Capacity-Equivocation Region of the Gaussian MIMO Wiretap Channel

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

***Abstract*—We study the Gaussian multiple-input multiple-output (MIMO) wiretap channel, which consists of a transmitter, a legitimate user, and an eavesdropper. In this channel, the transmitter sends a common message to both the legitimate user and the eavesdropper. In addition to this common message, the legitimate user receives a private message, which is desired to be kept hidden as much as possible from the eavesdropper. We obtain the entire capacity-equivocation region of the Gaussian MIMO wiretap channel. This region contains all achievable common message, private message, and private message's equivocation (secrecy) rates. In particular, we show the sufficiency of jointly Gaussian auxiliary random variables and channel input to evaluate the existing single-letter description of the capacity-equivocation region due to Csiszar–Korner.**

***Index Terms*—Capacity-equivocation region, Gaussian multiple-input multiple-output (MIMO) wiretap channel.**

I. Introduction

We consider the Gaussian multiple-input multiple-output (MIMO) wiretap channel, which consists of a transmitter, a legitimate user, and an eavesdropper. In this channel, the transmitter sends a common message to both the legitimate user and the eavesdropper in addition to a private message which is directed to only the legitimate user. There is a secrecy concern regarding this private message in the sense that the private message needs to be kept secret as much as possible from the eavesdropper. The secrecy of the private message is measured by its equivocation at the eavesdropper.

Here, we obtain the capacity-equivocation region of the Gaussian MIMO wiretap channel. This region contains all achievable rate triples (R_0, R_1, R_e), where R_0 denotes the common message rate, R_1 denotes the private message rate, and R_e denotes the private message's equivocation (secrecy) rate. In fact, this region is known in a single-letter form due to [1]. In this paper, we show that jointly Gaussian auxiliary random variables and channel input are sufficient to evaluate this single-letter description for the capacity-equivocation region of the Gaussian MIMO wiretap channel. We prove the sufficiency of the jointly Gaussian auxiliary random variables and channel input by using channel enhancement [2] and an extremal inequality from [3]. In our proof, we also use the equivalence between the Gaussian MIMO wiretap channel and the Gaussian MIMO wiretap channel with *public* messages [4, Problem33-c], [5]. In the latter channel model, the transmitter has three messages, a common, a confidential, and a public message. The common message is sent to both the legitimate user and the eavesdropper, while the confidential and public messages are directed to only the legitimate user. Here, the confidential message needs to be transmitted in perfect secrecy, whereas there is no secrecy constraint on the public message. Since the Gaussian MIMO wiretap channel and the Gaussian MIMO wiretap channel with public messages are equivalent, i.e., there is a one-to-one correspondence between the capacity regions of these two models, in our proof, we obtain the capacity region of the Gaussian MIMO wiretap channel with public messages, which, in turn, gives us the capacity-equivocation region of the Gaussian MIMO wiretap channel.

Our result subsumes the following previous findings about the capacity-equivocation region of the Gaussian MIMO wiretap channel: 1) The secrecy capacity of this channel, i.e., $\max R_1$ when $R_0 = 0, R_e = R_1$, is obtained in [6] and [7] for the general case, and in [8] for the 2–2-1 case. 2) The common and confidential rate region under perfect secrecy, i.e., (R_0, R_1) region with $R_e = R_1$, is obtained in [9]. 3) The capacity-equivocation region without a common message, i.e., (R_1, R_e) region with $R_0 = 0$, is obtained in [5]. 4) The capacity region of the Gaussian MIMO broadcast channel with degraded message sets without a secrecy concern, i.e., (R_0, R_1) region with no consideration on R_e, is obtained in [10]. Here, we obtain the entire (R_0, R_1, R_e) region. Our result as well as the previous results listed above hold when there is a covariance constraint on the channel input as well as when there is a total power constraint on the channel input.

II. Discrete Memoryless Wiretap Channels

The discrete memoryless wiretap channel consists of a transmitter, a legitimate user, and an eavesdropper. The channel transition probability is denoted by $p(y, z|x)$, where $x \in \mathcal{X}$ is the channel input, $y \in \mathcal{Y}$ is the legitimate user's observation, and $z \in \mathcal{Z}$ is the eavesdropper's observation. We consider the following scenario for the discrete memoryless wiretap channel: the transmitter sends a common message to both the legitimate user and the eavesdropper, and a private message to the legitimate user which is desired to be kept hidden as much as possible from the eavesdropper.

An $(n, 2^{nR_0}, 2^{nR_1})$ code for this channel consists of two message sets $\mathcal{W}_0 = \{1, \ldots, 2^{nR_0}\}$, $\mathcal{W}_1 = \{1, \ldots, 2^{nR_1}\}$, one encoder at the transmitter $f : \mathcal{W}_0 \times \mathcal{W}_1 \to \mathcal{X}^n$, one decoder at the legitimate user $g_u : \mathcal{Y}^n \to \mathcal{W}_0 \times \mathcal{W}_1$, and one decoder at the eavesdropper $g_e : \mathcal{Z}^n \to W_0$. The probability of error is defined as $P_e^n = \max\{P_{e,u}^n, P_{e,e}^n\}$, where $P_{e,u}^n = \Pr[g_u(Y^n) \neq (W_0, W_1)]$, $P_{e,e}^n = \Pr[g_e(Z^n) \neq W_0]$, and W_j is a uniformly

Manuscript received May 03, 2010, revised November 19, 2011; accepted March 21, 2012. Date of publication June 12, 2012; date of current version August 14, 2012. This work was supported by the National Science Foundation under Grants CCF 04-47613, CCF 05-14846, CNS 07-16311, and CCF 07-29127. This paper was presented in part at the 2010 IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications.

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).

Communicated by S. N. Diggavi, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2012.2204534

distributed random variable in $\mathcal{W}_j,\ \ j=0,1$. We note that W_0 corresponds to the common message that is transmitted to both the legitimate user and the eavesdropper, and W_1 denotes the private message sent only to the legitimate user, on which there is a secrecy constraint. The secrecy of the legitimate user's private message is measured by its equivocation at the eavesdropper [1], [11], i.e.,

$$\frac{1}{n}H(W_1|W_0,Z^n). \tag{1}$$

A rate triple (R_0,R_1,R_e) is said to be achievable if there exists an $(n,2^{nR_0},2^{nR_1})$ code such that $\lim_{n\to\infty}P_e^n=0$, and

$$R_e\le\lim_{n\to\infty}\frac{1}{n}H(W_1|W_0,Z^n). \tag{2}$$

The capacity-equivocation region of the discrete memoryless wiretap channel is defined as the convex closure of all achievable rate triples (R_0,R_1,R_e), and denoted by $\mathcal{C}$. The capacity-equivocation region of the discrete memoryless wiretap channel, which is obtained in [1], is stated in the following theorem.

Theorem 1 ([1, Theorem 1]): The capacity-equivocation region of the discrete memoryless wiretap channel $\mathcal{C}$ is given by the union of rate triples (R_0,R_1,R_e) satisfying

$$0\le R_e\le R_1 \tag{3}$$
$$R_e\le I(V;Y|U)-I(V;Z|U) \tag{4}$$
$$R_0+R_1\le I(V;Y|U)+\min\{I(U;Y),I(U;Z)\} \tag{5}$$
$$R_0\le\min\{I(U;Y),I(U;Z)\} \tag{6}$$

for some U,V,X such that

$$U\to V\to X\to(Y,Z). \tag{7}$$

We next provide an alternative description for $\mathcal{C}$. This alternative description will arise as the capacity region of a different, however related, communication scenario for the discrete memoryless wiretap channel. In this communication scenario, the transmitter has three messages, W_0,W_p,W_s, where W_0 is the common message sent to both the legitimate user and the eavesdropper, W_p is the public message sent only to the legitimate user on which there is no secrecy constraint, and W_s is the confidential message sent only to the legitimate user in perfect secrecy. In this scenario, since W_s needs to be transmitted in perfect secrecy, it needs to satisfy the following condition:

$$\lim_{n\to\infty}\frac{1}{n}I(W_s;Z^n,W_0)=0. \tag{8}$$

As we noted before, unlike W_s, there is no secrecy constraint on the *public* message W_p. We also note that the perfect secrecy on a message is attained when the equivocation of this message is equal to its rate, i.e., when we have $R_e=R_s$, which can be seen by comparing (2) and (8). To distinguish this communication scenario from the previous one, we call the channel model arising from this scenario the discrete memoryless wiretap channel with *public* messages. We note that this alternative description for wiretap channels has been previously considered in [4, Problem33-c], [5].

An $(n,2^{nR_0},2^{nR_p},2^{nR_s})$ code for this scenario consists of three message sets $\mathcal{W}_0=\{1,\dots,2^{nR_0}\},\mathcal{W}_p=\{1,\dots,2^{nR_p}\},\mathcal{W}_s=\{1,\dots,2^{nR_s}\}$, one encoder at the transmitter $f:\mathcal{W}_0\times\mathcal{W}_p\times\mathcal{W}_s\to\mathcal{X}^n$, one decoder at the legitimate user $g_u:\mathcal{Y}^n\to\mathcal{W}_0\times\mathcal{W}_p\times\mathcal{W}_s$, and one decoder at the eavesdropper $g_e:\mathcal{Z}^n\to\mathcal{W}_0$. The probability of error is defined as $P_e^n=\max\{P_{e,u}^n,P_{e,e}^n\}$, where $P_{e,u}^n=\Pr[g_u(Y^n)\neq(W_0,W_p,W_s)]$ and $P_{e,e}^n=\Pr[g_e(Z^n)\neq W_0]$. A rate triple (R_0,R_p,R_s) is said to be achievable if there exists an $(n,2^{nR_0},2^{nR_p},2^{nR_s})$ code such that $\lim_{n\to\infty}P_e^n=0$ and (8) is satisfied. The capacity region $\mathcal{C}_p$ of the discrete memoryless wiretap channel with *public* messages is defined as the convex closure of all achievable rate triples (R_0,R_p,R_s). The following lemma establishes the equivalence between $\mathcal{C}$ and $\mathcal{C}_p$.

Lemma 1: $(R_0,R_p,R_s)\in\mathcal{C}_p$ iff $(R_0,R_s+R_p,R_s)\in\mathcal{C}$.

The proof of this lemma is given in Appendix I. This proof consists of two steps. In the first step, we note that if $(R_0,R_p,R_s)\in\mathcal{C}_p$, then in the corresponding achievable scheme attaining this rate triple, we can combine the messages W_s,W_p to obtain $W_1=(W_s,W_p)$, whose equivocation will be as least R_s due to the perfect secrecy requirement on W_s. Hence, this argument proves the inclusion $\mathcal{C}_p\subseteq\mathcal{C}$. In the second step, we show the reverse inclusion $\mathcal{C}\subseteq\mathcal{C}_p$. To this end, we consider the achievable scheme that attains the entire region $\mathcal{C}$, and call this achievable scheme the optimal achievable scheme. If the rate triple $(R_0,R_1,R_e)\in\mathcal{C}$, in the corresponding optimal achievable scheme, the private message W_1 can be divided into two parts $W_1=(\tilde{W}_p,\tilde{W}_s)$ where the rate of $\tilde{W}_s$ is sufficiently close to R_e and satisfies the perfect secrecy requirement. Hence, this argument shows that $(R_0,R_1-R_e,R_e)\in\mathcal{C}_p$, i.e., $\mathcal{C}\subseteq\mathcal{C}_p$; completing the proof of Lemma 1. Using Lemma 1 and Theorem 1, we can express $\mathcal{C}_p$ as stated in the following theorem.

