gaussian BC, is the best known inner bound on $\mathcal{C}^{\text{fb}}_{\text{BC,MIMO}}$. In the two-user SISO case, the achievable region $\mathcal{C}^{\text{linfb}}_{\text{BC,SISO}}$ is generally very close to Ozarow & Leung's outer bound [21]. This can be seen at hands of Figures 3–5. Moreover, [13] shows that in the high-SNR limit as $P\to\infty$*

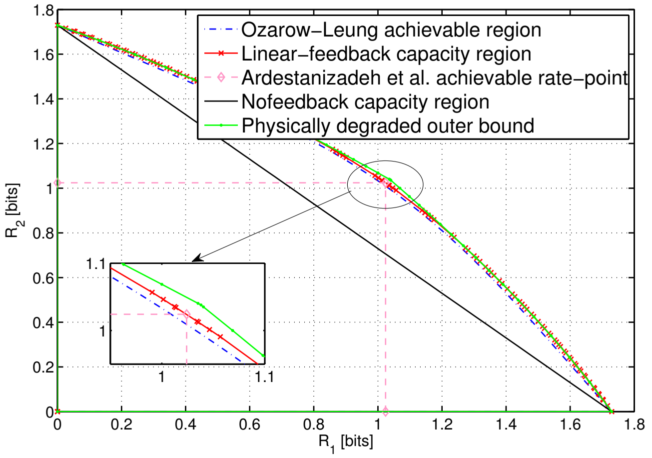


Fig. 3. Achievable regions and physically degraded outer bound for the symmetric SISO Gaussian BC with perfect feedback, with channel coefficients $h_1 = h_2 = 1$ and power constraint $P = 10$.

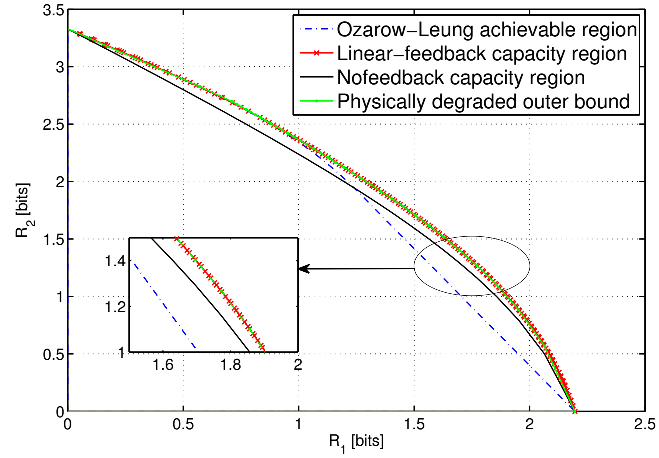


Fig. 5. Achievable regions and physically degraded outer bound for the non-symmetric SISO Gaussian BC with perfect feedback, with channel coefficients $h_1 = \frac{1}{\sqrt{5}}, h_2 = 1$ and power constraint $P = 100$.

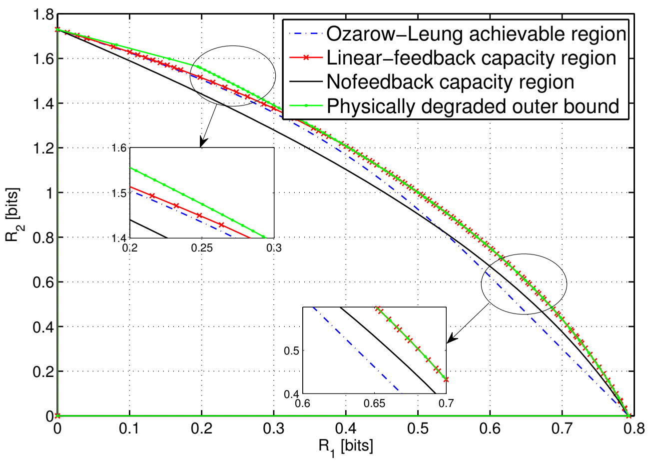


Fig. 4. Achievable regions and physically degarded outer bound for the non-symmetric SISO Gaussian BC with perfect feedback, with channel coefficients $h_1 = \frac{1}{\sqrt{5}}, h_2 = 1$ and power constraint $P = 10$.

*the difference the between linear-feedback sum-capacity and the actual sum-capacity vanishes.*

Using also (32), in the symmetric case we obtain:

**Corollary 2.** *If $h_1 = h_2 = h$, then*

$$\mathcal{C}^{\text{linfb}}_{\text{BC,SISO},\Sigma}(h,h;P) = \frac{1}{2}\log\left(1 + h^2P + h^2P \cdot \rho^\star(h,h;P/2,P/2)\right), \quad (51)$$

*where recall that $\rho^\star(h_1,h_2;P_1,P_2)$ is defined as the solution to the quartic equation in* (31).

The achievability of the sum-rate in (51) was already established by the control-theory-inspired scheme in [1]. Our result shows that for the symmetric scalar Gaussian BC the scheme in [1], [11], [30], is indeed sum-rate optimal among all linear-feedback coding schemes.

In the SIMO and the MISO case, Theorem 2 combined with (36) and (38) specialize to:

**Corollary 3.** *Consider the SIMO and MISO cases where the channel matrices reduce to vectors. Let $\mathbf{h}_1$ and $\mathbf{h}_2$ be $\kappa$-dimensional row-vectors. Then,*

$$\mathcal{C}^{\text{linfb}}_{\text{BC,MISO}}(\mathbf{h}_1,\mathbf{h}_2;P) = \mathcal{C}^{\text{fb}}_{\text{MAC,SIMO}}(\mathbf{h}_1^{\mathsf{T}},\mathbf{h}_2^{\mathsf{T}};P) \quad (52)$$

*Let now $\mathbf{h}_1$ and $\mathbf{h}_2$ be $\nu_1$ and $\nu_2$-dimensional column-vectors. Then,*

$$\mathcal{C}^{\text{linfb}}_{\text{BC,SIMO}}(\mathbf{h}_1,\mathbf{h}_2;P) = \mathcal{C}^{\text{fb}}_{\text{MAC,MISO}}(\mathbf{h}_1^{\mathsf{T}},\mathbf{h}_2^{\mathsf{T}};P) \quad (53)$$
$$= \mathcal{C}^{\text{fb}}_{\text{MAC,SISO}}(\|\mathbf{h}_1\|,\|\mathbf{h}_2\|;P). \quad (54)$$

*See* (29), (36), *and* (39) *for computable single-letter characterizations of $\mathcal{C}^{\text{fb}}_{\text{MAC,SISO}}$, $\mathcal{C}^{\text{fb}}_{\text{MAC,SIMO}}$, and $\mathcal{C}^{\text{fb}}_{\text{MAC,MISO}}$.*

### *B. Linear-Feedback Capacity-Achieving Coding Schemes for MAC and BC*

We first describe a class of linear-feedback coding schemes for the BC and the MAC that can achieve the linear-feedback capacity regions $\mathcal{C}^{\text{linfb}}_{\text{BC}}$ and $\mathcal{C}^{\text{linfb}}_{\text{MAC}}$. This allows us to find multi-letter expressions for these capacity regions. We then identify pairs of linear-feedback coding schemes for the BC and the MAC that are dual in the sense that they achieve the same rate-regions.

The idea of our schemes is to divide the blocklength $n$ into subblocks of equal length $\eta$ ($\eta$ is a design parameter of our schemes) and to apply an inner code that uses the feedback to transform each subblock of $\eta$ channel uses of the original MIMO BC or MAC into a single channel use of a new MIMO BC or MAC with more transmit and receive antennas. An outer code is then applied to communicate over the new MIMO BC or MAC without using the feedback.

We now explain this class of schemes in more detail.

*1) A Class of Linear-Feedback Coding Schemes for the BC:* Fix the blocklength $n$. The schemes in our class are characterized by the following parameters:

- a positive integer $\eta$;
- $\kappa$-by-$\nu_1$ matrices $\{\mathsf{A}_{1,\tau,\ell}\}$, for $\ell = 2,\dots,\eta$ and $\tau = 1,\dots,\ell-1$;
- $\kappa$-by-$\nu_2$ matrices $\{\mathsf{A}_{2,\tau,\ell}\}$, for $\ell = 2,\dots,\eta$ and $\tau = 1,\dots,\ell-1$;

- an encoder mapping $f^{(n')}\colon \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{R}^{(\kappa\eta)n'}$ that produces $n' \triangleq \lfloor \frac{n}{\eta} \rfloor$ codevectors (column-vectors) of size $\kappa\eta$ and
- two decoder mappings $g_1^{(n')}\colon \mathbb{R}^{(\nu_1\eta)n'} \to \mathcal{M}_1$ and $g_2^{(n')}\colon \mathbb{R}^{(\nu_2\eta)n'} \to \mathcal{M}_2$ that each decode a block of $n'$ output vectors (column-vectors) of size $\nu_1\eta$ and $\nu_2\eta$.

As already mentioned, the parameter $\eta$ characterizes the length of the subblocks in our scheme. That means, in our scheme the total blocklength $n$ is divided into $n'$ subblocks of equal length $\eta$.[1] The matrices $\{\mathsf{A}_{1,\tau,\ell}\}$ and $\{\mathsf{A}_{2,\tau,\ell}\}$ describe the inner code that is used within each of the $n'$ subblocks of length $\eta$. Finally, the parameters $f^{(n')}, g_1^{(n')}, g_2^{(n')}$ describe the outer code that is applied to code over the $n'$ subblocks without using the feedback.

Before describing how the inner code works and how we should choose the encoding and decoding functions of the outer code, we need some definitions. Let

$$\mathbf{X} \triangleq \begin{pmatrix} \mathbf{X}_1^\mathsf{T}, & \ldots, & \mathbf{X}_\eta^\mathsf{T} \end{pmatrix}^\mathsf{T}, \tag{55}$$

denote the $\eta\kappa$-dimensional column-vector that is obtained by stacking the first $\eta$ channel input vectors $\mathbf{X}_1, \ldots, \mathbf{X}_\eta$ (which are all $\kappa$-dimensional column-vectors) on top of each other. Similarly, for $i \in \{1,2\}$, let

$$\mathbf{Z}_i \triangleq \begin{pmatrix} \mathbf{Z}_{i,1}^\mathsf{T}, & \ldots, & \mathbf{Z}_{i,\eta}^\mathsf{T} \end{pmatrix}^\mathsf{T}, \tag{56}$$

$$\mathbf{Y}_i \triangleq \begin{pmatrix} \mathbf{Y}_1^\mathsf{T}, & \ldots, & \mathbf{Y}_\eta^\mathsf{T} \end{pmatrix}^\mathsf{T}, \tag{57}$$

denote the $\eta\nu_i$ dimensional column-vectors that are obtained by stacking the first $\eta$ noise vectors $\mathbf{Z}_{1,i}, \ldots, \mathbf{Z}_{i,\eta}$ or channel output vectors $\mathbf{Y}_{1,i}, \ldots, \mathbf{Y}_{i,\eta}$ on top of each other. Define for $i \in \{1,2\}$, the channel matrices of the $\eta$-length subblocks:

$$\mathsf{H}_i^\mathsf{B} \triangleq \mathsf{I}_\eta \otimes \mathsf{H}_i. \tag{58}$$

The input-output relation for the first block of $\eta$ channel uses is then summarized as

$$\mathbf{Y}_i = \mathsf{H}_i^\mathsf{B}\mathbf{X} + \mathbf{Z}_i, \qquad i \in \{1,2\}. \tag{59}$$

Let $\mathbf{U}$ denote the $\eta\kappa$-dimensional vector produced by outer encoder $f^{(n')}$ for this first block, and define, for $i \in \{1,2\}$, the $\eta\kappa$-by-$\eta\nu_i$ strictly-lower block-triangular matrix

$$\mathsf{A}_i^\mathsf{B} = \begin{bmatrix} 0 & & \cdots & & 0 \\ \mathsf{A}_{i,1,2} & 0 & & & \\ \mathsf{A}_{i,1,3} & \mathsf{A}_{i,2,3} & 0 & & \\ \vdots & & & \ddots & \\ \mathsf{A}_{i,1,\eta} & \mathsf{A}_{i,2,\eta} & \ldots & \mathsf{A}_{i,(\eta-1),\eta} & 0 \end{bmatrix}, \tag{60}$$

where here $0$ denotes the $\kappa$-by-$\nu_i$ matrix with all zero entries.

We now describe how the inner code—specified by the matrices $\{\mathsf{A}_{1,\tau,\ell}\}$ and $\{\mathsf{A}_{2,\tau,\ell}\}$— transforms the first block of $\eta$ channel uses of our original MIMO Gaussian BC into a single channel use of the new MIMO BC. All the other blocks are transformed in a similar way. In our scheme, we choose the encoder to produce the following $\eta$ channel inputs in the first block:

$$\mathbf{X} = \left(\mathsf{I} - \mathsf{A}_1^\mathsf{B}\mathsf{H}_1^\mathsf{B} - \mathsf{A}_2^\mathsf{B}\mathsf{H}_2^\mathsf{B}\right)\mathbf{U} + \mathsf{A}_1^\mathsf{B}\mathbf{Y}_1 + \mathsf{A}_2^\mathsf{B}\mathbf{Y}_2. \tag{61}$$

(This describes a general linear-feedback scheme. The only reason for precoding the codeword vector $\mathbf{U}$ by the matrix $\left(\mathsf{I} - \mathsf{A}_1^\mathsf{B}\mathsf{H}_1^\mathsf{B} - \mathsf{A}_2^\mathsf{B}\mathsf{H}_2^\mathsf{B}\right)$ is to simplify the calculations thereafter, see (64) ahead.) By (59), the inputs can also be written as

$$\begin{aligned}\mathbf{X} = &\left(\mathsf{I} - \mathsf{A}_1^\mathsf{B}\mathsf{H}_1^\mathsf{B} - \mathsf{A}_2^\mathsf{B}\mathsf{H}_2^\mathsf{B}\right)\mathbf{U} \\ &+ \mathsf{A}_1^\mathsf{B}(\mathsf{H}_1^\mathsf{B}\mathbf{X} + \mathbf{Z}_1) + \mathsf{A}_2^\mathsf{B}(\mathsf{H}_1^\mathsf{B}\mathbf{X} + \mathbf{Z}_2)\end{aligned} \tag{62}$$

and thus,

$$\begin{aligned}&\left(\mathsf{I} - \mathsf{A}_1^\mathsf{B}\mathsf{H}_1^\mathsf{B} - \mathsf{A}_2^\mathsf{B}\mathsf{H}_2^\mathsf{B}\right)\mathbf{X} \\ &\quad = \left(\mathsf{I} - \mathsf{A}_1^\mathsf{B}\mathsf{H}_1^\mathsf{B} - \mathsf{A}_2^\mathsf{B}\mathsf{H}_2^\mathsf{B}\right)\mathbf{U} + \mathsf{A}_1^\mathsf{B}\mathbf{Z}_1 + \mathsf{A}_2^\mathsf{B}\mathbf{Z}_2.\end{aligned} \tag{63}$$

Multiplying both sides of (63) from the left by the invertible matrix $\left(\mathsf{I} - \mathsf{A}_1^\mathsf{B}\mathsf{H}_1^\mathsf{B} - \mathsf{A}_2^\mathsf{B}\mathsf{H}_2^\mathsf{B}\right)^{-1}$ results in:

$$\mathbf{X} = \mathbf{U} + \mathsf{B}_1^\mathsf{B}\mathbf{Z}_1 + \mathsf{B}_2^\mathsf{B}\mathbf{Z}_2, \tag{64}$$

where we defined

$$\mathsf{B}_i^\mathsf{B} \triangleq \left(\mathsf{I} - \mathsf{A}_1^\mathsf{B}\mathsf{H}_1^\mathsf{B} - \mathsf{A}_2^\mathsf{B}\mathsf{H}_2^\mathsf{B}\right)^{-1}\mathsf{A}_i^\mathsf{B}, \quad i \in \{1,2\}. \tag{65}$$

By (59) the corresponding outputs can be written as

$$\mathbf{Y}_1 = \mathsf{H}_1^\mathsf{B}\mathbf{U} + (\mathsf{I} + \mathsf{H}_1^\mathsf{B}\mathsf{B}_1^\mathsf{B})\mathbf{Z}_1 + \mathsf{H}_1^\mathsf{B}\mathsf{B}_2^\mathsf{B}\mathbf{Z}_2, \tag{66a}$$

$$\mathbf{Y}_2 = \mathsf{H}_2^\mathsf{B}\mathbf{U} + (\mathsf{I} + \mathsf{H}_2^\mathsf{B}\mathsf{B}_2^\mathsf{B})\mathbf{Z}_2 + \mathsf{H}_2^\mathsf{B}\mathsf{B}_1^\mathsf{B}\mathbf{Z}_1. \tag{66b}$$

Inspecting (64), we see that the channel inputs $\{\mathbf{X}_t\}_{t=1}^n$ to our original MIMO BC satisfy the average block-power constraint (9) if

$$\mathrm{tr}\left(\mathsf{B}_1^\mathsf{B}(\mathsf{B}_1^\mathsf{B})^\mathsf{T}\right) + \mathrm{tr}\left(\mathsf{B}_2^\mathsf{B}(\mathsf{B}_2^\mathsf{B})^\mathsf{T}\right) \leq \eta P \tag{67}$$

and if the $n'$ codevectors produced by the outer encoder $f^{(n')}$ are average block-power constrained to power

$$\eta P - \mathrm{tr}\left(\mathsf{B}_1^\mathsf{B}(\mathsf{B}_1^\mathsf{B})^\mathsf{T}\right) - \mathrm{tr}\left(\mathsf{B}_2^\mathsf{B}(\mathsf{B}_2^\mathsf{B})^\mathsf{T}\right). \tag{68}$$

**Definition 2.** *Let* $\mathcal{R}_{\mathrm{BC}}\left(\eta, \mathsf{B}_1^\mathsf{B}, \mathsf{B}_2^\mathsf{B}, \mathsf{H}_1^\mathsf{B}, \mathsf{H}_2^\mathsf{B}; P\right)$ *denote the capacity region of the MIMO Gaussian BC in* (66) without feedback *when the vector-input* $\mathbf{U}$ *is average block-power constrained to* (68).

