Broadcasting Information subject to State Masking over a MIMO State Dependent Gaussian Channel

Michael Dikshtein*, Anelia Somekh-Baruch † and Shlomo Shamai (Shitz)*
*Department of EE, Technion, Haifa 32000, Israel, {michaeldic@campus,sshlomo@ee}.technion.ac.il
†Faculty of Engineering, Bar-Ilan University, Ramat-Gan, 52900, Israel, anelia.somekhbaruch@gmail.com

***Abstract*—The problem of channel coding over the Gaussian multiple-input multiple-output (MIMO) broadcast channel (BC) with additive independent Gaussian states is considered. The states are known in a noncausal manner to the encoder, and it wishes to minimize the amount of information that the receivers can learn from the channel outputs about the state sequence. The state leakage rate is measured as a normalized blockwise mutual information between the state sequence and the channel outputs' sequences. We employ a new version of a state-dependent extremal inequality and show that Gaussian input maximizes the state-dependent version of Marton's outer bound. Further, we show that our inner bound coincides with the outer bound. Our result generalizes previously studied scalar Gaussian BC with state and MIMO BC without the state.**

***Index Terms*—Dirty paper coding, Gelf'and-Pinsker scheme, noncausal CSI, Broadcast channel, state masking, extremal inequality, enhanced channel, entropy power inequality.**

I. Introduction

We consider the problem of reliable transmission of purely digital information over a two-user MIMO Gaussian BC with an additive interference, modeled as state, which is known in a noncausal manner to the encoder. In our setting, we impose an additional requirement to reduce the knowledge of the receivers regarding the state, measured as a normalized blockwise mutual information between the state sequence and the received sequences, as depicted in Figure 1. The problem under consideration can act as a simplified model to many evolving practical communication systems. Consider, for example, a base station (BS) which is equipped with multiple antennas while the cell is partitioned to various sectors. The BS wishes to prevent leakage of information between the sectors, that is, to minimize the knowledge of mobile users in a specific sector regarding the messages intended for other sectors. In such a scenario, the part of the BS signal intended for other sectors is modeled as a state sequence. The state sequence is known to the BS in a noncausal manner since it is the one that generates it.

Problems of information transmission over channels with a noncausal channel state information (CSI) have been the subject for extensive study. The single-letter expression for the capacity of the point-to-point discrete memoryless channel (DMC) with noncausal CSI at the encoder (the G-P channel) was derived in the seminal work of Gel'fand and Pinsker [1]. One of the most interesting special cases of the G-P channel is the Gaussian additive noise and interference setting in which the additive interference plays the role of the state sequence, which is known noncausally to the transmitter. Costa showed in [2] that the capacity of this channel is equal to the capacity of the same channel without additive interference. Cohen and Lapidoth [3] showed that any interference sequence can be totally removed when the channel noise is ergodic and Gaussian.

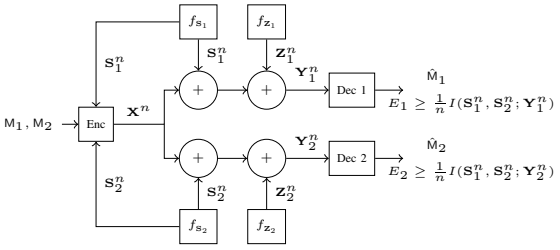

Fig. 1. System model for general BC subject to state masking constraints.

The general DM-BC was introduced by Cover [4]. The capacity region of the DM-BC is still an open problem. The largest known inner bound on the capacity region of the DM-BC was derived by Marton [5]. The best outer bound for DM-BC with common and private messages is due to Nair and El Gamal [6]. There are, however, some special cases in which the capacity region is fully characterized. For example, the capacity region of the Gaussian BC was derived by Bergmans [7], where the conditional version of the Entropy Power Inequality (EPI) was utilized. An interesting result is the capacity region of the Gaussian MIMO BC which was established by Weingarten et al. [8]. The authors showed that Bergmans' technique cannot be directly applied to the MIMO scenario since a certain proportionality condition is not always satisfied, hence they introduced a new notion of *an enhanced channel* and used it jointly with the EPI to show their result. Liu and Viswanath [9] developed *an extremal inequality* proof technique and showed that it can be used to establish a converse result in various Vector Gaussian multiterminal networks, including the Gaussian MIMO BC with private messages. The general DM-BC with a noncausal CSI at the encoder was studied in [10]. An inner bound was derived, and it was shown to be tight for the Gaussian BC with independent additive interference at both channels. Outer bounds for DM-BC with CSI at the encoder were derived in [11].

Leakage of information in terms of mutual information is a classical measure. The problem of state-masking and infor-

mation rate trade-off was introduced in [12]. In that work, the state sequence was treated as undesired information that leaks to the receiver and is known to the transmitter. The measure of the ability of the receiver to learn about the state from the received sequence was defined as the normalized blockwise mutual information between the state sequence S^n and the received sequence Y^n, that is, $I(\mathsf{S}^n;\mathsf{Y}^n)/n$. The concept of state amplification is a dual problem to state masking. Kim et al. [13] considered the problem of transmitting data at rate R over a DMC with random parameters and CSI at the encoder and simultaneously conveying the information about the channel state itself to the receiver. Courtade [14] considered a joint scenario, with two-encoder source coding setting where one source is to be amplified, while the other source is to be masked. Koyluoglu et al. [15] considered a state-dependent BC with state sequence known in a noncausal manner to Alice (the transmitter), and its goal is to effectively convey the state to Bob (receiver 1) while "masking" it from Eve (receiver 2). Liu and Chen [16] considered the problem of message transmission and state estimation over the Gaussian BC, where both received signals interfered by same additive Gaussian state. Grover and Sahai [17] related the problem of state masking to Witsenhausen's Counter-example [18]. A privacy-constrained information extraction problem was recently considered by Asoodeh et al. [19]. A good tutorial on channel coding in the presence of CSI that also covers the state masking setting can be found in [20].