Theorem 2: The capacity region of the discrete memoryless wiretap channel with public messages $\mathcal{C}_p$ is given by the union of rate triples (R_0,R_p,R_s) satisfying

$$0\le R_s\le I(V;Y|U)-I(V;Z|U) \tag{9}$$
$$R_0+R_p+R_s\le I(V;Y|U)+\min\{I(U;Y),I(U;Z)\} \tag{10}$$
$$R_0\le\min\{I(U;Y),I(U;Z)\} \tag{11}$$

for some (U,V,X) such that

$$U\to V\to X\to(Y,Z). \tag{12}$$

III. Gaussian MIMO Wiretap Channel

The Gaussian MIMO wiretap channel is defined by

$$\mathbf{Y}=\mathbf{H}_Y\mathbf{X}+\mathbf{N}_Y \tag{13}$$
$$\mathbf{Z}=\mathbf{H}_Z\mathbf{X}+\mathbf{N}_Z \tag{14}$$

where the channel input $\mathbf{X}$ is a $t\times 1$ vector, $\mathbf{Y}$ is an $r_Y\times 1$ column vector denoting the legitimate user's observation, $\mathbf{Z}$ is an $r_Z\times 1$ column vector denoting the eavesdropper's observation, $\mathbf{H}_Y,\mathbf{H}_Z$ are the channel gain matrices of sizes $r_Y\times t$, $r_Z\times t$, respectively, and $\mathbf{N}_Y,\mathbf{N}_Z$ are Gaussian random vectors with covariance matrices $\boldsymbol{\Sigma}_Y, \boldsymbol{\Sigma}_Z$,[1] respectively, which are assumed to be strictly positive-definite, i.e., $\boldsymbol{\Sigma}_Y \succ \mathbf{0}, \boldsymbol{\Sigma}_Z \succ \mathbf{0}$. We consider a covariance constraint on the channel input as follows:

$$E\left[\mathbf{X}\mathbf{X}^\top\right] \preceq \mathbf{S} \tag{15}$$

where $\mathbf{S} \succeq \mathbf{0}$. The capacity-equivocation region of the Gaussian MIMO wiretap channel is denoted by $\mathcal{C}(\mathbf{S})$ which contains all achievable rate triples (R_0, R_1, R_e). The main result of this paper is the characterization of the capacity-equivocation region $\mathcal{C}(\mathbf{S})$ as stated in the following theorem.

Theorem 3: The capacity-equivocation region of the Gaussian MIMO wiretap channel $\mathcal{C}(\mathbf{S})$ is given by the union of rate triples (R_0, R_1, R_e) satisfying

$$0 \le R_e \le \frac{1}{2}\log\frac{|\mathbf{H}_Y\mathbf{K}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|} - \frac{1}{2}\log\frac{|\mathbf{H}_Z\mathbf{K}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|} \tag{16}$$

$$\begin{aligned} R_0 + R_1 \le{}& \frac{1}{2}\log\frac{|\mathbf{H}_Y\mathbf{K}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|} \\ &+ \frac{1}{2}\min\left\{\log\frac{|\mathbf{H}_Y\mathbf{S}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\mathbf{H}_Y\mathbf{K}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}, \log\frac{|\mathbf{H}_Z\mathbf{S}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\mathbf{H}_Z\mathbf{K}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}\right\} \end{aligned} \tag{17}$$

$$R_0 \le \frac{1}{2}\min\left\{\log\frac{|\mathbf{H}_Y\mathbf{S}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\mathbf{H}_Y\mathbf{K}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}, \log\frac{|\mathbf{H}_Z\mathbf{S}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\mathbf{H}_Z\mathbf{K}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}\right\} \tag{18}$$

for some positive semidefinite matrix $\mathbf{K}$ such that $\mathbf{0} \preceq \mathbf{K} \preceq \mathbf{S}$.

Similar to what we did in the previous section, we can establish an alternative statement for Theorem 3 by considering the Gaussian MIMO wiretap channel with *public* messages, where the legitimate user's private message is divided into two parts such that one part (confidential message) needs to be transmitted in perfect secrecy and there is no secrecy constraint on the other part (public message). The capacity region for this alternative scenario is denoted by $\mathcal{C}_p(\mathbf{S})$. We note that Lemma 1 provides a one-to-one connection between the capacity regions $\mathcal{C}$ and $\mathcal{C}_p$, and this equivalence can be extended to the capacity regions $\mathcal{C}(\mathbf{S})$ and $\mathcal{C}_p(\mathbf{S})$ by incorporating the covariance constraint on the channel input in the proof of Lemma 1. Thus, using Lemma 1 and Theorem 3, $\mathcal{C}_p(\mathbf{S})$ can be obtained as follows.

Theorem 4: The capacity region of the Gaussian MIMO wiretap channel with public messages $\mathcal{C}_p(\mathbf{S})$ is given by the union of rate triples (R_0, R_p, R_s) satisfying

$$0 \le R_s \le \frac{1}{2}\log\frac{|\mathbf{H}_Y\mathbf{K}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|} - \frac{1}{2}\log\frac{|\mathbf{H}_Z\mathbf{K}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|} \tag{19}$$

$$\begin{aligned} R_0 + R_p + R_s \le{}& \frac{1}{2}\log\frac{|\mathbf{H}_Y\mathbf{K}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|} \\ &+ \frac{1}{2}\min\left\{\log\frac{|\mathbf{H}_Y\mathbf{S}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\mathbf{H}_Y\mathbf{K}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}, \log\frac{|\mathbf{H}_Z\mathbf{S}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\mathbf{H}_Z\mathbf{K}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}\right\} \end{aligned} \tag{20}$$

$$R_0 \le \frac{1}{2}\min\left\{\log\frac{|\mathbf{H}_Y\mathbf{S}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\mathbf{H}_Y\mathbf{K}\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}, \log\frac{|\mathbf{H}_Z\mathbf{S}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\mathbf{H}_Z\mathbf{K}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}\right\} \tag{21}$$

[1]Without loss of generality, we can set $\boldsymbol{\Sigma}_Y = \boldsymbol{\Sigma}_Z = \mathbf{I}$. However, we let $\boldsymbol{\Sigma}_Y, \boldsymbol{\Sigma}_Z$ be arbitrary for ease of presentation.

for some positive semidefinite matrix $\mathbf{K}$ such that $\mathbf{0} \preceq \mathbf{K} \preceq \mathbf{S}$.

We next define a subclass of Gaussian MIMO wiretap channels called the aligned Gaussian MIMO wiretap channel, which can be obtained from (13)–(14) by setting $\mathbf{H}_Y = \mathbf{H}_Z = \mathbf{I}$

$$\mathbf{Y} = \mathbf{X} + \mathbf{N}_Y \tag{22}$$

$$\mathbf{Z} = \mathbf{X} + \mathbf{N}_Z. \tag{23}$$

In this study, we first prove Theorems 3 and 4 for the aligned Gaussian MIMO wiretap channel. Then, we establish the capacity region for the general channel model in (13)–(14) by following the analysis in [2, Sec. V.B] and [12, Sec. 7.1] in conjunction with the capacity result we obtain for the aligned channel.

A. Capacity Region Under a Power Constraint

We note that the covariance constraint on the channel input in (15) is a rather general constraint that subsumes the average power constraint

$$E\left[\mathbf{X}^\top\mathbf{X}\right] = \operatorname{tr}\left(E\left[\mathbf{X}\mathbf{X}^\top\right]\right) \le P \tag{24}$$

as a special case, see Lemma 1 and [2, Corollary 1]. Therefore, using Theorem 3, the capacity-equivocation region arising from the average power constraint in (24), $\mathcal{C}(P)$, can be found as follows.

Corollary 1: The capacity-equivocation region of the Gaussian MIMO wiretap channel subject to an average power constraint P, $\mathcal{C}(P)$, is given by the union of rate triples (R_0, R_1, R_e) satisfying

$$R_e \le \frac{1}{2}\log\frac{|\mathbf{H}_Y\mathbf{K}_1\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|} - \frac{1}{2}\log\frac{|\mathbf{H}_Z\mathbf{K}_1\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|} \tag{25}$$

$$\begin{aligned} R_0 + R_1 \le{}& \frac{1}{2}\log\frac{|\mathbf{H}_Y\mathbf{K}_1\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|} \\ &+ \frac{1}{2}\min\left\{\log\frac{|\mathbf{H}_Y(\mathbf{K}_1+\mathbf{K}_2)\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\mathbf{H}_Y\mathbf{K}_1\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}, \right. \\ &\left. \log\frac{|\mathbf{H}_Z(\mathbf{K}_1+\mathbf{K}_2)\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\mathbf{H}_Z\mathbf{K}_1\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}\right\} \end{aligned} \tag{26}$$

$$\begin{aligned} R_0 \le{}& \frac{1}{2}\min\left\{\log\frac{|\mathbf{H}_Y(\mathbf{K}_1+\mathbf{K}_2)\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}{|\mathbf{H}_Y\mathbf{K}_1\mathbf{H}_Y^\top + \boldsymbol{\Sigma}_Y|}, \right. \\ &\left. \log\frac{|\mathbf{H}_Z(\mathbf{K}_1+\mathbf{K}_2)\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\mathbf{H}_Z\mathbf{K}_1\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}\right\} \end{aligned} \tag{27}$$

for some positive semidefinite matrices $\mathbf{K}_1, \mathbf{K}_2$ such that $\operatorname{tr}(\mathbf{K}_1 + \mathbf{K}_2) \le P$.