The outer code $\{f^{(n')}, g_1^{(n')}, g_2^{(n')}\}$ is designed to achieve the non-feedback capacity of the new MIMO Gaussian BC in (66) under average input-power constraint $\eta P - \mathrm{tr}(\mathsf{B}_1^\mathsf{B}(\mathsf{B}_1^\mathsf{B})^\mathsf{T}) - \mathrm{tr}(\mathsf{B}_2^\mathsf{B}(\mathsf{B}_2^\mathsf{B})^\mathsf{T})$.

Combining all this, we conclude that over the original MIMO Gaussian BC with feedback our overall scheme (consisting of inner and outer code) achieves the rate region $\mathcal{R}_{\mathrm{BC}}\left(\eta, \mathsf{B}_1^\mathsf{B}, \mathsf{B}_2^\mathsf{B}, \mathsf{H}_1^\mathsf{B}, \mathsf{H}_2^\mathsf{B}; P\right)$ scaled by a factor $\frac{1}{\eta}$. In view of the following Note 2, it thus follows that our schemes achieve the rate region in (70) ahead.

**Note 2.** *Let* $\mathcal{T} \triangleq \mathcal{T}_1 \times \mathcal{T}_2$ *where* $\mathcal{T}_i$*, for* $i \in \{1,2\}$*, denotes the set of strictly-lower block-triangular matrices with block matrices of size* $\kappa \times \nu_i$*. The mapping described by* (65) *has the form*

$$\omega\colon \quad \mathcal{T} \quad \to \quad \mathcal{T}$$

[1]For general blocklength $n$ there will be a few spare channel uses at the end of each block which we do ignore in our schemes. Since throughout we are interested in the performance limits as $n \to \infty$, this technicality does not influence our results and will therefore be ignored in the sequel.

$$(\mathsf{A}_1^{\mathsf{B}}, \mathsf{A}_2^{\mathsf{B}}) \mapsto (\mathsf{B}_1^{\mathsf{B}}, \mathsf{B}_2^{\mathsf{B}}), \tag{69}$$

*and is bijective.*

*Proof.* See Appendix A-B. □

**Proposition 1.** *The linear-feedback capacity region of the MIMO Gaussian BC with channel matrices $\mathsf{H}_1$ and $\mathsf{H}_2$ under a sum-power constraint $P$ is:*

$$\begin{aligned}\mathcal{C}_{\mathrm{BC}}^{\mathrm{linfb}}(\mathsf{H}_1, \mathsf{H}_2; P) \\ = \mathrm{cl}\left( \bigcup_{\eta, \mathsf{B}_1^{\mathsf{B}}, \mathsf{B}_2^{\mathsf{B}}} \frac{1}{\eta} \mathcal{R}_{\mathrm{BC}}\left(\eta, \mathsf{B}_1^{\mathsf{B}}, \mathsf{B}_2^{\mathsf{B}}, \mathsf{H}_1^{\mathsf{B}}, \mathsf{H}_2^{\mathsf{B}}; P\right)\right)\end{aligned} \tag{70}$$

*where the union is over all positive integers $\eta$ and all strictly-lower block-triangular $(\eta\kappa)$-by-$(\eta\nu_1)$ and $(\eta\kappa)$-by-$(\eta\nu_2)$ matrices $\mathsf{B}_1^{\mathsf{B}}$ and $\mathsf{B}_2^{\mathsf{B}}$ with blocks of sizes $\kappa \times \nu_1$ and $\kappa \times \nu_2$ that satisfy*

$$\mathrm{tr}\left(\mathsf{B}_1^{\mathsf{B}}(\mathsf{B}_1^{\mathsf{B}})^{\mathsf{T}}\right) + \mathrm{tr}\left(\mathsf{B}_2^{\mathsf{B}}(\mathsf{B}_2^{\mathsf{B}})^{\mathsf{T}}\right) \leq \eta P. \tag{71}$$

*Proof.* We already concluded the achievability part (see the paragraphs preceding Note 2). The converse is proved in Section X-A. □

*2) A Class of Linear-Feedback Coding Schemes for the MAC:* We fix the blocklength $n$. The schemes in our class are parametrized by

- a positive integer $\eta$;
- $\nu_1$-by-$\kappa$ matrices $\{\mathsf{C}_{1,\tau,\ell}\}$, for $\ell = 2, \ldots, \eta$ and $\tau = 1, \ldots, \ell - 1$;
- $\nu_2$-by-$\kappa$ matrices $\{\mathsf{C}_{2,\tau,\ell}\}$, for $\ell = 2, \ldots, \eta$ and $\tau = 1, \ldots, \ell - 1$;
- two encoder mappings $f_1^{(n')}\colon \mathcal{M}_1 \to \mathcal{R}^{(\nu_1\eta)n'}$ and $f_2^{(n')}\colon \mathcal{M}_2 \to \mathcal{R}^{(\nu_2\eta)n'}$ that produce $n' \triangleq \lfloor \frac{n}{\eta} \rfloor$ codevectors (column-vectors) of sizes $\nu_1\eta$ and $\nu_2\eta$, respectively; and
- a decoder mapping $g^{(n')}\colon \mathbb{R}^{(\kappa\eta)n'} \to \mathcal{M}_1 \times \mathcal{M}_2$ that decodes a block of $n'$ output vectors (column-vectors) of length $\kappa\eta$.

Similar to the BC schemes, the parameter $\eta$ characterizes the length of the subblocks in our scheme. That means, the total blocklength $n$ is again divided into $n'$ subblocks of equal length $\eta$. The matrices $\{\mathsf{C}_{1,\tau,\ell}\}$ and $\{\mathsf{C}_{2,\tau,\ell}\}$ describe the inner code that is used within each of the $n'$ subblocks of length $\eta$. Finally, the parameters $f_1^{(n')}, f_2^{(n')}, g^{(n')}$ describe the outer code that is applied to code over the $n'$ subblocks without using the feedback.

Before describing how the inner code works and how to design the outer code, we need to introduce some notation. Let, for $i \in \{1, 2\}$,

$$\mathbf{X}_i \triangleq \left(\mathbf{X}_{i,1}^{\mathsf{T}}, \quad \ldots, \quad \mathbf{X}_{i,\eta}^{\mathsf{T}}\right)^{\mathsf{T}}, \tag{72}$$

denote the $\eta\nu_i$-dimensional column-vector that is obtained by stacking the first $\eta$ channel input vectors $\mathbf{X}_{i,1}, \ldots, \mathbf{X}_{i,\eta}$ (which are all $\nu_i$-dimensional column-vectors) on top of each other. Similarly, let

$$\mathbf{Y} \triangleq \left(\mathbf{Y}_1^{\mathsf{T}}, \quad \ldots, \quad \mathbf{Y}_\eta^{\mathsf{T}}\right)^{\mathsf{T}} \tag{73}$$

$$\mathbf{Z} \triangleq \left(\mathbf{Z}_1^{\mathsf{T}}, \quad \ldots, \quad \mathbf{Z}_\eta^{\mathsf{T}}\right)^{\mathsf{T}} \tag{74}$$

denote the $\eta\kappa$-dimensional column vectors that are obtained by stacking the first $\eta$ noise vectors $\mathbf{Z}_1, \ldots, \mathbf{Z}_\eta$ and channel output vectors $\mathbf{Y}_1, \ldots, \mathbf{Y}_\eta$ on top of each other. Using the definition of the block channel matrices in (58), we can summarize the input-output relation for the first block of $\eta$ channel uses as

$$\mathbf{Y} = (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathbf{X}_1 + (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathbf{X}_2 + \mathbf{Z}. \tag{75}$$

Define the following matrices that we will use to describe the inner code. Let

$$\mathsf{C}_i^{\mathsf{B}} \triangleq \begin{bmatrix} 0 & & \cdots & & 0 \\ \mathsf{C}_{i,1,2} & 0 & & & \\ \mathsf{C}_{i,1,3} & \mathsf{C}_{i,2,3} & 0 & & \\ \vdots & & & \ddots & \\ \mathsf{C}_{i,1,\eta} & \mathsf{C}_{i,2,\eta} & \cdots & \mathsf{C}_{i,(\eta-1),\eta} & 0 \end{bmatrix}, \quad i \in \{1, 2\}, \tag{76}$$

where here $0$ denotes an $\nu_i$-by-$\kappa$ zero matrix, let

$$\mathsf{D}_i^{\mathsf{B}} \triangleq \mathsf{C}_i^{\mathsf{B}}\left(\mathsf{I} - (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_1^{\mathsf{B}} - (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_2^{\mathsf{B}}\right)^{-1}, \ i \in \{1, 2\}, \tag{77}$$

and let $\mathsf{Q}_1$ be the unique positive square root of the (positive-definite) $\nu_1\eta$-by-$\nu_1\eta$ matrix

$$\mathsf{M}_1 \triangleq (\mathsf{I} + \mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{I} + \mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}) + (\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}) \tag{78a}$$

and $\mathsf{Q}_2$ be the unique positive square root of the (positive-definite) $\nu_2\eta$-by-$\nu_2\eta$ matrix

$$\mathsf{M}_2 \triangleq (\mathsf{I} + \mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{I} + \mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}) + (\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}). \tag{78b}$$

We can now describe how the inner code—specified by the matrices $\{\mathsf{C}_{1,\tau,\ell}\}$ and $\{\mathsf{C}_{2,\tau,\ell}\}$—transforms the first block of $\eta$ channel uses into a single channel use of the new MIMO MAC. The transformation of the other blocks is done in a similar way. Let $\mathbf{U}_1$ and $\mathbf{U}_2$ denote the $\eta\nu_1$ and $\eta\nu_2$-length codevectors (column-vectors) produced by $f_1^{(n')}$ and $f_2^{(n')}$ for this first block.

Transmitter $i \in \{1, 2\}$, chooses its $\eta$ channel inputs in the first block as

$$\mathbf{X}_i = \mathsf{Q}_i^{-1}\mathbf{U}_i + \mathsf{C}_i^{\mathsf{B}}\mathbf{Y}. \tag{79}$$

(This describes a general linear-feedback scheme. The only reason for precoding the codeword vector $\mathbf{U}_i$ by the matrix $\mathsf{Q}_i^{-1}$ is to simplify the statement of the power constraint in Lemma 1 ahead.) By (75), the outputs $\mathbf{Y}$ in this first block can be written as:

$$\begin{aligned}\mathbf{Y} = (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{Q}_1^{-1}\mathbf{U}_1 + (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{Q}_2^{-1}\mathbf{U}_2 \\ + ((\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_1^{\mathsf{B}} + (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_2^{\mathsf{B}})\mathbf{Y} + \mathbf{Z}.\end{aligned} \tag{80}$$

Subtracting $((\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_1^{\mathsf{B}} + (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_2^{\mathsf{B}})\mathbf{Y}$ from both sides of (80) and then multiplying both sides from the left by the matrix $(\mathsf{I} - (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_1^{\mathsf{B}} - (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_2^{\mathsf{B}})^{-1}$, one sees that (80) is equivalent to

$$\begin{aligned}\mathbf{Y} = (\mathsf{I} - (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_1^{\mathsf{B}} - (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_2^{\mathsf{B}})^{-1} \\ \cdot \left((\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{Q}_1^{-1}\mathbf{U}_1 + (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{Q}_2^{-1}\mathbf{U}_2 + \mathbf{Z}\right).\end{aligned} \tag{81}$$

Using (77), we can now rewrite the inputs in (79) in a more explicit form:

$$\mathbf{X}_i = \mathsf{Q}_i^{-1}\mathbf{U}_i + \mathsf{D}_i^{\mathsf{B}}\big((\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{Q}_1^{-1}\mathbf{U}_1 + (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{Q}_2^{-1}\mathbf{U}_2 + \mathbf{Z}\big), \quad (82)$$

which allows us to rephrase the power constraints on the inputs as an equivalent power constraint on the codewords of the outer code $\mathbf{U}_1$ and $\mathbf{U}_2$:

**Lemma 1.** *In our scheme, the channel inputs* $\{\mathbf{X}_{1,t}\}_{t=1}^n$ *and* $\{\mathbf{X}_{2,t}\}_{t=1}^n$ *to the original MIMO Gaussian MAC satisfy the total average block-power constraint* (19) *whenever*

$$\text{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right) + \text{tr}\left(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}\right) \le \eta P \quad (83)$$

*and the codevectors produced by* $f_1^{(n')}$ *and* $f_2^{(n')}$ *are total average block-power constrained to power*

$$\eta P - \text{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right) - \text{tr}\left(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}\right). \quad (84)$$

*Proof.* See Section X-B. □

**Definition 3.** *Let* $\mathcal{R}_{\text{MAC}}\left(\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}, (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}, (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}; P\right)$ *denote the capacity region of the MIMO Gaussian MAC* without feedback *in* (81) *under average block-power constraint* (84) *on the input vectors* $\mathbf{U}_1$ *and* $\mathbf{U}_2$.

The outer code $\{f_1^{(n')}, f_2^{(n')}, g^{(n')}\}$ is designed so that it achieves the non-feedback capacity of the new MIMO Gaussian MAC in (81) under average input-power constraint $\eta P - \text{tr}(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}) - \text{tr}(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}})$.

Combining all this, we conclude that over the original MIMO Gaussian MAC our overall scheme (consisting of inner and outer code) achieves the rate region $\mathcal{R}_{\text{MAC}}\left(\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}, (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}, (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}; P\right)$ scaled by a factor $\frac{1}{\eta}$. In view of the following Note 3, it thus follows that our schemes achieve the rate region in (86).

**Note 3.** *Let* $\tilde{\mathcal{T}} \triangleq \tilde{\mathcal{T}}_1 \times \tilde{\mathcal{T}}_2$ *where, for* $i \in \{1,2\}$, $\tilde{\mathcal{T}}_i$ *denotes the set of strictly-lower block-triangular matrices with block matrices of size* $\nu_i \times \kappa$. *The mapping described in* (77) *is of the form*

$$\begin{aligned} \tilde{\omega}\colon \quad \tilde{\mathcal{T}} \quad &\to \quad \tilde{\mathcal{T}} \\ (\mathsf{C}_1^{\mathsf{B}}, \mathsf{C}_2^{\mathsf{B}}) &\mapsto (\mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}), \end{aligned} \quad (85)$$

*and is bijective.*

*Proof.* Analogous to the proof of Note 2. Details omitted. □

**Proposition 2.** *The linear-feedback capacity of the Gaussian MIMO MAC with channel matrices* $\mathsf{H}_1^{\mathsf{T}}$ *and* $\mathsf{H}_2^{\mathsf{T}}$ *under a sum-power constraint* $P$ *satisfies*

$$\begin{aligned} &\mathcal{C}_{\text{MAC}}^{\text{linfb}}\left(\mathsf{H}_1^{\mathsf{T}}, \mathsf{H}_2^{\mathsf{T}}; P\right) \\ &\quad = \text{cl}\left(\bigcup_{\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}} \frac{1}{\eta}\mathcal{R}_{\text{MAC}}\left(\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}, (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}, (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}; P\right)\right) \end{aligned} \quad (86)$$

*where the union is over all positive integers* $\eta$ *and all strictly-lower block-triangular* $(\eta\nu_1)$*-by-*$(\eta\kappa)$ *and* $(\eta\nu_2)$*-by-*$(\eta\kappa)$ *matrices* $\mathsf{D}_1^{\mathsf{B}}$ *and* $\mathsf{D}_2^{\mathsf{B}}$ *with blocks of sizes* $\nu_1 \times \kappa$ *and* $\nu_2 \times \kappa$ *that satisfy*

$$\text{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right) + \text{tr}\left(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}\right) \le \eta P. \quad (87)$$

*Proof.* The achievability follows from the considerations above. The converse is proved in Section X-C. □

*3) Dual Linear-Feedback Coding Schemes for MAC and BC:* Recall that for any matrix $\mathsf{M}$, we defined $\bar{\mathsf{M}} \triangleq \mathsf{E}\mathsf{M}^{\mathsf{T}}\mathsf{E}$, where $\mathsf{E}$ denotes the exchange matrix with appropriate dimensions.

**Proposition 3.** *If*

$$\mathsf{B}_i^{\mathsf{B}} = \bar{\mathsf{D}}_i^{\mathsf{B}}, \quad i \in \{1,2\}, \quad (88)$$

*then the following two regions coincide:*[2]

$$\begin{aligned} &\mathcal{R}_{\text{BC}}\left(\eta, \mathsf{B}_1^{\mathsf{B}}, \mathsf{B}_2^{\mathsf{B}}, \mathsf{H}_1^{\mathsf{B}}, \mathsf{H}_2^{\mathsf{B}}; P\right) \\ &\quad = \mathcal{R}_{\text{MAC}}\left(\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}, \bar{\mathsf{H}}_1^{\mathsf{B}}, \bar{\mathsf{H}}_2^{\mathsf{B}}; P\right). \end{aligned} \quad (89)$$

*Proof.* See Section X-D. □

When $\{\mathsf{A}_i^{\mathsf{B}}, \mathsf{B}_i^{\mathsf{B}}\}_{i=1}^2$ satisfy (65) and $\{\mathsf{C}_i^{\mathsf{B}}, \mathsf{D}_i^{\mathsf{B}}\}_{i=1}^2$ satisfy (77), Condition (88) is equivalent to

$$\mathsf{A}_i^{\mathsf{B}} = \bar{\mathsf{C}}_i^{\mathsf{B}}. \quad (90)$$

Combining Proposition 3, Equality (90), and Remark 1 we obtain:

**Corollary 4.** *Consider a MIMO Gaussian BC with channel matrices* $(\mathsf{H}_1, \mathsf{H}_2)$ *and its dual MAC with channel matrices* $(\mathsf{H}_1^{\mathsf{T}}, \mathsf{H}_2^{\mathsf{T}})$. *Fix the MAC-scheme parameters* $\eta$, $\{\mathsf{C}_{1,\tau,\ell}\}$, $\{\mathsf{C}_{2,\tau,\ell}\}$, *and let* $f_1^{(n')}$, $f_2^{(n')}$, $g^{(n')}$ *be an optimal outer code for these choices. Choose now the BC-scheme parameters*

$$\mathsf{A}_{i,\tau,\ell} = \bar{\mathsf{C}}_{i,\eta-\tau,\eta-\ell+2}, \quad (91)$$

*and an optimal outer code* $f^{(n')}$, $g_1^{(n')}$, *and* $g_2^{(n')}$ *as described in [28]. Then, our MAC and BC-schemes achieve the same rate regions:*

$$\begin{aligned} &\mathcal{R}_{\text{BC}}\left(\eta, \mathsf{B}_1^{\mathsf{B}}, \mathsf{B}_2^{\mathsf{B}}, \mathsf{H}_1^{\mathsf{B}}, \mathsf{H}_2^{\mathsf{B}}; P\right) \\ &\quad = \mathcal{R}_{\text{MAC}}\left(\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}, (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}, (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}; P\right). \end{aligned} \quad (92)$$

*In the SISO case, all conditions* (91) *are summarized by*

$$\mathsf{A}_i = \bar{\mathsf{C}}_i. \quad (93)$$

*In particular, when these matrices are Toeplitz, we have*

$$\mathsf{A}_i = \mathsf{C}_i. \quad (94)$$

*Proof.* See Section X-E. □

## VII. Constructive Sum-Rate Optimal Linear-Feedback Coding Scheme for the SISO Gaussian BC

Our duality result in Corollary 4 suggests to use the same feedback-matrices $\mathsf{A}_1$ and $\mathsf{A}_2$ on the two-user SISO BC as Ozarow [20] used on the MAC. Our duality results however do not give us an explicit construction of the codeword $\mathbf{U}$.