In our previous work [21], we extended the state masking scenario to the state-dependent DM-BC with noncausal CSI at the encoder. We developed inner and outer bounds and showed that they are tight for a special case of zero-rates transmission and the scalar Gaussian BC with additive state. As for this work, we consider the vector setting. Even though we use the same single-letter expressions, the evaluation of the capacity region for the MIMO case is not straightforward. We propose a new approach for choosing the auxiliary random variables in order to evaluate the inner bound. Furthermore, our technique to derive the optimal coefficients can be applied to other related problems in network information theory, where a Gaussian signal is considered. Moreover, we develop a new, conditional, extremal inequality, and show that a Gaussian input distribution is optimal for the vector setting.

II. Notations and Problem Formulation

Throughout the paper, random variables are denoted using a sans-serif font, e.g., X, their realizations are denoted by the respective lower case letters, e.g., x, and their alphabets are denoted by the respective calligraphic letter, e.g., $\mathcal{X}$. Let $\mathcal{X}^n$ stand for the set of all n-tuples of elements from $\mathcal{X}$. An element from $\mathcal{X}^n$ is denoted by $x^n = (x_1, x_2, \ldots, x_n)$. The probability density function of X, the joint density function of X and Y, and the conditional density of X given Y are denoted by f_{X}, f_{XY} and $f_{\mathsf{X}|\mathsf{Y}}$ respectively. The expectation of X is denoted by $\mathbb{E}[\mathsf{X}]$. The cross-covariance matrix of two random vectors $\mathbf{X}$ and $\mathbf{Y}$ is denoted as $\Sigma_{\mathbf{XY}} \triangleq \mathbb{E}\left[\mathbf{X}\mathbf{Y}^T\right]$. The probability of an event $\mathcal{E}$ is denoted as $\mathbb{P}(\mathcal{E})$. A set of consecutive integers starting at 1 and ending in 2^{nR} is denoted as $\mathcal{I}_R^{(n)} \triangleq \{1, 2, \ldots, 2^{nR}\}$.

An $(2^{nR_1}, 2^{nR_2}, n)$ code for the broadcast channel with state sequence S^n known noncausally at the encoder consists of

- two message sets $\mathcal{I}_{R_1}^{(n)}$ and $\mathcal{I}_{R_2}^{(n)}$,
- an encoder that assigns a codeword $x^n(m_1, m_2, s^n)$ to each message-state triple $(m_1, m_2, s^n) \in \mathcal{I}_{R_1}^{(n)} \times \mathcal{I}_{R_2}^{(n)} \times \mathcal{S}^n$,
- two decoders, where decoder k assigns an estimate $\hat{m}_k \in \mathcal{I}_{R_k}^{(n)}$ to each received sequence y_k^n, $k \in \{1, 2\}$.

Let $\hat{\mathsf{M}}_1$ and $\hat{\mathsf{M}}_2$ denote the outputs of decoder 1 and decoder 2, respectively. We assume that the message pair $(\mathsf{M}_1, \mathsf{M}_2)$ is uniformly distributed over $\mathcal{I}_{R_1}^{(n)} \times \mathcal{I}_{R_2}^{(n)}$. The average probability of error is defined as $P_e^{(n)} = \mathbb{P}\left(\hat{\mathsf{M}}_1 \neq \mathsf{M}_1, \hat{\mathsf{M}}_2 \neq \mathsf{M}_2\right)$.

We are interested in the interplay between reliable coding at rate pairs (R_1, R_2) which we would like to keep as high as possible and the (normalized) mutual informations $I(\mathbf{S}^n;\mathbf{Y}_1^n)/n$ and $I(\mathbf{S}^n;\mathbf{Y}_2^n)/n$, which we would like to make as small as possible.

Definition 1. For a given covariance matrix K, a quadruple (R_1, R_2, E_1, E_2) is said to be achievable if for every $\epsilon > 0$ and sufficiently large n, there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that the following conditions are simultaneously satisfied: $\frac{1}{n}\sum_{i=1}^n \mathbf{X}_i\mathbf{X}_i^T \preceq K$, $P_e^{(n)} \leq \epsilon$, and, $\frac{1}{n} I(\mathbf{S}^n;\mathbf{Y}_k^n) \leq E_k + \epsilon$, $k = 1, 2$. The achievable region $\mathcal{R}(K)$ is the closure of the set of all achievable quadruples (R_1, R_2, E_1, E_2).

III. Preliminaries

We use inner and outer bounds that were derived in [21] for the general DM-BC with random parameters and particularly utilize the private messages only case by setting $\mathsf{W} = \emptyset$.

Lemma 1 (Proposition 1 in [21])**.** *An achievable region $\mathcal{R}(K)$ consists of a quadruple (R_1, R_2, E_1, E_2) that satisfies the following conditions*

$$R_1 \leq I(\mathsf{U};\mathbf{Y}_1) - I(\mathsf{U};\mathbf{S}), \tag{1a}$$

$$R_2 \leq I(\mathsf{V};\mathbf{Y}_2) - I(\mathsf{V};\mathbf{S}), \tag{1b}$$

$$R_1 + R_2 \leq I(\mathsf{U};\mathbf{Y}_1) - I(\mathsf{U};\mathbf{S}) + I(\mathsf{V};\mathbf{Y}_2) - I(\mathsf{V};\mathbf{S}) - I(\mathsf{U};\mathsf{V}|\mathbf{S}) \tag{1c}$$

$$E_1 \geq I(\mathbf{S};\mathsf{U},\mathbf{Y}_1), \quad E_2 \geq I(\mathbf{S};\mathsf{V},\mathbf{Y}_2), \tag{1d}$$

for some pdf $f_{\mathbf{S}\mathsf{U}\mathsf{V}\mathbf{X}\mathbf{Y}_1\mathbf{Y}_2} = f_{\mathbf{S}} f_{\mathsf{U}\mathsf{V}\mathbf{X}|\mathbf{S}} f_{\mathbf{Y}_1\mathbf{Y}_2|\mathbf{X}\mathbf{S}}$.