IV. Proof of Theorem 3 for the Aligned Case

Instead of proving Theorem 3, here we prove Theorem 4, which implies Theorem 3 due to Lemma 1. Achievability of the region given in Theorem 4 can be shown by setting $\mathbf{V} = \mathbf{X}$ in Theorem 2, and using jointly Gaussian $(\mathbf{U}, \mathbf{X} = \mathbf{U} + \mathbf{T})$, where $\mathbf{U}, \mathbf{T}$ are independent Gaussian random vectors with covariance matrices $\mathbf{S} - \mathbf{K}, \mathbf{K}$, respectively. In the rest of this section, we provide the converse proof. To this end, we note that since $\mathcal{C}_p(\mathbf{S})$[2] is convex by definition, it can be characterized by solving the following optimization problem:[3]

$$f(R_0^*) = \max_{(R_0^*, R_p, R_s) \in \mathcal{C}_p(\mathbf{S})} \quad \mu_p R_p + \mu_s R_s \tag{29}$$

for all $\mu_p \in [0, \infty), \mu_s \in [0, \infty)$, and all possible common message rates R_0^*, which is bounded as follows:

$$0 \leq R_0^* \leq \min\{C_Y(\mathbf{S}), C_Z(\mathbf{S})\} \tag{30}$$

where $C_Y(\mathbf{S}), C_Z(\mathbf{S})$ are the single-user capacities for the legitimate user and the eavesdropper channels, respectively, i.e.,

$$C_Y(\mathbf{S}) = \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|} \tag{31}$$

$$C_Z(\mathbf{S}) = \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}. \tag{32}$$

We note that the optimization problem in (29) can be expressed in the following more explicit form:

$$f(R_0^*) = \max_{\substack{U \rightarrow V \rightarrow \mathbf{X} \rightarrow (\mathbf{Y}, \mathbf{Z}) \\ E[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{S}}} \quad \mu_p R_p + \mu_s R_s \tag{33}$$

$$\text{s.t.} \begin{cases} 0 \leq R_s \leq I(V; \mathbf{Y}|U) - I(V; \mathbf{Z}|U) \\ R_0^* + R_p + R_s \leq I(V; \mathbf{Y}|U) + \min\{I(U; \mathbf{Y}), I(U; \mathbf{Z})\} \\ R_0^* \leq \min\{I(U; \mathbf{Y}), I(U; \mathbf{Z})\}. \end{cases} \tag{34}$$

We also consider the Gaussian rate region $\mathcal{R}^G(\mathbf{S})$ which is defined by (35) at the bottom of the page, where $R_s(\mathbf{K}), R_p(\mathbf{K}), R_{0Y}(\mathbf{K}), R_{0Z}(\mathbf{K})$ are given as follows:

$$R_s(\mathbf{K}) = \frac{1}{2} \log \frac{|\mathbf{K} + \boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|} - \frac{1}{2} \log \frac{|\mathbf{K} + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|} \tag{36}$$

$$R_p(\mathbf{K}) = \frac{1}{2} \log \frac{|\mathbf{K} + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|} \tag{37}$$

$$R_{0Y}(\mathbf{K}) = \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_Y|}{|\mathbf{K} + \boldsymbol{\Sigma}_Y|} \tag{38}$$

$$R_{0Z}(\mathbf{K}) = \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_Z|}{|\mathbf{K} + \boldsymbol{\Sigma}_Z|}. \tag{39}$$

[2]Although $\mathcal{C}_p(\mathbf{S})$ is originally defined for the general, *not necessarily aligned*, Gaussian wiretap channel with public messages, here we use $\mathcal{C}_p(\mathbf{S})$ to denote the capacity region of the *aligned* Gaussian MIMO wiretap channel with public messages as well.

[3]Although characterizing $\mathcal{C}_p(\mathbf{S})$ by solving the following optimization problem:

$$\max_{(R_0, R_p, R_s) \in \mathcal{C}_p(\mathbf{S})} \mu_0 R_0 + \mu_p R_p + \mu_s R_s \tag{28}$$

for all μ_0, μ_p, μ_s seems to be more natural, we find working with (29) more convenient. Here, we characterize $\mathcal{C}_p(\mathbf{S})$ by solving (29) for all μ_p, μ_s, for all fixed feasible R_0^*.

To provide the converse proof, i.e., to prove the optimality of jointly Gaussian $(U, V = \mathbf{X})$ for the optimization problem in (33)–(34), we will show that

$$f(R_0^*) = g(R_0^*), \quad 0 \leq R_0^* \leq \min\{C_Y(\mathbf{S}), C_Z(\mathbf{S})\} \tag{40}$$

where $g(R_0^*)$ is defined as

$$g(R_0^*) = \max_{(R_0^*, R_p, R_s) \in \mathcal{R}^G(\mathbf{S})} \quad \mu_p R_p + \mu_s R_s. \tag{41}$$

We show (40) in two parts:
1) $\mu_s \leq \mu_p$
2) $\mu_p < \mu_s$.

A. $\mu_s \leq \mu_p$

In this case, $f(R_0^*)$ can be written as

$$f(R_0^*) = \max_{\substack{U \rightarrow V \rightarrow \mathbf{X} \rightarrow (\mathbf{Y}, \mathbf{Z}) \\ E[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{S}}} \quad \mu_p (R_p + R_s) \tag{42}$$

$$\text{s.t.} \begin{cases} R_0^* + R_p + R_s \leq I(\mathbf{X}; \mathbf{Y}|U) + \min\{I(U; \mathbf{Y}), I(U; \mathbf{Z})\} \\ R_0^* \leq \min\{I(U; \mathbf{Y}), I(U; \mathbf{Z})\} \end{cases} \tag{43}$$

where we use the fact that $\mu_s \leq \mu_p$, and the secret message rate R_s can be given up in favor of the private message rate R_p. In other words, we use the fact that when $\mu_p \geq \mu_s$, the maximum of $\mu_p R_p + \mu_s R_s$ is given by $\mu_p R_p'$, where $R_p' = R_s + R_p$ is an achievable public message rate since the secret message can be converted into a public message. This optimization problem gives us the capacity region of the two-user Gaussian MIMO broadcast channel with degraded message sets, where a common message is sent to both users, and a private message, on which there is no secrecy constraint, is sent to one of the two users [13]. The optimization problem for this case given in (42)–(43) is solved in [10] by showing the optimality of jointly Gaussian $(U, \mathbf{X})$, i.e., $f(R_0^*) = g(R_0^*)$. This completes the converse proof for the case $\mu_s \leq \mu_p$.

B. $\mu_p < \mu_s$

In this case, we first study the optimization problem in (41). We rewrite $g(R_0^*)$ as follows:

$$g(R_0^*) = \max_{\substack{\mathbf{0} \preceq \mathbf{K} \preceq \mathbf{S} \\ R_p}} \quad \mu_p R_p + \mu_s R_s(\mathbf{K}) \tag{44}$$

$$\text{s.t.} \begin{cases} R_0^* + R_p \leq R_p(\mathbf{K}) + \min\{R_{0Y}(\mathbf{K}), R_{0Z}(\mathbf{K})\} \\ R_0^* \leq \min\{R_{0Y}(\mathbf{K}), R_{0Z}(\mathbf{K})\} \end{cases} \tag{45}$$

where we use the fact that since $\mu_s > \mu_p$, the secret message rate should be set as high as possible to maximize $\mu_p R_p + \mu_s R_s$, i.e., we should set $R_s = R_s(\mathbf{K})$. Let $(\mathbf{K}^*, R_p^*)$ be the maximizer for this optimization problem. The necessary

$$\mathcal{R}^G(\mathbf{S}) = \left\{ (R_0, R_p, R_s) : \begin{array}{rcl} 0 \leq R_s & \leq & R_s(\mathbf{K}) \\ R_0 + R_p + R_s & \leq & R_s(\mathbf{K}) + R_p(\mathbf{K}) + \min\{R_{0Y}(\mathbf{K}), R_{0Z}(\mathbf{K})\} \\ R_0 & \leq & \min\{R_{0Y}(\mathbf{K}), R_{0Z}(\mathbf{K})\} \\ \text{for some } \mathbf{0} & \preceq & \mathbf{K} \preceq \mathbf{S} \end{array} \right\} \tag{35}$$

Karush–Kuhn–Tucker (KKT) conditions that $(\mathbf{K}^*, R_p^*)$ needs to satisfy are given in the following lemma.

Lemma 2: $\mathbf{K}^*$ needs to satisfy

$$(\mu_s - \mu_p\lambda - \beta_Y)(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1} + \mathbf{M} = (\mu_s - \mu_p\lambda + \beta_Z)(\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1} + \mathbf{M}_S \quad (46)$$

for some positive semidefinite matrices $\mathbf{M}, \mathbf{M}_S$ such that

$$\mathbf{K}^*\mathbf{M} = \mathbf{M}\mathbf{K}^* = \mathbf{0} \quad (47)$$

$$(\mathbf{S} - \mathbf{K}^*)\mathbf{M}_S = \mathbf{M}_S(\mathbf{S} - \mathbf{K}^*) = \mathbf{0} \quad (48)$$

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

$$\lambda \begin{cases} = 0, & \text{if} \quad R_{0Y}(\mathbf{K}^*) > R_{0Z}(\mathbf{K}^*) \\ = 1, & \text{if} \quad R_{0Y}(\mathbf{K}^*) < R_{0Z}(\mathbf{K}^*) \end{cases} \quad (49)$$

and (β_Y, β_Z) are given as follows:

$$(\beta_Y, \beta_Z) = \begin{cases} (0,0), & \text{if} \quad R_0^* < \min\{R_{0Y}(\mathbf{K}^*), R_{0Z}(\mathbf{K}^*)\} \\ (0, \geq 0), & \text{if} \quad R_0^* = R_{0Z}(\mathbf{K}^*) < R_{0Y}(\mathbf{K}^*) \\ (\geq 0, 0), & \text{if} \quad R_0^* = R_{0Y}(\mathbf{K}^*) < R_{0Z}(\mathbf{K}^*) \\ (\geq 0, \geq 0), & \text{if} \quad R_0^* = R_{0Y}(\mathbf{K}^*) = R_{0Z}(\mathbf{K}^*) \end{cases} \quad (50)$$

R_p^* needs to satisfy

$$R_p^* = R_p(\mathbf{K}^*) + \min\{R_{0Y}(\mathbf{K}^*), R_{0Z}(\mathbf{K}^*)\} - R_0^*. \quad (51)$$

The proof of Lemma 2 is given in Appendix II. We treat three cases separately:

1) $R_0^* < \min\{R_{0Y}(\mathbf{K}^*), R_{0Z}(\mathbf{K}^*)\}$;
2) $R_0^* = R_{0Y}(\mathbf{K}^*) \leq R_{0Z}(\mathbf{K}^*)$;
3) $R_0^* = R_{0Z}(\mathbf{K}^*) < R_{0Y}(\mathbf{K}^*)$.

1) $R_0^* < \min\{R_{0Y}(\mathbf{K}^*), R_{0Z}(\mathbf{K}^*)\}$: In this case, we have $\beta_Y = \beta_Z = 0$, see (50). Thus, the KKT condition in (46) reduces to

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1} + \mathbf{M} = (\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1} + \mathbf{M}_S. \quad (52)$$

We first note that $\mathbf{K}^*$ satisfying (52) achieves the secrecy capacity of this Gaussian MIMO wiretap channel [14], i.e.,

$$R_s^* = R_s(\mathbf{K}^*) \quad (53)$$

$$= C_S(\mathbf{S}) \quad (54)$$

$$= \max_{\mathbf{0} \preceq \mathbf{K} \preceq \mathbf{S}} \frac{1}{2}\log\frac{|\mathbf{K} + \boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|} - \frac{1}{2}\log\frac{|\mathbf{K} + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}. \quad (55)$$

Next, we define a new covariance matrix $\tilde{\boldsymbol{\Sigma}}_Z$ as follows:

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z)^{-1} = (\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1} + \mathbf{M}_S \quad (56)$$

which is similar to the channel enhancement done in [14]. This new covariance matrix $\tilde{\boldsymbol{\Sigma}}_Z$ has some useful properties which are listed in the following lemma, whose proof is given in Appendix III.

Lemma 3: We have the following facts:

1) $\mathbf{0} \preceq \tilde{\boldsymbol{\Sigma}}_Z$;
2) $\tilde{\boldsymbol{\Sigma}}_Z \preceq \boldsymbol{\Sigma}_Z$;
3) $\tilde{\boldsymbol{\Sigma}}_Z \preceq \boldsymbol{\Sigma}_Y$;
4) $(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z)^{-1}(\mathbf{S} + \tilde{\boldsymbol{\Sigma}}_Z) = (\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1}(\mathbf{S} + \boldsymbol{\Sigma}_Z)$.

