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

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

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

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

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

I. INTRODUCTION

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

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

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

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

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

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

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

In the symmetric case, the linear-feedback sum-capacity of the BC and MAC schemes than previously observed; and

Let us consider a channel with two receivers, where it is 1 and for a given power, error probability, and rate in the MAC and BC schemes.

Specifically, the BC encoder should run the operations

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

The transmitter wishes to convey a message $M_{1}$ to Receiver 1 and an independent message $M_{2}$ to Receiver 2. The messages are independent of the noise sequences $\{Z_{1,t}\}_{t=1}^{n}$ and $\{Z_{2,t}\}_{t=1}^{n}$ and uniformly distributed over the sets $\mathcal{M}_{1} \triangleq \{1, \ldots, 2^{nR_{1}}\}$ and $\mathcal{M}_{2} \triangleq \{1, \ldots, 2^{nR_{2}}\}$, where $R_{1}$ and $R_{2}$ denote the rates of transmission and $n$ the blocklength.

The transmitter observes causal noise-free channel output feedback from both receivers. Thus, the time-$t$ channel input $X_{t}$ can depend on all previous channel outputs $Y_{1,t-1}^{t-1}$ and $Y_{2,t-1}^{t-1}$ and messages $M_{1}$ and $M_{2}$:

$$X_{t} = \varphi^{(n)}_{t}(M_{1}, M_{2}, Y_{1,t-1}^{t-1}, Y_{2,t-1}^{t-1}), \quad t \in \{1, \ldots, n\},$$

for some encoding function $\varphi^{(n)}_{t}: \mathcal{M}_{1} \times \mathcal{M}_{2} \times \mathbb{R}^{t-1} \times \mathbb{R}^{t-1} \rightarrow \mathbb{R}$. The channel inputs have to satisfy an expected average block-power constraint

$$\frac{1}{n} \sum_{t=1}^{n} \mathbb{E}[X_{t}^2] \leq P,$$  \hspace{1cm} (3)

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

Each Receiver $i \in \{1, 2\}$ produces the guess $\hat{M}_{i}^{(n)} = \psi^{(n)}_{i}(Y_{i,t}^{t})$ for some decoding function $\psi^{(n)}_{i}: \mathbb{R}^{n} \rightarrow \mathcal{M}_{i}$. The average error probability is

$$P_{e,BC}^{(n)} \triangleq \Pr \{ (\hat{M}_{1}^{(n)} \neq M_{1}) \lor (\hat{M}_{2}^{(n)} \neq M_{2}) \}. \hspace{1cm} (4)$$

A rate-pair $(R_{1}, R_{2})$ is achievable over the Gaussian BC with feedback and power constraint $P$, if there exists a sequence of encoding and decoding functions $\{\varphi^{(n)}_{t}\}_{t=1}^{\infty}$, $\{\psi^{(n)}_{1}\}_{n=1}^{\infty}$, $\{\psi^{(n)}_{2}\}_{n=1}^{\infty}$ satisfying the power constraint (3) and such that the average probability of error $P_{e,BC}^{(n)}$ tends to zero as $n$ tends to infinity.

The closure of the union of all achievable regions is called capacity region and the supremum of the sum $R_{1} + R_{2}$ over all achievable pairs $(R_{1}, R_{2})$ is called sum-capacity.

This Gaussian BC is physically degraded and if only if

$$\lambda = \frac{h_{1}}{h_{2}} \quad \text{or} \quad \lambda = \frac{h_{2}}{h_{3}}. \hspace{1cm} (5)$$

For instance, if $h_{1} = h_{2} = \lambda = 1$ the two receivers observe the same sequence.

