Capacity Region of Gaussian MIMO Broadcast Channels With Common and Confidential Messages

Ersen Ekrem, *Student Member, IEEE*, and Sennur Ulukus, *Member, IEEE*

Abstract—We study the two-user Gaussian multiple-input multiple-output (MIMO) broadcast channel with common and confidential messages. In this channel, the transmitter sends a common message to both users, and a confidential message to each user which needs to be kept perfectly secret from the other user. We obtain the entire capacity region of this channel. We also explore the connections between the capacity region we obtain for the Gaussian MIMO broadcast channel with common and confidential messages and the capacity region of its nonconfidential counterpart, i.e., the Gaussian MIMO broadcast channel with common and private messages, which is not known completely.

Index Terms—Gaussian multiple-input multiple-output (MIMO) broadcast channel, secrecy capacity region.

Fig. 1. Gaussian MIMO broadcast channel with common and confidential messages.

I. INTRODUCTION

WE consider the two-user Gaussian multiple-input multiple-output (MIMO) broadcast channel, where each link between the transmitter and each user is modelled by a linear additive Gaussian channel. We study the two-user Gaussian MIMO broadcast channel for the following scenario: The transmitter sends a common message to both users, and a confidential message to each user which needs to be kept perfectly secret from the other user. In other words, in this channel model, there are three messages W_0, W_1, W_2, where W_0 denotes the common message sent to both users, W_1 denotes the first user's confidential message that needs to be kept hidden from the second user, and W_2 denotes the second user's confidential message that needs to be kept hidden from the first user. We call the corresponding channel model the *Gaussian MIMO broadcast channel with common and confidential messages* (see Fig. 1).

The Gaussian MIMO broadcast channel with common and confidential messages subsumes several other channel models as special cases. These special cases can be obtained from our channel model by disabling some of the messages W_0, W_1, W_2. The first such channel model is the Gaussian MIMO wiretap channel, where the transmitter has only one confidential
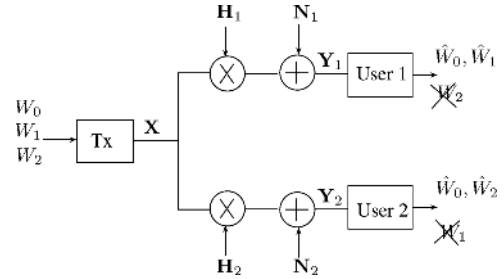
message for one (legitimate) user, which is kept perfectly secret from the other user (eavesdropper). This channel model can be obtained from our channel model by setting $W_0 = W_2 = \phi$. The secrecy capacity of the Gaussian MIMO wiretap channel is obtained in [1] and [2] for the general case, and in [3] for the 2-2-1 case. The second such channel model is the Gaussian MIMO wiretap channel with common message [4], in which the transmitter sends a common message to both the legitimate user and the eavesdropper, and a confidential message to the legitimate user that is kept perfectly secret from the eavesdropper. This channel model can be obtained from our channel model by setting $W_2 = \phi$. The capacity region of the Gaussian MIMO wiretap channel with common message is obtained in [4]. The third such channel model is the Gaussian MIMO broadcast channel with confidential messages [5], where the transmitter sends a confidential message to each user which is kept perfectly secret from the other user. This channel model can be obtained from our channel model by setting $W_0 = \phi$. The capacity region of the Gaussian MIMO broadcast channel with confidential messages is established in [5].

Here, we obtain the capacity region of the Gaussian MIMO broadcast channel with common and confidential messages.[1] In particular, we show that a variant of the secret dirty-paper coding (S-DPC) scheme proposed in [5] is capacity-achieving. Since the S-DPC scheme proposed in [5] is for the transmission of only two confidential messages, it is modified here to incorporate the transmission of a common message as well. Similar to [5], we also notice an invariance property of this achievable scheme with respect to the encoding order used in the S-DPC scheme. In other words, two achievable rate regions arising from two possible encoding orders used in the S-DPC scheme are identical, and equal to the capacity region. We provide the

The authors are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: ersen@umd.edu; ulukus@umd.edu).

[1]The same result is obtained independently and concurrently in [6] and [7]. The conference version [6] and the conference version of this paper [8] appeared concurrently at the IEEE ISIT 2010 as well as at [arXiv: 1001.2806] and [arXiv:1001.3297].

proof of this statement as well as the converse proof for the capacity region by using the channel enhancement technique [9] and an extremal inequality [10].

We also explore the connections between our channel model and its nonconfidential counterpart, i.e., the (two-user) Gaussian MIMO broadcast channel with common and private messages. In the Gaussian MIMO broadcast channel with common and private messages, the transmitter again sends a common message to both users, and a private message to each user, for which there is no secrecy constraint now, i.e., private message of each user does not need to be kept secret from the other user. Thus, the channel model we study here can be viewed as a constrained version of the Gaussian MIMO broadcast channel with common and private messages, where the constraint comes through forcing the private messages to be confidential. We note that although there are partial results for the Gaussian MIMO broadcast channel with common and private messages [11], [12], its capacity region is not known completely. However, here, we are able to obtain the entire capacity region for a constrained version of the Gaussian MIMO broadcast channel with common and private messages. We provide an intuitive explanation of this at-first-sight surprising point as well as the invariance property of the achievable rate region with respect to the encoding orders that can be used in the S-DPC scheme, by using a result from [12]. In particular, we use the following result from [12]: For a given common message rate, the private message sum rate capacity of the Gaussian MIMO broadcast channel with common and private messages is achieved by the dirty-paper coding (DPC) scheme in [13], and any one of the two possible encoding orders that can be used in DPC gives the private message sum rate capacity. Using this result, we show that there is a one-to-one correspondence between the points on the boundary of the achievable rate region of the Gaussian MIMO broadcast channel with common and confidential messages that are obtained by using a specific encoding order in the S-DPC scheme, and those points which are private message sum rate capacity-achieving for the Gaussian MIMO broadcast channel with common and private messages. This correspondence intuitively explains why the achievable rate regions arising from the use of different encoding orders in S-DPC are the same, and also why we can obtain the entire capacity region of the Gaussian MIMO broadcast channel with common and confidential messages although the capacity region of its nonconfidential counterpart is not known completely.

II. CHANNEL MODEL AND MAIN RESULT

We study the two-user Gaussian MIMO broadcast channel (see Fig. 1) which is defined by

$$\mathbf{Y}_1 = \mathbf{H}_1 \mathbf{X} + \mathbf{N}_1 \tag{1}$$
$$\mathbf{Y}_2 = \mathbf{H}_2 \mathbf{X} + \mathbf{N}_2 \tag{2}$$

where the channel input $\mathbf{X}$ is a $t \times 1$ vector, $\mathbf{H}_j$ is the channel gain matrix of size $r_j \times t$, the channel output of the jth user $\mathbf{Y}_j$ is a $r_j \times 1$ vector, and the Gaussian random vector $\mathbf{N}_j$ is of size $r_j \times 1$ with a covariance matrix $\boldsymbol{\Sigma}_j$ which is assumed to be

strictly positive definite, i.e., $\boldsymbol{\Sigma}_j \succ \mathbf{0}$. We consider a covariance constraint on the channel input as follows:

$$E\left[\mathbf{X}\mathbf{X}^\top\right] \preceq \mathbf{S} \tag{3}$$

where $\mathbf{S} \succeq \mathbf{0}$.

We study the following scenario for the Gaussian MIMO broadcast channel: There are three independent messages (W_0, W_1, W_2) with rates (R_0, R_1, R_2), respectively, where W_0 is the common message that needs to be delivered to both users, W_1 is the confidential message of the first user which needs to be kept perfectly secret from the second user, and similarly, W_2 is the confidential message of the second user which needs to be kept perfectly secret from the first user. The secrecy of the confidential messages is measured by the normalized mutual information rates [14], [15], i.e, we require

$$\frac{1}{n}I(W_1; W_0, W_2, \mathbf{Y}_2^n) \to 0 \text{ and } \frac{1}{n}I(W_2; W_0, W_1, \mathbf{Y}_1^n) \to 0 \tag{4}$$

as $n \to \infty$, where n denotes the number of channel uses. The closure of all achievable rate triples (R_0, R_1, R_2) is defined to be the capacity region, and will be denoted by $\mathcal{C}(\mathbf{S})$. We next define the following shorthand notations:

$$R_{0j}(\mathbf{K}_1, \mathbf{K}_2) = \frac{1}{2}\log\frac{|\mathbf{H}_j\mathbf{S}\mathbf{H}_j^\top + \boldsymbol{\Sigma}_j|}{|\mathbf{H}_j(\mathbf{K}_1 + \mathbf{K}_2)\mathbf{H}_j^\top + \boldsymbol{\Sigma}_j|}, \quad j = 1, 2 \tag{5}$$

$$R_1(\mathbf{K}_1, \mathbf{K}_2) = \frac{1}{2}\log\frac{|\mathbf{H}_1(\mathbf{K}_1 + \mathbf{K}_2)\mathbf{H}_1^\top + \boldsymbol{\Sigma}_1|}{|\mathbf{H}_1\mathbf{K}_2\mathbf{H}_1^\top + \boldsymbol{\Sigma}_1|}$$
$$- \frac{1}{2}\log\frac{|\mathbf{H}_2(\mathbf{K}_1 + \mathbf{K}_2)\mathbf{H}_2^\top + \boldsymbol{\Sigma}_2|}{|\mathbf{H}_2\mathbf{K}_2\mathbf{H}_2^\top + \boldsymbol{\Sigma}_2|} \tag{6}$$

$$R_2(\mathbf{K}_2) = \frac{1}{2}\log\frac{|\mathbf{H}_2\mathbf{K}_2\mathbf{H}_2^\top + \boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_2|}$$
$$- \frac{1}{2}\log\frac{|\mathbf{H}_1\mathbf{K}_2\mathbf{H}_1^\top + \boldsymbol{\Sigma}_1|}{|\boldsymbol{\Sigma}_1|} \tag{7}$$

using which, our main result can be stated as follows.

Theorem 1: The capacity region of the Gaussian MIMO broadcast channel with common and confidential messages $\mathcal{C}(\mathbf{S})$ is given by

$$\mathcal{C}(\mathbf{S}) = \mathcal{R}_{12}^{\text{S-DPC}}(\mathbf{S}) = \mathcal{R}_{21}^{\text{S-DPC}}(\mathbf{S}) \tag{8}$$

where $\mathcal{R}_{12}^{\text{S-DPC}}(\mathbf{S})$ is given by the union of rate triples (R_0, R_1, R_2) satisfying

$$R_0 \leq \min\{R_{01}(\mathbf{K}_1, \mathbf{K}_2), R_{02}(\mathbf{K}_1, \mathbf{K}_2)\} \tag{9}$$
$$R_1 \leq R_1(\mathbf{K}_1, \mathbf{K}_2) \tag{10}$$
$$R_2 \leq R_2(\mathbf{K}_2) \tag{11}$$

for some positive semidefinite matrices $\mathbf{K}_1, \mathbf{K}_2$ such that $\mathbf{K}_1 + \mathbf{K}_2 \preceq \mathbf{S}$, and $\mathcal{R}_{21}^{\text{S-DPC}}(\mathbf{S})$ can be obtained from $\mathcal{R}_{12}^{\text{S-DPC}}(\mathbf{S})$ by swapping the subscripts 1 and 2.