Thus, we have

$$R_{0Z}(\mathbf{K}^*) = \frac{1}{2}\log\frac{|\mathbf{S} + \boldsymbol{\Sigma}_Z|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_Z|} \quad (57)$$

$$= \frac{1}{2}\log\frac{|\mathbf{S} + \tilde{\boldsymbol{\Sigma}}_Z|}{|\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z|} \quad (58)$$

$$\geq \frac{1}{2}\log\frac{|\mathbf{S} + \boldsymbol{\Sigma}_Y|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_Y|} \quad (59)$$

$$= R_{0Y}(\mathbf{K}^*) \quad (60)$$

where (58) comes from the third part of Lemma 3, (59) is due to the fact that

$$\frac{|\mathbf{A} + \mathbf{B} + \boldsymbol{\Delta}|}{|\mathbf{B} + \boldsymbol{\Delta}|} \leq \frac{|\mathbf{A} + \mathbf{B}|}{|\mathbf{B}|} \quad (61)$$

for $\mathbf{A} \succeq \mathbf{0}, \boldsymbol{\Delta} \succeq \mathbf{0}, \mathbf{B} \succ \mathbf{0}$ by noting the second part of Lemma 3. Therefore, we have

$$R_{0Z}(\mathbf{K}^*) \geq R_{0Y}(\mathbf{K}^*) \quad (62)$$

where $\mathbf{K}^*$ satisfies (52). Using (62) in (51), we find R_p^* as follows:

$$R_p^* = R_p(\mathbf{K}^*) + R_{0Y}(\mathbf{K}^*) - R_0^*. \quad (63)$$

We also note the following:

$$R_0^* + R_p^* + R_s^* = R_{0Y}(\mathbf{K}^*) + R_p(\mathbf{K}^*) + R_s(\mathbf{K}^*) \quad (64)$$

$$= \frac{1}{2}\log\frac{|\mathbf{S} + \boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|} \quad (65)$$

$$= C_Y(\mathbf{S}). \quad (66)$$

Now, we show that

$$g(R_0^*) = f(R_0^*). \quad (67)$$

To this end, we assume that

$$g(R_0^*) < f(R_0^*) \quad (68)$$

which implies that there exists a rate triple $(R_0^*, R_p^o, R_s^o) \in \mathcal{C}_p(\mathbf{S})$ such that

$$\mu_p R_p^* + \mu_s R_s^* < \mu_p R_p^o + \mu_s R_s^o. \quad (69)$$

To prove (67), i.e., that (68) is not possible, we note the following bounds:

$$R_s^o \leq C_S(\mathbf{S}) = R_s^* \quad (70)$$

$$R_p^o + R_s^o \leq C_Y(\mathbf{S}) - R_0^* = R_p^* + R_s^* \quad (71)$$

where (70) comes from (55) and the fact that the rate of the confidential message, i.e., R_s, cannot exceed the secrecy capacity, and (71) is due to (66) and the fact that the sum rate $R_0+R_p+R_s$ cannot exceed the legitimate user's single-user capacity. Thus, in view of $\mu_s > \mu_p$, we can multiply (70) and (71) by $\mu_s - \mu_p$ and μ_p, respectively, and add the corresponding inequalities to obtain

$$\mu_p R_p^o + \mu_s R_s^o \leq \mu_p R_p^* + \mu_s R_s^* \quad (72)$$

which contradicts with (69), proving (67). This completes the converse proof for this case.

Before starting the proofs of the other two cases, we now recap our proof for the case $R_0^* < \min\{R_{0Y}(\mathbf{K}^*), R_{0Z}(\mathbf{K}^*)\}$. We note that we did not show the optimality of Gaussian signalling directly, instead, we prove it indirectly by showing the following:

$$g(R_0^*) = f(R_0^*). \tag{73}$$

First, we show that for the given common message rate R_0^*, we can achieve the secrecy capacity, i.e., $R_s^* = C_S(\mathbf{S})$, see (53)–(55). In other words, we show that $(R_0^*, 0, R_s^*)$ is on the boundary of the capacity region $\mathcal{C}_p(\mathbf{S})$. Secondly, we show that for the given common message rate R_0^*, (R_p^*, R_s^*) achieve the sum capacity of the public and confidential messages, i.e., $R_s^* + R_p^*$ is sum rate optimal for the given common message rate R_0^* [see (64)–(66) and (71)]. These two findings lead to the inequalities in (70)–(71). Finally, we use a time-sharing argument for these two inequalities in (70)–(71) to obtain (73), which completes the proof.

2) $R_0^* = R_{0Y}(\mathbf{K}^*) \leq R_{0Z}(\mathbf{K}^*)$: We first rewrite the KKT condition in (46) as follows:

$$(\mu_s - \mu_p\lambda - \mu_0\beta)(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1} + \mathbf{M} = (\mu_s - \mu_p\lambda + \mu_0\bar{\beta})(\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1} + \mathbf{M}_S \tag{74}$$

by defining $\mu_0 = \beta_Y + \beta_Z$, $\mu_0\beta = \beta_Y$, and $\mu_0\bar{\beta} = \beta_Z$. We note that if $R_{0Y}(\mathbf{K}^*) < R_{0Z}(\mathbf{K}^*)$, we have $\beta = \lambda = 1$, if $R_{0Y}(\mathbf{K}^*) = R_{0Z}(\mathbf{K}^*)$, we have $0 \leq \lambda \leq 1, 0 \leq \beta \leq 1$. The proof of these two cases are very similar, and we consider only the case $0 \leq \lambda \leq 1, 0 \leq \beta \leq 1$, i.e., we assume $R_{0Y}(\mathbf{K}^*) = R_{0Z}(\mathbf{K}^*)$. The other case can be proved similarly.

Similar to Section IV-B1, here also, we prove the desired identity

$$g(R_0^*) = f(R_0^*) \tag{75}$$

by contradiction. We first assume that

$$g(R_0^*) < f(R_0^*) \tag{76}$$

which implies that there exists a rate triple $(R_0^*, R_p^o, R_s^o) \in \mathcal{C}_p(\mathbf{S})$ such that

$$\mu_p R_p^* + \mu_s R_s^* < \mu_p R_p^o + \mu_s R_s^o \tag{77}$$

where we define $R_s^* = R_s(\mathbf{K}^*)$. Since the sum rate $R_0 + R_p + R_s$ needs to be smaller than the legitimate user's single user capacity, we have

$$R_0^* + R_p^o + R_s^o \leq C_Y(\mathbf{S}). \tag{78}$$

On the other hand, we have the following:

$$\begin{aligned} R_0^* + R_p^* + R_s^* &= \min\{R_{0Y}(\mathbf{K}^*), R_{0Z}(\mathbf{K}^*)\} + R_p(\mathbf{K}^*) \\ &\quad + R_s(\mathbf{K}^*) && (79) \\ &= R_{0Y}(\mathbf{K}^*) + R_p(\mathbf{K}^*) + R_s(\mathbf{K}^*) && (80) \\ &= C_Y(\mathbf{S}) && (81) \end{aligned}$$

where (79) comes from (51), and (80) is due to our assumption that $R_0^* = R_{0Y}(\mathbf{K}^*) = R_{0Z}(\mathbf{K}^*)$. Equations (78) and (81) imply that

$$R_p^o + R_s^o \leq R_p^* + R_s^*. \tag{82}$$

In the rest of this section, we prove that we have $R_s^o \leq R_s^*$ for the given common message rate R_0^*, which, in conjunction with (82), will yield a contradiction with (77); proving (75). To this end, we first define a new covariance matrix $\tilde{\boldsymbol{\Sigma}}_Y$ as follows:

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Y)^{-1} = (\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1} + \mathbf{M}. \tag{83}$$

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

Lemma 4: We have the following facts:
1) $\mathbf{0} \preceq \tilde{\boldsymbol{\Sigma}}_Y$;
2) $\tilde{\boldsymbol{\Sigma}}_Y \preceq \boldsymbol{\Sigma}_Y$;
3) $\tilde{\boldsymbol{\Sigma}}_Y \preceq \boldsymbol{\Sigma}_Z$;
4) $(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Y)^{-1}\tilde{\boldsymbol{\Sigma}}_Y = (\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1}\boldsymbol{\Sigma}_Y$.

The proof of this lemma is given in Appendix IV. Using this new covariance matrix, we define a random vector $\tilde{\mathbf{Y}}$ as

$$\tilde{\mathbf{Y}} = \mathbf{X} + \tilde{\mathbf{N}}_Y \tag{84}$$

where $\tilde{\mathbf{N}}_Y$ is a Gaussian random vector with covariance matrix $\tilde{\boldsymbol{\Sigma}}_Y$. Due to the first and second statements of Lemma 4, we have the following Markov chains:

$$U \rightarrow V \rightarrow \mathbf{X} \rightarrow \tilde{\mathbf{Y}} \rightarrow \mathbf{Y} \tag{85}$$

$$U \rightarrow V \rightarrow \mathbf{X} \rightarrow \tilde{\mathbf{Y}} \rightarrow \mathbf{Z}. \tag{86}$$

We next study the following optimization problem:

$$\begin{aligned} &\max_{(R_0,R_p,R_s)\in\mathcal{C}_p(\mathbf{S})} \mu_0 R_0 + (\mu_s - \mu_p\lambda)R_s \\ &= \max_{\substack{U\rightarrow V\rightarrow\mathbf{X}\rightarrow(\mathbf{Y},\mathbf{Z}) \\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} \mu_0 \min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\} \\ &\qquad + (\mu_s - \mu_p\lambda)\left[I(V;\mathbf{Y}|U) - I(V;\mathbf{Z}|U)\right] \end{aligned} \tag{87}$$

where the equality follows from the fact that the maximum of $\mu_0 R_0 + \mu_s R_s$ is obtained by selecting both R_0 and R_s to be individually maximum, i.e., by setting $R_0 = \min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\}, R_s = I(V;\mathbf{Y}|U) - I(V;\mathbf{Z}|U)$, since this is possible by simply setting $R_p = 0$.