In this paper, we focus on linear-feedback schemes where the transmitter’s channel input is a linear combination of the previous feedback signals and an information-carrying vector that depends only on the messages $(M_{1}, M_{2})$. Specifically, we assume the channel input vector $X \triangleq (X_{1}, \ldots, X_{n})^{\top}$ can be written as:

$$X = W + A_{1}Y_{1} + A_{2}Y_{2}, \hspace{1cm} (6)$$

where $Y_{i} \triangleq (Y_{i,1}, \ldots, Y_{i,n})^{\top}$ represents the channel output vector at Receiver $i$, $A_{1}$ and $A_{2}$ are $n$-by-$n$ strictly lower-triangular matrices, and $W$ is an $n$-dimensional information-carrying vector $W = \xi^{(n)}(M_{1}, M_{2})$. The strict lower-triangularity of $A_{1}$ and $A_{2}$ ensures that the feedback is used strictly causally. The mapping $\xi^{(n)}: \mathcal{M}_{1} \times \mathcal{M}_{2} \rightarrow \mathbb{R}^{n}$ and the decoding operations $\psi_{1}^{(n)}$ and $\psi_{2}^{(n)}$ can be arbitrary.

The set of all rate-pairs achieved by linear-feedback schemes is called linear-feedback capacity region and is
denoted by $c_{\text{linfb}}^t(h_1, h_2, \lambda; P)$. The largest sum-rate achieved by a linear-feedback scheme is called linear-feedback sum-capacity and is denoted by $c_{\text{linfb},\Sigma}^t(h_1, h_2, \lambda; P)$.

III. THE DUAL GAUSSIAN MULTI-ACCESS CHANNEL

![Diagram of two-user Gaussian memoryless MAC with perfect feedback.](image)

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

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

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

The goal of communication is that Transmitters 1 and 2 convey the independent messages $M_1$ and $M_2$ to the common receiver, where the pair $(M_1, M_2)$ is independent of the noise sequence $\{Z_t\}$. (Recall that $M_i$ is uniformly distributed over $M_i = \{1, \ldots, [2^n R_i]\}$)

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

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

for some encoding function $f_i^{(n)}: M_i \times \mathbb{R}^{t-1} \to \mathbb{R}$.

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

$$\mathbb{E}[\|X_1\|^2] + \mathbb{E}[\|X_2\|^2] + 2\lambda \mathbb{E}[\|X_1 X_2\|] \leq nP$$

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

The receiver produces a guess $(\hat{M}_1^{(n)}, \hat{M}_2^{(n)}) = \Phi^{(n)}(Y^n)$ by means of a decoding function $\Phi^{(n)}: \mathbb{R}^n \to M_1 \times M_2$.

The average probability of error is

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

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

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

$$X_i = W_i + C_i Y, \quad i \in \{1, 2\},$$ (11)

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

Linear-feedback capacity for the MAC is defined similarly to the BC. We denote it by $c_{\text{MAC, CorPower}}^t(h_1, h_2; \lambda, P)$, where we use the subscript “CorPower” to emphasize the non-standard power constraint in (9).

IV. MAC-BC DUALITY FOR BC WITH CORRELATED NOISES

**Theorem 1.**

$$c_{\text{linfb}}^t(h_1, h_2, \lambda; P) = c_{\text{MAC, CorPower}}^t(h_1, h_2; \lambda, P).$$ (12)

**Proof:** A sketch of proof is provided in Appendix A. $\blacksquare$

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

$$h_1^2 P + 2(h_1^2 - h_2^2) \mathbb{E}[\|X_1 X_2\|^2] \leq nP$$

The term $\mathbb{E}[\|h_1 X_1 + h_2 X_2\|^2]$ corresponds to the received power. By our assumption $|h_1| \geq |h_2|$, the term $(h_1^2 - h_2^2) \mathbb{E}[\|X_2\|^2]$ is nonnegative and thus, Inequality (13) implies

$$\mathbb{E}[\|h_1 X_1 + h_2 X_2\|^2] \leq h_1^2 P.$$ (14)

Even with feedback, the maximum sum-rate over the Gaussian MAC is upper bounded by

$$1/2 \log(1 + (1/n) \mathbb{E}[\|h_1 X_1 + h_2 X_2\|^2])$$

which cannot exceed $1/2 \log(1 + h_1^2 P)$. Thus, in this case the sum-capacity of the MAC under consideration is the same with and without feedback [9]. Notice that for these parameters the same also applies for the dual BC: feedback does not increase capacity.