Theorem 1 states that the common message, for which a covariance matrix $\mathbf{S} - \mathbf{K}_1 - \mathbf{K}_2$ is allotted, should be encoded by using a standard Gaussian codebook, and the confidential messages, for which covariance matrices $\mathbf{K}_1, \mathbf{K}_2$ are allotted, need to be encoded by using the S-DPC scheme proposed in

[5]. S-DPC is a modified version of DPC [16] to meet the secrecy requirements. The receivers first decode the common message by treating the confidential messages as noise, and then each receiver decodes the confidential message intended to itself. Depending on the encoding order used in S-DPC, one of the users gets a clean link for the transmission of its confidential message, where there is no interference originating from the other user's confidential message. Although one might expect that the two achievable regions arising from two possible encoding orders that can be used in S-DPC could be different, i.e., $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S}) \neq \mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S})$, and taking a convex closure of these two regions would yield a larger achievable rate region, Theorem 1 states that $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S}) = \mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S})$, i.e., the achievable rate region is invariant with respect to the encoding order used in S-DPC. This invariance property of S-DPC was first noticed in [5] for the case where there was no common message to be transmitted.

We acknowledge [6] and [7], where the authors obtain Theorem 1 (capacity region of the Gaussian MIMO broadcast channel with common and confidential messages) independently and concurrently. Their proof is identical to the one we present here.

A. Aligned Channel

We define a subclass of Gaussian MIMO broadcast channels called the aligned Gaussian MIMO broadcast channel, which can be obtained from (1)–(2) by setting $\mathbf{H}_1 = \mathbf{H}_2 = \mathbf{I}$, i.e.,

$$\mathbf{Y}_1 = \mathbf{X} + \mathbf{N}_1 \tag{12}$$
$$\mathbf{Y}_2 = \mathbf{X} + \mathbf{N}_2. \tag{13}$$

To distinguish the notation used for the aligned Gaussian MIMO broadcast channel from the one used for the general model in (1)–(2), we denote the capacity region of the aligned channel by $\mathcal{C}^{\mathrm{AL}}(\mathbf{S})$, the rate expressions in (5)–(7) for the special case $\mathbf{H}_1 = \mathbf{H}_2 = \mathbf{I}$ by $\{R_{0j}^{\mathrm{AL}}(\mathbf{K}_1, \mathbf{K}_2)\}_{j=1}^2$, $R_1^{\mathrm{AL}}(\mathbf{K}_1, \mathbf{K}_2)$, $R_2^{\mathrm{AL}}(\mathbf{K}_2)$, and the regions $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S})$, $\mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S})$ for the special case $\mathbf{H}_1 = \mathbf{H}_2 = \mathbf{I}$ by $\mathcal{R}_{12}^{\mathrm{S-DPC-AL}}(\mathbf{S})$, $\mathcal{R}_{21}^{\mathrm{S-DPC-AL}}(\mathbf{S})$.

In this paper, we first prove Theorem 1 for the aligned Gaussian MIMO broadcast channel. Then, we establish the capacity region for the general channel model in (1)–(2) by following the analysis in [9, Sec. V.B] and [17, Sec. 7.1] in conjunction with the capacity result we obtain for the aligned channel.

B. Capacity Region Under a Power Constraint

We note that the covariance constraint on the channel input in (3) is a rather general constraint that subsumes the power constraint

$$E\left[\mathbf{X}^\top \mathbf{X}\right] = \mathrm{tr}\left(E\left[\mathbf{X}\mathbf{X}^\top\right]\right) \leq P \tag{14}$$

as a special case, see Lemma 1 and [9, Corollary 1]. Therefore, using Theorem 1, the capacity region arising from the average power constraint in (14), $\mathcal{C}(P)$, can be found as follows.

Corollary 1: The capacity region of the Gaussian MIMO broadcast channel with common and confidential messages subject to a power constraint P, $\mathcal{C}(P)$, is given by

$$\mathcal{C}(P) = \mathcal{R}_{12}^{\mathrm{S-DPC}}(P) = \mathcal{R}_{21}^{\mathrm{S-DPC}}(P) \tag{15}$$

where $\mathcal{R}_{12}^{\mathrm{S-DPC}}(P)$ is given by the union of rate triples (R_0, R_1, R_2) satisfying

$$R_0 \leq \min\{R_{01}(\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_c), R_{02}(\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_c)\} \tag{16}$$
$$R_1 \leq R_1(\mathbf{K}_1, \mathbf{K}_2) \tag{17}$$
$$R_2 \leq R_2(\mathbf{K}_2) \tag{18}$$

for some positive semidefinite matrices $\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_c$ such that $\mathrm{tr}(\mathbf{K}_1 + \mathbf{K}_2 + \mathbf{K}_c) \leq P$, and $\{R_{0j}(\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_c)\}_{j=1}^2$ are defined as

$$R_{0j}(\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_c) = \frac{1}{2}\log\frac{|\mathbf{H}_j(\mathbf{K}_1+\mathbf{K}_2+\mathbf{K}_c)\mathbf{H}_j^\top + \boldsymbol{\Sigma}_j|}{|\mathbf{H}_j(\mathbf{K}_1+\mathbf{K}_2)\mathbf{H}_j^\top + \boldsymbol{\Sigma}_j|}. \tag{19}$$

Moreover, $\mathcal{R}_{21}^{\mathrm{S-DPC}}(P)$ can be obtained from $\mathcal{R}_{12}^{\mathrm{S-DPC}}(P)$ by swapping the subscripts 1 and 2.

III. PROOF OF THEOREM 1 FOR THE ALIGNED CASE

A. Achievability

Here, we prove the achievability of the regions $\mathcal{R}_{12}^{\mathrm{S-DPC-AL}}(\mathbf{S})$ and $\mathcal{R}_{21}^{\mathrm{S-DPC-AL}}(\mathbf{S})$. To this end, we consider the two-user discrete memoryless channel for the scenario where a common message is delivered to both users, and each user gets a confidential message which needs be kept perfectly secret from the other user. For this scenario, we have the following achievable rate region [18].

Lemma 1 [18, Theorem 1]: The rate triples (R_0, R_1, R_2) satisfying

$$R_0 \leq \min\{I(U; Y_1), I(U; Y_2)\} \tag{20}$$
$$R_1 \leq [I(V_1; Y_1|U) - I(V_1; Y_2, V_2|U)]^+ \tag{21}$$
$$R_2 \leq [I(V_2; Y_2|U) - I(V_2; Y_1, V_1|U)]^+ \tag{22}$$

for some (U, V_1, V_2) such that $(U, V_1, V_2) \to X \to (Y_1, Y_2)^2$ are achievable.

We now use Lemma 1 to show the achievability of the region $\mathcal{R}_{12}^{\mathrm{S-DPC-AL}}(\mathbf{S})$. We first introduce three independent Gaussian random vectors $\mathbf{U}_0, \mathbf{U}_1, \mathbf{U}_2$ with covariance matrices $\mathbf{S} - \mathbf{K}_1 - \mathbf{K}_2, \mathbf{K}_1, \mathbf{K}_2$, respectively. Using these Gaussian random vectors, we set the auxiliary random variables in Lemma 1 as follows:

$$U = \mathbf{U}_0 \tag{23}$$
$$V_1 = \mathbf{U}_1 + \mathbf{U}_0 \tag{24}$$
$$V_2 = \mathbf{U}_2 + \mathbf{A}\mathbf{U}_1 + \mathbf{U}_0 \tag{25}$$

where $\mathbf{A} = \mathbf{K}_2[\mathbf{K}_2 + \boldsymbol{\Sigma}_2]^{-1}$ is the precoding matrix for the second user to suppress the interference originating from $\mathbf{U}_1$ [16]. Furthermore, we set the channel input $\mathbf{X}$ as follows:

$$\mathbf{X} = \mathbf{U}_0 + \mathbf{U}_1 + \mathbf{U}_2. \tag{26}$$

---

²In [18], the necessary Markov chain that $(U, V_1, V_2, X, Y_1, Y_2)$ needs to satisfy is given by $U \to (V_1, V_2) \to X \to (Y_1, Y_2)$. However, their achievable rate region is valid for the looser Markov chain $(U, V_1, V_2) \to X \to (Y_1, Y_2)$ as well, which we use here.

Using the definitions in (23)–(26) for the common message rate given in Lemma 1, we get

$$R_0 = \min \left\{ \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_1|}{|\mathbf{K}_1 + \mathbf{K}_2 + \boldsymbol{\Sigma}_1|}, \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_2|}{|\mathbf{K}_1 + \mathbf{K}_2 + \boldsymbol{\Sigma}_2|} \right\}. \tag{27}$$

Next, we compute the confidential message rates. To this end, we note the following identity:

$$I(V_2; \mathbf{Y}_2|U) - I(V_2; V_1|U) = \frac{1}{2} \log \frac{|\mathbf{K}_2 + \boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_2|} \tag{28}$$

which is due to [16, Th. 1]. Now, we compute the second user's confidential message rate as follows:

$$R_2 = I(V_2; \mathbf{Y}_2|U) - I(V_2; \mathbf{Y}_1, V_1|U) \tag{29}$$

$$= I(V_2; \mathbf{Y}_2|U) - I(V_2; V_1|U) - I(V_2; \mathbf{Y}_1|U, V_1) \tag{30}$$

$$= \frac{1}{2} \log \frac{|\mathbf{K}_2 + \boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_2|} - I(V_2; \mathbf{Y}_1|U, V_1) \tag{31}$$

$$= \frac{1}{2} \log \frac{|\mathbf{K}_2 + \boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_2|} - \frac{1}{2} \log \frac{|\mathbf{K}_2 + \boldsymbol{\Sigma}_1|}{|\boldsymbol{\Sigma}_1|} \tag{32}$$

where (31) is due to (28). Next, we compute the first user's confidential message rate as follows:

$$R_1 = I(V_1; \mathbf{Y}_1|U) - I(V_1; \mathbf{Y}_2, V_2|U) \tag{33}$$

$$= I(V_1; \mathbf{Y}_1|U) - I(V_1; \mathbf{Y}_2|U, V_2) - I(V_1; V_2|U) \tag{34}$$

$$= I(V_1; \mathbf{Y}_1|U) - I(V_1, V_2; \mathbf{Y}_2|U) + I(V_2; \mathbf{Y}_2|U) - I(V_1; V_2|U) \tag{35}$$

$$= I(V_1; \mathbf{Y}_1|U) - I(V_1, V_2; \mathbf{Y}_2|U) + \frac{1}{2} \log \frac{|\mathbf{K}_2 + \boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_2|} \tag{36}$$

$$= \frac{1}{2} \log \frac{|\mathbf{K}_1 + \mathbf{K}_2 + \boldsymbol{\Sigma}_1|}{|\mathbf{K}_2 + \boldsymbol{\Sigma}_1|} - \frac{1}{2} \log \frac{|\mathbf{K}_1 + \mathbf{K}_2 + \boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_2|} + \frac{1}{2} \log \frac{|\mathbf{K}_2 + \boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_2|} \tag{37}$$

$$= \frac{1}{2} \log \frac{|\mathbf{K}_1 + \mathbf{K}_2 + \boldsymbol{\Sigma}_1|}{|\mathbf{K}_2 + \boldsymbol{\Sigma}_1|} - \frac{1}{2} \log \frac{|\mathbf{K}_1 + \mathbf{K}_2 + \boldsymbol{\Sigma}_2|}{|\mathbf{K}_2 + \boldsymbol{\Sigma}_2|} \tag{38}$$

where (36) is due to (28). Hence, we show the achievability of the region $\mathcal{R}_{12}^{\mathrm{S-DPC-AL}}(\mathbf{S})$. Due to the symmetry, achievability of $\mathcal{R}_{21}^{\mathrm{S-DPC-AL}}(\mathbf{S})$ follows.