Since we assume $(R_0^*, R_p^o, R_s^o) \in \mathcal{C}_p(\mathbf{S})$, we have the following lower bound for (87):

$$\mu_0 R_0^* + (\mu_s - \mu_p\lambda)R_s^o \leq \max_{(R_0,R_p,R_s)\in\mathcal{C}_p(\mathbf{S})} \mu_0 R_0 + (\mu_s - \mu_p\lambda)R_s. \tag{88}$$

Now we solve the optimization problem in (87) as follows:

$$\begin{aligned}
&\max_{(R_0,R_p,R_s)\in\mathcal{C}_p(\mathbf{S})} \mu_0 R_0 + (\mu_s-\mu_p\lambda)R_s \\
&= \max_{\substack{U\to V\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} \mu_0\min\{I(U;\mathbf{Y}),I(U;\mathbf{Z})\} \\
&\quad + (\mu_s-\mu_p\lambda)\left[I(V;\mathbf{Y}|U)-I(V;\mathbf{Z}|U)\right] && (89)\\
&\le \max_{\substack{U\to V\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} \mu_0\bar{\beta}I(U;\mathbf{Z})+\mu_0\beta I(U;\mathbf{Y}) \\
&\quad + (\mu_s-\mu_p\lambda)\left[I(V;\mathbf{Y}|U)-I(V;\mathbf{Z}|U)\right] && (90)\\
&\le \max_{\substack{U\to V\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} \mu_0\bar{\beta}I(U;\mathbf{Z})+\mu_0\beta I(U;\mathbf{Y}) \\
&\quad + (\mu_s-\mu_p\lambda)\left[I(V;\tilde{\mathbf{Y}}|U)-I(V;\mathbf{Z}|U)\right] && (91)\\
&\le \max_{\substack{U\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} \mu_0\bar{\beta}I(U;\mathbf{Z})+\mu_0\beta I(U;\mathbf{Y}) \\
&\quad + (\mu_s-\mu_p\lambda)\left[I(\mathbf{X};\tilde{\mathbf{Y}}|U)-I(\mathbf{X};\mathbf{Z}|U)\right] && (92)\\
&\le \frac{\mu_0\bar{\beta}}{2}\log\frac{|\mathbf{S}+\boldsymbol{\Sigma}_Z|}{|\mathbf{K}^*+\boldsymbol{\Sigma}_Z|}+\frac{\mu_0\beta}{2}\log\frac{|\mathbf{S}+\boldsymbol{\Sigma}_Y|}{|\mathbf{K}^*+\boldsymbol{\Sigma}_Y|} \\
&\quad +\frac{\mu_s-\mu_p\lambda}{2}\left[\log\frac{|\mathbf{K}^*+\tilde{\boldsymbol{\Sigma}}_Y|}{|\tilde{\boldsymbol{\Sigma}}_Y|}-\log\frac{|\mathbf{K}^*+\boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}\right] && (93)\\
&= \mu_0\bar{\beta}R_{0Z}(\mathbf{K}^*)+\mu_0\beta R_{0Y}(\mathbf{K}^*) \\
&\quad +\frac{\mu_s-\mu_p\lambda}{2}\left[\log\frac{|\mathbf{K}^*+\tilde{\boldsymbol{\Sigma}}_Y|}{|\tilde{\boldsymbol{\Sigma}}_Y|}-\log\frac{|\mathbf{K}^*+\boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}\right] && (94)\\
&= \mu_0\bar{\beta}R_{0Z}(\mathbf{K}^*)+\mu_0\beta R_{0Y}(\mathbf{K}^*) \\
&\quad +\frac{\mu_s-\mu_p\lambda}{2}\left[\log\frac{|\mathbf{K}^*+\boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|}-\log\frac{|\mathbf{K}^*+\boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}\right] && (95)\\
&= \mu_0\bar{\beta}R_{0Z}(\mathbf{K}^*)+\mu_0\beta R_{0Y}(\mathbf{K}^*)+(\mu_s-\mu_p\lambda)R_s(\mathbf{K}^*) && (96)\\
&= \mu_0R_0^*+(\mu_s-\mu_p\lambda)R_s^* && (97)
\end{aligned}$$

where (90) comes from the fact that $0\le\beta=1-\bar{\beta}\le 1$, (91)–(92) are due to the Markov chains in (85)–(86), respectively, (93) can be obtained by using the analysis in [9, eqs. (30)–(32)], which uses an extremal inequality from [3] to establish this result, (95) comes from the third part of Lemma 4, and (97) is due to our assumption that $R_0^*=R_{0Y}(\mathbf{K}^*)=R_{0Z}(\mathbf{K}^*)$. Thus, (97) implies

$$\max_{(R_0,R_p,R_s)\in\mathcal{C}_p(\mathbf{S})} \mu_0R_0+(\mu_s-\mu_p\lambda)R_s \le \mu_0R_0^*+(\mu_s-\mu_p\lambda)R_s^*. \quad (98)$$

Comparing (88) and (98) yields

$$R_s^o\le R_s^*. \quad (99)$$

Using (82) and (99) and noting $\mu_s>\mu_p$, we can get

$$\mu_pR_p^o+\mu_sR_s^o\le\mu_pR_p^*+\mu_sR_s^* \quad (100)$$

which contradicts with (77), proving (75). This completes the converse proof for this case.

Before providing the proof for the last case, we recap our proof for the case $R_0^*=R_{0Y}(\mathbf{K}^*)\le R_{0Z}(\mathbf{K}^*)$. Similar to Section IV-B1, here also, we prove the optimality of Gaussian signalling indirectly, i.e., we show the desired identity

$$g(R_0^*)=f(R_0^*) \quad (101)$$

indirectly. First, we show that for the given common message rate R_0^*, $R_s^*+R_p^*$ is sum rate optimal, i.e., (R_p^*,R_s^*) achieve the sum capacity of the public and confidential messages, by obtaining (82). Second, we show that $(R_0^*,0,R_s^*)$ is also on the boundary of the capacity region $\mathcal{C}_p(\mathbf{S})$ by obtaining (98). These two findings give us the inequalities in (82) and (99). Finally, we use a time-sharing argument for these two inequalities in (82) and (99) to establish (101), which completes the proof.

3) $R_0^*=R_{0Z}(\mathbf{K}^*)<R_{0Y}(\mathbf{K}^*)$: In this case, we have $\lambda=\beta_Y=0$, see (49)–(50). Hence, the KKT condition in (46) reduces to

$$\mu_s(\mathbf{K}^*+\boldsymbol{\Sigma}_Y)^{-1}+\mathbf{M}=(\mu_s+\beta_Z)(\mathbf{K}^*+\boldsymbol{\Sigma}_Z)^{-1}+\mathbf{M}_S. \quad (102)$$

We again prove the desired identity

$$g(R_0^*)=f(R_0^*) \quad (103)$$

by contradiction. We first assume that

$$g(R_0^*)<f(R_0^*) \quad (104)$$

which implies that there exists a rate triple $(R_0^*,R_p^o,R_s^o)\in\mathcal{C}_p(\mathbf{S})$ such that

$$\mu_pR_p^*+\mu_sR_s^*<\mu_pR_p^o+\mu_sR_s^o. \quad (105)$$

In the rest of the section, we show that

$$\mu_pR_p^*+\mu_sR_s^*\ge\mu_pR_p^o+\mu_sR_s^o \quad (106)$$

to reach a contradiction, and hence, prove (103). To this end, we define a new covariance matrix $\tilde{\boldsymbol{\Sigma}}_Y$ as follows:

$$\mu_s(\mathbf{K}^*+\tilde{\boldsymbol{\Sigma}}_Y)^{-1}=\mu_s(\mathbf{K}^*+\boldsymbol{\Sigma}_Y)^{-1}+\mathbf{M}. \quad (107)$$

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

Lemma 5: We have the following facts.

1) $\mathbf{0}\preceq\tilde{\boldsymbol{\Sigma}}_Y$;
2) $\tilde{\boldsymbol{\Sigma}}_Y\preceq\boldsymbol{\Sigma}_Y$;
3) $\tilde{\boldsymbol{\Sigma}}_Y\preceq\boldsymbol{\Sigma}_Z$;
4) $(\mathbf{K}^*+\tilde{\boldsymbol{\Sigma}}_Y)^{-1}\tilde{\boldsymbol{\Sigma}}_Y=(\mathbf{K}^*+\boldsymbol{\Sigma}_Y)^{-1}\boldsymbol{\Sigma}_Y$.

The proof of this lemma is very similar to the proof Lemma 4, and hence is omitted. Using this new covariance matrix $\tilde{\boldsymbol{\Sigma}}_Y$, we define a random vector $\tilde{\mathbf{Y}}$ as

$$\tilde{\mathbf{Y}}=\mathbf{X}+\tilde{\mathbf{N}}_Y \quad (108)$$

where $\tilde{\mathbf{N}}_Y$ is a Gaussian random vector with covariance matrix $\tilde{\boldsymbol{\Sigma}}_Y$. Due to the first and second statements of Lemma 5, we have the following Markov chains:

$$U \to V \to \mathbf{X} \to \tilde{\mathbf{Y}} \to \mathbf{Y} \tag{109}$$
$$U \to V \to \mathbf{X} \to \tilde{\mathbf{Y}} \to \mathbf{Z}. \tag{110}$$

Next, we study the following optimization problem:

$$\max_{(R_0,R_p,R_s)\in\mathcal{C}_p(\mathbf{S})} (\mu_p+\beta_Z)R_0+\mu_p R_p+\mu_s R_s. \tag{111}$$

We note that since $(R_0^*, R_p^o, R_s^o) \in \mathcal{C}_p(\mathbf{S})$, we have the following lower bound for the optimization problem in (111):

$$\begin{aligned}&(\mu_p+\beta_Z)R_0^*+\mu_p R_p^o+\mu_s R_s^o\\ &\quad\le \max_{(R_0,R_p,R_s)\in\mathcal{C}_p(\mathbf{S})} (\mu_p+\beta_Z)R_0+\mu_p R_p+\mu_s R_s.\end{aligned} \tag{112}$$

We next obtain the maximum for (111). To this end, we introduce the following lemma which provides an explicit form for this optimization problem.

Lemma 6: For $\mu_s > \mu_p$, we have

$$\begin{aligned}&\max_{(R_0,R_p,R_s)\in\mathcal{C}_p(\mathbf{S})} (\mu_p+\beta_Z)R_0+\mu_p R_p+\mu_s R_s\\ &=\max_{\substack{U\to V\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} (\mu_p+\beta_Z)\min\{I(U;\mathbf{Y}),I(U;\mathbf{Z})\}\\ &\quad+\mu_p I(V;\mathbf{Z}|U)+\mu_s\left[I(V;\mathbf{Y}|U)-I(V;\mathbf{Z}|U)\right].\end{aligned} \tag{113}$$

The proof of this lemma is given in Appendix V.

Next we introduce the following extremal inequality from [3], which will be used in the solution of (113).