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

$$P_1 + P_2 + 2 \lambda \rho \sqrt{P_1 P_2} \leq P,$$ (15a)

for some parameter

$$\rho \in [0, \rho^* (h_1, h_2, P_1, P_2)]$$ (15b)

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

$$(1 + h_1^2 P_1 (1 - x^2)) (1 + h_2^2 P_2 (1 - x^2))$$

$$= 1 + h_1^2 P_1 + h_2^2 P_2 + 2|h_2| \|x \sqrt{P_1 P_2}.$$


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

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

\[
R_1 \leq \frac{1}{2} \log(1 + h_1^2 P_1 (1 - \rho^2))
\]
\[
R_2 \leq \frac{1}{2} \log(1 + h_2^2 P_2 (1 - \rho^2))
\]
\[
R_1 + R_2 \leq \frac{1}{2} \log(1 + h_1^2 P_1 + h_2^2 P_2 + 2|h_1|\rho \sqrt{P_1 P_2})
\]

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

**V. CONSTRUCTIVE SUM-RATE OPTIMAL BC-SCHME FOR UNCORRELATED NOISES** (\(\lambda = 0\))

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

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

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

![Fig. 3. Ozarow’s MAC-scheme](image)

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

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

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

\[
\Theta_i(M_i) \triangleq - (M_i - 1) \Delta_i + \sqrt{P_i},
\]

where

\[
\Delta_i \triangleq \frac{2 \sqrt{P_i}}{(2 \pi h_i)}. \tag{17}
\]

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

\[
t = -2 : \quad X_{1,-2} = 0, \quad X_{2,-2} = \Theta_2(M_2), \quad (18a)
\]
\[
t = -1 : \quad X_{1,-1} = \Theta_1(M_1), \quad X_{2,-1} = 0. \quad (18b)
\]

Define for each \(i \in \{1, 2\}\),

\[
\Xi_i \triangleq \sqrt{1 - \rho^* Z_{-i}} + \sqrt{\rho^* Z_0}, \tag{19}
\]

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

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

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

\[
\sqrt{1 - \rho^* Y_{-i}} + \sqrt{\rho^* Z_0} = \sqrt{1 - \rho^* P_i \Theta_i(M_i) + \Xi_i}, \tag{20}
\]

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

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

where

\[
\gamma_1 \triangleq \sqrt{1 + h_1^2 P_1 (1 - (\rho^*)^2)}, \quad (22a)
\]
\[
\gamma_2 \triangleq -\sqrt{1 + h_2^2 P_2 (1 - (\rho^*)^2)}. \quad (22b)
\]
In the following Section VI we show that in the limit as $n \to \infty$ this linear-feedback coding scheme for the BC has average block-power tending to $P^1$, and that it achieves the same rates as Ozarow's MAC scheme, i.e., all rate-pairs $(R_1, R_2)$ satisfying

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

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

VI. PROOF OF PERFORMANCE OF THE BC SCHEME IN SECTION V

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

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

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

\[
X_1 \triangleq (X_{1,0}, \ldots, X_{1,n-1})^T \quad (31)
\]

\[
X_2 \triangleq (X_{2,0}, \ldots, X_{2,n-1})^T \quad (32)
\]

\[
Y \triangleq (Y_0, \ldots, Y_{n-1})^T. \quad (33)
\]

Recall the definitions of $\gamma_i$ and $\delta_i$ in (22) and (23), and define the matrix

\[
C_i = -\delta_i \begin{pmatrix} 0 & 0 & \ldots & 0 \\ \gamma_i & 0 & \gamma_i & 0 \\ \vdots & \ddots & \ddots & \ddots \\ \gamma_i^{n-2} & \ldots & \gamma_i & 0 \end{pmatrix}. \quad (34)
\]

as well as

\[
G \triangleq (I - h_1 C_1 - h_2 C_2)^{-1}, \quad (35)
\]

\[
D_i = C_i G. \quad (36)
\]