B. Converse

Since the capacity region $\mathcal{C}^{\mathrm{AL}}(\mathbf{S})$ is convex due to time-sharing, it can be characterized by the tangent planes to it, i.e., by the solution of

$$\max_{(R_0, R_1, R_2) \in \mathcal{C}^{\mathrm{AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2 \tag{39}$$

for $\mu_j \in [0, \infty)$, $j = 0, 1, 2$. We already have

$$\max_{(R_0, R_1, R_2) \in \mathcal{R}^{\mathrm{S-DPC-AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2$$
$$\leq \max_{(R_0, R_1, R_2) \in \mathcal{C}^{\mathrm{AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2 \tag{40}$$

due to achievability of $\mathcal{R}_{12}^{\mathrm{S-DPC-AL}}(\mathbf{S})$ and $\mathcal{R}_{21}^{\mathrm{S-DPC-AL}}(\mathbf{S})$, where $\mathcal{R}^{\mathrm{S-DPC-AL}}(\mathbf{S})$ is given by

$$\mathcal{R}^{\mathrm{S-DPC-AL}}(\mathbf{S})$$
$$= \mathrm{conv} \left(\mathcal{R}_{12}^{\mathrm{S-DPC-AL}}(\mathbf{S}) \bigcup \mathcal{R}_{21}^{\mathrm{S-DPC-AL}}(\mathbf{S}) \right) \tag{41}$$

and conv is the convex hull operator. Here, we show that

$$\max_{(R_0, R_1, R_2) \in \mathcal{C}^{\mathrm{AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2$$
$$\leq \max_{(R_0, R_1, R_2) \in \mathcal{R}_{12}^{\mathrm{S-DPC-AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2 \tag{42}$$
$$= \max_{(R_0, R_1, R_2) \in \mathcal{R}_{21}^{\mathrm{S-DPC-AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2 \tag{43}$$

to provide the converse proof. We first characterize the boundary of $\mathcal{R}_{12}^{\mathrm{S-DPC-AL}}(\mathbf{S})$ by studying the following optimization problem:

$$\max_{(R_0, R_1, R_2) \in \mathcal{R}_{12}^{\mathrm{S-DPC-AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2 \tag{44}$$

which can be written as

$$\max_{\substack{\mathbf{0} \preceq \mathbf{K}_j, \ j=1,2 \\ \mathbf{K}_1 + \mathbf{K}_2 \preceq \mathbf{S}}} \mu_0 \min \{ R_{01}^{\mathrm{AL}}(\mathbf{K}_1, \mathbf{K}_2), R_{02}^{\mathrm{AL}}(\mathbf{K}_1, \mathbf{K}_2) \}$$
$$+ \mu_1 R_1^{\mathrm{AL}}(\mathbf{K}_1, \mathbf{K}_2) + \mu_2 R_2^{\mathrm{AL}}(\mathbf{K}_2). \tag{45}$$

Let $\mathbf{K}_1^*$, $\mathbf{K}_2^*$ be the maximizer of (45). The necessary Karush–Kuhn–Tucker (KKT) conditions that $\mathbf{K}_1^*$, $\mathbf{K}_2^*$ need to satisfy are given in the following lemma, whose proof is given in Appendix I.

Lemma 2: $\mathbf{K}_1^*$, $\mathbf{K}_2^*$ need to satisfy

$$(\mu_1 + \mu_2)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \mathbf{M}_1$$
$$= (\mu_0 \lambda + \mu_2)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1}$$
$$+ (\mu_0 \bar{\lambda} + \mu_1)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} + \mathbf{M}_S \tag{46}$$
$$(\mu_1 + \mu_2)(\mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} + \mathbf{M}_2$$
$$= (\mu_1 + \mu_2)(\mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \mathbf{M}_1 \tag{47}$$

for some positive semidefinite matrices $\{\mathbf{M}_j\}_{j=1}^2$, $\mathbf{M}_S$ such that

$$\mathbf{K}_1^* \mathbf{M}_1 = \mathbf{M}_1 \mathbf{K}_1^* = \mathbf{0} \tag{48}$$
$$\mathbf{K}_2^* \mathbf{M}_2 = \mathbf{M}_2 \mathbf{K}_2^* = \mathbf{0} \tag{49}$$
$$(\mathbf{S} - \mathbf{K}_1^* - \mathbf{K}_2^*) \mathbf{M}_S = \mathbf{M}_S (\mathbf{S} - \mathbf{K}_1^* - \mathbf{K}_2^*) = \mathbf{0} \tag{50}$$

and for some $\lambda = 1 - \bar{\lambda}$ such that it satisfies $0 \leq \lambda \leq 1$ and

$$\lambda \begin{cases} = 0, & \text{if} \quad R_{01}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) > R_{02}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) \\ = 1, & \text{if} \quad R_{01}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) < R_{02}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) \\ \neq 0, 1, & \text{if} \quad R_{01}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) = R_{02}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*). \end{cases} \tag{51}$$

We now use channel enhancement [9] to define a new noise covariance matrix $\tilde{\boldsymbol{\Sigma}}$ as follows:

$$(\mu_1 + \mu_2)(\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1} = (\mu_1 + \mu_2)(\mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} + \mathbf{M}_2. \tag{52}$$

This new noise covariance matrix $\tilde{\boldsymbol{\Sigma}}$ has useful properties which are listed in the following lemma.

Lemma 3: We have the following facts:
1) $\mathbf{0} \prec \tilde{\boldsymbol{\Sigma}}$;
2) $\tilde{\boldsymbol{\Sigma}} \preceq \boldsymbol{\Sigma}_1, \tilde{\boldsymbol{\Sigma}} \preceq \boldsymbol{\Sigma}_2$;

3) $(\mu_1 + \mu_2)(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1} = (\mu_1 + \mu_2)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \mathbf{M}_1;$

4) $(\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1}\tilde{\boldsymbol{\Sigma}} = (\mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1}\boldsymbol{\Sigma}_2;$

5) $(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1}(\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}}) = (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1}(\mathbf{K}_2^* + \boldsymbol{\Sigma}_1).$

The proof of Lemma 3 is given in Appendix II. We now construct an enhanced channel using the new covariance matrix $\tilde{\boldsymbol{\Sigma}}$ as follows:

$$\tilde{\mathbf{Y}}_1 = \mathbf{X} + \tilde{\mathbf{N}} \tag{53}$$

$$\tilde{\mathbf{Y}}_2 = \mathbf{X} + \tilde{\mathbf{N}} \tag{54}$$

$$\mathbf{Y}_1 = \mathbf{X} + \mathbf{N}_1 \tag{55}$$

$$\mathbf{Y}_2 = \mathbf{X} + \mathbf{N}_2 \tag{56}$$

where $\tilde{\mathbf{N}}$ is a Gaussian random vector with a covariance matrix $\tilde{\boldsymbol{\Sigma}}$. In the enhanced channel defined by (53)–(56), the enhanced first and second users have the same observation, i.e., $\Pr[\tilde{\mathbf{Y}}_1 = \tilde{\mathbf{Y}}_2] = 1$. From now on, we denote the observations of the enhanced first and second users by a single random vector $\tilde{\mathbf{Y}}$. We now consider the following scenario for the enhanced channel in (53)–(56): There are three independent messages (W_0, W_1, W_2) with rates (R_0, R_1, R_2), respectively, where the common message W_0 is directed to all users, i.e., the users with observations $\tilde{\mathbf{Y}}_1, \tilde{\mathbf{Y}}_2, \mathbf{Y}_1, \mathbf{Y}_2$; W_1 is the confidential message of the enhanced first user, i.e., the one with observation $\tilde{\mathbf{Y}}$, which needs to be kept perfectly secret from the second user, i.e., the one with observation $\mathbf{Y}_2$; and W_2 is the confidential message of the enhanced second user, i.e., the one with observation $\tilde{\mathbf{Y}}$, which needs to be kept perfectly secret from the first user, i.e., the one with observation $\mathbf{Y}_1$. Here also, we measure the secrecy of the confidential messages by normalized equivocation rates, i.e., we require

$$\lim_{n \to \infty} \frac{1}{n} I(W_1; \mathbf{Y}_2^n, W_0) = 0 \tag{57}$$

$$\lim_{n \to \infty} \frac{1}{n} I(W_2; \mathbf{Y}_1^n, W_0) = 0. \tag{58}$$

We define the capacity region of the enhanced channel in (53)–(56) arising from this scenario as the convex closure of all achievable rate pairs (R_0, R_1, R_2) and denote it by $\tilde{C}(\mathbf{S})$.

We note that the process of obtaining a new enhanced channel from the original one by means of channel enhancement can be visualized as shown in Figs. 2 and 3. First, we provide an alternative view of the original channel model as depicted in Fig. 2. In this alternative view, each user is split into two identical users where one of them (user 11 for the first user and user 22 for the second user) gets a confidential message, and the other one (user 10 for the first user and user 20 for the second user) gets the common message and eavesdrops the other confidential message. Second, we enhance the users who are getting the confidential messages, i.e., user 11 and user 22, to improve their observations as shown in Fig. 3. This idea of splitting users and then enhancing them is also used in [4]. Since in the enhanced channel, the receivers to which only the common message is sent are identical to the receivers in the original channel in (12)–(13), and the receivers to which confidential messages are sent have better observations with respect to the receivers in
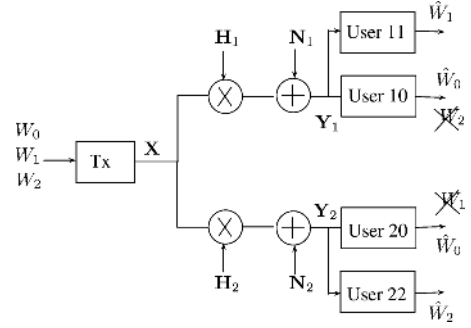

Fig. 2. Alternative view of the Gaussian MIMO broadcast channel with common and confidential messages.
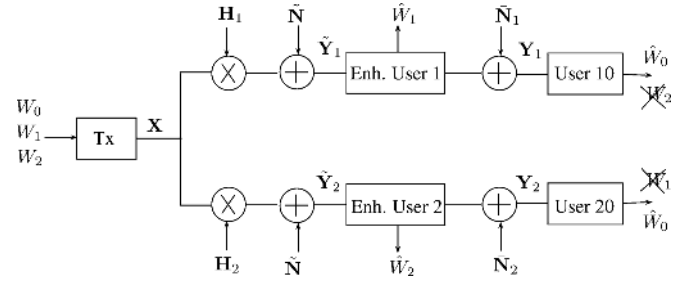

Fig. 3. New Gaussian MIMO broadcast channel obtained by channel enhancement.

the original channel in (12)–(13), we have $\mathcal{C}^{\mathrm{AL}}(\mathbf{S}) \subseteq \tilde{C}(\mathbf{S})$. We next introduce an outer bound on $\tilde{C}(\mathbf{S})$ in the following lemma.