Here, we describe a constructive coding scheme for the two-user SISO Gaussian BC that achieves the linear-feedback sum-capacity. In fact, we show that a simple rearrangement of

[2] By Note 1 matrices $\bar{\mathsf{H}}_1^{\mathsf{B}}$ and $\bar{\mathsf{H}}_2^{\mathsf{B}}$ can be understood both as $\mathsf{I}_\eta \otimes \bar{\mathsf{H}}_1$ and $\mathsf{I}_\eta \otimes \bar{\mathsf{H}}_2$ or as $\overline{(\mathsf{I}_\eta \otimes \mathsf{H}_1)}$ and $\overline{(\mathsf{I}_\eta \otimes \mathsf{H}_2)}$.

encoders and decoders of Ozarow's sum-capacity achieving coding scheme for the MAC [20] is also linear-feedback sum-rate optimal on the two-user SISO Gaussian BC. We propose that (see Figures 6 and 7):

- the single BC-Transmitter implements both Ozarow's MAC-encoders 1 and 2, and then sends the sum of the symbols produced by these encoders over the channel;
- 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 $M_2$.

(This is possible because in Ozarow's scheme the receiver decodes the two messages $M_1$ and $M_2$ separately, see Figure 6.)

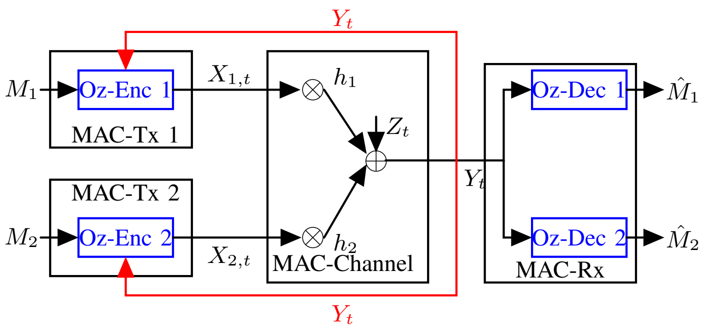


Fig. 6. Ozarow's MAC-scheme

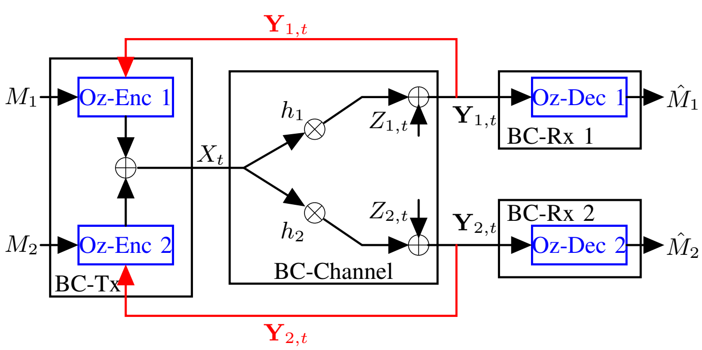


Fig. 7. New BC-scheme

More specifically, we propose the following constructive scheme for the two-user SISO Gaussian BC with independent noises.

Choose nonnegative powers $P_1$ and $P_2$ that sum up to $P$. Before transmission starts, map each message $M_i$, for $i \in \{1,2\}$, into the real-valued message point

$$\Theta_i(M_i) \triangleq -(M_i-1)\Delta_i + \sqrt{P_i}, \tag{95}$$

where

$$\Delta_i \triangleq \frac{2\sqrt{P_i}}{\lfloor 2^{nR_i} \rfloor}. \tag{96}$$

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,

$$t=-2: \qquad X_{-2} = \Theta_2(M_2), \tag{97a}$$
$$t=-1: \qquad X_{-1} = \Theta_1(M_1). \tag{97b}$$

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

$$\Xi_i \triangleq \sqrt{1-\rho^\star} Z_{i,-i} + \sqrt{\rho^\star} T_0, \tag{98}$$

where we write $\rho^\star$ as a short-hand notation for $\rho^\star(h_1,h_2;P_1,P_2)$ and where $T_0$ denotes a standard Gaussian random variable that acts as common randomness known at all terminals and is independent of the messages $M_1$ and $M_2$.

Through the feedback, the transmitter learns $Z_{1,-1}$ and $Z_{2,-2}$. It can thus compute $\Xi_1$ and $\Xi_2$, which it describes to receivers 1 and 2 during the remaining channel uses $0,\ldots,n-1$. Since

$$\sqrt{1-\rho^\star} Y_{i,-i} + \sqrt{\rho^\star} T_0 = \sqrt{1-\rho^\star} h_i \Theta_i(M_i) + \Xi_i, \tag{99}$$

for any estimate of $\Xi_i$, Receiver $i$ immediately obtains an estimate of $\Theta_i(M_i)$. To describe $\Xi_i$ to Receiver 1, the transmitter produces inputs

$$X_t = X_{1,t} + X_{2,t}, \qquad t \in \{0,\ldots,n-1\}, \tag{100}$$

where for $i \in \{1,2\}$:

$$X_{i,0} = \sqrt{P_i}\Xi_i \tag{101a}$$
$$X_{i,t} = \gamma_i \left( X_{i,t-1} - \delta_i Y_{i,t-1} \right), \quad t = 1,\ldots,n-1. \tag{101b}$$

We choose

$$\gamma_1 \triangleq \sqrt{1 + h_1^2 P_1 (1-(\rho^\star)^2)}, \tag{102a}$$
$$\gamma_2 \triangleq -\sqrt{1 + h_2^2 P_2 (1-(\rho^\star)^2)}. \tag{102b}$$

and

$$\delta_1 \triangleq \frac{h_1 P_1 + \rho^\star h_2 \sqrt{P_1 P_2}}{h_1^2 P_1 + h_2^2 P_2 + 2\rho^\star h_1 h_2 \sqrt{P_1 P_2} + 1}, \tag{103a}$$
$$\delta_2 \triangleq \frac{h_2 P_2 + \rho^\star h_1 \sqrt{P_1 P_2}}{h_1^2 P_1 + h_2^2 P_2 + 2\rho^\star h_1 h_2 \sqrt{P_1 P_2} + 1}. \tag{103b}$$

After reception of output symbols $Y_{i,-i},\ldots,Y_{i,n-1}$, for each $i \in \{1,2\}$, Receiver $i$ calculates the estimate $\hat{\Xi}_i^{(n-1)}$ of $\Xi_i$ based on $Y_{i,0},\ldots,Y_{i,n-1}$:

$$\hat{\Xi}_i^{(n-1)} \triangleq \delta_i (\sqrt{P_i})^{-1} \sum_{\tau=0}^{n-1} \gamma_i^{1-\tau} Y_{i,\tau}, \tag{104}$$

and forms its estimate of message point $\Theta_i$:

$$\begin{aligned} \hat{\Theta}_i &\triangleq \frac{1}{h_i} \left( Y_{i,-i} + \frac{\sqrt{\rho^\star}}{\sqrt{1-\rho^\star}} T_0 - \frac{1}{\sqrt{1-\rho^\star}} \hat{\Xi}_i^{(n-1)} \right) \\ &= \Theta_i(M_i) + \frac{1}{h_i \sqrt{1-\rho^\star}} \left( \Xi_i - \hat{\Xi}_i^{(n-1)} \right). \end{aligned} \tag{105}$$

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

$$\hat{M}_i = \underset{m_i \in \{1,\ldots,\lfloor 2^{nR_i} \rfloor\}}{\operatorname{argmin}} \left| \Theta_i(m_i) - \hat{\Theta}_i \right|. \tag{106}$$

In Appendix B we show that in the limit as $n \to \infty$ this linear-feedback coding scheme for the BC has average block-power[3] tending to $P$, and that it achieves the same rates

[3]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 Appendix B such a scaling does not change the set of achievable rates.

as Ozarow's MAC scheme [20], i.e., all rate-pairs $(R_1, R_2)$ satisfying

$$0 \leq R_1 < \frac{1}{2}\log\left(1 + h_1^2 P_1(1-\rho^{\star 2})\right), \tag{107a}$$

$$0 \leq R_2 < \frac{1}{2}\log\left(1 + h_2^2 P_2(1-\rho^{\star 2})\right). \tag{107b}$$

Analyzing the presented BC scheme directly seems difficult and cumbersome. We prove the desired bounds on the probability of error and the consumed power by drawing connections with Ozarow's sum-capacity achieving scheme for the two-user MAC with perfect feedback [20], which is much easier to analyze. The following remark will be key to establish the desired connections between the two schemes.

**Remark 3.** *If we replace everywhere the outputs $Y_{1,t}$ and $Y_{2,t}$ by $Y_t$ and ignore the sum in* (100)*, then above scheme describes Ozarow's sum-capacity achieving MAC coding scheme [20] with common randomness, where $X_{1,t}$ and $X_{2,t}$ denote the two transmitters' time-$t$ inputs.*

## VIII. Extension I: One-Sided Feedback

In this section we assume that there is feedback from only one side. That means, in the BC, there is feedback from only one of the two receivers, and in the MAC only one of the two transmitters has feedback.

### A. MIMO Gaussian BC with One-Sided Feedback

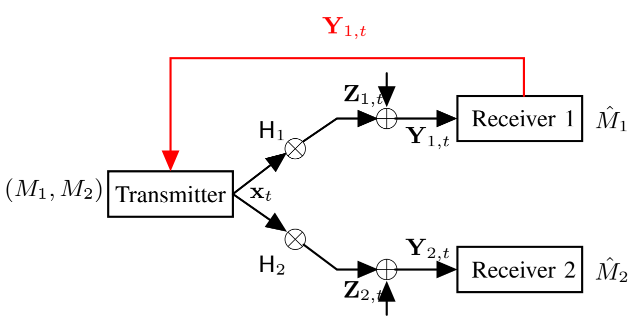


Fig. 8. Two-user MIMO Gaussian BC with one-sided perfect feedback.

Consider the Gaussian MIMO BC described in (6), Section III, but with feedback only from Receiver 1 (see Figure 8). The inputs are thus of the form

$$\mathbf{X}_t = \varphi_t^{(n)}(M_1, M_2, \mathbf{Y}_{1,1}, \ldots, \mathbf{Y}_{1,t-1}), \quad t \in \{1, \ldots, n\}. \tag{108}$$

We will again restrict to linear-feedback coding schemes where the inputs are generated as

$$\mathbf{X}_t = \mathbf{W}_t + \sum_{\tau=1}^{t-1} \mathsf{A}_{1,\tau,t}\mathbf{Y}_{1,\tau}, \quad t \in \{1, \ldots, n\}, \tag{109}$$

where $\mathbf{W}_t = \xi_t^{(n)}(M_1, M_2)$ for arbitrary functions $\xi_t^{(n)}$ and where $\mathsf{A}_{1,\tau,t}$ are arbitrary $\kappa$-by-$\nu_1$ matrices.

Decodings, power constraint, and the definitions of error probabilities and capacity regions are as described in Section III.

We denote the linear-feedback capacity region with one-sided feedback from Receiver 1 by $\mathcal{C}^{\text{linfb}}_{\text{BC,One}}(\mathsf{H}_1, \mathsf{H}_2; P)$. It is unknown to date. Inner bounds (i.e., achievable regions) have been proposed by Pillai [5] and Lapidoth, Steinberg and Wigger [17].

Analogous to Proposition 1, we can derive a multi-letter expression for the linear-feedback capacity region $\mathcal{C}^{\text{linfb}}_{\text{BC,One}}(\mathsf{H}_1, \mathsf{H}_2; P)$. Recall the definition of the regions $\mathcal{R}_{\text{BC}}$ in Definition 2.

**Proposition 4.**

$$\mathcal{C}^{\text{linfb}}_{\text{BC,One}}(\mathsf{H}_1, \mathsf{H}_2; P) = \text{cl}\left(\bigcup_{\eta, \mathsf{B}_1^{\mathsf{B}}} \frac{1}{\eta}\mathcal{R}_{\text{BC}}\left(\eta, \mathsf{B}_1^{\mathsf{B}}, 0, \mathsf{H}_1^{\mathsf{B}}, \mathsf{H}_2^{\mathsf{B}}; P\right)\right)$$

*where the union is over all positive integers $\eta$ and all strictly-lower block-triangular $(\eta\kappa)$-by-$(\eta\nu_1)$ matrices $\mathsf{B}_1^{\mathsf{B}}$ with blocks of sizes $\kappa \times \nu_1$ that satisfy* $\text{tr}\left(\mathsf{B}_1^{\mathsf{B}}(\mathsf{B}_1^{\mathsf{B}})^{\mathsf{T}}\right) \leq \eta P$*, and where* $0$ *denotes the $(\eta\kappa)$-by-$(\eta\nu_2)$ all-zero matrix.*

*Proof.* Analogous to the proof of Proposition 1, but where the matrix $\mathsf{B}_2^{\mathsf{B}}$ needs to be the $(\eta\kappa)$-by-$(\eta\nu_2)$ all-zero matrix, which by (65) implies that also $\mathsf{A}_{2,\tau,\ell} = 0$ for all $\tau, \ell$. □

### B. MIMO Gaussian MAC with One-Sided Feedback

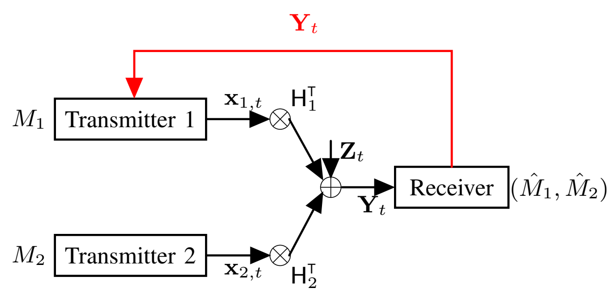


Fig. 9. Two-user MIMO Gaussian MAC with one-sided feedback.

Consider the Gaussian MIMO MAC described in (16), Section IV, but where only Transmitter 1 has feedback from the receiver (see Figure 9). The inputs are thus of the form

$$\mathbf{X}_{1,t} = \varphi_{1,t}^{(n)}(M_1, \mathbf{Y}_1, \ldots, \mathbf{Y}_{t-1}) \tag{110a}$$

$$\mathbf{X}_{2,t} = \varphi_{2,t}^{(n)}(M_2). \tag{110b}$$

We will again restrict to linear-feedback coding schemes where the inputs at Transmitter 1 are generated as

$$\mathbf{X}_{1,t} = \mathbf{W}_{1,t} + \sum_{\tau=1}^{t-1} \mathsf{C}_{1,\tau,t}\mathbf{Y}_\tau, \tag{111}$$

where $\mathbf{W}_{1,t}$ is a vector that only depends on the message $M_1$ but not on the feedback, $\mathbf{W}_{1,t} = \xi_{1,t}^{(n)}(M_1)$ for arbitrary functions $\xi_{1,t}^{(n)}$.

Decoding, power constraint, and the definitions of error probabilities and capacity regions are as described in Section IV.

We denote the linear-feedback capacity region of the Gaussian MIMO MAC with one-sided feedback to Transmitter 1 by $\mathcal{C}^{\text{linfb}}_{\text{MAC,One}}\left(\mathsf{H}_1^{\mathsf{T}},\mathsf{H}_2^{\mathsf{T}};P\right)$. It is unknown to date. Inner bounds (i.e., achievable regions) were presented in [6], [9], [18], [29].

Analogous to Proposition 2, we can derive a multi-letter expression for the linear-feedback capacity region $\mathcal{C}^{\text{linfb}}_{\text{MAC,One}}\left(\mathsf{H}_1^{\mathsf{T}},\mathsf{H}_2^{\mathsf{T}};P\right)$. Recall the definition of the regions $\mathcal{R}_{\text{MAC}}$ in Definition 3.