Define also the vectors

\[
u_i \triangleq \sqrt{P_i} (1, \gamma_i, \ldots, \gamma_i^{n-1})^T, \quad (37)
\]

\[
u_i \triangleq \delta_i (\sqrt{P_i})^{-1} (1, \gamma_i, \ldots, \gamma_i^{n-1})^T. \quad (38)
\]

With these definitions, we can write the channel inputs as

\[
X_i = u_i \Xi_i + C_i Y, \quad i \in \{1, 2\}, \quad (39)
\]

and the channel outputs as

\[
Y = h_1 X_1 + h_2 X_2 + Z \quad (40)
\]

\[
= h_1 u_1 \Xi_1 + h_2 u_2 \Xi_2 + (h_1 C_1 + h_2 C_2) Y + Z \quad (40)
\]

\[1\]To change our BC scheme to a scheme that satisfies the average block-power constraint $P$ for all sufficiently large $n$, it suffices to scale the first inputs appropriately. As can be verified at hand of the proof steps in Section VI such a scaling does not change the set of achievable rates.
Subtracting \((h_1C_1 + h_2C_2)Y\) from both sides of (40) and then multiplying both sides from the left by the matrix \(G\) in (35), we obtain
\[
Y = G(h_1u_1Ξ_1 + h_2u_2Ξ_2 + Z). \tag{41}
\]
The LMMSE estimate \(\hat{Ξ}_i^{(n-1)}\) in (24) can be expressed as
\[
\hat{Ξ}_i^{(n-1)} = v_i^T Y. \tag{42}
\]
Define now the two symmetric matrices \(Q_1 \) and \(Q_2\) as the unique positive square roots of the square matrices
\[
M_1 = (I + h_1D_1)T(I + h_1D_1) + (h_1D_2)T(h_1D_2), \quad M_2 = (h_2D_2)T(h_2D_2) + (I + h_2D_2). \tag{43a}
\]
Define also the vectors \(α_i, β_i, α_2, \) and \(β_2\) as:
\[
α_i = (I + h_1D_1)u_i, \quad i \in \{1, 2\} \tag{44a}
\]
\[
β_1 = h_1D_2u_1 \tag{44b}
\]
\[
β_2 = h_2D_1u_2. \tag{44c}
\]
We can rewrite the channel input vectors in (39) as
\[
X_i = α_iΞ_1 + β_1Ξ_2 + D_iZ_i \tag{45a}
\]
\[
X_2 = α_2Ξ_2 + β_1Ξ_1 + D_2Z. \tag{45b}
\]
Note that \(||α_i||^2 + ||β_i||^2 = ||Q_iu_i||^2,\) for \(i \in \{1, 2\}.
\]
We can show that the total power over the two transmitters uses after the initialization phase is
\[
\pi = ||Q_1u_1||^2 + ||Q_2u_2||^2 + \Omega + tr(D_1D_1^T) + tr(D_2D_2^T), \tag{46}
\]
where \(Ω = 2ρ^*(α_1^Tβ_2 + α_2^Tβ_1)\) and \(ρ^* \in (0, 1)\) denotes the correlation factor between \(Ξ_1\) and \(Ξ_2\) and is constant. By construction of the scheme, \(π = nP,\) see [7, 13].