The main idea behind the proof of the inner bound is an integration of the Marton and the G-P coding schemes, where for each message, a subcodebook is generated, whose size is large enough such that for every state sequence s^n, a jointly typical auxiliary codeword can be found in the subcodebook.

Next, we provide the outer bound on $\mathcal{R}(K)$.

Lemma 2 (Proposition 2 in [21])**.** *If a rate quadruple (R_1, R_2, E_1, E_2) is achievable for the DM-BC with random parameters and CSI known noncausally at the transmitter,*

then there exists a distribution $f_{\mathsf{U}\mathsf{V}\mathbf{X}|\mathbf{S}}$ such that the following inequalities are satisfied:

$$R_1 \le I(\mathsf{U};\mathbf{Y}_1|\mathbf{S}), \tag{2a}$$
$$R_2 \le I(\mathsf{V};\mathbf{Y}_2|\mathbf{S}), \tag{2b}$$
$$R_1+R_2 \le I(\mathsf{U};\mathbf{Y}_1|\mathbf{S})+I(\mathbf{X};\mathbf{Y}_2|\mathsf{U},\mathbf{S}), \tag{2c}$$
$$R_1+R_2 \le I(\mathbf{X};\mathbf{Y}_1|\mathsf{V},\mathbf{S})+I(\mathsf{V};\mathbf{Y}_2|\mathbf{S}), \tag{2d}$$
$$E_k \ge I(\mathbf{S};\mathbf{Y}_k) \quad k=1,2, \tag{2e}$$

where $f_{\mathbf{S}\mathsf{U}\mathsf{V}\mathbf{X}\mathbf{Y}_1\mathbf{Y}_2} = f_{\mathbf{S}} f_{\mathsf{U}\mathsf{V}\mathbf{X}|\mathbf{S}} f_{\mathbf{Y}_1\mathbf{Y}_2|\mathbf{X}\mathbf{S}}$.

IV. MIMO Gaussian Broadcast Channel

Our goal in this paper is to characterize the achievable region for the MIMO Gaussian State-Dependent BC with masking constraints. We show that a Gaussian input distribution maximizes the outer bound in Lemma 2. In order to show this a new proof technique is needed and we show the motivation to develop such technique.

The general two-user MIMO Gaussian BC with state [8], is an additive interference and noise channel where each time sample can be represented using the following equations:

$$\mathbf{Y}_k = \mathbf{X}+\mathbf{S}_k+\mathbf{Z}_k, \quad k\in\{1,2\},$$

where $\mathbf{X}$, $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{Z}_1$, $\mathbf{Z}_2$ are all real vectors of size $t\times 1$ and

- $\mathbf{X}$ is the input vector whose covariance matrix satisfies $\mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right]\preceq K$ for some $K\succeq 0$,
- $\mathbf{Y}_k$ is the output vector, $k\in\{1,2\}$,
- $\mathbf{S}_k$ is a real Gaussian random vector with zero mean and a covariance matrix $K_{\mathbf{S}_k}=\mathbb{E}\left[\mathbf{S}_k\mathbf{S}_k^T\right]\succeq 0$,
- $\mathbf{Z}_k$ is a real Gaussian random vector with zero mean and a covariance matrix $K_{\mathbf{Z}_k}=\mathbb{E}\left[\mathbf{Z}_k\mathbf{Z}_k^T\right]\succeq 0$,
- $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are mutually independent.

In the following we evaluate the bounds from Lemma 1 and Lemma 2 for the MIMO Gaussian BC setting. We first state our main result in the subsequent theorem.

Let,

$$K_{\mathbf{S}} \triangleq \begin{pmatrix} K_{\mathbf{S}_1} & 0 \\ 0 & K_{\mathbf{S}_2}\end{pmatrix}. \tag{3}$$

Theorem 1. *A rate-leakage region of the MIMO Gaussian State-Dependent BC with private messages is the quadruple (R_1,R_2,E_1,E_2) such that*

$$R_1 \le \frac{1}{2}\log\frac{|K_{\mathbf{X}_1}+K_{\mathbf{Z}_1}|}{|K_{\mathbf{Z}_1}|}, \tag{4}$$
$$R_2 \le \frac{1}{2}\log\frac{|K-\Sigma_{\mathbf{X}\mathbf{S}}K_{\mathbf{S}}^{-1}\Sigma_{\mathbf{X}\mathbf{S}}^T+K_{\mathbf{Z}_2}|}{|K_{\mathbf{X}_1}+K_{\mathbf{Z}_2}|}, \tag{5}$$
$$E_1 = \frac{1}{2}\log\frac{|K+\Sigma_{\mathbf{X}\mathbf{S}_1}+\Sigma_{\mathbf{X}\mathbf{S}_1}^T+K_{\mathbf{S}_1}+K_{\mathbf{Z}_1}|}{|K-\Sigma_{\mathbf{X}\mathbf{S}}K_{\mathbf{S}}^{-1}\Sigma_{\mathbf{X}\mathbf{S}}^T+K_{\mathbf{Z}_1}|}, \tag{6}$$
$$E_2 = \frac{1}{2}\log\frac{|K+\Sigma_{\mathbf{X}\mathbf{S}_2}+\Sigma_{\mathbf{X}\mathbf{S}_2}^T+K_{\mathbf{S}_2}+K_{\mathbf{Z}_2}|}{|K-\Sigma_{\mathbf{X}\mathbf{S}}K_{\mathbf{S}}^{-1}\Sigma_{\mathbf{X}\mathbf{S}}^T+K_{\mathbf{Z}_2}|}, \tag{7}$$

for some covariance matrices $(K_{\mathbf{X}_1},\Sigma_{\mathbf{X}\mathbf{S}_1},\Sigma_{\mathbf{X}\mathbf{S}_2})$, such that $0\preceq K_{\mathbf{X}_1}\preceq K-\Sigma_{\mathbf{X}\mathbf{S}}K_{\mathbf{S}}^{-1}\Sigma_{\mathbf{X}\mathbf{S}}^T$, where

$$\Sigma_{\mathbf{X}\mathbf{S}} \triangleq \begin{pmatrix}\Sigma_{\mathbf{X}\mathbf{S}_1} & \Sigma_{\mathbf{X}\mathbf{S}_2}\end{pmatrix}. \tag{8}$$

