# Degraded Compound Multi-Receiver Wiretap Channels

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

*Abstract*—We study the degraded compound multi-receiver wiretap channel (DCMRWC). DCMRWC consists of two groups of users and a group of eavesdroppers, where, if we pick an arbitrary user from each group of users and an arbitrary eavesdropper, they satisfy a certain Markov chain. We study two different communication scenarios for this channel. In the first scenario, the transmitter wants to send a confidential message to users in the first (stronger) group and a different confidential message to users in the second (weaker) group, where both messages need to be kept confidential from the eavesdroppers. For this scenario, we assume that there is only one eavesdropper. We obtain the secrecy capacity region for the discrete memoryless channel model, the parallel channel model, and the Gaussian parallel channel model. For the Gaussian multiple-input multiple-output (MIMO) channel model, we obtain the secrecy capacity region when there is only one user in the second group. In the second scenario we study, the transmitter sends a confidential message to users in the first group which needs to be kept confidential from the second group of users and the eavesdroppers. Moreover, the transmitter sends a different confidential message to users in the second group which needs to be kept confidential only from the eavesdroppers. For this scenario, we do not put any restriction on the number of eavesdroppers. As in the first scenario, we obtain the secrecy capacity region for the discrete memoryless channel model, the parallel channel model, and the Gaussian parallel channel model. For the Gaussian MIMO channel model, we establish the secrecy capacity region when there is only one user in the second group.

*Index Terms*—Compound wiretap channels, Gaussian multiple-input multiple-output (MIMO) compound wiretap channels, secrecy capacity region.

## I. INTRODUCTION

**I**NFORMATION-theoretic secrecy was initiated by Wyner in his seminal work [1], where he considered the degraded wiretap channel and established the capacity-equivocation rate region of this degraded channel model. Later, Csiszar and Korner generalized his result to arbitrary, not necessarily degraded, wiretap channels in [2]. In recent years, multiuser versions of the wiretap channel have attracted a considerable amount of research interest (see references [3]–[21] in [3]). Among all these extensions, two natural extensions of the wiretap channel to the multiuser setting are particularly of interest here: *secure broadcasting* and *compound wiretap channels*.

*Secure broadcasting* refers to the situation where a transmitter wants to communicate with several legitimate receivers confidentially in the presence of an external eavesdropper. We call this channel model the *multi-receiver wiretap channel*. Since the underlying channel model without an eavesdropper is the broadcast channel, which is not understood to the full extent even for the two-user case, most works on *secure broadcasting* have focused on some special classes of multi-receiver wiretap channels, where these classes are identified by certain degradation orders [4]–[9]. In particular, in [5]–[8], the authors consider the *degraded* multi-receiver wiretap channel, where observations of all users and the eavesdropper satisfy a certain Markov chain. In [5] and [6], the secrecy capacity region is derived for the two-user case, and in [7] and [8], the secrecy capacity region is established for an arbitrary number of legitimate users. The importance of this result lies in the fact that the Gaussian multi-receiver wiretap channel belongs to this class, and the secrecy capacity region of the degraded multi-receiver wiretap channel serves as a crucial step in establishing the secrecy capacity region of the Gaussian multiple-input multiple-output (MIMO) multi-receiver wiretap channel [3], [10], though the latter channel is not necessarily degraded. In [3], besides proving the secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel, we also present new optimization results regarding extremal properties of Gaussian random vectors, which we generalize here.

Another extension of the wiretap channel that we are particularly interested in here is the *compound wiretap channel*. In compound wiretap channels, there are a finite number of channel states determining the channel transition probability. The channel takes a certain fixed state for the entire duration of the transmission, and the transmitter does not have any knowledge about the channel state realization. Thus, the aim of the transmitter is to ensure the secrecy of messages irrespective of the channel state realization. In addition to this definition, the compound wiretap channel admits another interpretation. Consider the multi-receiver wiretap channel with several legitimate users and many eavesdroppers, where the transmitter wants to transmit a common confidential message to legitimate users while keeping all of the eavesdroppers totally ignorant of the message. Since each eavesdropper and legitimate user pair can be regarded as a different channel state realization, this channel is equivalent to a compound wiretap channel. Therefore, one can interpret a compound wiretap channel as *multicasting* a common confidential message to several legitimate receivers in the presence of one or more eavesdroppers [11]. In this study, we mostly refer to this interpretation, which is also the reason why we classify the compound wiretap channel as an extension of the wiretap channel to a multiuser setting.