Lemma 7 [3, 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 $\mathbf{N}_1,\mathbf{N}_2$ be Gaussian random vectors with covariance matrices $\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2$, respectively. They are independent of $(U,\mathbf{X})$. Furthermore, $\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2$ satisfy $\boldsymbol{\Sigma}_1\preceq\boldsymbol{\Sigma}_2$. Assume that there exists a covariance matrix $\mathbf{K}^*$ such that $\mathbf{K}^*\preceq\mathbf{S}$ and

$$\nu(\mathbf{K}^*+\boldsymbol{\Sigma}_1)^{-1}=\gamma(\mathbf{K}^*+\boldsymbol{\Sigma}_2)^{-1}+\mathbf{M}_S \tag{114}$$

where $\nu\ge 0,\gamma\ge 0$ 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

$$\begin{aligned}&\nu h(\mathbf{X}+\mathbf{N}_1|U)-\gamma h(\mathbf{X}+\mathbf{N}_2|U)\\ &\le\frac{\nu}{2}\log|(2\pi e)(\mathbf{K}^*+\boldsymbol{\Sigma}_1|-\frac{\gamma}{2}\log|(2\pi e)(\mathbf{K}^*+\boldsymbol{\Sigma}_2)|.\end{aligned} \tag{115}$$

Now we use Lemma 7. To this end, we note that using (107) in (102), we get

$$\mu_s(\mathbf{K}^*+\tilde{\boldsymbol{\Sigma}}_Y)^{-1}=(\mu_s+\beta_Z)(\mathbf{K}^*+\boldsymbol{\Sigma}_Z)^{-1}+\mathbf{M}_S. \tag{116}$$

In view of (116) and the fact that $\tilde{\boldsymbol{\Sigma}}_Y\preceq\boldsymbol{\Sigma}_Z$, Lemma 7 implies

$$\begin{aligned}&\mu_s h(\tilde{\mathbf{Y}}|U)-(\mu_s+\beta_Z)h(\mathbf{Z}|U)\\ &\quad\le\frac{\mu_s}{2}\log|(2\pi e)(\mathbf{K}^*+\tilde{\boldsymbol{\Sigma}}_Y)\\ &\qquad-\frac{\mu_s+\beta_Z}{2}\log|(2\pi e)(\mathbf{K}^*+\boldsymbol{\Sigma}_Z)|.\end{aligned} \tag{117}$$

We now consider the maximization in (113) as follows:

$$\begin{aligned}&\max_{(R_0,R_p,R_s)\in\mathcal{C}_p(\mathbf{S})} (\mu_p+\beta_Z)R_0+\mu_p R_p+\mu_s R_s\\
&=\max_{\substack{U\to V\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} (\mu_p+\beta_Z)\min\{I(U;\mathbf{Y}),I(U;\mathbf{Z})\}\\
&\quad+\mu_p I(V;\mathbf{Z}|U)+\mu_s\left[I(V;\mathbf{Y}|U)-I(V;\mathbf{Z}|U)\right] \qquad (118)\\
&\le\max_{\substack{U\to V\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} (\mu_p+\beta_Z)I(U;\mathbf{Z})+\mu_p I(V;\mathbf{Z}|U)\\
&\quad+\mu_s\left[I(V;\mathbf{Y}|U)-I(V;\mathbf{Z}|U)\right] \qquad (119)\\
&\le\max_{\substack{U\to V\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} (\mu_p+\beta_Z)I(U;\mathbf{Z})+\mu_p I(\mathbf{X};\mathbf{Z}|U)\\
&\quad+\mu_s\left[I(V;\mathbf{Y}|U)-I(V;\mathbf{Z}|U)\right] \qquad (120)\\
&\le\max_{\substack{U\to V\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} (\mu_p+\beta_Z)I(U;\mathbf{Z})+\mu_p I(\mathbf{X};\mathbf{Z}|U)\\
&\quad+\mu_s\left[I(V;\tilde{\mathbf{Y}}|U)-I(V;\mathbf{Z}|U)\right] \qquad (121)\\
&\le\max_{\substack{U\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} (\mu_p+\beta_Z)I(U;\mathbf{Z})+\mu_p I(\mathbf{X};\mathbf{Z}|U)\\
&\quad+\mu_s\left[I(\mathbf{X};\tilde{\mathbf{Y}}|U)-I(\mathbf{X};\mathbf{Z}|U)\right] \qquad (122)\\
&=\max_{\substack{U\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} (\mu_p+\beta_Z)h(\mathbf{Z})+\mu_s h(\tilde{\mathbf{Y}}|U)\\
&\quad-(\mu_s+\beta_Z)h(\mathbf{Z}|U)-\frac{\mu_s}{2}\log|(2\pi e)\tilde{\boldsymbol{\Sigma}}_Y|\\
&\quad+\frac{\mu_s-\mu_p}{2}\log|(2\pi e)\boldsymbol{\Sigma}_Z| \qquad (123)\\
&\le\frac{\mu_p+\beta_Z}{2}\log|(2\pi e)(\mathbf{S}+\boldsymbol{\Sigma}_Z)|\\
&\quad+\max_{\substack{U\to\mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq\mathbf{S}}} \mu_s h(\tilde{\mathbf{Y}}|U)-(\mu_s+\beta_Z)h(\mathbf{Z}|U)\\
&\quad-\frac{\mu_s}{2}\log|(2\pi e)\tilde{\boldsymbol{\Sigma}}_Y|+\frac{\mu_s-\mu_p}{2}\log|(2\pi e)\boldsymbol{\Sigma}_Z| \qquad (124)\\
&\le\frac{\mu_p+\beta_Z}{2}\log|(2\pi e)(\mathbf{S}+\boldsymbol{\Sigma}_Z)|\\
&\quad+\frac{\mu_s}{2}\log|(2\pi e)(\mathbf{K}^*+\tilde{\boldsymbol{\Sigma}}_Y)|\\
&\quad-\frac{\mu_s+\beta_Z}{2}\log|(2\pi e)(\mathbf{K}^*+\boldsymbol{\Sigma}_Z)|\\
&\quad-\frac{\mu_s}{2}\log|(2\pi e)\tilde{\boldsymbol{\Sigma}}_Y|+\frac{\mu_s-\mu_p}{2}\log|(2\pi e)\boldsymbol{\Sigma}_Z| \qquad (125)\\
&=\frac{\mu_p+\beta_Z}{2}\log\frac{|\mathbf{S}+\boldsymbol{\Sigma}_Z|}{|\mathbf{K}^*+\boldsymbol{\Sigma}_Z|}+\frac{\mu_p}{2}\log\frac{|\mathbf{K}^*+\boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}\\
&\quad+\frac{\mu_s}{2}\left[\log\frac{|\mathbf{K}^*+\tilde{\boldsymbol{\Sigma}}_Y|}{|\tilde{\boldsymbol{\Sigma}}_Y|}-\log\frac{|\mathbf{K}^*+\boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}\right] \qquad (126)\\
&=(\mu_p+\beta_Z)R_{0Z}(\mathbf{K}^*)+\mu_p R_p(\mathbf{K}^*)\\
&\quad+\frac{\mu_s}{2}\left[\log\frac{|\mathbf{K}^*+\tilde{\boldsymbol{\Sigma}}_Y|}{|\tilde{\boldsymbol{\Sigma}}_Y|}-\log\frac{|\mathbf{K}^*+\boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}\right] \qquad (127)\\
&=(\mu_p+\beta_Z)R_{0Z}(\mathbf{K}^*)+\mu_p R_p(\mathbf{K}^*)\\
&\quad+\frac{\mu_s}{2}\left[\log\frac{|\mathbf{K}^*+\boldsymbol{\Sigma}_Y|}{|\boldsymbol{\Sigma}_Y|}-\log\frac{|\mathbf{K}^*+\boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}\right] \qquad (128)\\
&=(\mu_p+\beta_Z)R_{0Z}(\mathbf{K}^*)+\mu_p R_p(\mathbf{K}^*)+\mu_s R_s(\mathbf{K}^*) \qquad (129)\\
&=(\mu_p+\beta_Z)R_0^*+\mu_p R_p^*+\mu_s R_s^* \qquad (130)
\end{aligned}$$

where (119) is due to $\min\{a,b\} \leq a$, (120) is due to the Markov chain in (110), (121)–(122) come from the Markov chains in (109)–(109), respectively, (124) is due to the maximum entropy theorem [15], (125) comes from (117), and (128) is due to the third part of Lemma 5. Comparing (130) and (112) yields

$$\mu_p R_p^o + \mu_s R_s^o \leq \mu_p R_p^* + \mu_s R_s^* \tag{131}$$

which contradicts with our assumption in (105); implying (103). This completes the converse proof for this case.

We note that contrary to Sections IV-B1 and IV-B2, here we prove the optimality of Gaussian signalling, i.e.,

$$g(R_0^*) = f(R_0^*) \tag{132}$$

directly. In other words, to show (132), we did not find any other points on the boundary of the capacity region $\mathcal{C}_p(\mathbf{S})$ and did not have to use a time-sharing argument between these points to reach (132). (This was our strategy in Sections IV-B1 and IV-B2.) Instead, we define a new optimization problem given in (113) whose solution yields (132).

V. Proof of Theorem 3 for the General Case

The achievability of the region given in Theorem 3 can be shown by computing the region in Theorem 1 with the following selection of $(U,V,\mathbf{X})$: $V=\mathbf{X}$, $\mathbf{X}=\mathbf{U}+\mathbf{T}$ where $\mathbf{T},\mathbf{U}$ are independent Gaussian random vectors with covariance matrices $\mathbf{K},\mathbf{S}-\mathbf{K}$, respectively, $U=\mathbf{U}$. In the rest of this section, we consider the converse proof. We first note that following the approaches in [2, Sec. V.B] and [12, Sec. 7.1], it can be shown that a new Gaussian MIMO wiretap channel can be constructed from any Gaussian MIMO wiretap channel described by (13)–(14) such that the new channel has the same capacity-equivocation region with the original one and in the new channel, both the legitimate user and the eavesdropper have the same number of antennas as the transmitter, i.e., $r_Y = r_Z = t$. Thus, without loss of generality, we assume that $r_Y = r_Z = t$. We next apply singular-value decomposition to the channel gain matrices $\mathbf{H}_Y, \mathbf{H}_Z$ as follows:

$$\mathbf{H}_Y = \mathbf{U}_Y \boldsymbol{\Lambda}_Y \mathbf{V}_Y^\top \tag{133}$$
$$\mathbf{H}_Z = \mathbf{U}_Z \boldsymbol{\Lambda}_Z \mathbf{V}_Z^\top \tag{134}$$

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

$$\overline{\mathbf{Y}} = \overline{\mathbf{H}}_Y \mathbf{X} + \mathbf{N}_Y \tag{135}$$
$$\overline{\mathbf{Z}} = \overline{\mathbf{H}}_Z \mathbf{X} + \mathbf{N}_Z \tag{136}$$

where $\overline{\mathbf{H}}_Y, \overline{\mathbf{H}}_Z$ are defined as

$$\overline{\mathbf{H}}_Y = \mathbf{U}_Y(\boldsymbol{\Lambda}_Y + \alpha\mathbf{I})\mathbf{V}_Y^\top \tag{137}$$
$$\overline{\mathbf{H}}_Z = \mathbf{U}_Z(\boldsymbol{\Lambda}_Z + \alpha\mathbf{I})\mathbf{V}_Z^\top \tag{138}$$

for some $\alpha > 0$. We denote the capacity-equivocation region of the Gaussian MIMO wiretap channel defined in (135)–(136) by $\mathcal{C}_\alpha(\mathbf{S})$. Since $\overline{\mathbf{H}}_Y, \overline{\mathbf{H}}_Z$ are invertible, the capacity-equivocation region of the channel in (135)–(136) is equal to the capacity-equivocation region of the following aligned channel:

$$\overline{\overline{\mathbf{Y}}} = \mathbf{X} + \overline{\mathbf{H}}_Y^{-1}\mathbf{N}_Y \tag{139}$$
$$\overline{\overline{\mathbf{Z}}} = \mathbf{X} + \overline{\mathbf{H}}_Z^{-1}\mathbf{N}_Z. \tag{140}$$