The information rate region in (4) and (5) is similar to the MIMO BC without state in [8]. The main difference is that part of the transmitted signal, reflected by the covariance matrix $\Sigma_{\mathbf{X}\mathbf{S}}$, is utilized to mask the state sequence.

Remark 1. *Our model and the definition of the leakage in terms of normalized mutual information defines in fact channels between the states and the receivers. Consider the structured states scenario where the sequences are drawn from some codebook with a given rate. The results of this work imply that if the rate of the codebook is higher than the leakage, reliable decoding of the structured states is impossible.*

V. Proof of Theorem 1

Denote $\mathbf{S}=(\mathbf{S}_1^T,\mathbf{S}_2^T)^T$.

A. Proof of the converse part of Theorem 1

Consider the RHS of (2b),

$$\begin{aligned} I(\mathsf{V};\mathbf{Y}_2|\mathbf{S}) &= I(\mathsf{V},\mathbf{X};\mathbf{Y}_2|\mathbf{S})-I(\mathbf{X};\mathbf{Y}_2|\mathsf{V},\mathbf{S}) \\ &\overset{(a)}{=} I(\mathbf{X};\mathbf{Y}_2|\mathbf{S})-I(\mathbf{X};\mathbf{Y}_2|\mathsf{V},\mathbf{S}),\end{aligned}$$

where (a) follows since $\mathbf{Y}_2=\mathbf{X}+\mathbf{S}_2+\mathbf{Z}_2$ and hence $h(\mathbf{Y}_2|\mathsf{V},\mathbf{X},\mathbf{S})=h(\mathbf{Z}_2)=h(\mathbf{Y}_2|\mathbf{X},\mathbf{S})$. Define

$$\mathcal{A}\triangleq\{f_{\mathsf{v}\mathbf{x}|\mathbf{s}}:\mathsf{V}\to(\mathbf{X},\mathbf{S})\to(\mathbf{Y}_1,\mathbf{Y}_2),\mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right]\preceq K\}.$$

The weighted sum rate upper bound (2d) can be written as

$$\begin{aligned} &R_1+\mu R_2 \\ &\le \mu I(\mathbf{X};\mathbf{Y}_2|\mathbf{S})+I(\mathbf{X};\mathbf{Y}_1|\mathsf{V},\mathbf{S})-\mu I(\mathbf{X};\mathbf{Y}_2|\mathsf{V},\mathbf{S}) \\ &\le \sup_{\mathcal{A}}\{\mu I(\mathbf{X};\mathbf{Y}_2|\mathbf{S})+I(\mathbf{X};\mathbf{Y}_1|\mathsf{V},\mathbf{S})-\mu I(\mathbf{X};\mathbf{Y}_2|\mathsf{V},\mathbf{S})\} \\ &\le \sup_{\mathcal{A}}\{\mu I(\mathbf{X};\mathbf{Y}_2|\mathbf{S})\}+\sup_{\mathcal{A}}\{I(\mathbf{X};\mathbf{Y}_1|\mathsf{V},\mathbf{S})-\mu I(\mathbf{X};\mathbf{Y}_2|\mathsf{V},\mathbf{S})\},\end{aligned}$$

where $\mu>1$.

The first term can be upper bounded as follows:

$$I(\mathbf{X};\mathbf{Y}_2|\mathbf{S}) \le \frac{1}{2}\log\frac{\left|K-\Sigma_{\mathbf{X}\mathbf{S}}K_{\mathbf{S}}^{-1}\Sigma_{\mathbf{X}\mathbf{S}}^T+K_{\mathbf{Z}_2}\right|}{|K_{\mathbf{Z}_2}|}. \tag{9}$$

As for the difference between mutual informations, we obtain

$$\begin{aligned} &I(\mathbf{X};\mathbf{Y}_1|\mathsf{V},\mathbf{S})-\mu I(\mathbf{X};\mathbf{Y}_2|\mathsf{V},\mathbf{S}) \\ &=h(\mathbf{X}+\mathbf{Z}_1|\mathsf{V},\mathbf{S})-h(\mathbf{Z}_1)-\mu\left(h(\mathbf{X}+\mathbf{Z}_2|\mathsf{V},\mathbf{S})-h(\mathbf{Z}_2)\right).\end{aligned} \tag{10}$$

Consider the following optimization problem, denoted as P:

$$\sup_{\mathcal{A}}\{h(\mathbf{X}+\mathbf{Z}_1|\mathsf{V},\mathbf{S})-\mu h(\mathbf{X}+\mathbf{Z}_2|\mathsf{V},\mathbf{S})\}. \tag{11}$$