**Proposition 5.**

$$\begin{aligned}&\mathcal{C}^{\text{linfb}}_{\text{MAC,One}}\left(\mathsf{H}_1^{\mathsf{T}},\mathsf{H}_2^{\mathsf{T}};P\right)\\&\quad=\text{cl}\left(\bigcup_{\eta,\mathsf{D}_1^{\mathsf{B}}}\frac{1}{\eta}\mathcal{R}_{\text{MAC}}\left(\eta,\mathsf{D}_1^{\mathsf{B}},\mathsf{0},(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}},(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}};P\right)\right)\end{aligned}$$

*where the union is over all positive integers $\eta$ and all strictly-lower block-triangular $(\eta\nu_1)$-by-$(\eta\kappa)$ matrices $\mathsf{D}_1^{\mathsf{B}}$ with block sizes $\nu_1\times\kappa$ that satisfy $\operatorname{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right)\leq\eta P$, and where $\mathsf{0}$ denotes the $(\eta\nu_2)$-by-$(\eta\kappa)$ all-zero matrix.*

*Proof.* Analogous to the proof of Proposition 2, but where the matrix $\mathsf{D}_2^{\mathsf{B}}$ needs to be the $(\eta\nu_2)$-by-$(\eta\kappa)$ all-zero matrix which implies that $\{\mathsf{C}_{2,\tau,\ell}\}$ are all equal to the $\nu_2$-by-$\kappa$ all-zero matrix. □

### C. Duality Result

**Theorem 3.**

$$\mathcal{C}^{\text{linfb}}_{\text{BC,One}}\left(\mathsf{H}_1,\mathsf{H}_2;P\right)=\mathcal{C}^{\text{linfb}}_{\text{MAC,One}}\left(\mathsf{H}_1^{\mathsf{T}},\mathsf{H}_2^{\mathsf{T}};P\right).\tag{112}$$

*Proof.* Follows from Propositions 4 and 5 and Remark 1 which continues to hold in the one-sided feedback setup, and because $\bar{\mathsf{0}}=\mathsf{0}$ and Propositon 3 imply the following:

If $\mathsf{B}_1^{\mathsf{B}}=\bar{\mathsf{D}}_1^{\mathsf{B}}$, then

$$\mathcal{R}_{\text{BC}}\left(\eta,\mathsf{B}_1^{\mathsf{B}},\mathsf{0},\mathsf{H}_1^{\mathsf{B}},\mathsf{H}_2^{\mathsf{B}};P\right)=\mathcal{R}_{\text{MAC}}\left(\eta,\mathsf{D}_1^{\mathsf{B}},\mathsf{0},\bar{\mathsf{H}}_1^{\mathsf{B}},\bar{\mathsf{H}}_2^{\mathsf{B}};P\right).\tag{113}$$

□

## IX. Extension II: $K\geq 2$ Users

In this section we consider the $K$-user Gaussian BC and MAC with feedback, when $K\geq 2$.

### A. $K\geq 2$-User MIMO Gaussian BC with Feedback

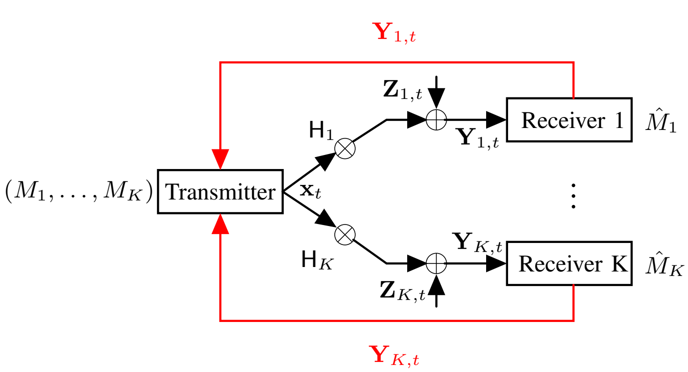


Fig. 10. $K$-user MIMO Gaussian BC with perfect feedback.

We consider the $K\geq 2$-receiver Gaussian BC with perfect output-feedback depicted in Figure 10. At each time $t\in\mathbb{N}$, if $\mathbf{x}_t$ denotes the real vector-valued input symbol sent by the transmitter, Receiver $i$, for $i\in\{1,\ldots,K\}$, observes the real vector-valued channel output

$$\mathbf{Y}_{i,t}=\mathsf{H}_i\mathbf{x}_t+\mathbf{Z}_{i,t},\tag{114}$$

where $\mathsf{H}_i$ is a deterministic nonzero real $\nu_i$-by-$\kappa$ channel matrix known to transmitter and receivers, and the sequence of noises $\{(\mathbf{Z}_{1,t},\ldots,\mathbf{Z}_{K,t})\}_{t=1}^n$ is a sequence of i.i.d. centered Gaussian random vectors, each of identity covariance matrix.

We will again restrict to linear-feedback schemes where the inputs, at each time $t\in\{1,\ldots,n\}$, are generated as

$$\mathbf{X}_t=\mathbf{W}_t+\sum_{i=1}^{K}\sum_{\tau=1}^{t-1}\mathsf{A}_{i,\tau,t}\mathbf{Y}_{i,\tau},\tag{115}$$

where $\mathbf{W}_t=\xi_t^{(n)}(M_1,\ldots,M_K)$, for an arbitrary function $\xi_t^{(n)}$, is thus a vector that only depends on the messages but not on the feedback.

Decodings, power constraint, and the definitions of error probabilities and capacity regions are similar to Section III when we consider $K$ instead of two users.

We denote the linear-feedback capacity region for this setup by $\mathcal{C}^{\text{linfb}}_{\text{BC}}(\mathsf{H}_1,\ldots,\mathsf{H}_K;P)$. It is unknown to date. Achievable regions are presented in [1] and [21].

Analogous to the definition of the regions $\mathcal{R}_{\text{BC}}$ in Definition 2, we define $\mathcal{R}_{\text{BC}}\left(\eta,\mathsf{B}_1^{\mathsf{B}},\ldots,\mathsf{B}_K^{\mathsf{B}},\mathsf{H}_1^{\mathsf{B}},\ldots,\mathsf{H}_K^{\mathsf{B}};P\right)$ as the capacity region of the MIMO BC

$$\mathbf{Y}_i=\mathsf{H}_i^{\mathsf{B}}\mathbf{U}+\mathsf{H}_i^{\mathsf{B}}\left(\sum_{j=1}^{K}\mathsf{B}_j^{\mathsf{B}}\mathbf{Z}_j\right)+\mathbf{Z}_i,\;\; i\in\{1,\ldots,K\},\tag{116}$$

when the channel inputs $\mathbf{U}$ is average block-power constrained to

$$\eta P-\sum_{j=1}^{K}\operatorname{tr}\left(\mathsf{B}_j^{\mathsf{B}}(\mathsf{B}_j^{\mathsf{B}})^{\mathsf{T}}\right).\tag{117}$$

**Proposition 6.**

$$\begin{aligned}&\mathcal{C}^{\text{linfb}}_{\text{BC}}(\mathsf{H}_1,\ldots,\mathsf{H}_K;P)\\&\quad=\text{cl}\left(\bigcup_{\eta,\mathsf{B}_1^{\mathsf{B}},\ldots,\mathsf{B}_K^{\mathsf{B}}}\frac{1}{\eta}\mathcal{R}_{\text{BC}}\left(\eta,\mathsf{B}_1^{\mathsf{B}},\ldots,\mathsf{B}_K^{\mathsf{B}},\mathsf{H}_1^{\mathsf{B}},\ldots,\mathsf{H}_K^{\mathsf{B}};P\right)\right)\end{aligned}$$

*where the union is over all positive integers $\eta$ and all strictly-lower block-triangular $(\eta\kappa)$-by-$(\eta\nu_i)$ matrices $\mathsf{B}_i^{\mathsf{B}}$ with blocks of sizes $\kappa\times\nu_i$, for $i\in\{1,\ldots,K\}$, that satisfy $\sum_{j=1}^{K}\operatorname{tr}\left(\mathsf{B}_j^{\mathsf{B}}(\mathsf{B}_j^{\mathsf{B}})^{\mathsf{T}}\right)\leq\eta P$.*

*Proof.* Similar to the proof of Proposition 1 if the linear-feedback coding schemes described in Section VI-B1 and the converse are modified so as to allow for an arbitrary number $K\geq 2$ of users. Details omitted. □

### B. $K\geq 2$-User MIMO Gaussian MAC with Feedback

We consider the $K\geq 2$-transmitter Gaussian MAC with perfect output-feedback depicted in Figure 11. At each time

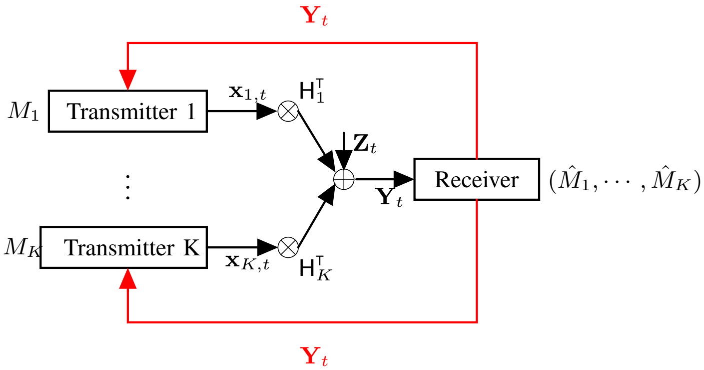


Fig. 11. $K$-user MIMO Gaussian MAC with perfect feedback.

$t \in \mathbb{N}$, if $\mathbf{x}_{i,t}$, for $i \in \{1, \ldots, K\}$ denotes the real vector-valued input symbol sent by Transmitter $i$, the receiver observes the real vector-valued channel output

$$\mathbf{Y}_t = \sum_{i=1}^{K} \mathsf{H}_i^\mathsf{T} \mathbf{x}_{i,t} + \mathbf{Z}_t, \tag{118}$$

where $\mathsf{H}_i$, for $i \in \{1, \ldots, K\}$, is a constant nonzero real $\nu_i$-by-$\kappa$ channel matrix and the sequence of noises $\{\mathbf{Z}_t\}_{t=1}^n$ is a sequence of i.i.d. centered Gaussian random vectors of identity covariance matrices.

We will again restrict to linear-feedback coding schemes where the inputs at Transmitter $i$, for $i \in \{1, \ldots, K\}$, are generated as

$$\mathbf{X}_{i,t} = \mathbf{W}_{i,t} + \sum_{i=1}^{K} \sum_{\tau=1}^{t-1} \mathsf{C}_{i,\tau,t} \mathbf{Y}_\tau, \tag{119}$$

where $\mathbf{W}_{i,t} = \xi_{i,t}^{(n)}(M_i)$ for an arbitrary function $\xi_{i,t}^{(n)}$ is thus a vector that only depends on the message $M_i$ but not on the feedback.

Decoding, power constraint, and the definitions of error probabilities and capacity regions are as described in Section IV extended to $K \geq 2$ users. The linear-feedback capacity region is denoted by $\mathcal{C}_{\mathrm{MAC}}^{\mathrm{linfb}}(\mathsf{H}_1^\mathsf{T}, \ldots, \mathsf{H}_K^\mathsf{T}; P)$. It is unknown when $K > 2$.

We will be specially interested in the SISO case ($\nu_i = \kappa = 1$) when the channel matrices $\mathsf{H}_1, \ldots, \mathsf{H}_K$ reduce to scalars $h_1, \ldots, h_K$. We denote the linear-feedback capacity region for this case by $\mathcal{C}_{\mathrm{MAC,SISO},\Sigma}^{\mathrm{linfb}}(h_1, \ldots, h_K; P)$. Also this SISO capacity region is unknown when $K > 2$. However, for equal channel coefficients $h_1 = \ldots, h_K = h$, the results by Kramer [16] and Ardestanizadeh *et al.* [2] combined with a symmetry argument as presented in Appendix A-A immediately yield:

$$\mathcal{C}_{\mathrm{MAC,SISO},\Sigma}^{\mathrm{linfb}}(h, \ldots, h; P) = \frac{1}{2} \log\left(1 + h^2 P \phi(K, h, P)\right), \tag{120}$$

where $\phi(K, h, P)$ is the unique solution in $[1, K]$ to the following equation in $\phi$:

$$\left(1 + h^2 P \phi\right)^{K-1} = \left(1 + \frac{h^2 P}{K} \phi(K - \phi)\right)^K. \tag{121}$$

Analogous to the definition of the regions $\mathcal{R}_{\mathrm{MAC}}$ in Definition 3, we define $\mathcal{R}_{\mathrm{MAC}}\left(\eta, \mathsf{D}_1^\mathsf{B}, \ldots, \mathsf{D}_K^\mathsf{B}, (\mathsf{H}_1^\mathsf{B})^\mathsf{T}, \ldots, (\mathsf{H}_K^\mathsf{B})^\mathsf{T}; P\right)$ as the capacity region of the MIMO MAC

$$\mathbf{Y} = \left(\mathsf{I} + \sum_{i=1}^{K} (\mathsf{H}_i^\mathsf{B})^\mathsf{T} \mathsf{D}_i^\mathsf{B}\right) \cdot \left(\sum_{i=1}^{K} (\mathsf{H}_i^\mathsf{B})^\mathsf{T} \mathsf{Q}_i^{-1} \mathbf{U}_i + \mathbf{Z}\right), \tag{122}$$

when the inputs $\mathbf{U}_1, \ldots, \mathbf{U}_K$ are average block-sum-power constrained to

$$\eta P - \sum_{j=1}^{K} \operatorname{tr}\left(\mathsf{D}_j^\mathsf{B} (\mathsf{D}_j^\mathsf{B})^\mathsf{T}\right), \tag{123}$$

where $\mathsf{Q}_i$, for $i \in \{1, \ldots, K\}$, is the unique positive square root of

$$\begin{aligned} \mathsf{M}_i = {} & (\mathsf{I} + \mathsf{D}_i^\mathsf{B} (\mathsf{H}_i^\mathsf{B})^\mathsf{T})^\mathsf{T} (\mathsf{I} + \mathsf{D}_i^\mathsf{B} (\mathsf{H}_i^\mathsf{B})^\mathsf{T}) \\ & + \sum_{j=1; j \neq i}^{K} (\mathsf{D}_j^\mathsf{B} (\mathsf{H}_j^\mathsf{B})^\mathsf{T})^\mathsf{T} (\mathsf{D}_j^\mathsf{B} (\mathsf{H}_j^\mathsf{B})^\mathsf{T}). \end{aligned} \tag{124}$$

**Proposition 7.**

$$\mathcal{C}_{\mathrm{MAC}}^{\mathrm{linfb}}(\mathsf{H}_1^\mathsf{T}, \ldots, \mathsf{H}_K^\mathsf{T}; P) = \mathrm{cl}\left(\bigcup_{\eta, \mathsf{D}_1^\mathsf{B}, \ldots, \mathsf{D}_K^\mathsf{B}} \frac{1}{\eta} \mathcal{R}_{\mathrm{MAC}}\left(\eta, \mathsf{D}_1^\mathsf{B}, \ldots, \mathsf{D}_K^\mathsf{B}, (\mathsf{H}_1^\mathsf{B})^\mathsf{T}, \ldots, (\mathsf{H}_K^\mathsf{B})^\mathsf{T}; P\right)\right),$$

*where the union is over all positive integers $\eta$ and all strictly-lower block-triangular $(\eta\nu_i)$-by-$(\eta\kappa)$ matrices $\mathsf{D}_i^\mathsf{B}$ of blocks with sizes $\nu_i \times \kappa$, for $i \in \{1, \ldots, K\}$, that satisfy $\sum_{j=1}^{K} \operatorname{tr}\left(\mathsf{D}_j^\mathsf{B} (\mathsf{D}_j^\mathsf{B})^\mathsf{T}\right) \leq \eta P$.*

*Proof.* Analogous to the proof of Proposition 2, but where the linear-feedback coding schemes described in Section VI-B2 and the converse need to be modified so as to allow for an arbitrary number $K \geq 2$ of users. Details omitted. □

### C. Duality Result

Our main result on duality can also be extended to the MIMO BC and MAC with more than two users.

**Theorem 4.** *The linear-feedback capacity regions of the $K \geq 2$-user MIMO Gaussian BC with channel matrices $\mathsf{H}_1, \ldots, \mathsf{H}_K$ under sum-power constraint $P$ and the $K \geq 2$-user MIMO Gaussian MAC with channel matrices $\mathsf{H}_1^\mathsf{T}, \ldots, \mathsf{H}_K^\mathsf{T}$ under sum-power constraint $P$ coincide:*

$$\mathcal{C}_{\mathrm{BC}}^{\mathrm{linfb}}(\mathsf{H}_1, \ldots, \mathsf{H}_K; P) = \mathcal{C}_{\mathrm{MAC}}^{\mathrm{linfb}}(\mathsf{H}_1^\mathsf{T}, \ldots, \mathsf{H}_K^\mathsf{T}; P). \tag{125}$$

*Proof.* The proof follows by Proposition 6 and 7, Remark 1 which continues to hold for this setup, and Proposition 3 which can be extended to $K \geq 2$ users since the non-feedback MAC-BC duality holds for $K \geq 2$ users [26].

Specializing this theorem to the SISO case under equal channel gains $h_1 = \ldots, h_K = h$, we obtain:

**Corollary 5.**

$$C_{\mathrm{BC,SISO},\Sigma}^{\mathrm{linfb}}(h, \ldots, h; P) = C_{\mathrm{MAC,SISO},\Sigma}^{\mathrm{linfb}}(h, \ldots, h; P) \tag{126}$$

*where a computable expression for $C_{\mathrm{MAC,SISO},\Sigma}^{\mathrm{linfb}}(h, \ldots, h; P)$ is given in* (120).

The achievability of the sum-rate in (126) for the $K$-user scalar Gaussian BC with equal channel gains was already established by the control-theory-inspired scheme in [1]. Our result here establishes that for the symmetric scalar Gaussian BC and arbitrary number of users $K > 2$ this scheme is indeed sum-rate optimal among all linear-feedback coding schemes.