Thus, using the capacity result for the aligned case, we obtain $\mathcal{C}_\alpha(\mathbf{S})$ as the union of nonnegative rate triples (R_0, R_1, R_e) satisfying

$$R_e \leq \frac{1}{2}\log\frac{\left|\overline{\mathbf{H}}_Y\mathbf{K}\overline{\mathbf{H}}_Y^\top + \boldsymbol{\Sigma}_Y\right|}{|\boldsymbol{\Sigma}_Y|} - \frac{1}{2}\log\frac{\left|\overline{\mathbf{H}}_Z\mathbf{K}\overline{\mathbf{H}}_Z^\top + \boldsymbol{\Sigma}_Z\right|}{|\boldsymbol{\Sigma}_Z|} \tag{141}$$

$$R_0 + R_1 \leq \frac{1}{2}\log\frac{\left|\overline{\mathbf{H}}_Y\mathbf{K}\overline{\mathbf{H}}_Y^\top + \boldsymbol{\Sigma}_Y\right|}{|\boldsymbol{\Sigma}_Y|} + \frac{1}{2}\min\left\{\log\frac{\left|\overline{\mathbf{H}}_Y\mathbf{S}\overline{\mathbf{H}}_Y^\top + \boldsymbol{\Sigma}_Y\right|}{\left|\overline{\mathbf{H}}_Y\mathbf{K}\overline{\mathbf{H}}_Y^\top + \boldsymbol{\Sigma}_Y\right|}, \log\frac{\left|\overline{\mathbf{H}}_Z\mathbf{S}\overline{\mathbf{H}}_Z^\top + \boldsymbol{\Sigma}_Z\right|}{\left|\overline{\mathbf{H}}_Z\mathbf{K}\overline{\mathbf{H}}_Z^\top + \boldsymbol{\Sigma}_Z\right|}\right\} \tag{142}$$

$$R_0 \leq \frac{1}{2}\min\left\{\log\frac{\left|\overline{\mathbf{H}}_Y\mathbf{S}\overline{\mathbf{H}}_Y^\top + \boldsymbol{\Sigma}_Y\right|}{\left|\overline{\mathbf{H}}_Y\mathbf{K}\overline{\mathbf{H}}_Y^\top + \boldsymbol{\Sigma}_Y\right|}, \log\frac{\left|\overline{\mathbf{H}}_Z\mathbf{S}\overline{\mathbf{H}}_Z^\top + \boldsymbol{\Sigma}_Z\right|}{\left|\overline{\mathbf{H}}_Z\mathbf{K}\overline{\mathbf{H}}_Z^\top + \boldsymbol{\Sigma}_Z\right|}\right\} \tag{143}$$

for some positive semidefinite matrix $\mathbf{0} \preceq \mathbf{K} \preceq \mathbf{S}$.

We next obtain an outer bound for the capacity-equivocation region of the original Gaussian MIMO wiretap channel in (13)–(14) in terms of $\mathcal{C}_\alpha(\mathbf{S})$. To this end, we first note the following Markov chains:

$$\mathbf{X} \to \overline{\mathbf{Y}} \to \mathbf{Y} \tag{144}$$
$$\mathbf{X} \to \overline{\mathbf{Z}} \to \mathbf{Z} \tag{145}$$

which imply that if the messages (W_0, W_1) with rates (R_0, R_1) are transmitted with a vanishingly small probability of error in the original Gaussian MIMO wiretap channel given by (13)–(14), they will be transmitted with a vanishingly small probability of error in the new Gaussian MIMO wiretap channel given by (135)–(136) as well. However, as opposed to the rates R_0, R_1, we cannot immediately conclude that if an equivocation rate R_e is achievable in the original Gaussian MIMO wiretap channel given in (13)–(14), it is also achievable in the new Gaussian MIMO wiretap channel in (135)–(136). The reason for this is that both the legitimate user's and the eavesdropper's channel gain matrices are enhanced in the new channel given by (135)–(136) [see (137)–(138) and/or (144)–(145)], and consequently, it is not clear what the overall effect of these two enhancements on the equivocation rate will be. However, in the sequel, we show that if $(R_0, R_1, R_e) \in \mathcal{C}(\mathbf{S})$, then we have $(R_0, R_1, R_e - \gamma) \in \mathcal{C}_\alpha(\mathbf{S})$. This will let us write down an outer bound for $\mathcal{C}(\mathbf{S})$ in terms of $\mathcal{C}_\alpha(\mathbf{S})$. To this end, we note that if $(R_0, R_1, R_e) \in \mathcal{C}(\mathbf{S})$, we need to have a random vector $(U,V,\mathbf{X})$ such that the inequalities given in Theorem 1 hold. Assume that we use the same random vector $(U,V,\mathbf{X})$ for the new Gaussian MIMO wiretap channel in (135)–(136), and achieve the rate triple $(\overline{R}_0, \overline{R}_1, \overline{R}_e)$. Due to the Markov chains in (144)–(145), we already have $R_1 \leq \overline{R}_1, R_0 \leq \overline{R}_0$. Furthermore, following the analysis in [9, Sec. 4], we can bound the gap between R_e and $\overline{R}_e$, i.e., γ, as follows:

$$\gamma = R_e - \overline{R}_e \tag{146}$$

$$\leq \frac{1}{2}\log\frac{\left|\overline{\mathbf{H}}_Z\mathbf{S}\overline{\mathbf{H}}_Z^\top + \boldsymbol{\Sigma}_Z\right|}{|\boldsymbol{\Sigma}_Z|} - \frac{1}{2}\log\frac{|\mathbf{H}_Z\mathbf{S}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}. \tag{147}$$

Thus, we have

$$\mathcal{C}(\mathbf{S}) \subseteq \mathcal{C}_\alpha(\mathbf{S}) + \mathcal{G}(\mathbf{S}) \tag{148}$$

where $\mathcal{G}(\mathbf{S})$ is given by (149), which is given at the bottom of the page. Taking $\alpha \to 0$ in (148), we get

$$\mathcal{C}(\mathbf{S}) \subseteq \lim_{\alpha\to 0}\mathcal{C}_\alpha(\mathbf{S}) \tag{150}$$

where we use the fact that

$$\lim_{\alpha\to 0}\frac{1}{2}\log\frac{\left|\overline{\mathbf{H}}_Z\mathbf{S}\overline{\mathbf{H}}_Z^\top + \boldsymbol{\Sigma}_Z\right|}{|\boldsymbol{\Sigma}_Z|} - \frac{1}{2}\log\frac{|\mathbf{H}_Z\mathbf{S}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|} = 0 \tag{151}$$

which follows from the continuity of $\log|\cdot|$ in positive semidefinite matrices, and the fact that $\lim_{\alpha\to 0}\overline{\mathbf{H}}_Z = \mathbf{H}_Z$. Finally, we note that

$$\lim_{\alpha\to 0}\mathcal{C}_\alpha(\mathbf{S}) \tag{152}$$

converges to the region given in Theorem 3 due to the continuity of $\log|\cdot|$ in positive semidefinite matrices and $\lim_{\alpha\to 0}\overline{\mathbf{H}}_Y = \mathbf{H}_Y, \lim_{\alpha\to 0}\overline{\mathbf{H}}_Z = \mathbf{H}_Z$; completing the proof.

VI. Conclusion

We study the Gaussian MIMO wiretap channel in which a common message is sent to both the legitimate user and the eavesdropper in addition to the private message sent only to the legitimate user. We first establish an equivalence between this original definition of the wiretap channel and the wiretap channel with public messages, in which the private message is divided into two parts as the confidential message, which needs to be transmitted in perfect secrecy, and public message, on which there is no secrecy constraint. We next obtain capacity regions for both cases. We show that it is sufficient to consider jointly Gaussian auxiliary random variables and channel input to evaluate the single-letter description of the capacity-equivocation region due to [1]. We prove this by using channel enhancement [2] and an extremal inequality from [3].

Appendix I
Proof of Lemma 1

The proof of this lemma for $R_0 = 0$ is outlined in [4, Problem 33-c], [5]. We extend their proof to the general case of interest here. We first note the inclusion $\mathcal{C}_p \subseteq \mathcal{C}$, which follows from the fact that if $(R_0, R_p, R_s) \in \mathcal{C}_p$, we can attain the rate triple $(R_0, R_1 = R_s + R_p, R_e = R_s)$, i.e., $(R_0, R_s + R_p, R_s) \in \mathcal{C}$. To show the reverse inclusion, we use the achievability proof for Theorem 1 given in [1]. According to this achievable scheme, W_1 can be divided into two parts as $W_1 = (W_p, W_s)$ with rates $(R_1 - R_e, R_e)$, respectively, and we have

$$H(W_1|W_0, Z^n) = H(W_p, W_s|Z^n, W_0) \tag{153}$$

$$\geq H(W_s|Z^n, W_0) \tag{154}$$

$$\geq H(W_s) - n\gamma_n \tag{155}$$

for some γ_n which satisfies $\lim_{n\to\infty}\gamma_n = 0$. Hence, using this capacity achieving scheme for $\mathcal{C}$, we can attain the rate triple $(R_0, R_p = R_1 - R_e, R_s = R_e) \in \mathcal{C}_p$. This implies $\mathcal{C} \subseteq \mathcal{C}_p$; completing the proof of the lemma.

Appendix II
Proof of Lemma 2

Since the program in (44)–(45) is not necessarily convex, the KKT conditions are necessary but not sufficient. The Lagrangian for this optimization problem is given by

$$\begin{aligned}\mathcal{L} &= \mu_s R_s(\mathbf{K}) + \mu_p R_p + \lambda_Y\left[R_p(\mathbf{K}) + R_{0Y}(\mathbf{K}) - R_p - R_0^*\right] \\ &+ \lambda_Z\left[R_p(\mathbf{K}) + R_{0Z}(\mathbf{K}) - R_p - R_0^*\right] + \beta_Y\left[R_{0Y}(\mathbf{K}) - R_0^*\right] \\ &+ \beta_Z\left[R_{0Z}(\mathbf{K}) - R_0^*\right] + \mathrm{tr}(\mathbf{K}\mathbf{M}) + \mathrm{tr}((\mathbf{S}-\mathbf{K})\mathbf{M}_S)\end{aligned} \tag{156}$$

where $\mathbf{M}, \mathbf{M}_S$ are positive semidefinite matrices, and $\lambda_Y \geq 0, \lambda_Z \geq 0, \beta_Y \geq 0, \beta_Z \geq 0$.