We would like to show using a conditional version of an extremal inequality that a conditional Gaussian distribution $f_{\mathbf{X}|\mathsf{V}\mathbf{S}}\sim\mathcal{N}(0,K_{\mathbf{X}_1})$ maximizes (11). Assume $K_{\mathbf{Z}_1}\preceq K_{\mathbf{Z}_2}$. Let $\mathbf{Z}$ be such that $\mathbf{Z}_2=\mathbf{Z}_1+\mathbf{Z}$, where $\mathbf{Z}\sim\mathcal{N}(0,K_{\mathbf{Z}})$ and $K_{\mathbf{Z}}=K_{\mathbf{Z}_2}-K_{\mathbf{Z}_1}$. The main tool used in the proof is the conditional EPI by Bergmans [7], for which equality in

$$e^{\frac{2}{t}h(\mathbf{X}+\mathbf{Z}_1+\mathbf{Z}|\mathsf{V},\mathbf{S})} \ge e^{\frac{2}{t}h(\mathbf{X}+\mathbf{Z}_1|\mathsf{V},\mathbf{S})}+e^{\frac{2}{t}h(\mathbf{Z})} \tag{12}$$

holds iff $f_{\mathbf{X}|\mathsf{V}\mathbf{S}}\sim\mathcal{N}(0,K_{\mathbf{X}_1})$ with the same $K_{\mathbf{X}_1}$ for every $(\mathsf{V}=v,\mathbf{S}=\mathbf{s})$ and $K_{\mathbf{X}_1}+K_{\mathbf{Z}_1}$ is proportional to $K_{\mathbf{Z}}$. The

problem is that the proportionality condition is not always satisfied.

We start by restricting the solution space of P to be Gaussian. Denote the respective optimization problem as P_G. Hence (11) becomes

$$\max_{K_{\mathbf{X}} \preceq K} \frac{1}{2}\log\left((2\pi e)^t|K_{\mathbf{X}}+K_{\mathbf{Z}_1}|\right)-\frac{\mu}{2}\log\left((2\pi e)^t|K_{\mathbf{X}}+K_{\mathbf{Z}_2}|\right). \quad (13)$$

The optimal solution of P_G, $K^\star_{\mathbf{X}}$, must satisfy the following KKT-like conditions

$$\frac{1}{2}\left(K^\star_{\mathbf{X}}+K_{\mathbf{Z}_1}\right)^{-1}+M_1=\frac{\mu}{2}\left(K^\star_{\mathbf{X}}+K_{\mathbf{Z}_2}\right)^{-1}+M_2, \quad (14)$$
$$M_1K^\star_{\mathbf{X}}=0, \quad M_2(K-K^\star_{\mathbf{X}})=0.$$

Obviously $(P_G) \leq (P)$. We now introduce the enhanced channel with additive state

$$\tilde{\mathbf{Y}}_k=\mathbf{X}+\mathbf{S}_k+\tilde{\mathbf{Z}}_k, \quad k\in\{1,2\},$$

where $\tilde{\mathbf{Z}}_k \sim \mathcal{N}(0,K_{\tilde{\mathbf{Z}}_k})$, $k\in\{1,2\}$ are constructed such that $K_{\tilde{\mathbf{Z}}_1}$ and $K_{\tilde{\mathbf{Z}}_2}$, be two real symmetric matrices satisfying

$$\frac{1}{2}\left(K^\star_{\mathbf{X}}+K_{\mathbf{Z}_1}\right)^{-1}+M_1=\frac{1}{2}\left(K^\star_{\mathbf{X}}+K_{\tilde{\mathbf{Z}}_1}\right)^{-1}, \quad (15a)$$

$$\frac{\mu}{2}\left(K^\star_{\mathbf{X}}+K_{\mathbf{Z}_2}\right)^{-1}+M_2=\frac{\mu}{2}\left(K^\star_{\mathbf{X}}+K_{\tilde{\mathbf{Z}}_2}\right)^{-1}. \quad (15b)$$

We define the following auxiliary optimization problem $\tilde{P}$ with optimum value $(\tilde{P})$

$$\sup_{\mathcal{A}} h(\mathbf{X}+\tilde{\mathbf{Z}}_1|\mathsf{V},\mathbf{S})-\mu h(\mathbf{X}+\tilde{\mathbf{Z}}_2|\mathsf{V},\mathbf{S})+F. \quad (16)$$

The constant F is defined as

$$F\triangleq h(\mathbf{Z}_1)-h(\tilde{\mathbf{Z}}_1)+\mu(h(\mathbf{X}_{G_K}+\tilde{\mathbf{Z}}_2)-h(\mathbf{X}_{G_K}+\mathbf{Z}_2)),$$

where $\mathbf{X}_{G_K}\sim\mathcal{N}(0,K)$, $\tilde{\mathbf{Z}}_1\sim\mathcal{N}(0,K_{\tilde{\mathbf{Z}}_1})$, $\tilde{\mathbf{Z}}_2\sim\mathcal{N}(0,K_{\tilde{\mathbf{Z}}_1})$, and $K_{\tilde{\mathbf{Z}}_2}\succeq K_{\tilde{\mathbf{Z}}_1}$.

It was shown in [9] that $0 \preceq K_{\tilde{\mathbf{Z}}_1} \preceq K_{\mathbf{Z}_1}$ and $K_{\tilde{\mathbf{Z}}_1} \preceq K_{\tilde{\mathbf{Z}}_2} \preceq K_{\mathbf{Z}_2}$, hence we can write $\tilde{\mathbf{Z}}_2=\tilde{\mathbf{Z}}_1+\tilde{\mathbf{Z}}$, where $\tilde{\mathbf{Z}}\sim\mathcal{N}(0,K_{\tilde{\mathbf{Z}}})$ and $K_{\tilde{\mathbf{Z}}}=K_{\tilde{\mathbf{Z}}_2}-K_{\tilde{\mathbf{Z}}_1}$. By substituting (15) into the KKT-like condition (14) we have

$$K^\star_{\mathbf{X}}+K_{\tilde{\mathbf{Z}}_1}=(\mu-1)^{-1}K_{\tilde{\mathbf{Z}}},$$