Using Cauchy-Schwarz inequality, we can show that
\[
Ω \leq −ρ^*(||α_1||^2 + ||β_2||^2 + ||α_2||^2 + ||β_1||^2). \tag{47}
\]
Hence, we have
\[
nP \geq (1 − ρ^*) (||Q_1u_1||^2 + ||Q_2u_2||^2) + tr(D_1D_1^T) + tr(D_2D_2^T). \tag{48}
\]
Thus, by the positivity of matrices \(D_1\) and \(D_2\) and because \(||Q_1u_1||^2 + ||Q_2u_2||^2 > 0\) and \(1 − ρ^* > 0\),
\[
tr(D_1D_1^T) + tr(D_2D_2^T) < nP, \tag{49a}
\]
\[
||Q_1u_1||^2 + ||Q_2u_2||^2 < \frac{nP}{1 − ρ^*}. \tag{49b}
\]
From the analysis in [7, 13], we also know that the receiver’s error variances about \(Ξ_1\) and \(Ξ_2\) satisfy
\[
Var(Ξ_i − \hat{Ξ}_i^{(n-1)}) = γ^n_i, \quad i \in \{1, 2\}. \tag{50}
\]
By (40) and (42)
\[
\hat{Ξ}_i^{(n-1)} = h_1v_1^TGu_1Ξ_1 + h_2v_2^TGu_2Ξ_2 + ν_i^T GZ. \tag{51}
\]
Using the independence of the vector \(Z\) with the message points \(Ξ_1\) and \(Ξ_2\), for \(i \in \{1, 2\},
\[
Var(Ξ_i − \hat{Ξ}_i^{(n-1)}) = Var(Ξ_i − h_1v_1^TGu_1Ξ_1 − h_2v_2^TGu_2Ξ_2 + ||ν_i^T G||^2, \tag{52}
\]
which combines with (50) to
\[
Var(Ξ_i − h_1v_1^TGu_1Ξ_1 − h_2v_2^TGu_2Ξ_2) \leq γ^{-2n}_i, \tag{53a}
\]
\[
||ν_i^T G||^2 \leq γ^{-2n}_i. \tag{53b}
\]
Inequalities (49) and (53) will be key to analyze our BC scheme.

B. Vector-Description of our BC scheme & Analysis

We now rewrite our BC scheme of Subsection V-B in vector notation. To this end, collect the inputs, the outputs, and the noise symbols after the initialization phase in the vectors
\[
X_i = (X_0, ..., X_{n-1})^T, \tag{54}
\]
\[
Y_i = (Y_{i,0}, ..., Y_{i,n-1})^T, \quad i \in \{1, 2\}, \tag{55}
\]
\[
Z_i = (Z_{i,0}, ..., Z_{i,n-1})^T, \quad i \in \{1, 2\}. \tag{56}
\]
Also, let
\[
A_i = \bar{C}_i \begin{bmatrix} a_1 \end{bmatrix}, \quad \text{and} \tag{57}
\]
\[
B_i = \bar{D}_i \begin{bmatrix} b_1 \end{bmatrix}, \quad i \in \{1, 2\} \tag{58}
\]
where (a) and (b) hold because the matrices \(C_1, C_2\) and \(D_1, D_2\) are Toeplitz.