Lemma 4: The capacity region of the enhanced channel in (53)–(56), $\tilde{C}(\mathbf{S})$, is contained in the union of rate triples (R_0, R_1, R_2) satisfying

$$R_0 \leq \min\{I(U; \mathbf{Y}_1), I(U; \mathbf{Y}_2)\} \tag{59}$$

$$R_1 \leq I(\mathbf{X}; \tilde{\mathbf{Y}}|U) - I(\mathbf{X}; \mathbf{Y}_2|U) \tag{60}$$

$$R_2 \leq I(\mathbf{X}; \tilde{\mathbf{Y}}|U) - I(\mathbf{X}; \mathbf{Y}_1|U) \tag{61}$$

for some $(U, \mathbf{X})$ such that

$$U \to \mathbf{X} \to \tilde{\mathbf{Y}} \to (\mathbf{Y}_1, \mathbf{Y}_2) \tag{62}$$

and $E\left[\mathbf{X}\mathbf{X}^\top\right] \preceq \mathbf{S}$.

The proof of this lemma is given in Appendix III. We also introduce the following extremal inequality from [10]:

Lemma 5 [10, Corollary 4]: Let $(U, \mathbf{X})$ be an arbitrarily correlated random vector, where $\mathbf{X}$ has a covariance constraint $E\left[\mathbf{X}\mathbf{X}^\top\right] \preceq \mathbf{S}$ and $\mathbf{S} \succ \mathbf{0}$. Let $\tilde{\mathbf{N}}, \mathbf{N}_1, \mathbf{N}_2$ be Gaussian random vectors with covariance matrices $\tilde{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2$, respectively. They are independent of $(U, \mathbf{X})$. Furthermore, $\tilde{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2$ satisfy $\tilde{\boldsymbol{\Sigma}} \preceq \boldsymbol{\Sigma}_j$, $j = 1, 2$. Assume that there exists a covariance matrix $\mathbf{K}^*$ such that $\mathbf{K}^* \preceq \mathbf{S}$ and

$$\beta(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}})^{-1} = \sum_{j=1}^{2} \gamma_j (\mathbf{K}^* + \boldsymbol{\Sigma}_j)^{-1} + \mathbf{M}_S \tag{63}$$

where $\beta \geq 0$, $\gamma_j \geq 0$, $j = 1, 2$ and $\mathbf{M}_S$ is positive semidefinite matrix such that $(\mathbf{S} - \mathbf{K}^*)\mathbf{M}_S = \mathbf{0}$. Then, for any $(U, \mathbf{X})$, we have

$$\beta h(\mathbf{X} + \tilde{\mathbf{N}}|U) - \sum_{j=1}^{2} \gamma_j h(\mathbf{X} + \mathbf{N}_j|U)$$

$$\leq \frac{\beta}{2} \log |(2\pi e)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}})| - \sum_{j=1}^{2} \frac{\gamma_j}{2} \log |(2\pi e)(\mathbf{K}^* + \boldsymbol{\Sigma}_j)|. \quad (64)$$

We now use this lemma. For that purpose, we note that using the second statement of Lemma 3 in (46) yields

$$(\mu_1 + \mu_2)(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1}$$
$$= (\mu_0 \lambda + \mu_2)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1}$$
$$+ (\mu_0 \bar{\lambda} + \mu_1)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} + \mathbf{M}_S \quad (65)$$

using which in conjunction with Lemma 5, we get

$$(\mu_1 + \mu_2)h(\tilde{\mathbf{Y}}|U) - (\mu_0 \lambda + \mu_2)h(\mathbf{Y}_1|U)$$
$$- (\mu_0 \bar{\lambda} + \mu_1)h(\mathbf{Y}_2|U)$$
$$\leq \frac{\mu_1 + \mu_2}{2} \log |(2\pi e)(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})|$$
$$- \frac{\mu_0 \lambda + \mu_2}{2} \log |(2\pi e)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)|$$
$$- \frac{\mu_0 \bar{\lambda} + \mu_1}{2} \log |(2\pi e)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2)| \quad (66)$$

which will be used subsequently.

We are now ready to complete the converse proof as follows:

$$\max_{(R_0, R_1, R_2) \in \mathcal{C}^{\mathrm{AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2$$
$$\leq \max_{(R_0, R_1, R_2) \in \tilde{\mathcal{C}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2 \quad (67)$$
$$\leq \max_{(U, \mathbf{X}) \in \mathcal{F}_{U, \mathbf{X}}} \mu_0 \min\{I(U; \mathbf{Y}_1), I(U; \mathbf{Y}_2)\}$$
$$+ \mu_1 \left[I(\mathbf{X}; \tilde{\mathbf{Y}}|U) - I(\mathbf{X}; \mathbf{Y}_2|U) \right]$$
$$+ \mu_2 \left[I(\mathbf{X}; \tilde{\mathbf{Y}}|U) - I(\mathbf{X}; \mathbf{Y}_1|U) \right] \quad (68)$$
$$\leq \max_{(U, \mathbf{X}) \in \mathcal{F}_{U, \mathbf{X}}} \mu_0 \lambda I(U; \mathbf{Y}_1) + \mu_0 \bar{\lambda} I(U; \mathbf{Y}_2)$$
$$+ \mu_1 \left[I(\mathbf{X}; \tilde{\mathbf{Y}}|U) - I(\mathbf{X}; \mathbf{Y}_2|U) \right]$$
$$+ \mu_2 \left[I(\mathbf{X}; \tilde{\mathbf{Y}}|U) - I(\mathbf{X}; \mathbf{Y}_1|U) \right] \quad (69)$$
$$= \max_{(U, \mathbf{X}) \in \mathcal{F}_{U, \mathbf{X}}} \mu_0 \lambda h(\mathbf{Y}_1) + \mu_0 \bar{\lambda} h(\mathbf{Y}_2) + (\mu_1 + \mu_2)h(\tilde{\mathbf{Y}}|U)$$
$$- (\mu_0 \lambda + \mu_2)h(\mathbf{Y}_1|U) - (\mu_0 \bar{\lambda} + \mu_1)h(\mathbf{Y}_2|U)$$
$$- \frac{\mu_1}{2} \log \frac{|\tilde{\boldsymbol{\Sigma}}|}{|\boldsymbol{\Sigma}_2|} - \frac{\mu_2}{2} \log \frac{|\tilde{\boldsymbol{\Sigma}}|}{|\boldsymbol{\Sigma}_1|} \quad (70)$$
$$\leq \frac{\mu_0 \lambda}{2} \log |(2\pi e)(\mathbf{S} + \boldsymbol{\Sigma}_1)| + \frac{\mu_0 \bar{\lambda}}{2} \log |(2\pi e)(\mathbf{S} + \boldsymbol{\Sigma}_2)|$$
$$+ \max_{(U, \mathbf{X}) \in \mathcal{F}_{U, \mathbf{X}}} (\mu_1 + \mu_2)h(\tilde{\mathbf{Y}}|U) - (\mu_0 \lambda + \mu_2)h(\mathbf{Y}_1|U)$$
$$- (\mu_0 \bar{\lambda} + \mu_1)h(\mathbf{Y}_2|U) - \frac{\mu_1}{2} \log \frac{|\tilde{\boldsymbol{\Sigma}}|}{|\boldsymbol{\Sigma}_2|} - \frac{\mu_2}{2} \log \frac{|\tilde{\boldsymbol{\Sigma}}|}{|\boldsymbol{\Sigma}_1|} \quad (71)$$
$$\leq \frac{\mu_0 \lambda}{2} \log |(2\pi e)(\mathbf{S} + \boldsymbol{\Sigma}_1)| + \frac{\mu_0 \bar{\lambda}}{2} \log |(2\pi e)(\mathbf{S} + \boldsymbol{\Sigma}_2)|$$
$$+ \frac{(\mu_1 + \mu_2)}{2} \log |(2\pi e)(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})|$$

$$- \frac{(\mu_0 \lambda + \mu_2)}{2} \log |(2\pi e)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)|$$
$$- \frac{(\mu_0 \bar{\lambda} + \mu_1)}{2} \log |(2\pi e)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2)|$$
$$- \frac{\mu_1}{2} \log \frac{|\tilde{\boldsymbol{\Sigma}}|}{|\boldsymbol{\Sigma}_2|} - \frac{\mu_2}{2} \log \frac{|\tilde{\boldsymbol{\Sigma}}|}{|\boldsymbol{\Sigma}_1|} \quad (72)$$
$$= \frac{\mu_0 \lambda}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_1|}{|\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1|} + \frac{\mu_0 \bar{\lambda}}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_2|}{|\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2|}$$
$$+ \frac{\mu_1}{2} \log \frac{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})\boldsymbol{\Sigma}_2|}{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2)\tilde{\boldsymbol{\Sigma}}|}$$
$$+ \frac{\mu_2}{2} \log \frac{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})\boldsymbol{\Sigma}_1|}{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)\tilde{\boldsymbol{\Sigma}}|} \quad (73)$$
$$= \mu_0 \min\{R_{01}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*), R_{02}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*)\}$$
$$+ \frac{\mu_1}{2} \log \frac{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})\boldsymbol{\Sigma}_2|}{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2)\tilde{\boldsymbol{\Sigma}}|}$$
$$+ \frac{\mu_2}{2} \log \frac{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})\boldsymbol{\Sigma}_1|}{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)\tilde{\boldsymbol{\Sigma}}|} \quad (74)$$
$$= \mu_0 \min\{R_{01}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*), R_{02}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*)\}$$
$$+ \mu_1 R_1^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) + \mu_2 R_2^{\mathrm{AL}}(\mathbf{K}_2^*) \quad (75)$$

where (67) comes from the fact that $\mathcal{C}^{\mathrm{AL}}(\mathbf{S}) \subseteq \tilde{\mathcal{C}}(\mathbf{S})$, the set $\mathcal{F}_{U, \mathbf{X}}$ denotes the set of feasible random vectors $(U, \mathbf{X})$, i.e.,

$$\mathcal{F}_{U, \mathbf{X}}$$
$$= \{(U, \mathbf{X}) : U \to \mathbf{X} \to \tilde{\mathbf{Y}} \to \mathbf{Y}_1, \mathbf{Y}_2 \text{ and } E\left[\mathbf{X}\mathbf{X}^\top\right] \preceq \mathbf{S}\} \quad (76)$$

(68) is due to Lemma 4, (69) results from the fact that $0 \leq \lambda = 1 - \bar{\lambda} \leq 1$, (71) is due to the maximum entropy theorem, (72) comes from (66), (74) results from

$$\lambda R_{01}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) + \bar{\lambda} R_{02}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*)$$
$$= \min\{R_{01}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*), R_{02}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*)\} \quad (77)$$

and (75) will be shown next. We first note the following:

$$R_1^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*)$$
$$= \frac{1}{2} \log \frac{|\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1|}{|\mathbf{K}_2^* + \boldsymbol{\Sigma}_1|} - \frac{1}{2} \log \frac{|\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2|}{|\mathbf{K}_2^* + \boldsymbol{\Sigma}_2|} \quad (78)$$
$$= \frac{1}{2} \log \frac{|\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}}|}{|\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}}|} - \frac{1}{2} \log \frac{|\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2|}{|\mathbf{K}_2^* + \boldsymbol{\Sigma}_2|} \quad (79)$$
$$= \frac{1}{2} \log \frac{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})\boldsymbol{\Sigma}_2|}{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2)\tilde{\boldsymbol{\Sigma}}|} \quad (80)$$

where (79) is due to the fourth statement of Lemma 3 and (80) comes from the third statement of Lemma 3. We next note the following identity:

$$R_2^{\mathrm{AL}}(\mathbf{K}_2^*) = \frac{1}{2} \log \frac{|\mathbf{K}_2^* + \boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_2|} - \frac{1}{2} \log \frac{|\mathbf{K}_2^* + \boldsymbol{\Sigma}_1|}{|\boldsymbol{\Sigma}_1|} \quad (81)$$
$$= \frac{1}{2} \log \frac{|\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}}|}{|\tilde{\boldsymbol{\Sigma}}|} - \frac{1}{2} \log \frac{|\mathbf{K}_2^* + \boldsymbol{\Sigma}_1|}{|\boldsymbol{\Sigma}_1|} \quad (82)$$
$$= \frac{1}{2} \log \frac{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})\boldsymbol{\Sigma}_1|}{|(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)\tilde{\boldsymbol{\Sigma}}|} \quad (83)$$

where (82) is due to the third statement of Lemma 3, and (83) comes from the fourth statement of Lemma 3. Identities in (80) and (83) give (75).