and hence the proportionality condition in EPI is satisfied for the enhanced channel. The difference of differential entropy terms in (16) can be upper bounded using EPI as

$$\begin{aligned}
&h(\mathbf{X}+\tilde{\mathbf{Z}}_1|\mathsf{V},\mathbf{S})-\mu h(\mathbf{X}+\tilde{\mathbf{Z}}_2|\mathsf{V},\mathbf{S})\\
&\overset{(a)}{\leq} f\left(h(\mathbf{X}+\tilde{\mathbf{Z}}_1|\mathsf{V},\mathbf{S}),h(\tilde{\mathbf{Z}})\right)\\
&\overset{(b)}{\leq} f\left(h(\tilde{\mathbf{Z}})-\frac{t}{2}\log(\mu-1),h(\tilde{\mathbf{Z}})\right),
\end{aligned}$$

where $f(a,b)=a-\frac{\mu t}{2}\log\left(e^{\frac{2}{t}a}+e^{\frac{2}{t}b}\right)$. Equality in (a) and (b) holds if $f_{\mathbf{X}|\mathsf{V}\mathbf{S}}\sim\mathcal{N}(0,K_{\mathbf{X}_1})$ with same $K_{\mathbf{X}_1}$ for all $(\mathsf{V}=v,\mathbf{S}=\mathrm{s})$ and by the proportionality condition. Hence the optimizing distribution of (16) is Gaussian and the optimal value equals to

$$\max_{K_{\mathbf{X}}\preceq K}\frac{1}{2}\log\left((2\pi e)^t|K_{\mathbf{X}}+K_{\mathbf{Z}_1}|\right)-\frac{\mu}{2}\log\left((2\pi e)^t|K_{\mathbf{X}}+K_{\mathbf{Z}_2}|\right),$$

which is the same optimal value as was found for P_G.

The last step is to show that the objective function of the original channel (11) is less or equal than the objective function of (16) for any fixed distribution $f_{\mathbf{X}|\mathsf{V}\mathbf{S}}$. Consider the difference of those functions

$$\begin{aligned}
&h(\mathbf{X}+\mathbf{Z}_1|\mathsf{V},\mathbf{S})-h(\mathbf{Z}_1)-h(\mathbf{X}+\tilde{\mathbf{Z}}_1|\mathsf{V},\mathbf{S})+h(\tilde{\mathbf{Z}}_1)\\
&\quad-\mu\left(h(\mathbf{X}+\mathbf{Z}_2|\mathsf{V},\mathbf{S})-h(\mathbf{X}_{G_K}+\mathbf{Z}_2)\right)\\
&\quad+\mu\left(h(\mathbf{X}+\tilde{\mathbf{Z}}_2|\mathsf{V},\mathbf{S})-h(\mathbf{X}_{G_K}+\tilde{\mathbf{Z}}_2)\right).
\end{aligned}$$

Since given $(\mathsf{V},\mathbf{S})$, we have the Markov chain $\mathbf{X}\to\mathbf{X}+\tilde{\mathbf{Z}}_1\to\mathbf{X}+\tilde{\mathbf{Z}}_1+\hat{\mathbf{Z}}_1$ and applying Data Processing Inequality [22, p. 24], we have

$$\begin{aligned}
&h(\mathbf{X}+\mathbf{Z}_1|\mathsf{V},\mathbf{S})-h(\mathbf{Z}_1)-h(\mathbf{X}+\tilde{\mathbf{Z}}_1|\mathsf{V},\mathbf{S})+h(\tilde{\mathbf{Z}}_1)\\
&=I(\mathbf{X};\mathbf{X}+\mathbf{Z}_1|\mathsf{V},\mathbf{S})-I(\mathbf{X};\mathbf{X}+\tilde{\mathbf{Z}}_1|\mathsf{V},\mathbf{S})\\
&=I(\mathbf{X};\mathbf{X}+\tilde{\mathbf{Z}}_1+\hat{\mathbf{Z}}_1|\mathsf{V},\mathbf{S})-I(\mathbf{X};\mathbf{X}+\tilde{\mathbf{Z}}_1|\mathsf{V},\mathbf{S})\leq 0,
\end{aligned}$$

Using $K_{\tilde{\mathbf{Z}}_2}\preceq K_{\mathbf{Z}_2}$, denote $\hat{\mathbf{Z}}_2\triangleq\mathbf{Z}_2-\tilde{\mathbf{Z}}_2$, hence

$$\begin{aligned}
&h(\mathbf{X}+\mathbf{Z}_2|\mathsf{V},\mathbf{S})-h(\mathbf{X}+\tilde{\mathbf{Z}}_2|\mathsf{V},\mathbf{S})-h(\mathbf{X}_{G_K}+\mathbf{Z}_2)+h(\mathbf{X}_{G_K}+\tilde{\mathbf{Z}}_2)\\
&=I(\hat{\mathbf{Z}}_2;\mathbf{X}+\tilde{\mathbf{Z}}_2+\hat{\mathbf{Z}}_2|\mathsf{V},\mathbf{S})-I(\hat{\mathbf{Z}}_2;\mathbf{X}_{G_K}+\tilde{\mathbf{Z}}_2+\hat{\mathbf{Z}}_2)\\
&\overset{(a)}{\geq} I(\hat{\mathbf{Z}}_2;\mathbf{X}_G+\tilde{\mathbf{Z}}_2+\hat{\mathbf{Z}}_2|\mathsf{V},\mathbf{S})-I(\hat{\mathbf{Z}}_2;\mathbf{X}_{G_K}+\tilde{\mathbf{Z}}_2+\hat{\mathbf{Z}}_2)\\
&\overset{(b)}{=} I(\hat{\mathbf{Z}}_2;\mathbf{X}_G+\tilde{\mathbf{Z}}_2+\hat{\mathbf{Z}}_2|\mathsf{V},\mathbf{S})-I(\hat{\mathbf{Z}}_2;\mathbf{X}_G+\hat{\mathbf{X}}_G+\tilde{\mathbf{Z}}_2+\hat{\mathbf{Z}}_2)\\
&\overset{(c)}{\geq} 0,
\end{aligned}$$