Keeping this interpretation in mind, first works about the compound wiretap channel are due to Yamamoto [12], [13].

Yamamoto [12], [13] considers the parallel wiretap channel with two subchannels where each subchannel is wiretapped by a different eavesdropper. [12] and [13] establish capacity-equivocation rate region for the situation where in each subchannel, the legitimate receiver is less noisy with respect to the eavesdropper of this subchannel. Other works which implicitly study the compound wiretap channel are [4], [7]–[9], and [14], where works [4], [7], and [8] consider the transmission of a common confidential message to many legitimate receivers in the presence of a single eavesdropper, work in [9] focuses on two legitimate receivers one eavesdropper and one legitimate receiver two eavesdroppers scenarios, and [14] studies the fading wiretap channel with many receivers. Liang *et al.* [11] consider the general discrete compound wiretap channel and provides inner and outer bounds for the secrecy capacity. In addition to these inner and outer bounds, [11] also establishes the secrecy capacity of the degraded compound wiretap channel as well as its degraded Gaussian MIMO instance. Another work on the compound wiretap channel is [15] where the secrecy capacity of a class of nondegraded Gaussian parallel compound wiretap channels is established.

In this paper, we consider compound broadcast channels from a secrecy point of view, which enables us to study the *secure broadcasting* problem over *compound channels*. We note that the current literature regarding the compound wiretap channel considers the transmission of only one confidential message, whereas here, we study the transmission of multiple confidential messages, where each of these messages needs to be delivered to a different group of users in perfect secrecy. Hereafter, we call this channel model the *compound multi-receiver wiretap channel* to emphasize the presence of more than one confidential message. The compound multi-receiver wiretap channel we study here consists of two groups of users and a group of eavesdroppers, as shown in Fig. 1. We focus on a special class of compound multi-receiver wiretap channels which exhibits a certain degradation order. If we consider an arbitrary user from each group and an arbitrary eavesdropper, they satisfy a certain Markov chain. In particular, we assume that there exist two fictitious users. The first fictitious user is degraded with respect to any user from the first group, and any user from the second group is degraded with respect to the first fictitious user. There exists a similar degradedness structure for the second fictitious user in the sense that it is degraded with respect to any user from the second group, and any eavesdropper is degraded with respect to it. Without eavesdroppers, this channel model reduces to the degraded compound broadcast channel studied in [16]. Adapting their terminology, we call our channel model the *degraded compound multi-receiver wiretap channel* (DCMRWC). Here, we consider the general discrete memoryless version of DCMRWC as well as its specializations to the parallel DCMRWC, the Gaussian parallel DCMRWC, and the Gaussian MIMO DCMRWC. We study two different communication scenarios for each version of DCMRWC.

In the first scenario, which is illustrated in Fig. 2, the transmitter wants to send a confidential message to users in the first group, and a different confidential message to users in the second group, where both messages need to be kept confidential from the eavesdroppers. For this scenario, we assume that there
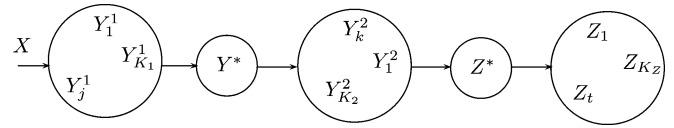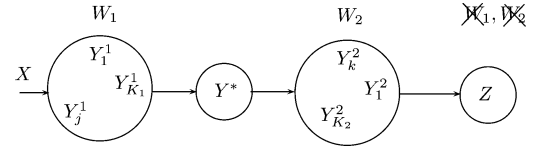


Fig. 1. DCMRWC.
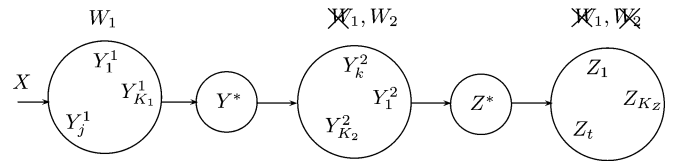


Fig. 2. First scenario for the DCMRWC.