The inputs of our BC scheme after the initialization phase satisfy
\[
X = u_1Ξ_1 + u_2Ξ_2 + A_1Y_1 + A_2Y_2, \tag{59}
\]
\[
= u_1Ξ_1 + u_2Ξ_2 + (h_1A_1 + h_2A_2)X + A_1Z_1 + A_2Z_2, \tag{60}
\]
Subtracting \((h_1A_1 + h_2A_2)X\) from both sides of (60) and then multiplying both sides from the left by the matrix \((I − h_1A_1 − h_2A_2)^{-1}\), by equivalence (57) and by the definition of the matrix \(G\) in (35), we obtain
\[
X = Gu_1Ξ_1 + Gu_2Ξ_2 + B_1Z_1 + B_2Z_2. \tag{61}
\]
The channel outputs are thus
\[
Y_1 = h_1(Gu_1Ξ_1 + Gu_2Ξ_2) + (I + h_1B_1)Z_1 + h_1B_2Z_2, \tag{62a}
\]
\[
Y_2 = h_2(Gu_1Ξ_1 + Gu_2Ξ_2) + (I + h_2B_2)Z_2 + h_2B_1Z_1, \tag{62b}
\]
Receiver \(i\) estimates \(Ξ_i\) as
\[
\hat{Ξ}_i^{(n-1)} = v_i^T Y_i, \quad i \in \{1, 2\}. \tag{63}
\]
The following will be key in our analysis: since
\[
u_i = \frac{P_i}{\delta_i} s_i^{n-1}Ev_i, \tag{64}
\]
for any \(n\)-by-\(n\) matrix \(M\), and for \(i \in \{1, 2\},
\[
||v_i^TM||^2 = ||M^iv_i||^2 = ||M^iEv_i||^2 \frac{\delta_{i}^2}{P_i} \gamma_i^{2(n-1)} = \frac{\delta_{i}^2}{P_i} \gamma_i^{2(n-1)}. \tag{65}
\]
In particular, when the matrix \(M\) is Toeplitz,
\[
||v_i^TM||^2 = ||Mu_i||^2 \frac{\delta_{i}^2}{P_i} \gamma_i^{2(n-1)}. \tag{66}
\]
We first analyze the blockpower used by our BC scheme. By (61), in channel uses 0, \ldots, n−1 the transmitter uses power
\[ \|G_1\|^2 + \|G_2\|^2 + 2\rho^* \text{tr}(G_1 u_1 G_1^T) + \text{tr}(B_1^* B_1^*) + \text{tr}(B_2^* B_2^*) \]
which can be upper-bounded using Cauchy-Schwarz and because 0 < \rho^* < 1 by
\[ 2\|G_1\|^2 + 2\|G_2\|^2 + \text{tr}(B_1^* B_1^*) + \text{tr}(B_2^* B_2^*) . \]  
(67)
By (66) and since \( G \) is Toeplitz, (67) equals
\[ 2\|v_1\|^2 + 2\|v_2\|^2 + \text{tr}(B_1^* B_1^*) + \text{tr}(B_2^* B_2^*) \]
\[ \leq \frac{2P_1^2}{\delta_1} + \frac{2P_2^2}{\delta_2} + \text{tr}(B_1^* B_1^*) + \text{tr}(B_2^* B_2^*) , \]  
where the inequality follows by Upper bound (53b). By (58), we have
\[ \text{tr}(B_i B_i^*) = \text{tr}(D_i D_i^*) , \quad i \in \{1, 2\} . \]  
(69)
Hence, by (49a)
\[ \text{tr}(B_1^* B_1^*) + \text{tr}(B_2^* B_2^*) < nP . \]  
(70)
Since the term \( \frac{P_1^2}{\delta_1} + \frac{P_2^2}{\delta_2} \) 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 blockpower constraint asymptotically as \( n \to \infty \).

We now analyze the variance of Receiver \( i \)’s estimation error \( \hat{\Xi}_i(n) \). By (62) and (63), we have
\[ \hat{\Xi}_1(n-1) = h_1 v_1^T G_1 u_1 \Xi_1 + h_1 v_1^T G_2 u_2 \Xi_2 \]
\[ + v_1^T ((1 + h_1 B_1^*) Z_1 + h_2 B_2 Z_2) , \]  
(71)
\[ \hat{\Xi}_2(n-1) = h_2 v_2^T G_1 u_1 \Xi_1 + h_2 v_2^T G_2 u_2 \Xi_2 \]
\[ + v_2^T ((1 + h_2 B_2^*) Z_2 + h_2 B_1 Z_1) . \]  
(72)
Define
\[ S_i = \text{EQ}(i) , \quad i \in \{1, 2\} . \]  
(73)
It can be shown that the product \( S_1 S_1 \) equals the covariance matrix of the noise term \((1 + h_1 B_1^*) Z_1 + h_1 B_2 Z_2\), and the product \( S_2 S_2 \) equals the covariance matrix of the noise term \((1 + h_2 B_2^*) Z_2 + h_2 B_1 Z_1\). Thus, by the independence of the symbols \( \Xi_1 \) and \( \Xi_2 \) with the noise vectors \( Z_1 \) and \( Z_2 \),
\[ \text{Var}(\hat{\Xi}_i(n-1)) = \text{Var}(\Xi_1 - \hat{\Xi}_i(n-1)) + \|v_i S_i\|^2 \]
\[ \leq \gamma_i^{-2n} + \|S_i u_i\|^2 \frac{\delta_i^2}{P_i} \gamma_i^{-2(n-1)} \]
\[ \leq \gamma_i^{-2n} + \|Q_i u_i\|^2 \frac{\delta_i^2}{P_i} \gamma_i^{-2(n-1)} \]
\[ \leq \gamma_i^{-2n} + nP \frac{\delta_i^2}{P_i} \gamma_i^{-2(n-1)} \]
\[ = \left( 1 + nP \frac{\delta_i^2}{P_i} \right) \gamma_i^{-2n} , \]  
(74)
where in the first inequality we used (53a) and (65); in the second inequality we used that since \( Q_i \) and \( S_i \) are symmetric, \( Q_i = Q_i^T \), and thus by (73): \( S_i = Q_i \); and in the last inequality we used (49b).