## X. Proofs

### A. Converse Proof to Proposition 1

We wish to prove

$$\begin{aligned}&\mathcal{C}_{\text{BC}}^{\text{linfb}}(\mathsf{H}_1,\mathsf{H}_2;P)\\ &\quad\subseteq \text{cl}\left(\bigcup_{(\eta,\mathsf{B}_1^{\mathsf{B}},\mathsf{B}_2^{\mathsf{B}})}\frac{1}{\eta}\mathcal{R}_{\text{BC}}\left(\eta,\mathsf{B}_1^{\mathsf{B}},\mathsf{B}_2^{\mathsf{B}},\mathsf{H}_1^{\mathsf{B}},\mathsf{H}_2^{\mathsf{B}};P\right)\right). \qquad (127)\end{aligned}$$

Fix $(R_1,R_2)\in\mathcal{C}_{\text{BC}}^{\text{linfb}}(\mathsf{H}_1,\mathsf{H}_2;P)$ and for these rates and for each blocklength $n$ we fix encoding and decoding functions $\tilde{\xi}^{(n)},\phi_1^{(n)},\phi_2^{(n)}$ and linear-feedback matrices $\{\mathsf{B}_{i,\tau,\ell}^{(n)}\}$ such that the sequence of probabilities of error $P_{\text{e,BC}}^{(n)}\to 0$ as $n\to\infty$ and the power constraint (9) is satisfied for each $n$. (Thus, we use the form in (15) to describe the channel inputs.)

Applying Fano's inequality, we obtain that for each $i\in\{1,2\}$ and for each positive integer $n$,

$$nR_i \le I(M_i;\mathbf{Y}_i^{(n)})+\epsilon_n, \qquad (128)$$

where $\frac{\epsilon_n}{n}\to 0$ as $n\to\infty$ and where $\mathbf{Y}_i^{(n)}$ denotes the $n\nu_i$-dimensional column-vector that is obtained by stacking on top of each other all the $n$ vectors observed at Receiver $i$ when the blocklength-$n$ scheme is applied.

Letting $n\to\infty$, we have

$$R_i \le \varlimsup_{n\to\infty}\frac{1}{n}I(M_i;\mathbf{Y}_i^{(n)}),\quad i\in\{1,2\}. \qquad (129)$$

Since the RHS of (127) is closed, it suffices to prove that for all $\delta>0$, the pair $(R_1',R_2')$,

$$R_1' \triangleq \eta(R_1-\delta), \qquad (130a)$$
$$R_2' \triangleq \eta(R_2-\delta), \qquad (130b)$$

lies in $\mathcal{R}_{\text{BC}}\left(\eta,\mathsf{B}_1^{\mathsf{B}},\mathsf{B}_2^{\mathsf{B}},\mathsf{H}_1^{\mathsf{B}},\mathsf{H}_2^{\mathsf{B}};P\right)$ for some positive integer $\eta$ and strictly-lower block-triangular $\eta\kappa$-by-$\eta\nu_1$ and $\eta\kappa$-by-$\eta\nu_2$ matrices $\mathsf{B}_1^{\mathsf{B}}$ and $\mathsf{B}_2^{\mathsf{B}}$ of block sizes $\kappa\times\nu_1$ and $\kappa\times\nu_2$.

By (129) and (130), there exists a finite blocklength $n$ such that

$$R_1' \le I(M_1;\mathbf{Y}_1^{(n)}), \qquad (131a)$$
$$R_2' \le I(M_2;\mathbf{Y}_2^{(n)}). \qquad (131b)$$

In the sequel, let $n$ be so that (131) holds. Also, based on the parameters $\{\mathsf{B}_{i,\tau,\ell}^{(n)}\}$ of the blocklength-$n$ scheme, define

$$\mathsf{B}_i^{\mathsf{B}}=\begin{bmatrix}0 & & \cdots & & 0\\ \mathsf{B}_{i,1,2}^{(n)} & 0 & & & \\ \mathsf{B}_{i,1,3}^{(n)} & \mathsf{B}_{i,2,3}^{(n)} & 0 & & \\ \vdots & & & \ddots & \\ \mathsf{B}_{i,1,n}^{(n)} & \mathsf{B}_{i,2,n}^{(n)} & \cdots & \mathsf{B}_{i,(n-1),n}^{(n)} & 0\end{bmatrix},\quad i\in\{1,2\}. \qquad (132)$$

The corresponding channel outputs $\mathbf{Y}_1^{(n)}$ and $\mathbf{Y}_2^{(n)}$ are

$$\mathbf{Y}_1^{(n)}=\mathsf{H}_1^{\mathsf{B}}\tilde{\mathbf{W}}^{(n)}+(\mathsf{I}+\mathsf{H}_1^{\mathsf{B}}\mathsf{B}_1^{\mathsf{B}})\mathbf{Z}_1^{(n)}+\mathsf{H}_1^{\mathsf{B}}\mathsf{B}_2^{\mathsf{B}}\mathbf{Z}_2^{(n)}, \qquad (133a)$$
$$\mathbf{Y}_2^{(n)}=\mathsf{H}_2^{\mathsf{B}}\tilde{\mathbf{W}}^{(n)}+(\mathsf{I}+\mathsf{H}_2^{\mathsf{B}}\mathsf{B}_2^{\mathsf{B}})\mathbf{Z}_2^{(n)}+\mathsf{H}_2^{\mathsf{B}}\mathsf{B}_1^{\mathsf{B}}\mathbf{Z}_1^{(n)}, \qquad (133b)$$

where $\mathbf{Z}_1^{(n)}=\left(\mathbf{Z}_{1,1}^{\mathsf{T}},\ldots,\mathbf{Z}_{1,n}^{\mathsf{T}}\right)^{\mathsf{T}}$, $\mathbf{Z}_2^{(n)}=\left(\mathbf{Z}_{2,1}^{\mathsf{T}},\ldots,\mathbf{Z}_{2,n}^{\mathsf{T}}\right)^{\mathsf{T}}$, and $\tilde{\mathbf{W}}^{(n)}$ is the $n\kappa$-dimensional vector that is obtained when stacking on top of each other all the $n$ codevectors ($\kappa$-dimensional column-vectors) that are produced by the encoding function $\tilde{\xi}^{(n)}$. Notice that the power-constraint (9) is equivalent to requiring that

$$\mathsf{E}\left[\|\tilde{\mathbf{W}}^{(n)}\|^2\right]\le nP-\text{tr}\left(\mathsf{B}_1^{\mathsf{B}}(\mathsf{B}_1^{\mathsf{B}})^{\mathsf{T}}\right)-\text{tr}\left(\mathsf{B}_2^{\mathsf{B}}(\mathsf{B}_2^{\mathsf{B}})^{\mathsf{T}}\right). \qquad (134)$$

Let now $\eta=n$ and consider the BC in (133) where the transmitter is equipped with $\eta\kappa$ antennas and Receiver $i$ with $\eta\nu_i$ antennas, for $i\in\{1,2\}$, and where $\tilde{\mathbf{W}}^{(\eta)}$ denotes the $\eta\kappa$-dimensional input-vector. Recall that we denoted by $\mathcal{R}_{\text{BC}}(\eta,\mathsf{B}_1^{\mathsf{B}},\mathsf{B}_2^{\mathsf{B}},\mathsf{H}_1^{\mathsf{B}},\mathsf{H}_2^{\mathsf{B}};P)$ the capacity region of this channel under an expected average block-power constrained $(\eta P-\text{tr}\left(\mathsf{B}_1^{\mathsf{B}}(\mathsf{B}_1^{\mathsf{B}})^{\mathsf{T}}\right)-\text{tr}\left(\mathsf{B}_2^{\mathsf{B}}(\mathsf{B}_2^{\mathsf{B}})^{\mathsf{T}}\right))$ on the input $\tilde{\mathbf{W}}^{(\eta)}$. Using random coding and joint typicality decoding, it can be shown that the nonnegative rate-pair $(\tilde{R}_1,\tilde{R}_2)$ lies in this capacity region $\mathcal{R}_{\text{BC}}(\eta,\mathsf{B}_1^{\mathsf{B}},\mathsf{B}_2^{\mathsf{B}},\mathsf{H}_1^{\mathsf{B}},\mathsf{H}_2^{\mathsf{B}};P)$ if it satisfies

$$\tilde{R}_1\le I(\Theta_1;\mathbf{Y}_1^{(\eta)}) \qquad (135a)$$
$$\tilde{R}_2\le I(\Theta_2;\mathbf{Y}_2^{(\eta)}) \qquad (135b)$$

for some independent auxiliary random variables $\Theta_1$ and $\Theta_2$ and a choice of $\tilde{\mathbf{W}}^{(\eta)}$ such that $(\Theta_1,\Theta_2,\tilde{\mathbf{W}}^{(\eta)})$ are independent of $(\mathbf{Z}_1^{(\eta)},\mathbf{Z}_2^{(\eta)})$.

Specializing this last argument to $\Theta_1=M_1$ and $\Theta_2=M_2$, by (131), we conclude that for any $\delta>0$ the rate-pair $(R_1',R_2')$ defined in (130) lies in $\mathcal{R}_{\text{BC}}(\eta,\mathsf{B}_1^{\mathsf{B}},\mathsf{B}_2^{\mathsf{B}},\mathsf{H}_1^{\mathsf{B}},\mathsf{H}_2^{\mathsf{B}};P)$, which concludes the proof.

### B. Proof of Lemma 1

For the inputs transmitted in the first $\eta$-length block and described by (79), we have

$$\begin{aligned}&\mathsf{E}\left[\|\mathbf{X}_1\|^2\right]\\ &=\mathsf{E}\left[\mathbf{U}_1^{\mathsf{T}}\mathsf{Q}_1^{-\mathsf{T}}(\mathsf{I}+\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{I}+\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})\mathsf{Q}_1^{-1}\mathbf{U}_1\right]\\ &\quad+\mathsf{E}\left[\mathbf{U}_2^{\mathsf{T}}\mathsf{Q}_2^{-\mathsf{T}}(\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})\mathsf{Q}_2^{-1}\mathbf{U}_2\right]+\text{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right)\end{aligned}$$

and

$$\begin{aligned}&\mathsf{E}\left[\|\mathbf{X}_2\|^2\right]\\ &=\mathsf{E}\left[\mathbf{U}_2^{\mathsf{T}}\mathsf{Q}_2^{-\mathsf{T}}(\mathsf{I}+\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{I}+\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})\mathsf{Q}_2^{-1}\mathbf{U}_2\right]\\ &\quad+\mathsf{E}\left[\mathbf{U}_1^{\mathsf{T}}\mathsf{Q}_1^{-\mathsf{T}}(\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})\mathsf{Q}_1^{-1}\mathbf{U}_1\right]+\text{tr}\left(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}\right)\end{aligned}$$

By the definitions of $\mathsf{M}_1$ and $\mathsf{M}_2$ in (78), and because we defined $\mathsf{Q}_1$ and $\mathsf{Q}_2$ as being their positive square roots, we obtain the sequence of equalities in (136) on top of the next page.

From (136) we conclude that the input sequences $\{X_{1,t}\}_{t=1}^n$ and $\{X_{2,t}\}_{t=1}^n$ satisfy the average total input-power constraint $P$ whenever $\eta P-\text{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right)-\text{tr}\left(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}\right)\ge 0$ and the vectors $\mathbf{U}_1$ and $\mathbf{U}_2$ produced by the outer code satisfy the average total input-power constraint $\eta P-\text{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right)-\text{tr}\left(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}\right)$.

$$
\begin{aligned}
&\mathsf{E}\left[\|\mathbf{X}_1\|^2\right] + \mathsf{E}\left[\|\mathbf{X}_2\|^2\right] \\
&= \mathsf{E}\left[\mathbf{U}_1^{\mathsf{T}}\mathsf{Q}_1^{-\mathsf{T}}\Big((\mathsf{I}+\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{I}+\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}) + (\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})\Big)\mathsf{Q}_1^{-1}\mathbf{U}_1\right] \\
&\quad + \mathsf{E}\left[\mathbf{U}_2^{\mathsf{T}}\mathsf{Q}_2^{-\mathsf{T}}\Big((\mathsf{I}+\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{I}+\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}) + (\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})\Big)\mathsf{Q}_2^{-1}\mathbf{U}_2\right] + \mathrm{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right) + \mathrm{tr}\left(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}\right) \\
&= \mathsf{E}\left[\mathbf{U}_1^{\mathsf{T}}\mathsf{Q}_1^{-\mathsf{T}}\mathsf{M}_1\mathsf{Q}_1^{-1}\mathbf{U}_1\right] + \mathsf{E}\left[\mathbf{U}_2^{\mathsf{T}}\mathsf{Q}_2^{-\mathsf{T}}\mathsf{M}_2\mathsf{Q}_2^{-1}\mathbf{U}_2\right] + \mathrm{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right) + \mathrm{tr}\left(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}\right) \\
&= \mathsf{E}\left[\|\mathbf{U}_1\|^2\right] + \mathsf{E}\left[\|\mathbf{U}_2\|^2\right] + \mathrm{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right) + \mathrm{tr}\left(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}\right) \qquad (136)
\end{aligned}
$$

### C. Converse Proof to Proposition 2

We wish to prove

$$
\begin{aligned}
&\mathcal{C}_{\mathrm{MAC}}^{\mathrm{linfb}}\left(\mathsf{H}_1^{\mathsf{T}}, \mathsf{H}_2^{\mathsf{T}}; P\right) \\
&\quad \subseteq \mathrm{cl}\left(\bigcup_{(\eta,\mathsf{D}_1^{\mathsf{B}},\mathsf{D}_2^{\mathsf{B}})} \frac{1}{\eta}\mathcal{R}_{\mathrm{MAC}}\left(\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}, (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}, (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}; P\right)\right). \qquad (137)
\end{aligned}
$$

Fix $(R_1, R_2) \in \mathcal{C}_{\mathrm{MAC}}^{\mathrm{linfb}}\left(\mathsf{H}_1^{\mathsf{T}}, \mathsf{H}_2^{\mathsf{T}}; P\right)$ and for these rates and for each blocklength $n$ we fix encoding and decoding functions $\xi_1^{(n)}, \xi_2^{(n)}, \phi^{(n)}$, and linear-feedback matrices $\{\mathsf{C}_{i,\tau,\ell}^{(n)}\}$ such that the sequence of probabilities of error $P_{\mathrm{e,MAC}}^{(n)} \to 0$ as $n \to \infty$ and the power constraint (19) is satisfied.

Applying Fano's inequality, we obtain that for each positive integer $n$,

$$nR_1 \leq I(M_1; \mathbf{Y}^{(n)}) + \epsilon_n, \qquad (138a)$$

$$nR_2 \leq I(M_2; \mathbf{Y}^{(n)}) + \epsilon_n, \qquad (138b)$$

where $\frac{\epsilon_n}{n} \to 0$ as $n \to \infty$ and where $\mathbf{Y}^{(n)}$ denotes the $n\kappa$-dimensional column-vector that is obtained by stacking on top of each other all the $n$ vectors observed at the receiver when the blocklength-$n$ scheme is applied.

Letting $n \to \infty$, we have

$$R_1 \leq \varlimsup_{n\to\infty} \frac{1}{n} I(M_1; \mathbf{Y}^{(n)}) \qquad (139a)$$

$$R_2 \leq \varlimsup_{n\to\infty} \frac{1}{n} I(M_2; \mathbf{Y}^{(n)}). \qquad (139b)$$

Since the right-hand-side of (137) is closed, it suffices to prove that $\forall \delta > 0$, the pair $(R_1', R_2')$,

$$R_1' \triangleq \eta(R_1 - \delta) \qquad (140a)$$

$$R_2' \triangleq \eta(R_2 - \delta), \qquad (140b)$$

lies in $\mathcal{R}_{\mathrm{MAC}}(\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}, (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}, (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}; P)$ for some positive integer $\eta$ and strictly-lower block-triangular $\eta\nu_1$-by-$\eta\kappa$ and $\eta\nu_2$-by-$\eta\kappa$ matrices $\mathsf{D}_1^{\mathsf{B}}$ and $\mathsf{D}_2^{\mathsf{B}}$ of block sizes $\nu_1 \times \kappa$ and $\nu_2 \times \kappa$, respectively.

By (139) and (140), there exists a finite blocklength $n$ such that

$$R_1' \leq I(M_1; \mathbf{Y}^{(n)}), \qquad (141a)$$

$$R_2' \leq I(M_2; \mathbf{Y}^{(n)}). \qquad (141b)$$

In the sequel, let $n$ be fixed and so that (141) holds. Also, based on the parameters $\{\mathsf{C}_{i,\tau,\ell}^{(n)}\}$ of the blocklength-$n$ scheme, let

$$
\mathsf{C}_i^{\mathsf{B}} = \begin{bmatrix}
0 & & \cdots & & 0 \\
\mathsf{C}_{i,1,2}^{(n)} & 0 & & & \\
\mathsf{C}_{i,1,3}^{(n)} & \mathsf{C}_{i,2,3}^{(n)} & 0 & & \\
\vdots & & & \ddots & \\
\mathsf{C}_{i,1,n}^{(n)} & \mathsf{C}_{i,2,n}^{(n)} & \cdots & \mathsf{C}_{i,(n-1),n}^{(n)} & 0
\end{bmatrix}, \quad i \in \{1,2\}, \qquad (142)
$$

and

$$\mathsf{D}_i^{\mathsf{B}} = \mathsf{C}_i^{\mathsf{B}}\left(\mathsf{I} - (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_1^{\mathsf{B}} - (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{C}_2^{\mathsf{B}}\right)^{-1}, \quad i \in \{1,2\}. \qquad (143)$$

Let moreover, $\mathsf{Q}_1$ and $\mathsf{Q}_2$ be the unique positive square roots of the (positive-definite) matrices

$$
\begin{aligned}
\mathsf{M}_1 &= (\mathsf{I}+\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{I}+\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}) + (\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}} \\
\mathsf{M}_2 &= (\mathsf{I}+\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}(\mathsf{I}+\mathsf{D}_2^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}) + (\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}})^{\mathsf{T}}\mathsf{D}_1^{\mathsf{B}}(\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}
\end{aligned}
$$

and define

$$\mathbf{U}_1^{(n)} \triangleq \mathsf{Q}_1\mathbf{W}_2^{(n)} \qquad (145)$$

$$\mathbf{U}_2^{(n)} \triangleq \mathsf{Q}_2\mathbf{W}_2^{(n)} \qquad (146)$$

where $\mathbf{W}_i^{(n)}$ denotes the $n\nu_i$-dimensional column-vector that is obtained by stacking on top of each other all the $n$ vectors produced by the encoding function $\xi_i^{(n)}$.