Thus, in the view of (75), we have shown that

$$\max_{(R_0,R_1,R_2)\in\mathcal{C}^{\mathrm{AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2$$
$$= \max_{(R_0,R_1,R_2)\in\mathcal{R}_{12}^{\mathrm{S-DPC-AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2. \quad (84)$$

Similarly, we can show the following:

$$\max_{(R_0,R_1,R_2)\in\mathcal{C}^{\mathrm{AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2$$
$$= \max_{(R_0,R_1,R_2)\in\mathcal{R}_{21}^{\mathrm{S-DPC-AL}}(\mathbf{S})} \mu_0 R_0 + \mu_1 R_1 + \mu_2 R_2 \quad (85)$$

completing the converse proof.

IV. PROOF OF THEOREM 1 FOR THE GENERAL CASE

We now prove Theorem 1 for the general channel model in (1)–(2). Achievability of Theorem 1 for the general channel model in (1)–(2) can be shown as we did for the aligned case in the previous section. In particular, the only difference of the achievability proof for the general channel model in (1)–(2) from the achievability proof for the aligned case will be the selection of the precoding matrix $\mathbf{A}$, which needs to be chosen as $\mathbf{A} = \mathbf{K}_2\mathbf{H}_2^\top(\boldsymbol{\Sigma}_2 + \mathbf{H}_2\mathbf{K}_2\mathbf{H}_2^\top)^{-1}\mathbf{H}_2$ in this general case. Thus, in the rest of this section, we consider the converse proof. For that purpose, we follow the analysis in [9, Sec. V.B] and [17, Sec. 7.1] in conjunction with the capacity result obtained for the aligned case in the previous section. To this end, we first note that, following the approaches in [9, Sec. V.B] and [17, Sec. 7.1], it can be shown that a new channel can be constructed from any channel described by (1)–(2), such that the new channel has the same capacity region as the original one, and in the new channel, both receivers have the same number of antennas as the transmitter, i.e., $r_1 = r_2 = t$. Thus, without loss of generality, we assume that $r_1 = r_2 = t$. We next apply singular-value decomposition to the channel gain matrices $\mathbf{H}_1$, $\mathbf{H}_2$ as follows:

$$\mathbf{H}_j = \mathbf{U}_j\boldsymbol{\Lambda}_j\mathbf{V}_j^\top, \quad j = 1, 2 \quad (86)$$

where $\mathbf{U}_j$, $\mathbf{V}_j$ are $t \times t$ orthogonal matrices, and $\boldsymbol{\Lambda}_j$ is a diagonal matrix. We now define a new Gaussian MIMO broadcast channel as follows:

$$\overline{\mathbf{Y}}_1 = \overline{\mathbf{H}}_1\mathbf{X} + \mathbf{N}_1 \quad (87)$$
$$\overline{\mathbf{Y}}_2 = \overline{\mathbf{H}}_2\mathbf{X} + \mathbf{N}_2 \quad (88)$$

where $\overline{\mathbf{H}}_j$ is defined as

$$\overline{\mathbf{H}}_j = \mathbf{U}_j(\boldsymbol{\Lambda}_j + \alpha\mathbf{I})\mathbf{V}_j^\top \quad (89)$$

for some $\alpha > 0$. We denote the capacity region of the channel defined in (87)–(88) by $\mathcal{C}_\alpha(\mathbf{S})$, and achievable rate regions for this channel by $\mathcal{R}_{12,\alpha}^{\mathrm{S-DPC}}(\mathbf{S})$, $\mathcal{R}_{21,\alpha}^{\mathrm{S-DPC}}(\mathbf{S})$. Since $\overline{\mathbf{H}}_1$, $\overline{\mathbf{H}}_2$ are invertible, the capacity region of the channel in (87)–(88) is equal to the capacity region of the following aligned channel:

$$\overline{\overline{\mathbf{Y}}}_1 = \mathbf{X} + \overline{\mathbf{H}}_1^{-1}\mathbf{N}_1 \quad (90)$$
$$\overline{\overline{\mathbf{Y}}}_2 = \mathbf{X} + \overline{\mathbf{H}}_2^{-1}\mathbf{N}_2. \quad (91)$$

Thus, using the capacity result for the aligned case, which was proved in the previous section, we get

$$\mathcal{C}_\alpha(\mathbf{S}) = \mathcal{R}_{12,\alpha}^{\mathrm{S-DPC}}(\mathbf{S}) = \mathcal{R}_{21,\alpha}^{\mathrm{S-DPC}}(\mathbf{S}). \quad (92)$$

We next show the following inclusion:

$$\mathcal{C}(\mathbf{S}) \subseteq \lim_{\alpha\to 0}\mathcal{C}_\alpha(\mathbf{S}). \quad (93)$$

To this end, assume that (R_0, R_1, R_2) is achievable in the channel given by (1)–(2), i.e., $(R_0, R_1, R_2) \in \mathcal{C}(\mathbf{S})$. To prove the inclusion in (93), we need to show that $(R_0, R_1, R_2) \in \lim_{\alpha\to 0}\mathcal{C}_\alpha(\mathbf{S})$. To this end, we note the following Markov chains:

$$\mathbf{X} \to \overline{\mathbf{Y}}_j \to \mathbf{Y}_j, \quad j = 1, 2 \quad (94)$$

which imply that if the message triple (W_0, W_1, W_2) with rates (R_0, R_1, R_2) is transmitted with a vanishingly small probability of error in the original channel given by (1)–(2), they will be transmitted with a vanishingly small probability of error in the channel given by (87)–(88) as well. In other words, each receiver in the channel given by (87)–(88) will decode the messages intended to itself. However, we still need to check the secrecy requirements on the confidential messages W_1, W_2. We first check the secrecy of the first user's confidential message as follows:

$$\lim_{n\to\infty}\frac{1}{n}I(W_1;\overline{\mathbf{Y}}_2^n, W_0, W_2)$$
$$= \lim_{n\to\infty}\frac{1}{n}I(W_1;\overline{\mathbf{Y}}_2^n, W_0, W_2) - \frac{1}{n}I(W_1;\mathbf{Y}_2^n, W_0, W_2) \quad (95)$$

where we used the fact that since $(R_0, R_1, R_2) \in \mathcal{C}(\mathbf{S})$, we have

$$\lim_{n\to\infty}\frac{1}{n}I(W_1;\mathbf{Y}_2^n, W_0, W_2) = 0. \quad (96)$$

We now bound the term on the right hand-side of as follows (95):

$$I(W_1;\overline{\mathbf{Y}}_2^n, W_0, W_2) - I(W_1;\mathbf{Y}_2^n, W_0, W_2)$$
$$= I(W_1;\overline{\mathbf{Y}}_2^n|W_0, W_2) - I(W_1;\mathbf{Y}_2^n|W_0, W_2) \quad (97)$$
$$= I(W_1;\overline{\mathbf{Y}}_2^n|W_0, W_2, \mathbf{Y}_2^n) \quad (98)$$
$$= \sum_{i=1}^n I(W_1;\overline{\mathbf{Y}}_{2,i}|W_0, W_2, \mathbf{Y}_2^n, \overline{\mathbf{Y}}_2^{i-1}) \quad (99)$$
$$\leq \sum_{i=1}^n h(\overline{\mathbf{Y}}_{2,i}|\mathbf{Y}_{2,i}) - h(\overline{\mathbf{Y}}_{2,i}|W_0, W_2, \mathbf{Y}_2^n, \overline{\mathbf{Y}}_2^{i-1}, W_1, \mathbf{X}_i) \quad (100)$$
$$= \sum_{i=1}^n I(\mathbf{X}_i;\overline{\mathbf{Y}}_{2,i}|\mathbf{Y}_{2,i}) \quad (101)$$
$$= \sum_{i=1}^n I(\mathbf{X}_i;\overline{\mathbf{Y}}_{2,i}) - I(\mathbf{X}_i;\mathbf{Y}_{2,i}) \quad (102)$$
$$= \sum_{i=1}^n h(\overline{\mathbf{Y}}_{2,i}) - h(\mathbf{Y}_{2,i}) \quad (103)$$
$$\leq \sum_{i=1}^n \frac{1}{2}\log\frac{\left|\overline{\mathbf{H}}_2 E\left[\mathbf{X}_i\mathbf{X}_i^\top\right]\overline{\mathbf{H}}_2^\top + \boldsymbol{\Sigma}_2\right|}{\left|\mathbf{H}_2 E\left[\mathbf{X}_i\mathbf{X}_i^\top\right]\mathbf{H}_2^\top + \boldsymbol{\Sigma}_2\right|} \quad (104)$$
$$\leq \frac{n}{2}\log\frac{\left|\overline{\mathbf{H}}_2\left(\sum_{i=1}^n\frac{1}{n}E\left[\mathbf{X}_i\mathbf{X}_i^\top\right]\right)\overline{\mathbf{H}}_2^\top + \boldsymbol{\Sigma}_2\right|}{\left|\mathbf{H}_2\left(\sum_{i=1}^n\frac{1}{n}E\left[\mathbf{X}_i\mathbf{X}_i^\top\right]\right)\mathbf{H}_2^\top + \boldsymbol{\Sigma}_2\right|} \quad (105)$$
$$\leq \frac{n}{2}\log\frac{\left|\overline{\mathbf{H}}_2\mathbf{S}\overline{\mathbf{H}}_2^\top + \boldsymbol{\Sigma}_2\right|}{\left|\mathbf{H}_2\mathbf{S}\mathbf{H}_2^\top + \boldsymbol{\Sigma}_2\right|} \quad (106)$$

where (98) is due to the Markov chain in (94), (100) comes from the fact that conditioning cannot increase entropy, (101) is due to the fact that the channel is memoryless, (102) results from the Markov chain in (94), and (104) can be shown by using the worst additive noise lemma in [19] and [20]. Before showing the steps in (105) and (106), we note that the following function:

$$\log \frac{\left| \overline{\mathbf{H}}_2 \mathbf{K} \overline{\mathbf{H}}_2^\top + \boldsymbol{\Sigma}_2 \right|}{\left| \mathbf{H}_2 \mathbf{K} \mathbf{H}_2^\top + \boldsymbol{\Sigma}_2 \right|} \tag{107}$$

is concave and monotonically increasing in positive semidefinite matrices $\mathbf{K}$ (see [21, Lemma 4]). Thus, (105) follows from Jensen's inequality by noting the concavity of the function in (107) and (106) comes from the monotonicity of the function in (107) and the covariance constraint on the channel input. Hence, using (106) in (95), we have

$$\lim_{n \to \infty} \frac{1}{n} I(W_1; \overline{\mathbf{Y}}_2^n, W_0, W_2) \le \frac{1}{2} \log \frac{\left| \overline{\mathbf{H}}_2 \mathbf{S} \overline{\mathbf{H}}_2^\top + \boldsymbol{\Sigma}_2 \right|}{\left| \mathbf{H}_2 \mathbf{S} \mathbf{H}_2^\top + \boldsymbol{\Sigma}_2 \right|} \tag{108}$$

where the right-hand side vanishes as $\alpha \to 0$, i.e.,

$$\lim_{\alpha \to 0} \frac{1}{2} \log \frac{\left| \overline{\mathbf{H}}_2 \mathbf{S} \overline{\mathbf{H}}_2^\top + \boldsymbol{\Sigma}_2 \right|}{\left| \mathbf{H}_2 \mathbf{S} \mathbf{H}_2^\top + \boldsymbol{\Sigma}_2 \right|} = 0 \tag{109}$$

due to the continuity of $\log | \cdot |$ in positive semidefinite matrices and $\lim_{\alpha \to 0} \overline{\mathbf{H}}_2 = \mathbf{H}_2$. Thus, we have shown that if a confidential message W_1 with rate R_1 can be transmitted in perfect secrecy in the original channel given by (1)–(2), we have

$$\lim_{\alpha \to 0} \lim_{n \to \infty} \frac{1}{n} I(W_1; \overline{\mathbf{Y}}_2^n, W_0, W_2) = 0. \tag{110}$$