where inequality in (a) follows from Worst Additive Noise Lemma [23, Lemma II.2], (b) is due to $\mathbf{X}_{G_K}=\mathbf{X}_G+\hat{\mathbf{X}}_G$, $\hat{\mathbf{X}}_G$ is independent of $\mathbf{X}_G$. Inequality in (c) is again due to the Markov chain $\hat{\mathbf{Z}}_2\to\mathbf{X}_G+\tilde{\mathbf{Z}}_2+\hat{\mathbf{Z}}_2\to\mathbf{X}_G+\hat{\mathbf{X}}_G+\tilde{\mathbf{Z}}_2+\hat{\mathbf{Z}}_2$, and Data Processing Inequality [22, p. 24]. Thus, we have shown that the objective function of P is less or equal to the objective function of $\tilde{P}$ for any choice of $f_{\mathbf{X}|\mathsf{V}\mathbf{S}}$. To conclude, we have shown that $(P_G)=(\tilde{P})$. Furthermore $P_G\leq P\leq\tilde{P}$ for any choice of $f_{\mathbf{X}|\mathsf{V}\mathbf{S}}$. Thus, $f_{\mathbf{X}|\mathsf{V}\mathbf{S}}\sim\mathcal{N}(0,K_{\mathbf{X}_1})$ also solves P.

Finally, by collecting (9), (10) and (13), we obtain

$$\begin{aligned}
&R_1+\mu R_2\\
&\leq\frac{1}{2}\log\frac{|K_{\mathbf{X}_1}+K_{\mathbf{Z}_1}|}{|K_{\mathbf{Z}_1}|}+\frac{\mu}{2}\log\frac{\left|K-\Sigma_{\mathbf{X}\mathbf{S}}K_{\mathbf{S}}^{-1}\Sigma_{\mathbf{X}\mathbf{S}}^T+K_{\mathbf{Z}_2}\right|}{|K_{\mathbf{X}_1}+K_{\mathbf{Z}_2}|}.
\end{aligned}$$

Next we proceed to lower bound the leakage rates for $k\in\{1,2\}$,

$$I(\mathbf{S};\mathbf{Y}_k)=h(\mathbf{S})-h(\mathbf{S}|\mathbf{Y}_k). \quad (17)$$

The conditional differential entropy can be upper bounded as follows

$$h(\mathbf{S}|\mathbf{Y}_k)\leq\frac{1}{2}\log(2\pi e)^2\left|K_{\mathbf{S}}-\Sigma_{\mathbf{S}\mathbf{Y}_k}\Sigma_{\mathbf{Y}_k}^{-1}\Sigma_{\mathbf{S}\mathbf{Y}_k}^T\right|, \quad (18)$$

where $\Sigma_{\mathbf{S}\mathbf{Y}_k}=\Sigma_{\mathbf{X}\mathbf{S}}^T+\Sigma_{\mathbf{S}\mathbf{S}_k}$ and $\Sigma_{\mathbf{Y}_k}=K+\Sigma_{\mathbf{X}\mathbf{S}_k}+\Sigma_{\mathbf{X}\mathbf{S}_k}^T+K_{\mathbf{S}_k}+K_{\mathbf{Z}_k}$. Next, by applying Sylvester's Identity Theorem, it can be shown that

$$|K_{\mathrm{s}}-\Sigma_{\mathbf{S}\mathbf{Y}_k}\Sigma_{\mathbf{Y}_k}^{-1}\Sigma_{\mathbf{S}\mathbf{Y}_k}^T|=|K_{\mathrm{s}}||K-\Sigma_{\mathbf{x}\mathbf{s}}K_{\mathrm{s}}^{-1}\Sigma_{\mathbf{x}\mathbf{s}}^T+K_{\mathbf{z}_k}||\Sigma_{\mathbf{Y}_k}^{-1}|. \quad (19)$$

Gathering (17), (18), and (19), we obtain

$$I(\mathbf{S};\mathbf{Y}_k) \geq \frac{1}{2}\log\frac{|K+\Sigma_{\mathbf{XS}_k}+\Sigma^T_{\mathbf{XS}_k}+K_{\mathbf{S}_k}+K_{\mathbf{Z}_k}|}{|K-\Sigma_{\mathbf{XS}}K^{-1}_{\mathbf{S}}\Sigma^T_{\mathbf{XS}}+K_{\mathbf{Z}_k}|},$$

which concludes the proof of the converse part of Theorem 1.

B. Proof of the direct part of Theorem 1

Let

$$\begin{aligned}\mathbf{X} &= \mathbf{X}_1+\mathbf{X}_2+B_1\mathbf{S}_1+B_2\mathbf{S}_2,\\ \mathbf{U} &= \mathbf{X}_1+A_{10}\mathbf{X}_2+A_{11}\mathbf{S}_1+A_{12}\mathbf{S}_2,\\ \mathbf{V} &= \mathbf{X}_2+A_{21}\mathbf{S}_1+A_{22}\mathbf{S}_2,\end{aligned}$$

such that $\mathbf{X}_1 \sim \mathcal{N}(0,K_{\mathbf{X}_1})$, $\mathbf{X}_2 \sim \mathcal{N}(0,K_{\mathbf{X}_2})$, $\mathbf{S}_1 \sim \mathcal{N}(0,K_{\mathbf{S}_1})$ and $\mathbf{S}_2 \sim \mathcal{N}(0,K_{\mathbf{S}_2})$ are mutually independent. The achievability of Theorem 1 follows by evaluating (1) with the above choice of Gaussian random vectors and the following choice of matrix coefficients