Fig. 3. Second scenario for the DCMRWC.

exists only one eavesdropper and obtain the secrecy capacity region in a single-letter form. While obtaining this result, the presence of the fictitious user between the two groups of users plays a crucial role in the converse proof by providing a conditional independence structure in the channel, which enables us to define an auxiliary random variable that yields a tight outer bound. After establishing single-letter expressions for the secrecy capacity region, we consider the parallel DCMRWC. For the parallel DCMRWC, we obtain the secrecy capacity region in a single-letter form as well. Though the general discrete memoryless DCMRWC encompasses the parallel DCMRWC as a special case, we still need a converse proof to establish the optimality of independent signalling in each subchannel. After we obtain the secrecy capacity region of the parallel DCMRWC, we consider the Gaussian parallel DCMRWC. In particular, we evaluate the secrecy capacity region of the parallel DCMRWC for the Gaussian case, which is tantamount to finding the optimal joint distribution of auxiliary random variables and channel inputs, which is shown to be Gaussian. We accomplish this by using Costa's entropy power inequality [17]. Finally, we consider the Gaussian MIMO DCMRWC and evaluate its secrecy capacity region when there is only one user in the second group. We show the optimality of a jointly Gaussian distribution for auxiliary random variables and channel inputs by generalizing our optimization results in [3].

In the second scenario we study here, which is illustrated in Fig. 3, the transmitter wants to send a confidential message to users in the first group which needs to be kept confidential from users in the second group and eavesdroppers. Moreover, the transmitter sends a different confidential message to users in the second group, which needs to be kept confidential from the eavesdroppers. If there were only one user in each group and one eavesdropper, this channel model would reduce to the channel model that was studied in [18]. However, here, there are an arbitrary number of users in each group and an arbitrary number of eavesdroppers. Hence, our model can be viewed as a

generalization of [18] to a compound setting. Adapting their terminology, we call this channel model the *DCMRWC with layered messages*. We first obtain the secrecy capacity region in a single-letter form for a general discrete memoryless setting, where again the presence of fictitious users plays a key role in the converse proof. Next, we consider the parallel DCMRWC with layered messages and establish its secrecy capacity region in a single-letter form. In this case as well, we provide the converse proof which is again necessary to show the optimality of independent signalling in each subchannel. After we obtain the secrecy capacity region of the parallel DCMRWC with layered messages, we evaluate it for the Gaussian parallel DCMRWC with layered messages by showing the optimality of a jointly Gaussian distribution for auxiliary random variables and channel inputs. For that purpose, we again use Costa's entropy power inequality [17]. Finally, we consider the Gaussian MIMO DCMRWC with layered messages and evaluate its secrecy capacity region when there is only one user in the second group. To this end, we show that jointly Gaussian auxiliary random variables and channel inputs are optimal by extending our optimization results in [3].

## II. SYSTEM MODEL

In this paper, we consider DCMRWC (see Fig. 1) which consists of two groups of users and a group of eavesdroppers. There are $K_1$ users in the first group, $K_2$ users in the second group, and $K_Z$ eavesdroppers. The channel is assumed to be memoryless with a transition probability

$$p(y_1^1, \ldots, y_{K_1}^1, y_1^2, \ldots, y_{K_2}^2, z_1, \ldots, z_{K_Z} | x) \qquad (1)$$

where $X \in \mathcal{X}$ is the channel input, $Y_j^1 \in \mathcal{Y}_j^1$ is the channel output of the $j$th user in the first group, $j = 1, \ldots, K_1$, $Y_k^2 \in \mathcal{Y}_k^2$ is the channel output of the $k$th user in the second group, $k = 1, \ldots, K_2$, and $Z_t \in \mathcal{Z}_t$ is the channel output of the $t$th eavesdropper, $t = 1, \ldots, K_Z$.