Similarly, if a confidential message W_2 with rate R_2 can be transmitted in perfect secrecy in the original channel given by (1)–(2), we have

$$\lim_{\alpha \to 0} \lim_{n \to \infty} \frac{1}{n} I(W_2; \overline{\mathbf{Y}}_1^n, W_0, W_1) = 0. \tag{111}$$

These two conditions in (110) and (111) enable us to conclude that if $(R_0, R_1, R_2) \in \mathcal{C}(\mathbf{S})$, we also have $(R_0, R_1, R_2) \in \lim_{\alpha \to 0} \mathcal{C}_\alpha(\mathbf{S})$. Thus, we have shown that

$$\mathcal{C}(\mathbf{S}) \subseteq \lim_{\alpha \to 0} \mathcal{C}_\alpha(\mathbf{S}) = \lim_{\alpha \to 0} \mathcal{R}_{12,\alpha}^{\mathrm{S-DPC}}(\mathbf{S}) = \lim_{\alpha \to 0} \mathcal{R}_{21,\alpha}^{\mathrm{S-DPC}}(\mathbf{S}) \tag{112}$$

where we have

$$\lim_{\alpha \to 0} \mathcal{R}_{12,\alpha}^{\mathrm{S-DPC}}(\mathbf{S}) = \mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S}) \tag{113}$$

$$\lim_{\alpha \to 0} \mathcal{R}_{21,\alpha}^{\mathrm{S-DPC}}(\mathbf{S}) = \mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S}) \tag{114}$$

due to the continuity of the rate expressions in $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S})$ and $\mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S})$ in α. Since $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S})$ and $\mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S})$ are achievable in the channel defined by (1)–(2), we have

$$\mathcal{C}(\mathbf{S}) = \mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S}) = \mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S}) \tag{115}$$

in the view of (112)–(114), completing the proof.

V. CONNECTIONS TO THE GAUSSIAN MIMO BROADCAST CHANNEL WITH COMMON AND PRIVATE MESSAGES

Here, we provide intuitive explanations for the two facts that Theorem 1 reveals: 1) the achievable rate region does not depend on the encoding order used in S-DPC, i.e., $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S}) = \mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S})$; and 2) the capacity region of the Gaussian MIMO broadcast channel with common and confidential messages can be completely characterized, although the capacity region of the its nonconfidential counterpart, i.e., the Gaussian MIMO broadcast channel with common and private messages, is not known completely.

In the Gaussian MIMO broadcast channel with common and private messages, there are again three messages W_0, W_1, W_2 with rates R_0, R_1, R_2, respectively, such that W_0 is again sent to both users, W_1 (respectively, W_2) is again directed to only the first (respectively, second) user, however, there are no secrecy constraints on W_1, W_2. The capacity region of the Gaussian MIMO broadcast channel with common and private messages will be denoted by $\mathcal{C}^{\mathrm{NS}}(\mathbf{S})$. The achievable rate region for the Gaussian MIMO broadcast channel with common and private messages that can be obtained by using DPC will be denoted by $\mathcal{R}_{12}^{\mathrm{NS-DPC}}(\mathbf{S}), \mathcal{R}_{21}^{\mathrm{NS-DPC}}(\mathbf{S})$ (depending on the encoding order), where $\mathcal{R}_{12}^{\mathrm{NS-DPC}}(\mathbf{S})$ is given by the rate triples (R_0, R_1, R_2) satisfying

$$R_0 \le \min\{R_{01}^{\mathrm{NS}}(\mathbf{K}_1, \mathbf{K}_2), R_{02}^{\mathrm{NS}}(\mathbf{K}_1, \mathbf{K}_2)\} \tag{116}$$

$$R_1 \le R_1^{\mathrm{NS}}(\mathbf{K}_1, \mathbf{K}_2) \tag{117}$$

$$R_2 \le R_2^{\mathrm{NS}}(\mathbf{K}_2) \tag{118}$$

for some positive semidefinite matrices $\mathbf{K}_1, \mathbf{K}_2$ such that $\mathbf{K}_1 + \mathbf{K}_2 \preceq \mathbf{S}$, and $\{R_{0j}^{\mathrm{NS}}(\mathbf{K}_1, \mathbf{K}_2)\}_{j=1}^2, R_1^{\mathrm{NS}}(\mathbf{K}_1, \mathbf{K}_2), R_2^{\mathrm{NS}}(\mathbf{K}_1, \mathbf{K}_2)$ are defined as

$$R_{0j}^{\mathrm{NS}}(\mathbf{K}_1, \mathbf{K}_2) = \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_j|}{|\mathbf{K}_1 + \mathbf{K}_2 + \boldsymbol{\Sigma}_j|}, \quad j = 1, 2 \tag{119}$$

$$R_1^{\mathrm{NS}}(\mathbf{K}_1, \mathbf{K}_2) = \frac{1}{2} \log \frac{|\mathbf{K}_1 + \mathbf{K}_2 + \boldsymbol{\Sigma}_1|}{|\mathbf{K}_2 + \boldsymbol{\Sigma}_1|} \tag{120}$$

$$R_2^{\mathrm{NS}}(\mathbf{K}_1, \mathbf{K}_2) = \frac{1}{2} \log \frac{|\mathbf{K}_2 + \boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_2|}. \tag{121}$$

Moreover, $\mathcal{R}_{21}^{\mathrm{NS-DPC}}(\mathbf{S})$ can be obtained from $\mathcal{R}_{12}^{\mathrm{NS-DPC}}(\mathbf{S})$ by swapping the subscripts 2 and 1. We now state a result of [12] on the capacity region of the Gaussian MIMO broadcast channel with common and private messages: For a given common message rate R_0, the private message sum rate capacity, i.e., $R_1 + R_2$, is achieved by both $\mathcal{R}_{12}^{\mathrm{NS}}(\mathbf{S})$ and $\mathcal{R}_{21}^{\mathrm{NS}}(\mathbf{S})$. This result can also be stated as follows:

$$\max_{(R_0, R_1, R_2) \in \mathcal{C}^{\mathrm{NS}}(\mathbf{S})} \mu_0' R_0 + \mu_1' R_1 + \mu_2' R_2$$

$$= \max_{(R_0, R_1, R_2) \in \mathcal{R}_{12}^{\mathrm{NS-DPC}}(\mathbf{S})} \mu_0' R_0 + \mu_1' R_1 + \mu_2' R_2 \tag{122}$$

$$= \max_{(R_0, R_1, R_2) \in \mathcal{R}_{21}^{\mathrm{NS-DPC}}(\mathbf{S})} \mu_0' R_0 + \mu_1' R_1 + \mu_2' R_2 \tag{123}$$

for $\mu_1' = \mu_2' = \mu'$. This result is crucial to understand the aforementioned two points suggested by Theorem 1, which will be explained next using (122)–(123).

In the proof of Theorem 1, first, we characterize the boundary of $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S})$ by finding the properties of the covariance matrices that achieve the boundary of $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S})$, see Lemma 2. According to Lemma 2, the boundary of $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S})$ can be achieved by using the covariance matrices $\mathbf{K}_1^*, \mathbf{K}_2^*$ satisfying

$$
\begin{aligned}
(\mu_1 + \mu_2)&(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \mathbf{M}_1 \\
&= (\mu_0 \lambda + \mu_2)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} \\
&\quad + (\mu_0 \bar{\lambda} + \mu_1)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} + \mathbf{M}_S \quad (124)
\end{aligned}
$$

$$
\begin{aligned}
(\mu_1 + \mu_2)&(\mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} + \mathbf{M}_2 \\
&= (\mu_1 + \mu_2)(\mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \mathbf{M}_1. \quad (125)
\end{aligned}
$$

On the other hand, using these covariance matrices, we can also achieve the boundary points of $\mathcal{R}_{12}^{\mathrm{NS-DPC}}(\mathbf{S})$, which are actually on the boundary of the capacity region $\mathcal{C}^{\mathrm{NS}}(\mathbf{S})$ as well, and are the private message sum rate capacity points for a given common message rate. To see this point, we define $\mu' = \mu_1 + \mu_2, \mu_0' = \mu_0 + \mu_1 + \mu_2$ and $\gamma = 1 - \bar{\gamma} = \frac{\mu_0 \lambda + \mu_2}{\mu_0 + \mu_1 + \mu_2}$ Thus, the conditions in (124)–(125) can be written as

$$
\begin{aligned}
\mu'&(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \mathbf{M}_1 \\
&= \mu_0' \gamma (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \mu_0' \bar{\gamma} (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} \\
&\quad + \mathbf{M}_S \quad (126)
\end{aligned}
$$

$$
\mu'(\mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} + \mathbf{M}_2 = \mu'(\mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \mathbf{M}_1 \quad (127)
$$

which are the necessary conditions that the following problem needs to satisfy

$$
\max_{(R_0, R_1, R_2) \in \mathcal{R}_{12}^{\mathrm{NS-DPC}}(\mathbf{S})} \mu_0' R_0 + \mu'(R_1 + R_2). \quad (128)
$$

On the other hand, due to (122)–(123), we know that the solution of (128) gives us the private message sum rate capacity for a given common message rate, i.e., the points that achieve the maximum in (128) are on the boundary of the capacity region $\mathcal{C}^{\mathrm{NS}}(\mathbf{S})$. Furthermore, the maximum value in (128) can also be achieved by using the other possible encoding order, i.e.,

$$
\begin{aligned}
\max_{(R_0, R_1, R_2) \in \mathcal{R}_{12}^{\mathrm{NS-DPC}}(\mathbf{S})}& \mu_0' R_0 + \mu'(R_1 + R_2) \\
&= \max_{(R_0, R_1, R_2) \in \mathcal{R}_{21}^{\mathrm{NS-DPC}}(\mathbf{S})} \mu_0' R_0 + \mu'(R_1 + R_2). \quad (129)
\end{aligned}
$$

Thus, this discussion reveals that there is a one-to-one correspondence between any rate triple on the boundary of $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S})$ and the private message sum rate capacity points on $\mathcal{C}^{\mathrm{NS}}(\mathbf{S})$. Hence, the boundary of $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S})$, similarly $\mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S})$, can be constructed by considering the private message sum rate capacity points on $\mathcal{C}^{\mathrm{NS}}(\mathbf{S})$. This connection between the private message sum rate capacity points and the boundaries of $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S})$, $\mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S})$ intuitively explains the two facts suggested by Theorem 1: 1) the achievable rate region for the Gaussian MIMO broadcast channel with common and confidential messages is invariant with respect to the encoding order, i.e., $\mathcal{R}_{12}^{\mathrm{S-DPC}}(\mathbf{S}) = \mathcal{R}_{21}^{\mathrm{S-DPC}}(\mathbf{S})$ because the boundaries of these two regions correspond to those points on the DPC region for the Gaussian MIMO broadcast channel with common and private messages, for which encoding order does

not matter either; and 2) we can obtain the entire capacity region of the Gaussian MIMO broadcast channel with common and confidential messages, although the capacity region of its nonconfidential counterpart is not known completely. The reason is that the boundary of the capacity region of the Gaussian MIMO broadcast channel with common and confidential messages comes from those points on the boundary of the DPC region of its nonconfidential counterpart, which are known to be tight, i.e., which are known to be on the boundary of the capacity region of the Gaussian MIMO broadcast channel with common and private messages.