By (29) we conclude that Receiver \( i \)’s estimation error on message point \( \Theta_i(M_i) \) is of variance
\[ \text{Var}\left(\Theta_i(M_i) - \hat{\Theta}_i\right) \leq \frac{1}{h_i^2(1 - \rho^*)} \left( 1 + nP \frac{\delta_i^2}{P_i} \right) \gamma_i^{-2n} . \]

The nearest-neighbor decoding rule (26) produces the correct estimate whenever \( |\Theta_i(M_i) - \hat{\Theta}_i| < \Delta_i/2 \), where \( \Delta_i \) is defined in (17). Since by (29), the estimation error \( \Theta_i(M_i) - \hat{\Theta}_i \) is zero-mean Gaussian, this happens with probability
\[ 1 - 2Q\left( \frac{\Delta_i/2}{\text{Var}(\Theta_i(M_i) - \hat{\Theta}_i)} \right) \]
and thus Receiver \( i \)’s probability of error is bounded by
\[ \text{Pr}\left[M_i \neq M_i\right] \leq 2Q\left( \frac{\sqrt{P_i}}{2|\alpha n R_i|} \left( 1 + nP \frac{\delta_i^2}{P_i} \right)^{-1} \gamma_i^{-2n} \right) . \]

This probability tends to 0 as \( n \to \infty \) whenever
\[ R_i < \frac{1}{2} \log(\gamma_i^2) , \quad i \in \{1, 2\} , \]  
(75)
which is equivalent to (30) and concludes the proof.

**Appendix A**: Sketch of Proof of Theorem 1

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

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

The idea of the schemes is to divide the blocklength \( n \) into subblocks of length \( \eta \) and to apply an inner linear-feedback code parametrized by \( (A_1, A_2) \) (or by \( (C_1, C_2) \) to transform each subblock of \( \eta \) channel uses of the scalar channel into a single use of a new super MIMO channel. An optimal MIMO nocode code is then used to communicate over the resulting super BC (or MAC) without using the feedback.

Now, choose for the BC scheme and the MAC scheme the same parameter \( \eta \) and feedback matrices that satisfy
\[ A_i = \tilde{C}_i , \quad i \in \{1, 2\} . \]  
(76)
We shall show, using nocode MIMO MAC-BC duality [16], [17], that the resulting super MIMO MAC and BC have the same capacity regions. By taking the union over all possible parameters, this will conclude the proof of the theorem.

Consider first the BC and define the matrices
\[ B_i \triangleq (1 - h_1 A_1 - h_2 A_2)^{-1} A_i , \quad i \in \{1, 2\} . \]  
(77)
The channel inputs in a subblock can be written as
\[
X = (I - h_1 A_1 - h_2 A_2) U + A_1 Y_1 + A_2 Y_2
\]
\[= U + B_1 Z_1 + B_2 Z_2\]  \hfill (78)
for some code vector \(U\) that is average power constrained to
\[
\eta P - \text{tr}(B_1^* B_1^T) - \text{tr}(B_2^* B_2^T) - 2 \lambda \text{tr}(B_1^* B_2^T)
\].  \hfill (80)

Note that the preceding factor \((I - h_1 A_1 - h_2 A_2)\) in (78) is used only to simplify calculations. This definition remains general.