The necessary KKT conditions that they need to satisfy are given as follows:

$$\frac{\partial\mathcal{L}}{\partial R_p}\Big|_{R_p = R_p^*} = 0 \tag{157}$$

$$\nabla_{\mathbf{K}}\mathcal{L}\,|_{\mathbf{K}=\mathbf{K}^*} = \mathbf{0} \tag{158}$$

$$\mathrm{tr}(\mathbf{K}^*\mathbf{M}) = 0 \tag{159}$$

$$\mathrm{tr}((\mathbf{S}-\mathbf{K}^*)\mathbf{M}_S) = 0 \tag{160}$$

$$\lambda_Y\left[R_p(\mathbf{K}^*) + R_{0Y}(\mathbf{K}^*) - R_p^* - R_0^*\right] = 0 \tag{161}$$

$$\lambda_Z\left[R_p(\mathbf{K}^*) + R_{0Z}(\mathbf{K}^*) - R_p^* - R_0^*\right] = 0 \tag{162}$$

$$\beta_Y(R_{0Y}(\mathbf{K}^*) - R_0^*) = 0 \tag{163}$$

$$\beta_Z(R_{0Z}(\mathbf{K}^*) - R_0^*) = 0. \tag{164}$$

$$\mathcal{G}(\mathbf{S}) = \left\{(0,0,R_e) : 0 \leq R_e \leq \frac{1}{2}\log\frac{\left|\overline{\mathbf{H}}_Z\mathbf{S}\overline{\mathbf{H}}_Z^\top + \boldsymbol{\Sigma}_Z\right|}{|\boldsymbol{\Sigma}_Z|} - \frac{1}{2}\log\frac{|\mathbf{H}_Z\mathbf{S}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|}\right\} \tag{149}$$

The first KKT condition in (157) implies $\lambda_Y + \lambda_Z = \mu_p$. We define $\lambda_Y = \mu_p\lambda, \lambda_Z = \mu_p\bar{\lambda}$ and consequently, we have $0 \leq \bar{\lambda} = 1 - \lambda \leq 1$. The second KKT condition in (158) implies (46). Since $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ and $\text{tr}(\mathbf{AB}) \geq 0$ for $\mathbf{A} \succeq \mathbf{0}, \mathbf{B} \succeq \mathbf{0}$, (159)–(160) imply (47)–(48). The KKT conditions in (161)–(162) imply (51). Furthermore, the KKT conditions in (161)–(162) state the conditions that if $R_{0Y}(\mathbf{K}^*) > R_{0Z}(\mathbf{K}^*)$, $\lambda = 0$, if $R_{0Y}(\mathbf{K}^*) < R_{0Z}(\mathbf{K}^*)$, $\lambda = 1$, and if $R_{0Y}(\mathbf{K}^*) = R_{0Z}(\mathbf{K}^*)$, λ is arbitrary, i.e., $0 \leq \lambda \leq 1$. Similarly, the KKT conditions in (163)–(164) imply (50).

APPENDIX III
PROOF OF LEMMA 3

We note the following identities:

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z)^{-1} = (\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1} + \mathbf{M}_S \tag{165}$$

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z)^{-1} = (\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1} + \mathbf{M} \tag{166}$$

where (165) is due to (56), and (166) is obtained by plugging (165) into (52). Since $\mathbf{M} \succeq \mathbf{0}, \mathbf{M}_S \succeq \mathbf{0}$, (165)–(166) implies

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z)^{-1} \succeq (\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1} \tag{167}$$

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z)^{-1} \succeq (\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1}. \tag{168}$$

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 (167)–(168), we can get the second and third parts of Lemma 3. Next, we prove the first part of the lemma as follows:

$$\tilde{\boldsymbol{\Sigma}}_Z = \left[(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1} + \frac{1}{\mu_s - \mu_p\lambda}\mathbf{M}\right]^{-1} - \mathbf{K}^* \tag{169}$$

$$= \left[\mathbf{I} + \frac{1}{\mu_s - \mu_p\lambda}(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)\mathbf{M}\right]^{-1} (\mathbf{K}^* + \boldsymbol{\Sigma}_Y) - \mathbf{K}^* \tag{170}$$

$$= \left[\mathbf{I} + \frac{1}{\mu_s - \mu_p\lambda}\boldsymbol{\Sigma}_Y\mathbf{M}\right]^{-1} (\mathbf{K}^* + \boldsymbol{\Sigma}_Y) - \mathbf{K}^* \tag{171}$$

$$= \left[\boldsymbol{\Sigma}_Y^{-1} + \frac{1}{\mu_s - \mu_p\lambda}\mathbf{M}\right]^{-1} \boldsymbol{\Sigma}_Y^{-1}(\mathbf{K}^* + \boldsymbol{\Sigma}_Y) - \mathbf{K}^* \tag{172}$$

$$= \left[\boldsymbol{\Sigma}_Y^{-1} + \frac{1}{\mu_s - \mu_p\lambda}\mathbf{M}\right]^{-1} \left[\boldsymbol{\Sigma}_Y^{-1} + \frac{1}{\mu_s - \mu_p\lambda}\mathbf{M}\right]\mathbf{K}^* + \left[\boldsymbol{\Sigma}_Y^{-1} + \frac{1}{\mu_s - \mu_p\lambda}\mathbf{M}\right]^{-1} - \mathbf{K}^* \tag{173}$$

$$= \mathbf{K}^* + \left[\boldsymbol{\Sigma}_Y^{-1} + \frac{1}{\mu_s - \mu_p\lambda}\mathbf{M}\right]^{-1} - \mathbf{K}^* \tag{174}$$

$$= \left[\boldsymbol{\Sigma}_Y^{-1} + \frac{1}{\mu_s - \mu_p\lambda}\mathbf{M}\right]^{-1} \tag{175}$$

$$\succeq \mathbf{0} \tag{176}$$

where (169) comes from (166), (171) and (173) follow from the KKT condition in (47).

Finally, we show the fourth part of Lemma 3 as follows:

$$(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z)^{-1}(\mathbf{S} + \tilde{\boldsymbol{\Sigma}}_Z)$$

$$= (\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z)^{-1}(\mathbf{S} + \mathbf{K}^* - \mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z) \tag{177}$$

$$= \mathbf{I} + (\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Z)^{-1}(\mathbf{S} - \mathbf{K}^*) \tag{178}$$

$$= \mathbf{I} + \left[(\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1} + \frac{1}{\mu_s - \mu_p\lambda}\mathbf{M}_S\right](\mathbf{S} - \mathbf{K}^*) \tag{179}$$

$$= \mathbf{I} + (\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1}(\mathbf{S} - \mathbf{K}^*) \tag{180}$$

$$= (\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1}(\mathbf{K}^* + \boldsymbol{\Sigma}_Z) + (\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1}(\mathbf{S} - \mathbf{K}^*) \tag{181}$$

$$= (\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1}(\mathbf{S} + \boldsymbol{\Sigma}_Z) \tag{182}$$

where (179) is due to (165), and (180) comes from (48); completing the proof.

APPENDIX IV
PROOF OF LEMMA 4

We note the following:

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Y)^{-1} = (\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1} + \mathbf{M} \tag{183}$$

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Y)^{-1} = (\mu_s - \mu_p\lambda + \mu_0\bar{\beta})(\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1} + \mu_0\beta(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1} + \mathbf{M}_S \tag{184}$$

where (183) is (83), and (184) comes from plugging (183) into (74). Since $\mathbf{M} \succeq \mathbf{0}$, (183) implies

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Y)^{-1} \succeq (\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1}. \tag{185}$$

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 (185) yields the second statement of the lemma. Since $0 \leq \beta = 1 - \bar{\beta} \leq 1$ and $\mathbf{M}_S \succeq \mathbf{0}$, (184) implies

$$(\mu_s - \mu_p\lambda)(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Y)^{-1} \succeq (\mu_s - \mu_p\lambda)(\mathbf{K}^* + \boldsymbol{\Sigma}_Z)^{-1}. \tag{186}$$

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 (186) yields the third statement of the lemma. To prove the first statement of the lemma, we note that (183) implies

$$\tilde{\boldsymbol{\Sigma}}_Y = \left[(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1} + \frac{1}{\mu_s - \mu_p\lambda}\mathbf{M}\right]^{-1} - \mathbf{K}^* \tag{187}$$

which is already shown to be positive semidefinite as done through (169)–(176) in Appendix III.

Finally, we consider the fourth statement of this lemma as follows:

$$(\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Y)^{-1}\tilde{\boldsymbol{\Sigma}}_Y$$

$$= (\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Y)^{-1}(\mathbf{K}^* - \mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Y) \tag{188}$$

$$= \mathbf{I} - (\mathbf{K}^* + \tilde{\boldsymbol{\Sigma}}_Y)^{-1}\mathbf{K}^* \tag{189}$$

$$= \mathbf{I} - \left[(\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1} + \frac{1}{\mu_s - \mu_p\lambda}\mathbf{M}\right]\mathbf{K}^* \tag{190}$$

$$= \mathbf{I} - (\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1}\mathbf{K}^* \tag{191}$$

$$= (\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1}(\mathbf{K}^* + \boldsymbol{\Sigma}_Y) - (\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1}\mathbf{K}^* \tag{192}$$

$$= (\mathbf{K}^* + \boldsymbol{\Sigma}_Y)^{-1}\boldsymbol{\Sigma}_Y \tag{193}$$

where (190) is due to (183) and (191) comes from (47).

APPENDIX V
PROOF OF LEMMA 6

The optimization problem in (113) can be written as

$$\max_{\substack{U\to V\to \mathbf{X}\to(\mathbf{Y},\mathbf{Z})\\ E[\mathbf{X}\mathbf{X}^\top]\preceq \mathbf{S}}} \quad \mu_s R_s + \mu_p R_p + (\mu_p+\beta_Z)R_0 \tag{194}$$

$$\text{s.t.}\begin{cases} 0\le R_s \le I(V;\mathbf{Y}|U) - I(V;\mathbf{Z}|U)\\ R_s+R_p+R_0 \le I(V;\mathbf{Y}|U) + \min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\}\\ R_0 \le \min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\}.\end{cases} \tag{195}$$

For a given $(U,V,\mathbf{X})$, we can rewrite the cost function in (194) as follows:

$$\begin{aligned}
&\mu_s R_s + \mu_p R_p + (\mu_p+\beta_Z)R_0\\
&\le \mu_s R_s\\
&\quad + \mu_p[I(V;\mathbf{Y}|U) + \min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\} - R_s - R_0]\\
&\quad + (\mu_p+\beta_Z)R_0 && (196)\\
&= (\mu_s-\mu_p)R_s + \mu_p[I(V;\mathbf{Y}|U) + \min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\}]\\
&\quad + \beta_Z R_0 && (197)\\
&\le (\mu_s-\mu_p)[I(V;\mathbf{Y}|U) - I(V;\mathbf{Z}|U)]\\
&\quad + \mu_p[I(V;\mathbf{Y}|U) + \min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\}] + \beta_Z R_0 && (198)\\
&= \mu_s[I(V;\mathbf{Y}|U) - I(V;\mathbf{Z}|U)]\\
&\quad + \mu_p[I(V;\mathbf{Z}|U) + \min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\}] + \beta_Z R_0 && (199)\\
&\le \mu_s[I(V;\mathbf{Y}|U) - I(V;\mathbf{Z}|U)]\\
&\quad + \mu_p[I(V;\mathbf{Z}|U) + \min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\}]\\
&\quad + \beta_Z \min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\} && (200)\\
&= \mu_s[I(V;\mathbf{Y}|U) - I(V;\mathbf{Z}|U)] + \mu_p I(V;\mathbf{Z}|U)\\
&\quad + (\mu_p+\beta_Z)\min\{I(U;\mathbf{Y}), I(U;\mathbf{Z})\} && (201)
\end{aligned}$$

where (196) comes from the second constraint in (195), (198) is due to the first constraint in (195) and the assumption $\mu_s > \mu_p$, and (200) comes from the third constraint in (195). The proof can be concluded by noting that the upper bound on the cost function given in (201) is attainable.

REFERENCES

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

[2] 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.

[3] 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.

[4] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1982.

[5] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "The capacity-equivocation region of the MIMO Gaussian wiretap channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 2568–2572.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

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

[12] 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.

[13] J. Korner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 60–64, Jan. 1977.

[14] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multi-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[15] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

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.