VI. CONCLUSION

We study the Gaussian MIMO broadcast channel with common and confidential messages, and obtain the entire capacity region. We show that a variant of the S-DPC scheme proposed in [5] is capacity-achieving. We provide the converse proof by using channel enhancement [9] and an extremal inequality from [10]. We also uncover the connections between the Gaussian MIMO broadcast channel with common and confidential messages and its nonconfidential counterpart, i.e., the Gaussian MIMO broadcast channel with common and private messages, to provide further insight into capacity result we obtained.

APPENDIX I
PROOF OF LEMMA 2

Since the program in (45) is not necessarily convex, the KKT conditions are necessary but not sufficient. We first rewrite the program in (45) as follows:

$$
\begin{aligned}
\max_{\substack{\mathbf{0} \preceq \mathbf{K}_j, \, j=1,2 \\ \mathbf{K}_1 + \mathbf{K}_2 \preceq \mathbf{S} \\ a}}& \quad \mu_0 a + \mu_1 R_1^{\mathrm{AL}}(\mathbf{K}_1, \mathbf{K}_2) + \mu_2 R_2^{\mathrm{AL}}(\mathbf{K}_2) \\
\text{s.t.}& \quad R_{01}^{\mathrm{AL}}(\mathbf{K}_1, \mathbf{K}_2) \geq a \\
& \quad R_{02}^{\mathrm{AL}}(\mathbf{K}_1, \mathbf{K}_2) \geq a \quad (130)
\end{aligned}
$$

where we introduce an additional variable a. Thus, the optimization in (130) is over three variables $a, \mathbf{K}_1, \mathbf{K}_2$. The Lagrangian of (130) is given by

$$
\begin{aligned}
\mathcal{L} &= \mu_0 a + \mu_1 R_1^{\mathrm{AL}}(\mathbf{K}_1, \mathbf{K}_2) + \mu_2 R_2^{\mathrm{AL}}(\mathbf{K}_2) \\
&\quad + \mu_0 \sum_{j=1}^{2} \lambda_j (R_{0j}^{\mathrm{AL}}(\mathbf{K}_1, \mathbf{K}_2) - a) + \operatorname{tr}(\mathbf{K}_1 \mathbf{M}_1) \\
&\quad + \operatorname{tr}(\mathbf{K}_2 \mathbf{M}_2) + \operatorname{tr}((\mathbf{S} - \mathbf{K}_1 - \mathbf{K}_2)\mathbf{M}_S) \quad (131)
\end{aligned}
$$

where $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_S$ are positive semidefinite matrices and $\lambda_j \geq 0, \, j = 1, 2$. Let $(a^*, \mathbf{K}_1^*, \mathbf{K}_2^*)$ be the maximizer for (130). The necessary KKT conditions that they need to satisfy are given as follows:

$$
\frac{\partial \mathcal{L}}{\partial a} \bigg|_{a=a^*} = 0 \quad (132)
$$

$$
\nabla_{\mathbf{K}_1} \mathcal{L} \big|_{\mathbf{K}_1 = \mathbf{K}_1^*} = \mathbf{0} \quad (133)
$$

$$
\nabla_{\mathbf{K}_2} \mathcal{L} \big|_{\mathbf{K}_2 = \mathbf{K}_2^*} = \mathbf{0} \quad (134)
$$

$$
\operatorname{tr}(\mathbf{K}_1^* \mathbf{M}_1) = 0 \quad (135)
$$

$$
\operatorname{tr}(\mathbf{K}_2^* \mathbf{M}_2) = 0 \quad (136)
$$

$$
\operatorname{tr}((\mathbf{S} - \mathbf{K}_1^* - \mathbf{K}_2^*)\mathbf{M}_S) = 0 \quad (137)
$$

$$
\lambda_j (R_{0j}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) - a^*) = 0, \quad j = 1, 2. \quad (138)
$$

The first KKT condition in (132) implies $\lambda_1 + \lambda_2 = 1$. We define $\lambda = \lambda_1$ and consequently $\bar{\lambda} = 1 - \lambda = \lambda_2$. The second KKT condition in (133) implies

$$
\mu_1 (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \mathbf{M}_1
$$
$$
= \mu_0 \lambda (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + (\mu_0 \bar{\lambda} + \mu_1)(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1}
$$
$$
+ \mathbf{M}_S. \tag{139}
$$

Adding $\mu_2 (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1}$ to both sides yields (46). Subtracting (133) from (134) yields (47). Since $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$ and $\mathrm{tr}(\mathbf{AB}) \geq 0$ for $\mathbf{A} \succeq \mathbf{0}, \mathbf{B} \succeq \mathbf{0}$, (135)–(137) imply (48)–(50). Furthermore, (138) states the conditions if $R_{01}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) > R_{02}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*)$, $\lambda = 0$, if $R_{01}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) < R_{02}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*)$, $\lambda = 1$, and if $R_{01}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*) = R_{02}^{\mathrm{AL}}(\mathbf{K}_1^*, \mathbf{K}_2^*)$, λ is arbitrary, i.e., $0 < \lambda < 1$.

APPENDIX II
PROOF OF LEMMA 3

We first note the following identities:

$$
(\mu_1 + \mu_2)(\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1} = (\mu_1 + \mu_2)(\mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} + \mathbf{M}_2 \tag{140}
$$

$$
(\mu_1 + \mu_2)(\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1} = (\mu_1 + \mu_2)(\mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \mathbf{M}_1 \tag{141}
$$

where (140) is the definition of the new noise covariance matrix in (52) and (141) comes from plugging (52) in (47). Using the fact that for $\mathbf{A} \succ \mathbf{0}, \mathbf{B} \succ \mathbf{0}$, if $\mathbf{A} \preceq \mathbf{B}$, then $\mathbf{A}^{-1} \succeq \mathbf{B}^{-1}$ in (140)–(141) yields the second statement of the lemma.

Now, we prove the first statement of the lemma as follows:

$$
\tilde{\boldsymbol{\Sigma}} = \left[(\mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_2 \right]^{-1} - \mathbf{K}_2^* \tag{142}
$$

$$
= \left[\mathbf{I} + \frac{1}{\mu_1 + \mu_2} (\mathbf{K}_2^* + \boldsymbol{\Sigma}_2) \mathbf{M}_2 \right]^{-1} (\mathbf{K}_2^* + \boldsymbol{\Sigma}_2) - \mathbf{K}_2^* \tag{143}
$$

$$
= \left[\mathbf{I} + \frac{1}{\mu_1 + \mu_2} \boldsymbol{\Sigma}_2 \mathbf{M}_2 \right]^{-1} (\mathbf{K}_2^* + \boldsymbol{\Sigma}_2) - \mathbf{K}_2^* \tag{144}
$$

$$
= \left[\boldsymbol{\Sigma}_2^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_2 \right]^{-1} \boldsymbol{\Sigma}_2^{-1} (\mathbf{K}_2^* + \boldsymbol{\Sigma}_2) - \mathbf{K}_2^* \tag{145}
$$

$$
= \left[\boldsymbol{\Sigma}_2^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_2 \right]^{-1}
$$
$$
+ \left[\boldsymbol{\Sigma}_2^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_2 \right]^{-1} \boldsymbol{\Sigma}_2^{-1} \mathbf{K}_2^* - \mathbf{K}_2^* \tag{146}
$$

$$
= \left[\boldsymbol{\Sigma}_2^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_2 \right]^{-1}
$$
$$
+ \left[\boldsymbol{\Sigma}_2^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_2 \right]^{-1} \left[\boldsymbol{\Sigma}_2^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_2 \right] \mathbf{K}_2^*
$$
$$
- \mathbf{K}_2^* \tag{147}
$$

$$
= \left[\boldsymbol{\Sigma}_2^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_2 \right]^{-1} \tag{148}
$$

$$
\succ \mathbf{0} \tag{149}
$$

where (142) is due to (140), and (144) and (147) follow from (49).

We next show the third statement of the lemma as follows:

$$
\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}} = \left[(\mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_1 \right]^{-1} \tag{150}
$$

$$
= \left[\mathbf{I} + \frac{1}{\mu_1 + \mu_2} (\mathbf{K}_2^* + \boldsymbol{\Sigma}_1) \mathbf{M}_1 \right]^{-1} (\mathbf{K}_2^* + \boldsymbol{\Sigma}_1) \tag{151}
$$

$$
= \left[\mathbf{I} + \frac{1}{\mu_1 + \mu_2} (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1) \mathbf{M}_1 \right]^{-1} (\mathbf{K}_2^* + \boldsymbol{\Sigma}_1) \tag{152}
$$

$$
= \left[(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_1 \right]^{-1}
$$
$$
\times (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} (\mathbf{K}_2^* + \boldsymbol{\Sigma}_1) \tag{153}
$$

$$
= \left[(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_1 \right]^{-1}
$$
$$
\times (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1 - \mathbf{K}_1^*) \tag{154}
$$

$$
= \left[(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_1 \right]^{-1}
$$
$$
- \left[(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_1 \right]^{-1}
$$
$$
\times (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} \mathbf{K}_1^* \tag{155}
$$

$$
= \left[(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_1 \right]^{-1}
$$
$$
- \left[(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_1 \right]^{-1}
$$
$$
\times \left[(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_1 \right] \mathbf{K}_1^* \tag{156}
$$

$$
= \left[(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_1 \right]^{-1} - \mathbf{K}_1^* \tag{157}
$$

where (150) is due to (141), (152) and (156) come from (48).

We now show the fourth statement of the lemma as follows:

$$
(\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1} \tilde{\boldsymbol{\Sigma}} = \mathbf{I} - (\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1} \mathbf{K}_2^* \tag{158}
$$

$$
= \mathbf{I} - \left[(\mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_2 \right] \mathbf{K}_2^* \tag{159}
$$

$$
= \mathbf{I} - (\mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} \mathbf{K}_2^* \tag{160}
$$

$$
= (\mathbf{K}_2^* + \boldsymbol{\Sigma}_2)^{-1} \boldsymbol{\Sigma}_2 \tag{161}
$$

where (159) comes from (140), and (160) is due to (49).

We finally show the last, i.e., fifth, statement of the lemma as follows:

$$
(\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1} (\mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})
$$
$$
= \mathbf{I} - (\mathbf{K}_1^* + \mathbf{K}_2^* + \tilde{\boldsymbol{\Sigma}})^{-1} \mathbf{K}_1^* \tag{162}
$$

$$
= \mathbf{I} - \left[(\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} + \frac{1}{\mu_1 + \mu_2} \mathbf{M}_1 \right] \mathbf{K}_1^* \tag{163}
$$

$$
= \mathbf{I} - (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} \mathbf{K}_1^* \tag{164}
$$

$$
= (\mathbf{K}_1^* + \mathbf{K}_2^* + \boldsymbol{\Sigma}_1)^{-1} (\mathbf{K}_2^* + \boldsymbol{\Sigma}_1) \tag{165}
$$

where (163) comes from the second statement of this lemma, and (164) is due to (48).