The corresponding channel outputs are
\[
Y_1 = h_1 U + (I + h_1 B_1) Z_1 + h_1 B_2 Z_2, \quad (81a)
\]
\[
Y_2 = h_2 U + (I + h_2 B_2) Z_2 + h_2 B_1 Z_1. \quad (81b)
\]

Let \(S_1\) and \(S_2\) denote the unique positive square-roots of the the positive definite noise covariance matrices
\[
N_1 \triangleq (I + h_1 B_1^T)(I + h_1 B_1) + h_1 B_2^T(h_1 B_2)^T + \lambda (I + h_1 B_1)(h_1 B_1^T)^T + \lambda (h_1 B_2)(I + h_1 B_1^T)^T, \quad (82a)
\]
\[
N_2 \triangleq (I + h_2 B_2^T)(I + h_2 B_2) + h_2 B_1^T(h_2 B_1)^T + \lambda (I + h_2 B_2)(h_2 B_2^T)^T + \lambda (h_2 B_1)(I + h_2 B_2^T)^T. \quad (82b)
\]

Define \(\tilde{U} = EU\) and notice that \(\|\tilde{U}\|^2 = \|U\|^2\). Since \(ES_1\) and \(ES_2\) are invertible, the following MIMO BC has the same capacity region as the MIMO BC in (81):
\[
Y_i = h_i ES_i^{-1} E\tilde{U} + \tilde{Z}_i, \quad i \in \{1, 2\}, \quad (83)
\]
where \(\tilde{Z}_1\) and \(\tilde{Z}_2\) denote independent centered Gaussian vectors of identity covariance matrix \(E\) and \(U\) are constrained to (80).

Now we turn to the MAC and choose feedback matrices that satisfy (76). Define the matrices
\[
D_i \triangleq C_i (I_\eta - h_i C_1 - h_2 C_2)^{-1}, \quad (84)
\]
and let \(Q_1\) and \(Q_2\) denote the unique positive square roots of the positive definite matrices
\[
M_1 \triangleq (I + h_1 D_1^T)(I + h_1 D_1) + (h_1 D_2)^T(h_1 D_2) + \lambda (I + h_1 D_1)(h_1 D_2^T)^T + \lambda (h_1 D_2)(I + h_1 D_1^T)^T, \quad (85a)
\]
\[
M_2 \triangleq (h_2 D_2^T)(h_2 D_2) + (I + h_2 D_2^T)(I + h_2 D_2) + \lambda (I + h_2 D_2^T)(h_2 D_2^T)^T + \lambda (h_2 D_1)(I + h_2 D_2^T)^T. \quad (85b)
\]

The channel inputs in a subblock are given by
\[
X_i = Q_i^{-1} U_i + C_i Y_i, \quad i \in \{1, 2\}, \quad (86)
\]
for some code-vectors \(U_1\) and \(U_2\) that are average sum-power constrained to
\[
\eta P - \text{tr}(D_1 D_1^T) - \text{tr}(D_2 D_2^T) - 2 \lambda \text{tr}(D_1 D_2^T). \quad (87)
\]
The corresponding channel output vector can be written as
\[
Y = (I + h_1 D_1 + h_2 D_2)(h_1 Q_1^{-1} U_1 + h_2 Q_2^{-1} U_2 + Z). \quad (88)
\]
Since \((I + h_1 D_1 + h_2 D_2)\) is invertible, under the same power constraint, the MIMO MAC described by (88) has the same capacity region as
\[
\tilde{Y} = h_1 Q_1^{-1} U_1 + h_2 Q_2^{-1} U_2 + Z. \quad (89)
\]