Using similar algebraic manipulations as leading to (81), we can write $\mathbf{Y}^{(n)}$ as

$$
\begin{aligned}
\mathbf{Y}^{(n)} &= (\mathsf{I} + (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{D}_1^{\mathsf{B}} + (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{D}_2^{\mathsf{B}}) \\
&\quad \cdot\left((\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{Q}_1^{-1}\mathbf{U}_1^{(n)} + (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}\mathsf{Q}_2^{-1}\mathbf{U}_2^{(n)} + \mathbf{Z}^{(n)}\right), \qquad (147)
\end{aligned}
$$

where $\mathbf{Z}^{(n)} = \left(\mathbf{Z}_1^{\mathsf{T}}, \ \ldots, \ \mathbf{Z}_n^{\mathsf{T}}\right)^{\mathsf{T}}$. In the same way as in Lemma 1 it can be shown that the power constraint (19) is equivalent to requiring that

$$
\begin{aligned}
&\mathsf{E}\left[\|\mathbf{U}_1^{(n)}\|^2\right] + \mathsf{E}\left[\|\mathbf{U}_2^{(n)}\|^2\right] \\
&\qquad \leq \eta P - \mathrm{tr}(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}) - \mathrm{tr}(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}). \qquad (148)
\end{aligned}
$$

Let now $\eta = n$ and consider the MIMO MAC (147), where Transmitter $i$, for $i \in \{1,2\}$, is equipped with $\eta\nu_i$ antennas, the receiver is equipped with $\eta\kappa$ antennas, and where $\mathbf{U}_1^{(\eta)}$ and $\mathbf{U}_2^{(\eta)}$ denote the $\eta\nu_1$ and $\eta\nu_2$-dimensional independent input-vectors. Recall that we denoted by $\mathcal{R}_{\mathrm{MAC}}(\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}, \mathsf{H}_1^{\mathsf{B}}, \mathsf{H}_2^{\mathsf{B}}; P)$ the capacity region of this channel under an expected total average block-power constraint $(\eta P - \mathrm{tr}\left(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}\right) - \mathrm{tr}\left(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}\right))$ on the inputs $\mathbf{U}_1^{(\eta)}$ and $\mathbf{U}_2^{(\eta)}$. Using random coding and joint typicality decoding, it can be shown that the nonnegative rate-pair $(\tilde{R}_1, \tilde{R}_2)$ lies in $\mathcal{R}_{\text{MAC}}\left(\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}, (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}, (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}; P\right)$ if it satisfies

$$\tilde{R}_1 \leq I(\Theta_1; \mathbf{Y}^{(\eta)}), \tag{149a}$$
$$\tilde{R}_2 \leq I(\Theta_2; \mathbf{Y}^{(\eta)}) \tag{149b}$$

for some auxiliary random variables $\Theta_1$ and $\Theta_2$ and some choice of the inputs $\mathbf{U}_1^{(\eta)}$ and $\mathbf{U}_2^{(\eta)}$ such that the pairs $(\Theta_1, \mathbf{U}_1^{(\eta)})$ and $(\Theta_2, \mathbf{U}_2^{(\eta)})$ are independent of each other and of the noise vectors $\mathbf{Z}_1^{(\eta)}, \mathbf{Z}_2^{(\eta)}$.

Specializing this last argument to $\Theta_1 = M_1$ and $\Theta_2 = M_2$, by (141), we conclude that the rate-pair $(R_1', R_2')$ defined in (140) lies in $\mathcal{R}_{\text{MAC}}\left(\eta, \mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}, (\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}, (\mathsf{H}_2^{\mathsf{B}})^{\mathsf{T}}; P\right)$, which establishes the desired proof.

### *D. Proof of Proposition 3*

Fix $\eta$, channel matrices $\mathsf{H}_1$ and $\mathsf{H}_2$ and strictly-lower block-triangular matrices $\mathsf{A}_1^{\mathsf{B}}, \mathsf{A}_2^{\mathsf{B}}$ as in (60). Let $\mathsf{H}_1^{\mathsf{B}}$ and $\mathsf{H}_2^{\mathsf{B}}$ be defined by (58) and for $i \in \{1,2\}$ let $\bar{\mathsf{H}}_i^{\mathsf{B}} = \mathsf{I}_\eta \otimes \bar{\mathsf{H}}_i$. Also, define $\mathsf{B}_1^{\mathsf{B}}, \mathsf{B}_2^{\mathsf{B}}$ as in (65) and let $\mathsf{D}_1^{\mathsf{B}}, \mathsf{D}_2^{\mathsf{B}}$ be given as in (88). Notice that since $\mathsf{A}_1^{\mathsf{B}}$ and $\mathsf{A}_2^{\mathsf{B}}$ are strictly-lower block-triangular, so are $\mathsf{B}_1^{\mathsf{B}}$, $\mathsf{B}_2^{\mathsf{B}}$, $\mathsf{D}_1^{\mathsf{B}}$, and $\mathsf{D}_2^{\mathsf{B}}$.

We shall show that the MIMO MAC in (81) subject to power constraint (84) and where everywhere $\mathsf{H}_i^{\mathsf{T}}$ and $(\mathsf{H}_i^{\mathsf{B}})^{\mathsf{T}}$ are replaced by $\bar{\mathsf{H}}_i$ and $\bar{\mathsf{H}}_i^{\mathsf{B}}$, and the MIMO BC in (66) subject to power constraint (68) have equal capacity regions.

Consider the MIMO MAC:

$$\begin{aligned}\mathbf{Y}' = {} & (\mathsf{I} + \bar{\mathsf{H}}_1^{\mathsf{B}}\mathsf{D}_1^{\mathsf{B}} + \bar{\mathsf{H}}_2^{\mathsf{B}}\mathsf{D}_2^{\mathsf{B}}) \\ & \cdot(\bar{\mathsf{H}}_1^{\mathsf{B}}\mathsf{Q}_1^{-1}\mathbf{U}_1 + \bar{\mathsf{H}}_2^{\mathsf{B}}\mathsf{Q}_2^{-1}\mathbf{U}_2 + \mathbf{Z}),\end{aligned} \tag{150}$$

where now $\mathsf{Q}_1$ and $\mathsf{Q}_2$ are the unique positive-definite square-roots of the matrices

$$\mathsf{M}_1 = (\mathsf{I} + \mathsf{D}_1^{\mathsf{B}}\bar{\mathsf{H}}_1^{\mathsf{B}})^{\mathsf{T}}(\mathsf{I} + \mathsf{D}_1^{\mathsf{B}}\bar{\mathsf{H}}_1^{\mathsf{B}}) + (\mathsf{D}_2^{\mathsf{B}}\bar{\mathsf{H}}_1^{\mathsf{B}})^{\mathsf{T}}(\mathsf{D}_2^{\mathsf{B}}\bar{\mathsf{H}}_1^{\mathsf{B}}), \tag{151a}$$
$$\mathsf{M}_2 = (\mathsf{I} + \mathsf{D}_2^{\mathsf{B}}\bar{\mathsf{H}}_2^{\mathsf{B}})^{\mathsf{T}}(\mathsf{I} + \mathsf{D}_2^{\mathsf{B}}\bar{\mathsf{H}}_2^{\mathsf{B}}) + (\mathsf{D}_1^{\mathsf{B}}\bar{\mathsf{H}}_2^{\mathsf{B}})^{\mathsf{T}}(\mathsf{D}_1^{\mathsf{B}}\bar{\mathsf{H}}_2^{\mathsf{B}}). \tag{151b}$$

That means $\mathsf{Q}_1$ and $\mathsf{Q}_2$ are the unique positive-definite symmetric matrices that satisfy

$$\mathsf{Q}_1\mathsf{Q}_1 = \mathsf{M}_1 \tag{152a}$$
$$\mathsf{Q}_2\mathsf{Q}_2 = \mathsf{M}_2. \tag{152b}$$

Since the matrix $(\mathsf{I} + \bar{\mathsf{H}}_1^{\mathsf{B}}\mathsf{D}_1^{\mathsf{B}} + \bar{\mathsf{H}}_2^{\mathsf{B}}\mathsf{D}_2^{\mathsf{B}})$ is invertible, the capacity region of the MAC in (150) under any input power constraint equals the capacity region of the MAC

$$\mathbf{Y}'_{\text{MAC}} = \bar{\mathsf{H}}_1^{\mathsf{B}}\mathsf{Q}_1^{-1}\mathbf{U}_1 + \bar{\mathsf{H}}_2^{\mathsf{B}}\mathsf{Q}_2^{-1}\mathbf{U}_2 + \mathbf{Z} \tag{153}$$

under the same input power constraint. This holds because the receiver can multiply its output vectors by an invertible matrix without changing the capacity region of the MAC.

We now turn to the BC (66). Let $\mathsf{S}_1$ and $\mathsf{S}_2$ be the positive square roots of the positive-definite matrices

$$\mathsf{N}_1 \triangleq (\mathsf{I} + \mathsf{H}_1^{\mathsf{B}}\mathsf{B}_1^{\mathsf{B}})(\mathsf{I} + \mathsf{H}_1^{\mathsf{B}}\mathsf{B}_1^{\mathsf{B}})^{\mathsf{T}} + (\mathsf{H}_1^{\mathsf{B}}\mathsf{B}_2^{\mathsf{B}})(\mathsf{H}_1^{\mathsf{B}}\mathsf{B}_2)^{\mathsf{T}} \tag{154a}$$
$$\mathsf{N}_2 \triangleq (\mathsf{H}_2^{\mathsf{B}}\mathsf{B}_1^{\mathsf{B}})(\mathsf{H}_2^{\mathsf{B}}\mathsf{B}_1^{\mathsf{B}})^{\mathsf{T}} + (\mathsf{I} + \mathsf{H}_2^{\mathsf{B}}\mathsf{B}_2^{\mathsf{B}})(\mathsf{I} + \mathsf{H}_2^{\mathsf{B}}\mathsf{B}_2^{\mathsf{B}})^{\mathsf{T}} \tag{154b}$$

That means, $\mathsf{S}_1$ and $\mathsf{S}_2$ are the unique positive-definite symmetric matrices that satisfy

$$\mathsf{S}_1\mathsf{S}_1 = \mathsf{N}_1 \tag{155a}$$
$$\mathsf{S}_2\mathsf{S}_2 = \mathsf{N}_2. \tag{155b}$$

The matrices $\mathsf{S}_1$ and $\mathsf{S}_2$ are invertible. Therefore, since in a MIMO BC each receiver can multiply its output vectors by an invertible matrix (here $\mathsf{E}\mathsf{S}_i^{-1}$) without changing the capacity of the BC, under any power constraint on the input vector $\mathbf{U}$, the MIMO BC in (66) has the same capacity region as the MIMO BC

$$\mathbf{Y}_i' \triangleq \mathsf{E}\mathsf{S}_i^{-1}\mathsf{H}_i^{\mathsf{B}}\mathbf{U} + \tilde{\mathbf{Z}}_i, \quad i \in \{1,2\}, \tag{156}$$

where $\tilde{\mathbf{Z}}_1$ and $\tilde{\mathbf{Z}}_2$ denote independent centered Gaussian vectors of identity covariance matrices.

Define now the new input-vector $\breve{\mathbf{U}}$ which is obtained from $\mathbf{U}$ by reversing the order of the elements:

$$\breve{\mathbf{U}} \triangleq \mathsf{E}\mathbf{U}. \tag{157}$$

Notice that $\|\breve{\mathbf{U}}\|^2$ and $\|\mathbf{U}\|^2$ are equal. Thus, when the input vectors $\mathbf{U}$ are average block-power constrained to

$$\eta P - \operatorname{tr}(\mathsf{B}_1^{\mathsf{B}}(\mathsf{B}_1^{\mathsf{B}})^{\mathsf{T}}) - \operatorname{tr}(\mathsf{B}_2^{\mathsf{B}}(\mathsf{B}_2^{\mathsf{B}})^{\mathsf{T}}), \tag{158}$$

the MIMO BC in (156) has the same capacity region as the MIMO BC

$$\mathbf{Y}_{i,\text{BC}}' \triangleq \mathsf{E}\mathsf{S}_i^{-1}\mathsf{H}_i^{\mathsf{B}}\mathsf{E}\breve{\mathbf{U}} + \tilde{\mathbf{Z}}_i, \quad i \in \{1,2\}, \tag{159}$$

when the input vectors $\breve{\mathbf{U}}$ are average block-power constrained to the same power (158).

We conclude the proof by showing that the capacity region of the MIMO BC in (159) under average input power constraint (158) and the capacity region of the MIMO MAC (153) under average input-power constraint

$$\eta P - \operatorname{tr}(\mathsf{D}_1^{\mathsf{B}}(\mathsf{D}_1^{\mathsf{B}})^{\mathsf{T}}) - \operatorname{tr}(\mathsf{D}_2^{\mathsf{B}}(\mathsf{D}_2^{\mathsf{B}})^{\mathsf{T}}) \tag{160}$$

coincide. To this end, notice that by Assumption (88), by Properties 2.) and 4.) of Note 1, and because for any matrices A and B we have $\operatorname{tr}(\mathsf{AB}) = \operatorname{tr}(\mathsf{BA})$,

$$\operatorname{tr}(\mathsf{B}_i^{\mathsf{B}}(\mathsf{B}_i^{\mathsf{B}})^{\mathsf{T}}) = \operatorname{tr}(\mathsf{D}_i^{\mathsf{B}}(\mathsf{D}_i^{\mathsf{B}})^{\mathsf{T}}), \qquad i \in \{1,2\}, \tag{161}$$

and the two power constraints (158) and (160) coincide. Also, by Assumption (88) and because $\mathsf{E} = \mathsf{E}^{\mathsf{T}}$ and $\mathsf{EE} = \mathsf{I}$,

$$\begin{aligned}\mathsf{E}\mathsf{M}_1\mathsf{E} &= \mathsf{E}(\mathsf{I} + \mathsf{D}_1^{\mathsf{B}}\bar{\mathsf{H}}_1^{\mathsf{B}})^{\mathsf{T}}(\mathsf{I} + \mathsf{D}_1^{\mathsf{B}}\bar{\mathsf{H}}_1^{\mathsf{B}})\mathsf{E} + \mathsf{E}(\mathsf{D}_2^{\mathsf{B}}\bar{\mathsf{H}}_1^{\mathsf{B}})^{\mathsf{T}}(\mathsf{D}_2^{\mathsf{B}}\bar{\mathsf{H}}_1^{\mathsf{B}})\mathsf{E} \\ &= \mathsf{E}(\mathsf{I} + \mathsf{E}(\mathsf{B}_1^{\mathsf{B}})^{\mathsf{T}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{E})^{\mathsf{T}}(\mathsf{I} + \mathsf{E}(\mathsf{B}_1^{\mathsf{B}})^{\mathsf{T}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{E})\mathsf{E} \\ &\quad + \mathsf{E}(\mathsf{E}(\mathsf{B}_2^{\mathsf{B}})^{\mathsf{T}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{E})^{\mathsf{T}}(\mathsf{E}(\mathsf{B}_2^{\mathsf{B}})^{\mathsf{T}}(\mathsf{H}_1^{\mathsf{B}})^{\mathsf{T}}\mathsf{E})\mathsf{E} \\ &= (\mathsf{I} + \mathsf{H}_1^{\mathsf{B}}\mathsf{B}_1^{\mathsf{B}})(\mathsf{I} + \mathsf{H}_1^{\mathsf{B}}\mathsf{B}_1^{\mathsf{B}})^{\mathsf{T}} + (\mathsf{H}_1^{\mathsf{B}}\mathsf{B}_2^{\mathsf{B}})(\mathsf{H}_1^{\mathsf{B}}\mathsf{B}_2^{\mathsf{B}})^{\mathsf{T}} \\ &= \mathsf{N}_1.\end{aligned} \tag{162}$$

By (152), (155), and the uniqueness of $\mathsf{S}_1$ this yields

$$\mathsf{S}_1 = \mathsf{E}\mathsf{Q}_1\mathsf{E}. \tag{163a}$$

In a similar way we can also prove

$$\mathsf{S}_2 = \mathsf{E}\mathsf{Q}_2\mathsf{E}. \tag{163b}$$

We conclude that for each $i \in \{1,2\}$:

$$\begin{aligned}\mathsf{E}\mathsf{S}_i^{-1}\mathsf{H}_i^{\mathsf{B}}\mathsf{E} &= \mathsf{Q}_i^{-1}\mathsf{E}\mathsf{H}_i^{\mathsf{B}}\mathsf{E} \\ &= \mathsf{Q}_i^{-1}(\bar{\mathsf{H}}_i^{\mathsf{B}})^{\mathsf{T}} \\ &= (\bar{\mathsf{H}}_i^{\mathsf{B}}\mathsf{Q}_i^{-\mathsf{T}})^{\mathsf{T}}\end{aligned}$$

$$= (\bar{\mathsf{H}}_i^{\mathrm{B}} \mathsf{Q}_i^{-1})^{\mathsf{T}}, \tag{164}$$

where here in the last equality we used that $\mathsf{Q}_i$ is symmetric and thus $\mathsf{Q}_i^{-1} = \mathsf{Q}_i^{-\mathsf{T}}$.

The MIMO BC in (159) and the MIMO MAC in (153) are thus dual and the desired equality (89) in the proposition follows from the non-feedback duality of the MIMO Gaussian MAC and BC [26], [27], [28].