APPENDIX III

PROOF OF LEMMA 4

We prove this lemma for a discrete memoryless broadcast channel with a transition probability $p(\tilde{y}_1, \tilde{y}_2, y_1, y_2 | x)$ which satisfies $p(\tilde{y}_1 | x) = p(\tilde{y}_2 | x) = p(\tilde{y} | x)$ and

$$X \to \tilde{Y} \to (Y_1, Y_2). \tag{166}$$

Consequently, Lemma 4 can be concluded from the proof for this discrete memoryless broadcast channel. We note that if (R_0, R_1, R_2) is achievable, we need to have ϵ_n, γ_n such that both ϵ_n and γ_n vanish as $n \to \infty$, and

$$H(W_0 | Y_j^n) \leq n\epsilon_n, \quad j = 1, 2 \tag{167}$$

$$H(W_j | \tilde{Y}^n, W_0) \leq n\epsilon_n, \quad j = 1, 2 \tag{168}$$

$$I(W_1; Y_2^n, W_0) \leq n\gamma_n \tag{169}$$

$$I(W_2; Y_1^n, W_0) \leq n\gamma_n \tag{170}$$

where (167)–(168) are due to Fano's lemma, and (169)–(170) comes from the perfect secrecy conditions in (58). We define the following auxiliary random variables:

$$U_i = W_0 \tilde{Y}^{i-1}, \quad i = 1, \ldots, n \tag{171}$$

which satisfy the following Markov chains for all i:

$$U_i \to X_i \to \tilde{Y}_i \to (Y_{1i}, Y_{2i}) \tag{172}$$

since the channel is memoryless, and degraded, i.e., satisfies the Markov chain in (166).

We first bound the common message rate R_0 as follows:

$$nR_0 = H(W_0) \tag{173}$$

$$\leq I(W_0; Y_1^n) + n\epsilon_n \tag{174}$$

$$= \sum_{i=1}^{n} I(W_0; Y_{1i} | Y_1^{i-1}) + n\epsilon_n \tag{175}$$

$$\leq \sum_{i=1}^{n} I(W_0, \tilde{Y}^{i-1}, Y_1^{i-1}; Y_{1i}) + n\epsilon_n \tag{176}$$

$$= \sum_{i=1}^{n} I(W_0, \tilde{Y}^{i-1}; Y_{1i}) + n\epsilon_n \tag{177}$$

$$= \sum_{i=1}^{n} I(U_i; Y_{1i}) + n\epsilon_n \tag{178}$$

where (177) comes from the Markov chain

$$Y_1^{i-1} \to \tilde{Y}^{i-1} \to (W_0, Y_{1i}) \tag{179}$$

which is a consequence of the fact that the channel is degraded, i.e., satisfies the Markov chain in (166). Similarly, we can get

$$nR_0 \leq \sum_{i=1}^{n} I(U_i; Y_{2i}) + n\epsilon_n. \tag{180}$$

We next bound the confidential message rate of the enhanced first user, i.e., R_1, as follows:

$$nR_1 = H(W_1 | W_0) \tag{181}$$

$$\leq I(W_1; \tilde{Y}^n | W_0) - I(W_1; Y_2^n | W_0) + n(\epsilon_n + \gamma_n) \tag{182}$$

$$\leq I(W_1; \tilde{Y}^n | W_0, Y_2^n) + n(\epsilon_n + \gamma_n) \tag{183}$$

$$= \sum_{i=1}^{n} I(W_1; \tilde{Y}_i | W_0, Y_2^n, \tilde{Y}^{i-1}) + n(\epsilon_n + \gamma_n) \tag{184}$$

$$= \sum_{i=1}^{n} I(W_1; \tilde{Y}_i | W_0, Y_{2i}^n, \tilde{Y}^{i-1}) + n(\epsilon_n + \gamma_n) \tag{185}$$

$$\leq \sum_{i=1}^{n} I(W_1, X_i; \tilde{Y}_i | W_0, Y_{2i}^n, \tilde{Y}^{i-1}) + n(\epsilon_n + \gamma_n) \tag{186}$$

$$= \sum_{i=1}^{n} I(X_i; \tilde{Y}_i | W_0, Y_{2i}^n, \tilde{Y}^{i-1}) + n(\epsilon_n + \gamma_n) \tag{187}$$

$$= \sum_{i=1}^{n} H(\tilde{Y}_i | W_0, Y_{2i}^n, \tilde{Y}^{i-1}) - H(\tilde{Y}_i | W_0, Y_{2i}^n, \tilde{Y}^{i-1}, X_i)$$
$$+ n(\epsilon_n + \gamma_n) \tag{188}$$

$$\leq \sum_{i=1}^{n} H(\tilde{Y}_i | W_0, Y_{2i}, \tilde{Y}^{i-1}) - H(\tilde{Y}_i | W_0, Y_{2i}^n, \tilde{Y}^{i-1}, X_i)$$
$$+ n(\epsilon_n + \gamma_n) \tag{189}$$

$$= \sum_{i=1}^{n} H(\tilde{Y}_i | W_0, Y_{2i}, \tilde{Y}^{i-1}) - H(\tilde{Y}_i | W_0, Y_{2i}, \tilde{Y}^{i-1}, X_i)$$
$$+ n(\epsilon_n + \gamma_n) \tag{190}$$

$$= \sum_{i=1}^{n} I(X_i; \tilde{Y}_i | U_i, Y_{2i}) + n(\epsilon_n + \gamma_n) \tag{191}$$

$$= \sum_{i=1}^{n} I(X_i; \tilde{Y}_i | U_i) - I(X_i; Y_{2i} | U_i) + n(\epsilon_n + \gamma_n) \tag{192}$$

where (185) comes from the Markov chain

$$W_0, W_1, Y_{2i}^n \to \tilde{Y}^{i-1} \to Y_2^{i-1} \tag{193}$$

which is a consequence of the fact that the channel is degraded, i.e., satisfies the Markov chain in (166), (187) comes from the Markov chain

$$W_0, W_1, \tilde{Y}^{i-1}, Y_{2(i+1)}^n \to X_i \to \tilde{Y}_i, Y_{2i} \tag{194}$$

which is due to the fact that the channel is memoryless, (189) comes from the fact that conditioning cannot increase entropy, (190) results from the Markov chain in (194), and (192) stems from the Markov chain in (172). Similarly, we can get the following bound on the confidential message rate of the enhanced second user R_2:

$$nR_2 \leq \sum_{i=1}^{n} I(X_i; \tilde{Y}_i | U_i) - I(X_i; Y_{2,i} | U_i) + n(\epsilon_n + \gamma_n). \tag{195}$$

The bounds in (178), (180), (192) and (195) can be single-letterized yielding the following bounds:

$$R_0 \leq \min\{I(U; Y_1), I(U; Y_2)\} \tag{196}$$

$$R_1 \leq I(X; \tilde{Y} | U) - I(X; Y_2 | U) \tag{197}$$

$$R_2 \leq I(X; \tilde{Y} | U) - I(X; Y_1 | U) \tag{198}$$

from which, Lemma 4 can be concluded.

REFERENCES

[1] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[2] A. Khisti and G. Wornell, "Secure transmission with multiple antennas II: The MIMOME channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[3] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4033–4039, Sep. 2009.

[4] H. D. Ly, T. Liu, and Y. Liang, "Multiple-input multiple-output Gaussian broadcast channels with common and confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5477–5487, Nov. 2010.

[5] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4215–4227, Sep. 2010.

[6] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "MIMO Gaussian broadcast channels with confidential and common messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 2578–2582.

[7] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "New results on multiple-input multiple-output broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, Jan. 2011, also available at [arXiv: 1101:2007], submitted for publication.

[8] E. Ekrem and S. Ulukus, "Gaussian MIMO broadcast channels with common and confidential messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 2583–2587.

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

[10] H. Weingarten, T. Liu, S. Shamai (Shitz), Y. Steinberg, and P. Viswanath, "The capacity region of the degraded multiple-input multiple-output compound broadcast channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5011–5023, Nov. 2009.

[11] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "On the capacity region of the multi-antenna broadcast channel with common messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 2195–2199.

[12] H. Weingarten, "Multiple-input multiple-output broadcast systems," Ph.D. dissertation, Technion, Haifa, Israel, 2007.

[13] N. Jindal and A. Goldsmith, "Optimal power allocation for parallel broadcast channels with independent and common information," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2004, p. 215.

[14] A. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Jan. 1975.

[15] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[16] W. Yu and J. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1875–1892, Sep. 2004.

[17] E. Ekrem and S. Ulukus, "The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2083–2114, Apr. 2011.

[18] J. Xu, Y. Cao, and B. Chen, "Capacity bounds for broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4529–4542, Oct. 2009.

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

[20] S. Ihara, "On the capacity of channels with additive non-Gaussian noise," *Inf. Control*, vol. 37, no. 1, pp. 34–39, Apr. 1978.

[21] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wire-tap channels," *EURASIP J. Wireless Commun. Netw. (Special Issue Wireless Physical Layer Security)*, vol. 2049, p. 142375, 2009.

Ersen Ekrem (S'08) received the B.S. and M.S. degrees in electrical and electronics engineering from Bogaziçi University, Istanbul, Turkey, in 2006 and 2007, respectively. Currently, he is working toward the Ph.D. degree in the department of electrical and computer engineering at the University of Maryland, College Park.

He received the Distinguished Dissertation Fellowship from the ECE Department at the University of Maryland, College Park, in 2012. His research interests include information theory and wireless communications.

Sennur Ulukus (S'90–M'98) is a Professor of Electrical and Computer Engineering at the University of Maryland at College Park, where she also holds a joint appointment with the Institute for Systems Research (ISR). Prior to joining UMD, she was a Senior Technical Staff Member at AT&T Labs-Research. She received her Ph.D. degree in Electrical and Computer Engineering from Wireless Information Network Laboratory (WINLAB), Rutgers University, and B.S. and M.S. degrees in Electrical and Electronics Engineering from Bilkent University. Her research interests are in wireless communication theory and networking, network information theory for wireless communications, signal processing for wireless communications, information-theoretic physical-layer security, and energy-harvesting communications.

Dr. Ulukus received the 2003 IEEE Marconi Prize Paper Award in Wireless Communications, the 2005 NSF CAREER Award, and the 2010–2011 ISR Outstanding Systems Engineering Faculty Award. She served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY between 2007–2010, as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS between 2003–2007, as a Guest Editor for the *Journal of Communications and Networks* for the special issue on energy harvesting in wireless networks, as a Guest Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY for the special issue on interference networks, as a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS for the special issue on multiuser detection for advanced communication systems and networks. She served as the TPC co-chair of the Communication Theory Symposium at the 2007 IEEE Global Telecommunications Conference, the Medium Access Control (MAC) Track at the 2008 IEEE Wireless Communications and Networking Conference, the Wireless Communications Symposium at the 2010 IEEE International Conference on Communications, the 2011 Communication Theory Workshop, the Physical-Layer Security Workshop at the 2011 IEEE International Conference on Communications, the Physical-Layer Security Workshop at the 2011 IEEE Global Telecommunications Conference. She was the Secretary of the IEEE Communication Theory Technical Committee (CTTC) in 2007–2009.