$$\begin{aligned}A_{10} &= K_{\mathbf{X}_1}(K_{\mathbf{X}_1}+K_{\mathbf{Z}_1})^{-1},\\ A_{11} &= K_{\mathbf{X}_1}(K_{\mathbf{X}_1}+K_{\mathbf{Z}_1})^{-1}(B_1+I),\\ A_{12} &= K_{\mathbf{X}_1}(K_{\mathbf{X}_1}+K_{\mathbf{Z}_1})^{-1}B_2,\\ A_{21} &= K_{\mathbf{X}_2}(K_{\mathbf{X}_1}+K_{\mathbf{X}_2}+K_{\mathbf{Z}_2})^{-1}B_1,\\ A_{22} &= K_{\mathbf{X}_2}(K_{\mathbf{X}_1}+K_{\mathbf{X}_2}+K_{\mathbf{Z}_2})^{-1}(B_2+I)\\ B_k &= \Sigma_{\mathbf{XS}_k}\Sigma^{-1}_{\mathbf{S}_k} \qquad k\in\{1,2\}.\end{aligned}$$

The main idea for this choice of coefficients is to eliminate the state variables from the mutual information terms in Lemma 1. The complete proof is given in Appendix A of the extended version of this paper [24].

VI. Conclusions

We considered the problem of reliable communication over the Gaussian MIMO BC where each receiver is also interfered by the Gaussian independent interferences which are known to the encoder in a non-causal manner. We added an additional requirement to reduce the amount of information that *leaks* to the receivers regarding the states. The main contribution here is the optimization of the outer bound. While in the scalar setting [21], we showed that Gaussian signaling is optimal using the EPI, in the vector case, however, EPI can no longer be used since the covariance matrices of the noise vectors are not degraded in general. This is similar to calculating the capacity region of the MIMO BC, which is substantially different from the evaluation of the scalar broadcast channel. Hence, we derived a new state-dependent extremal inequality. Although this inequality is based on the original extremal inequality of [9] , it is far from being trivial, and it provides the tool needed to approach this MIMO setting. Moreover, it has not been used for state-dependent channels to the best of our knowledge. Furthermore, the standard results of point-to-point masking [12] and state-dependent BC [10] (no masking requirements), emerge as special cases of the bounds here. An extension to the MIMO Gaussian BC with an additional common message is under current study.

Acknowledgment

The work of M. Dikshtein and S. Shamai (Shitz) has been supported by the European Union's Horizon 2020 Research And Innovation Programme, grant agreement no. 694630. The work of A. Somekh-Baruch and S. Shamai (Shitz) was also supported by the Heron consortium via the Israel ministary of economy and science.

References

[1] S. Gel'fand and M. Pinsker, "Coding for channels with ramdom parameters," *Probl. Contr. Inf. Theory*, vol. 9, no. 1, pp. 19–31, Jan 1980.

[2] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[3] A. S. Cohen and A. Lapidoth, "Generalized writing on dirty paper," in *Proc. IEEE ISIT*, Jun/Jul 2002, p. 227.

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

[5] K. Marton, "A coding theorem for the discrete memoryless BC," *IEEE Trans. Inform. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.

[6] C. Nair and A. E. Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 350–355, Jan 2007.

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

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

[9] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *IEEE Trans. Inform. Theory*, vol. 53, no. 5, pp. 1839–1851, May 2007.

[10] Y. Steinberg and S. Shamai, "Achievable rates for the broadcast channel with states known at the transmitter," in *Proc. IEEE ISIT*, Sept 2005, pp. 2184–2188.

[11] R. Khosravi-Farsani and F. Marvasti, "Capacity bounds for multiuser channels with non-causal channel state information at the transmitters," in *Proc. IEEE Information Theory Workshop*, Oct 2011, pp. 195–199.

[12] N. Merhav and S. Shamai, "Information rates subject to state masking," *IEEE Trans. Inform. Theory*, vol. 53, no. 6, pp. 2254–2261, June 2007.

[13] Y. H. Kim, A. Sutivong, and T. M. Cover, "State amplification," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 1850–1859, May 2008.

[14] T. A. Courtade, "Information masking and amplification: The source coding setting," in *Proc. IEEE ISIT*, July 2012, pp. 189–193.

[15] O. O. Koyluoglu, R. Soundararajan, and S. Vishwanath, "State amplification subject to masking constraints," *IEEE Trans. Inform. Theory*, vol. 62, no. 11, pp. 6233–6250, Nov 2016.

[16] W. Liu and B. Chen, "Message transmission and state estimation over Gaussian broadcast channels," in *2009 43rd Annual Conference on Information Sciences and Systems*, March 2009, pp. 147–151.

[17] P. Grover and A. Sahai, "Witsenhausen's counterexample as assisted interference suppression," *International Journal of Systems, Control and Communications*, vol. 2, no. 1-3, pp. 197–237, 2010.

[18] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

[19] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Information extraction under privacy constraints," *Information*, vol. 7, no. 1, pp. 15:1–37, 2016.

[20] G. Keshet, Y. Steinberg, and N. Merhav, "Channel coding in the presence of side information," *Foundations and Trends® in Communications and Information Theory*, vol. 4, no. 6, pp. 445–586, 2008.

[21] M. Dikshtein and S. Shamai, "Broadcasting information subject to state masking," in *2018 IEEE International Conference on the Science of Electrical Engineering (ICSEE)*, Dec 2018.

[22] A. El Gamal and Y. Kim, *Network information theory.* Cambridge : Cambridge University Press, c2011., 2011.

[23] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint," *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 3072–3081, Nov 2001.

[24] M. Dikshtein, A. Somekh-Baruch, and S. Shamai, "Broadcasting Information subject to State Masking over a MIMO State Dependent Gaussian Channel," *CoRR*, vol. abs/1901.03377, 2019.