### E. Proof of Corollary 4

As a first step, define the matrices

$$\mathsf{C}'_{i,\tau,\ell} \triangleq \mathsf{E}\mathsf{C}_{i,\tau,\ell}\mathsf{E}, \tag{165}$$

and construct the strictly-lower block-triangular matrices $\mathsf{C}_1^{\mathrm{B}\prime}$ and $\mathsf{C}_2^{\mathrm{B}\prime}$ similarly to (76)

$$\mathsf{C}_i^{\mathrm{B}\prime} = \begin{bmatrix} 0 & & \cdots & & 0 \\ \mathsf{C}'_{i,1,2} & 0 & & & \\ \mathsf{C}'_{i,1,3} & \mathsf{C}'_{i,2,3} & 0 & & \\ \vdots & & & \ddots & \\ \mathsf{C}'_{i,1,\eta} & \mathsf{C}'_{i,2,\eta} & \cdots & \mathsf{C}'_{i,(\eta-1),\eta} & 0 \end{bmatrix}, \quad i \in \{1,2\}, \tag{166}$$

Also, let

$$\mathsf{D}_i^{\mathrm{B}\prime} \triangleq \mathsf{C}_i^{\mathrm{B}\prime} \left(\mathsf{I} - \bar{\mathsf{H}}_1^{\mathrm{B}}\mathsf{C}_1^{\mathrm{B}\prime} - \bar{\mathsf{H}}_2^{\mathrm{B}}\mathsf{C}_2^{\mathrm{B}\prime}\right)^{-1}, \; i \in \{1,2\}. \tag{167}$$

We now show that under Assumption (91),

$$\begin{aligned} &\mathcal{R}_{\mathrm{BC}}(\eta, \mathsf{B}_1^{\mathrm{B}}, \mathsf{B}_2^{\mathrm{B}}, \mathsf{H}_1^{\mathrm{B}}, \mathsf{H}_2^{\mathrm{B}}; P) \\ &\quad = \mathcal{R}_{\mathrm{MAC}}(\eta, \mathsf{D}_1^{\mathrm{B}\prime}, \mathsf{D}_2^{\mathrm{B}\prime}, \bar{\mathsf{H}}_1^{\mathrm{B}}, \bar{\mathsf{H}}_2^{\mathrm{B}}; P) \end{aligned} \tag{168}$$

and moreover,

$$\begin{aligned} &\mathcal{R}_{\mathrm{MAC}}(\eta, \mathsf{D}_1^{\mathrm{B}}, \mathsf{D}_2^{\mathrm{B}}, (\mathsf{H}_1^{\mathrm{B}})^{\mathsf{T}}, (\mathsf{H}_2^{\mathrm{B}})^{\mathsf{T}}; P) \\ &\quad = \mathcal{R}_{\mathrm{MAC}}(\eta, \mathsf{D}_1^{\mathrm{B}\prime}, \mathsf{D}_2^{\mathrm{B}\prime}, \bar{\mathsf{H}}_1^{\mathrm{B}}, \bar{\mathsf{H}}_2^{\mathrm{B}}; P), \end{aligned} \tag{169}$$

which combined establish the desired proof.

Equation (169) follows by Remark 1 and because through the operation (165) the encoders transform the channel matrix $\mathsf{H}_i^{\mathsf{T}}$ into $\bar{\mathsf{H}}_i$. The multiplication from the left by $\mathsf{E}$ makes that the inputs are premultiplied by $\mathsf{E}$ before they are sent over the channel and the multiplication from the right makes that the feedback outputs are first multiplied by $\mathsf{E}$ before further use, see (79). (See also the proof of Remark 1.)

To prove (168), we shall show that

$$\bar{\mathsf{D}}_i^{\mathrm{B}\prime} = \mathsf{B}_i^{\mathrm{B}}, \tag{170}$$

which by Proposition 3 establishes (168). Notice first that Condition (91) implies

$$\bar{\mathsf{A}}_i^{\mathrm{B}} = \mathsf{C}_i^{\mathrm{B}\prime}. \tag{171}$$

Therefore, by (167), and by the properties in Note 1,

$$\begin{aligned} \bar{\mathsf{D}}_i^{\mathrm{B}\prime} &= \overline{\bar{\mathsf{A}}_i^{\mathrm{B}} \left(\mathsf{I} - \bar{\mathsf{H}}_1^{\mathrm{B}}\bar{\mathsf{A}}_1^{\mathrm{B}} - \bar{\mathsf{H}}_2^{\mathrm{B}}\bar{\mathsf{A}}_2^{\mathrm{B}}\right)^{-1}} \\ &= \left(\mathsf{I} - \mathsf{A}_1^{\mathrm{B}}\mathsf{H}_1^{\mathrm{B}} - \mathsf{A}_2^{\mathrm{B}}\mathsf{H}_2^{\mathrm{B}}\right)^{-1}\mathsf{A}_i^{\mathrm{B}} \\ &= \mathsf{B}_i^{\mathrm{B}} \end{aligned} \tag{172}$$

and thus concludes the proof.

## Appendix A
## Proofs of Auxiliary Results

### A. Proof of (32)

Fix a nonzero real number $h$ and a positive real number $P$. By (30),

$$\begin{aligned} &C_{\mathrm{MAC,SISO},\Sigma}^{\mathrm{fb}}(h,h;P) \\ &= \max_{\substack{P_1,P_2\geq 0:\\ P_1+P_2=P}} \frac{1}{2}\log\left(1 + h^2P + 2h^2\sqrt{P_1P_2}\rho^\star(h,h;P_1,P_2)\right), \\ &= \max_{\alpha\in[0,1]} \frac{1}{2}\log(1 + h^2P + 2h^2P\zeta_{P,h}(\alpha)) \end{aligned} \tag{173}$$

where the function $\zeta_{P,h}$ is defined as

$$\begin{aligned} \zeta_{P,h} \colon [0,1] &\to \left[0, \frac{1}{4}\right] \\ \alpha &\mapsto \sqrt{\alpha(1-\alpha)}\rho^\star(h,h;\alpha P,(1-\alpha)P). \end{aligned} \tag{174}$$

We argue in the following that irrespective of the values of $h$ and $P$:

$$\operatorname*{argmax}_{\alpha\in[0,1]} \zeta_{P,h}(\alpha) = \frac{1}{2}, \tag{175}$$

and thus the sum-capacity $C_{\mathrm{MAC,SISO},\Sigma}^{\mathrm{fb}}(h,h;P)$ is as in (32). More specifically, we show that if (175) was violated, then the sum-capacity of the scalar Gaussian MAC with symmetric channel gains $h$ and symmetric individual power constraints $P/2$ differs from $\frac{1}{2}\log(1 + h^2P + +2h^2P\zeta_{P,h}(1/2))$, which contradicts the results in [20]. In fact, let's assume for contradiction that there exists a $\alpha^\star \in [0,1]$ such that

$$\zeta_{P,h}(\alpha^\star) > \zeta_{P,h}(1/2). \tag{176}$$

By symmetry of the function $\zeta_{P,h}$, also

$$\zeta_{P,h}(1-\alpha^\star) > \zeta_{P,h}(1/2). \tag{177}$$

We consider the following time-sharing scheme over the scalar Gaussian MAC with symmetric channel gains and power constraints. During the first half of the channel uses we apply Ozarow's scheme [20] where Transmitter 1 uses average power $\alpha^\star P$ and Transmitter 2 uses average power $(1-\alpha^\star)P$. During the second half we again apply Ozarow's scheme, but now Transmitter 1 uses average power $(1-\alpha^\star)P$ and Transmitter 2 uses average power $\alpha^\star P$. Over the entire block of transmission, each transmitter thus uses average power $P/2$ and satisfies the individual average power constraint. The described scheme achieves a sum-rate of

$$\begin{aligned} R_\Sigma &= \frac{1}{4}\log(1 + h^2P + 2h^2P\zeta_{P,h}(\alpha^\star)) \\ &\quad + \frac{1}{4}\log(1 + h^2P + 2h^2P\zeta_{P,h}(1-\alpha^\star)) \qquad (178) \\ &= \frac{1}{2}\log(1 + h^2P + 2h^2P\zeta_{P,h}(\alpha^\star)). \qquad (179) \end{aligned}$$

By (176) and (179) the rate of our scheme thus exceeds the sum-capacity of the channel under symmetric individual power constraints, which establishes the desired contradiction.

### B. Proof of Note 2

Recall the mapping $\omega$ defined by (65)

$$\mathsf{B}_i^{\mathsf{B}} \triangleq \left(\mathsf{I} - \mathsf{A}_1^{\mathsf{B}}\mathsf{H}_1^{\mathsf{B}} - \mathsf{A}_2^{\mathsf{B}}\mathsf{H}_2^{\mathsf{B}}\right)^{-1} \mathsf{A}_i^{\mathsf{B}}, \quad i \in \{1,2\}. \tag{180}$$

One can verify that

$$\mathsf{A}_i^{\mathsf{B}} \triangleq \left(\mathsf{I} + \mathsf{B}_1^{\mathsf{B}}\mathsf{H}_1^{\mathsf{B}} + \mathsf{B}_2^{\mathsf{B}}\mathsf{H}_2^{\mathsf{B}}\right)^{-1} \mathsf{B}_i^{\mathsf{B}}, \quad i \in \{1,2\}. \tag{181}$$

Observe now that:

- If a matrix $\mathsf{A}$ is strictly-lower block-triangular with block sizes $\kappa_1 \times \kappa_2$ and a matrix $\mathsf{B}$ is lower block-triangular with block sizes $\kappa_2 \times \kappa_3$, then the product $\mathsf{AB}$ is strictly-lower block-triangular with block sizes $\kappa_1 \times \kappa_3$.
- The inverse of a lower block-triangular matrix with block sizes $\kappa$-by-$\kappa$ is again lower block-triangular with the same block sizes.

With these observations and inspecting the expressions in (180) and (181), the lemma follows.

## Appendix B
## Proof of Performance of the BC Scheme in Section VII

We rewrite our BC scheme in vector notation. To this end, collect inputs, outputs, and noise symbols at times $t = 0, \ldots, n-1$ in vectors:

$$\mathbf{X} \triangleq (X_0, \ldots, X_{n-1})^{\mathsf{T}}, \tag{182}$$

$$\mathbf{Y}_i \triangleq (Y_{i,0}, \ldots, Y_{i,n-1})^{\mathsf{T}}, \qquad i \in \{1,2\}, \tag{183}$$

$$\mathbf{Z}_i \triangleq (Z_{i,0}, \ldots, Z_{i,n-1})^{\mathsf{T}}, \qquad i \in \{1,2\}. \tag{184}$$

Recall the definitions of $\gamma_i$ and $\delta_i$ in (102) and (103), and define $n \times n$ matrices

$$\mathsf{C}_i = -\delta_i \begin{pmatrix} 0 & 0 & \ldots & & 0 \\ \gamma_i & 0 & & & \\ \gamma_i^2 & \gamma_i & 0 & & \\ \vdots & & & & \\ \gamma_i^{n-1} & \ldots & & \gamma_i & 0 \end{pmatrix}, \quad i \in \{1,2\}, \tag{185}$$

and

$$\mathsf{G} \triangleq (\mathsf{I} - h_1\mathsf{C}_1 - h_2\mathsf{C}_2)^{-1}, \tag{186}$$

$$\mathsf{D}_i = \mathsf{C}_i\mathsf{G}, \qquad i \in \{1,2\}. \tag{187}$$

Define now the two symmetric matrices $\mathsf{Q}_1$ and $\mathsf{Q}_2$ as the unique positive square roots of the square matrices

$$\mathsf{M}_1 = (\mathsf{I} + h_1\mathsf{D}_1)^{\mathsf{T}}(\mathsf{I} + h_1\mathsf{D}_1) + (h_1\mathsf{D}_2)^{\mathsf{T}}(h_1\mathsf{D}_2), \tag{188a}$$

$$\mathsf{M}_2 = (h_2\mathsf{D}_1)^{\mathsf{T}}(h_2\mathsf{D}_1) + (\mathsf{I} + h_2\mathsf{D}_2)^{\mathsf{T}}(\mathsf{I} + h_2\mathsf{D}_2). \tag{188b}$$

Define also the $n$-length column vectors

$$\mathbf{u}_i \triangleq \sqrt{P_i}\left(1, \gamma_i, \ldots, \gamma_i^{n-1}\right)^{\mathsf{T}}, \tag{189}$$

and

$$\mathbf{v}_i \triangleq \delta_i(\sqrt{P_i})^{-1}(1, \gamma_i^{-1}, \ldots, \gamma_i^{-n+1})^{\mathsf{T}}, \quad i \in \{1,2\}. \tag{190}$$

Notice that

$$\mathbf{u}_i = \frac{P_i}{\delta_i}\gamma_i^{n-1}\mathsf{E}\mathbf{v}_i, \tag{191}$$

and thus for any $n \times n$ Toeplitz-matrix $\mathsf{M}$ and for $i \in \{1,2\}$,

$$\begin{aligned} \|\mathbf{v}_i^{\mathsf{T}}\mathsf{M}\|^2 &= \|\mathsf{M}^{\mathsf{T}}\mathbf{v}_i\|^2 = \|\mathsf{E}\mathsf{M}^{\mathsf{T}}\mathbf{v}_i\|^2 = \|\mathsf{E}\mathsf{M}^{\mathsf{T}}\mathsf{E}\mathbf{u}_i\|^2\frac{\delta_i^2}{P_i^2}\gamma_i^{2(-n+1)} \\ &= \|\mathsf{M}\mathbf{u}_i\|^2\frac{\delta_i^2}{P_i^2}\gamma_i^{2(-n+1)} \end{aligned} \tag{192}$$

where in the last equality we used Property 3.) of Note 1.

We can now write the inputs of our BC scheme at times $t = 0, \ldots, n-1$ as

$$\begin{aligned} \mathbf{X} &= \mathbf{u}_1\Xi_1 + \mathbf{u}_2\Xi_2 + \mathsf{C}_1\mathbf{Y}_1 + \mathsf{C}_2\mathbf{Y}_2, \\ &= \mathbf{u}_1\Xi_1 + \mathbf{u}_2\Xi_2 + (h_1\mathsf{C}_1 + h_2\mathsf{C}_2)\mathbf{X} + \mathsf{C}_1\mathbf{Z}_1 + \mathsf{C}_2\mathbf{Z}_2 \\ &= \mathsf{G}\mathbf{u}_1\Xi_1 + \mathsf{G}\mathbf{u}_2\Xi_2 + \mathsf{D}_1\mathbf{Z}_1 + \mathsf{D}_2\mathbf{Z}_2 \end{aligned} \tag{193}$$

and the channel outputs as

$$\mathbf{Y}_1 = h_1\left(\mathsf{G}\mathbf{u}_1\Xi_1 + \mathsf{G}\mathbf{u}_2\Xi_2\right) + (\mathsf{I} + h_1\mathsf{D}_1)\mathbf{Z}_1 + h_1\mathsf{D}_2\mathbf{Z}_2, \tag{194a}$$

$$\mathbf{Y}_2 = h_2\left(\mathsf{G}\mathbf{u}_1\Xi_1 + \mathsf{G}\mathbf{u}_2\Xi_2\right) + (\mathsf{I} + h_2\mathsf{D}_2)\mathbf{Z}_2 + h_2\mathsf{D}_1\mathbf{Z}_1. \tag{194b}$$

Receiver 1's total noise vector $(I + h_1\mathsf{D}_1)\mathbf{Z}_1 + h_1\mathsf{D}_2\mathbf{Z}_2$ has covariance matrix $\mathsf{M}_1 = \mathsf{Q}_1\mathsf{Q}_1$ and Receiver 2's total noise vector $(I + h_2\mathsf{D}_2)\mathbf{Z}_2 + h_2\mathsf{D}_1\mathbf{Z}_1$ has covariance matrix $\mathsf{M}_2 = \mathsf{Q}_2\mathsf{Q}_2$.

The receivers' estimates can be written as

$$\hat{\Xi}_i^{(n-1)} \triangleq \mathbf{v}_i^{\mathsf{T}}\mathbf{Y}_i, \quad i \in \{1,2\}. \tag{195}$$

We now analyze the probability of error of our BC scheme with the help of Lemma 2 presented on the next page.

The nearest-neighbor decoding rule (106) produces the correct estimate whenever $\left|\Theta_i(M_i) - \hat{\Theta}_i\right| < \Delta_i/2$, where $\Delta_i$ is defined in (96). Since the estimation error $\Theta_i(M_i) - \hat{\Theta}_i$ is zero-mean Gaussian, Receiver $i$'s probability of error is bounded by

$$\begin{aligned} \Pr\left[\hat{M}_i \neq M_i\right] &\leq 2\mathcal{Q}\left(\frac{\Delta_i/2}{\sqrt{\mathsf{Var}\left(\Theta_i(M_i) - \hat{\Theta}_i\right)}}\right) \\ &= 2\mathcal{Q}\left(\frac{\sqrt{P_i}|h_i|\sqrt{1-\rho^\star}}{\lfloor 2^{nR_i}\rfloor\sqrt{\mathsf{Var}\left(\Xi_i - \hat{\Xi}_i^{(n-1)}\right)}}\right), \end{aligned}$$

where the equality follows by (96) and (105). Since the numerator is constant, the probability of error of our BC scheme tends to 0 whenever

$$R_i < \varliminf_{n\to\infty}\frac{1}{2n}\log\left(\mathsf{Var}\left(\Xi_i - \hat{\Xi}_i^{(n-1)}\right)\right), \ i \in \{1,2\}. \tag{196}$$

By the independence of the symbols $\Xi_1$ and $\Xi_2$ with the noise vectors $\mathbf{Z}_1$ and $\mathbf{Z}_2$,

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

$$\le \gamma_i^{-2n} + \frac{nP}{1-\rho^\star}\frac{\delta_i^2}{P_i^2}\gamma_i^{2(1-n)}$$
$$= \left(1 + nP\frac{\delta_i^2\gamma_i^2}{P_i^2}\right)\gamma_i^{-2n}, \tag{197}$$

where the second equality follows by (192) and the inequality by Inequalities (200a) and (200c) of Lemma 2 ahead. Combined with (196), it follows that the probability of error of our BC scheme tends to 0 as $n$ tends to infinity, whenever

$$R_i < \frac{1}{2}\log\left(\gamma_i^2\right), \qquad i \in \{1,2\}, \tag{198}$$

which coincides with the desired rate constraints in (107).

It remains to analyze the consumed power. The total power in channel uses $t = 0, \ldots, n-1$ is

$$\begin{aligned}\|\mathbf{X}\|^2 &= \|\mathsf{G}\mathbf{u}_1\|^2 + \|\mathsf{G}\mathbf{u}_2\|^2 + 2\rho^\star \mathrm{tr}\left(\mathsf{G}\mathbf{u}_1\mathbf{u}_2^\mathsf{T}\mathsf{G}^\mathsf{T}\right) \\ &\quad + \mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^\mathsf{T}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^\mathsf{T}\right) \\ &\le 2\|\mathsf{G}\mathbf{u}_1\|^2 + 2\|\mathsf{G}\mathbf{u}_2\|^2 + \mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^\mathsf{T}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^\mathsf{T}\right) \\ &= 2\frac{P_1^2}{\delta_1^2}\gamma_1^{2(n-1)}\|\mathbf{v}_1^\mathsf{T}\mathsf{G}\|^2 + 2\frac{P_2^2}{\delta_2^2}\gamma_2^{2(n-1)}\|\mathbf{v}_2^\mathsf{T}\mathsf{G}\|^2 \\ &\quad + \mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^\mathsf{T}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^\mathsf{T}\right) \\ &\le 2\frac{P_1^2}{\delta_1^2}\gamma_1^{-2} + 2\frac{P_2^2}{\delta_2^2}\gamma_2^{-2} + nP \end{aligned} \tag{199}$$

where the first inequality follows by Cauchy-Schwarz Inequality and because $0 < \rho^\star < 1$, the first equality by (192), and the second inequality by Inequalities (200d) and (200b) in Lemma 2 ahead. Since the term $P_1^2\delta_1^{-2}\gamma_1^{-2} + P_2^2\delta_2^{-2}\gamma_2^{-2}$ 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 $n \to \infty$.

**Lemma 2.** *Consider Ozarow's sum-capacity achieving scheme for the MAC with perfect feedback and common randomness as described by Remark 3 or in [20], [10]. Assume power constraints $P_1$ and $P_2$ at the two transmitters, as chosen in our BC scheme. (Thus, in particular, $P_1 + P_2 = P$.)*

*Analyzing the transmit powers in Ozarow's MAC scheme, one can deduce:*

$$\|\mathsf{Q}_1\mathbf{u}_1\|^2 + \|\mathsf{Q}_2\mathbf{u}_2\|^2 \le \frac{nP}{1-\rho^\star} \tag{200a}$$
$$\mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^\mathsf{T}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^\mathsf{T}\right) \le nP. \tag{200b}$$

*Analyzing the variances of the estimation errors in Ozarow's scheme, one can deduce:*

$$\mathsf{Var}(\Xi_i - h_1\mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{u}_1\Xi_1 - h_2\mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{u}_2\Xi_2) \le \gamma_i^{-2n}, \tag{200c}$$
$$\|\mathbf{v}_i^\mathsf{T}\mathsf{G}\|^2 \le \gamma_i^{-2n}. \tag{200d}$$

*Proof:* We use vector notation to write inputs and outputs of Ozarow's MAC-scheme as described by Remark 3 (see also [20], [10]) in channel uses $t = 0, \ldots, n-1$. Let

$$\mathbf{X}_1 \triangleq (X_{1,0}, \ldots, X_{1,n-1})^\mathsf{T} \tag{201}$$
$$\mathbf{X}_2 \triangleq (X_{2,0}, \ldots, X_{2,n-1})^\mathsf{T} \tag{202}$$
$$\mathbf{Y} \triangleq (Y_0, \ldots, Y_{n-1})^\mathsf{T}. \tag{203}$$

Then,

$$\mathbf{X}_i = \mathbf{u}_i\Xi_i + \mathsf{C}_i\mathbf{Y}, \quad i \in \{1,2\}, \tag{204}$$

and

$$\begin{aligned}\mathbf{Y} &= h_1\mathbf{X}_1 + h_2\mathbf{X}_2 + \mathbf{Z} \\ &= h_1\mathbf{u}_1\Xi_1 + h_2\mathbf{u}_2\Xi_2 + (h_1\mathsf{C}_1 + h_2\mathsf{C}_2)\mathbf{Y} + \mathbf{Z} \\ &= \mathsf{G}\big(h_1\mathbf{u}_1\Xi_1 + h_2\mathbf{u}_2\Xi_2 + \mathbf{Z}\big),\end{aligned} \tag{205}$$

and for the estimate $\hat{\Xi}_i^{(n-1)}$:

$$\begin{aligned}\hat{\Xi}_i^{(n-1)} &= \mathbf{v}_i^\mathsf{T}\mathbf{Y} \\ &= h_1\mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{u}_1\Xi_1 + h_2\mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{u}_2\Xi_2 + \mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{Z}.\end{aligned} \tag{206}$$

We analyze the variance of the estimation errors $(\Xi_i - \hat{\Xi}_i^{(n-1)})$. By [20], [10],

$$\mathsf{Var}\Big(\Xi_i - \hat{\Xi}_i^{(n-1)}\Big) = \gamma_i^{-2n}, \quad i \in \{1,2\}. \tag{207}$$

(In this MAC scheme, $\hat{\Xi}_i^{(n-1)}$ is the LMMSE-estimate of $\Xi_i$ given $\mathbf{Y}$, and the variance $\mathsf{Var}\Big(\Xi_i - \hat{\Xi}_i^{(n-1)}\Big)$ can be obtained through a relatively simple calculation. This is not the case in our BC scheme.)

On the other hand, by the independence of the vector $\mathbf{Z}$ with $\Xi_1$ and $\Xi_2$, for $i \in \{1,2\}$,

$$\begin{aligned}&\mathsf{Var}\Big(\Xi_i - \hat{\Xi}_i^{(n-1)}\Big) \\ &= \mathsf{Var}(\Xi_i - h_1\mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{u}_1\Xi_1 - h_2\mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{u}_2\Xi_2) + \|\mathbf{v}_i^\mathsf{T}\mathsf{G}\|^2\end{aligned} \tag{208}$$

Thus, for $i \in \{1,2\}$,

$$\mathsf{Var}(\Xi_i - h_1\mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{u}_1\Xi_1 - h_2\mathbf{v}_i^\mathsf{T}\mathsf{G}\mathbf{u}_2\Xi_2) + \|\mathbf{v}_i^\mathsf{T}\mathsf{G}\|^2 = \gamma_i^{-2n}. \tag{209}$$

Since variances and norms cannot be negative, (209) establishes Inequalities (200c) and (200d) in the lemma.

To prove Inequalities (200a) and (200b) in the lemma, we consider the power of input vectors $\mathbf{X}_1$ and $\mathbf{X}_2$. By construction [20], [10],

$$\mathsf{E}\left[\|\mathbf{X}_1\|^2\right] + \mathsf{E}\left[\|\mathbf{X}_2\|^2\right] = nP. \tag{210}$$

On the other hand, following similar steps as in the proof of Lemma 1, we can express the total power $\mathsf{E}\left[\|\mathbf{X}_1\|^2\right] + \mathsf{E}\left[\|\mathbf{X}_2\|^2\right]$ in terms of the matrices $\mathsf{Q}_1$, $\mathsf{Q}_2$, etc. By (204) and (205),

$$\mathbf{X}_1 = (\mathsf{I} + h_1\mathsf{D}_1)\mathbf{u}_1\Xi_1 + h_2\mathsf{D}_1\mathbf{u}_2\Xi_2 + \mathsf{D}_1\mathbf{Z} \tag{211a}$$
$$\mathbf{X}_2 = (\mathsf{I} + h_2\mathsf{D}_2)\mathbf{u}_2\Xi_2 + h_1\mathsf{D}_2\mathbf{u}_1\Xi_1 + \mathsf{D}_2\mathbf{Z}. \tag{211b}$$

Thus,

$$\begin{aligned}\mathsf{E}\left[\|\mathbf{X}_1\|^2\right] &= \mathbf{u}_1^\mathsf{T}(\mathsf{I} + h_1\mathsf{D}_1)^\mathsf{T}(\mathsf{I} + h_1\mathsf{D}_1)\mathbf{u}_1 + h_2^2\mathbf{u}_2^\mathsf{T}\mathsf{D}_1^\mathsf{T}\mathsf{D}_1\mathbf{u}_2 \\ &\quad + 2\rho^\star\mathbf{u}_1^\mathsf{T}(\mathsf{I} + h_1\mathsf{D}_1)^\mathsf{T}\mathsf{D}_1\mathbf{u}_2h_2 + \mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^\mathsf{T}\right)\end{aligned} \tag{212}$$

and

$$\begin{aligned}\mathsf{E}\left[\|\mathbf{X}_2\|^2\right] &= \mathbf{u}_2^\mathsf{T}(\mathsf{I} + h_2\mathsf{D}_2)^\mathsf{T}(\mathsf{I} + h_2\mathsf{D}_2)\mathbf{u}_2 + h_1^2\mathbf{u}_1^\mathsf{T}\mathsf{D}_2^\mathsf{T}\mathsf{D}_2\mathbf{u}_1 \\ &\quad + 2\rho^\star\mathbf{u}_2^\mathsf{T}(\mathsf{I} + h_2\mathsf{D}_2)^\mathsf{T}\mathsf{D}_2\mathbf{u}_1h_1 + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^\mathsf{T}\right)\end{aligned} \tag{213}$$

Applying Cauchy-Schwarz Inequality to the pair of vectors $(\mathsf{I} + h_1\mathsf{D}_1)\mathbf{u}_1$ and $\mathsf{D}_1\mathbf{u}_2h_2$ and to the pair of vectors $(\mathsf{I} + h_2\mathsf{D}_2)\mathbf{u}_2$ and $\mathsf{D}_2\mathbf{u}_1h_1$, gives the two bounds:

$$\begin{aligned}&2\rho^\star\mathbf{u}_1^\mathsf{T}(\mathsf{I} + h_1\mathsf{D}_1)^\mathsf{T}\mathsf{D}_1\mathbf{u}_2h_2 \\ &\ge -\rho^\star\left(\mathbf{u}_1^\mathsf{T}(\mathsf{I} + h_1\mathsf{D}_1)^\mathsf{T}(\mathsf{I} + h_1\mathsf{D}_1)\mathbf{u}_1 + h_2^2\mathbf{u}_2^\mathsf{T}\mathsf{D}_1^\mathsf{T}\mathsf{D}_1\mathbf{u}_2\right)\end{aligned}$$

$$2\rho^\star \mathbf{u}_2^\mathsf{T}(\mathsf{I}+h_2\mathsf{D}_2)^\mathsf{T}\mathsf{D}_2\mathbf{u}_1 h_1 \tag{214a}$$

$$\geq -\rho^\star\big(\mathbf{u}_2^\mathsf{T}(\mathsf{I}+h_2\mathsf{D}_2)^\mathsf{T}(\mathsf{I}+h_2\mathsf{D}_2)\mathbf{u}_2 + h_1^2\mathbf{u}_1^\mathsf{T}\mathsf{D}_2^\mathsf{T}\mathsf{D}_2\mathbf{u}_1\big) \tag{214b}$$

By (212)–(214),

$$\begin{aligned}
&\mathsf{E}\left[\|\mathbf{X}_1\|^2\right] + \mathsf{E}\left[\|\mathbf{X}_2\|^2\right] \\
&\geq (1-\rho^\star)\mathbf{u}_1^\mathsf{T}\Big((\mathsf{I}+h_1\mathsf{D}_1)^\mathsf{T}(\mathsf{I}+h_1\mathsf{D}_1) + h_1^2\mathsf{D}_2^\mathsf{T}\mathsf{D}_2\Big)\mathbf{u}_1 \\
&\quad +(1-\rho^\star)\mathbf{u}_2^\mathsf{T}\Big((\mathsf{I}+h_2\mathsf{D}_2)^\mathsf{T}(\mathsf{I}+h_2\mathsf{D}_2) + h_2^2\mathsf{D}_1^\mathsf{T}\mathsf{D}_1\Big)\mathbf{u}_2 \\
&\quad +\mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^\mathsf{T}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^\mathsf{T}\right) \\
&\geq (1-\rho^\star)\mathbf{u}_1^\mathsf{T}\mathsf{M}_1\mathbf{u}_1 + (1-\rho^\star)\mathbf{u}_2^\mathsf{T}\mathsf{M}_2\mathbf{u}_2 \\
&\quad +\mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^\mathsf{T}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^\mathsf{T}\right) \\
&= (1-\rho^\star)\|\mathsf{Q}_1\mathbf{u}_1\|^2 + (1-\rho^\star)\|\mathsf{Q}_2\mathbf{u}_2\|^2 \\
&\quad +\mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^\mathsf{T}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^\mathsf{T}\right)
\end{aligned} \tag{215}$$

where in the last two equalities we used the definitions of $\mathsf{M}_1$ and $\mathsf{M}_2$ in (188) and that $\mathsf{Q}_1$ and $\mathsf{Q}_2$ are the positive square roots of $\mathsf{M}_1$ and $\mathsf{M}_2$.

Equality (210) and (215) yield

$$\begin{aligned}
(1-\rho^\star)\|\mathsf{Q}_1\mathbf{u}_1\|^2 + (1-\rho^\star)\|\mathsf{Q}_2\mathbf{u}_2\|^2 \\
+\mathrm{tr}\left(\mathsf{D}_1\mathsf{D}_1^\mathsf{T}\right) + \mathrm{tr}\left(\mathsf{D}_2\mathsf{D}_2^\mathsf{T}\right) \leq nP
\end{aligned} \tag{216}$$

which together with the positivity of the matrices $\mathsf{D}_1$ and $\mathsf{D}_2$, the nonnegativity of the norm, and the bounds $0 < \rho^\star < 1$ imply the missing two inequalities (200a) and (200b) in the lemma. ■

## 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] E. Ardestanizadeh, M. Wigger, Y. H. Kim, and T. Javidi, "Linear-feedback sum-capacity for Gaussian multiple access channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 224–236, 2012.

[3] A. Ben-Yishai and O. Shayevitz, "The AWGN BC with MAC feedback: a reduction to noiseless feedback via interaction," in *Proc. of IEEE ITW (invited)*, 2015.

[4] P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise (corresp.)," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 279–280, 1974.

[5] S. R. Bhaskaran, "Gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5252–5257, 2008.

[6] A. B. Carleial, "Multiple-access channels with different generalized feedback signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 841–850, 1982.

[7] R. S. Cheng and S. Verdu, "Gaussian multiaccess channels with ISI: capacity region and multiuser water-filling," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 773–785, 1993.

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

[9] T. M. Cover and C. S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. Inf. Theory*, vol. 27, no. 3, pp. 292–298, 1981.

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

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

[12] 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.

[13] ——, "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.

[14] Y. Geng and C. Nair, "The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages," *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 2087–2104, April 2014.

[15] S. Jafar and A. Goldsmith, "On the capacity of the vector MAC with feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3259–3264, 2006.

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

[17] 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*, 2010, p. 641.

[18] A. Lapidoth and M. Wigger, "On the AWGN MAC with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5432–5476, 2010.

[19] Y. Murin, Y. Kaspi, and R. Dabora, "On the Ozarow-Leung scheme for the Gaussian broadcast channel with feedback," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 948–952, 2015.

[20] 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.

[21] 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.

[22] S. Pillai and V. Prabhakaran, "On the noisy feedback capacity of Gaussian broadcast channels," in *Proc. of IEEE ITW*, 2015.

[23] 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.

[24] O. Shayevitz and M. Wigger, "On the capacity of the discrete memoryless broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1329–1345, 2013.

[25] R. Venkataramanan and S. S. Pradhan, "An achievable rate region for the broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6175–6191, 2013.

[26] 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.

[27] 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.

[28] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, 2006.

[29] F. M. J. Willems, E. C. van der Meulen, and J. P. M. Schalkwijk, "A coding scheme for the additive white Gaussian noise multiple access channel with semi-feedback," *Tijdschrft van het Nederlands Elektronica-en Radiogenootschap*, vol. 48, no. 3, pp. 103–107, 1983.

[30] 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.

[31] 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.

[32] Y. Wu and M. Wigger, "Any positive feedback rate increases the capacity of strictly less-noisy broadcast channels," in *Proc. of IEEE ITW*, 2013, pp. 1–5.



**Selma Belhadj Amor** (S'13–M'15) received the B.Sc. and M.Eng. degrees in electrical engineering (with distinction) in 2009 and 2011, respectively, from Grenoble-INP Ensimag, Grenoble, France. She also received the M.Sc. degree in mathematics and computer science (with distinction) in 2011 jointly from Grenoble-INP Ensimag and Université Joseph Fourier (UJF), Grenoble, France. She received the Ph.D. degree in electrical engineering in 2015 from Télécom ParisTech, Paris, France.

Since June 2015, she is a Postdoctoral Researcher at the Institut National de Recherche en Informatique et Automatique (Inria), Lyon, France. Her research interests are in information theory and communications theory.

**Yossef Steinberg** (M'96–SM'09–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. After completing his Ph.D., he was a Lady Davis Fellow in the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa, Israel. During the years 1991-1993, Dr. Steinberg held a post-doctoral appointment in the Department of Electrical Engineering at Princeton University, Princeton, NJ. In 1994, I was a research visiting professor at the $C^3$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, Dr. Steinberg joined the Department of Electrical Engineering at the Technion. He served in the years 2004-2007 as Associate Editor for Shannon Theory, in 2010-2013 as Associate Editor at large for the *IEEE Transactions on Information Theory*, and currently serves in the editorial board of *Foundation and Trends in Communications and Information Theory*. He served as general co-chair of ITW 2015. He won the 2007 best paper award of the IEEE Transactions on Information Theory, jointly with Hanan Weingarten and Shlomo Shamai. Dr. Steinberg is an IEEE Fellow (2011 class), for contributions to Information Theory. His research interests are in Digital Communications, Information Theory, and Estimation.

**Michèle Wigger** (S'05–M'09–SM'14) 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, USA. Since December 2009, she has been an Assistant Professor at Telecom ParisTech, in Paris, France.

She has been an associate editor of the IEEE Communication Letters since December 2012. Her main research interests are in multi-terminal information theory, in particular in distributed source coding, and capacities of networks with states, feedback, and user cooperation.