We assume that there exist two fictitious users with observations $Y^* \in \mathcal{Y}^*$, $Z^* \in \mathcal{Z}^*$ such that they satisfy the Markov chain

$$X \to Y_j^1 \to Y^* \to Y_k^2 \to Z^* \to Z_t \quad \forall (j, k, t). \qquad (2)$$

This Markov chain is the reason why we call the compound multi-receiver wiretap channel we study the DCMRWC. Actually, there is a slight inexactness in the terminology here because the Markov chain in (2) is more restrictive than the Markov chain

$$X \to Y_j^1 \to Y_k^2 \to Z_t \quad \forall (j, k, t) \qquad (3)$$

and it might be more natural to define the degradedness of the compound multi-receiver wiretap channel by the Markov chain in (3). However, in this study, we adapt the terminology of the previous work on compound broadcast channels [16], and call the channel satisfying (2) the DCMRWC. Finally, we note that when there are no eavesdroppers, this channel reduces to the degraded compound broadcast channel that was studied in [16].

### A. Parallel DCMRWC

The parallel DCMRWC, where each user's and each eavesdropper's channel consists of $L$ independent subchannels, i.e.,

$$Y_j^1 = (Y_{j1}^1, \ldots, Y_{jL}^1), \quad j = 1, \ldots, K_1 \qquad (4)$$

$$Y_k^2 = (Y_{k1}^2, \ldots, Y_{kL}^2), \quad k = 1, \ldots, K_2 \qquad (5)$$

$$Z_t = (Z_{t1}, \ldots, Z_{tL}), \quad t = 1, \ldots, K_Z \qquad (6)$$

has the following overall transition probability:

$$p\left(\{y_j^1\}_{j=1}^{K_1}, \{y_k^2\}_{k=1}^{K_2}, \{z_t\}_{t=1}^{K_Z} | x\right)$$
$$= \prod_{\ell=1}^{L} p\left(\{y_{j\ell}^1\}_{j=1}^{K_1}, \{y_{k\ell}^2\}_{k=1}^{K_2}, \{z_{t\ell}\}_{t=1}^{K_Z} | x_\ell\right) \quad (7)$$

where $X_\ell$, $\ell = 1, \ldots, L$, is the $\ell$th subchannel's input. We define the degradedness of the parallel compound multi-receiver wiretap channel in a similar fashion. In particular, we call a parallel compound multi-receiver wiretap channel degraded, if there exist two sequences of random variables

$$Y^* = (Y_1^*, \ldots, Y_L^*) \qquad (8)$$

$$Z^* = (Z_1^*, \ldots, Z_L^*) \qquad (9)$$

which satisfy Markov chains

$$X_\ell \to Y_{j\ell}^1 \to Y_\ell^* \to Y_{k\ell}^2 \to Z_\ell^* \to Z_{t\ell} \quad \forall (j, k, t, \ell). \quad (10)$$

### B. Gaussian Parallel DCMRWC

The Gaussian parallel DCMRWC is defined by

$$\mathbf{Y}_j^1 = \mathbf{X} + \mathbf{N}_j^1, \quad j = 1, \ldots, K_1 \qquad (11)$$

$$\mathbf{Y}_k^2 = \mathbf{X} + \mathbf{N}_k^2, \quad k = 1, \ldots, K_2 \qquad (12)$$

$$\mathbf{Z}_t = \mathbf{X} + \mathbf{N}_t^Z, \quad t = 1, \ldots, K_Z \qquad (13)$$

where all column vectors $\{\mathbf{Y}_j^1\}_{j=1}^{K_1}$, $\{\mathbf{Y}_k^1\}_{k=1}^{K_2}$, $\{\mathbf{Z}_t\}_{t=1}^{K_Z}$, $\mathbf{X}$, $\{\mathbf{N}_j^1\}_{j=1}^{K_1}$, $\{\mathbf{N}_k^2\}_{k=1}^{K_2}$, $\{\mathbf{N}_t^Z\}_{t=1}^{K_Z}$ are of dimensions $L \times 1$. $\{\mathbf{N}_j^1\}_{j=1}^{K_1}$, $\{\mathbf{N}_k^2\}_{k=1}^{K_2}$, $\{\mathbf{N}_t^Z\}_{t=1}^{K_Z}$ are Gaussian random vectors with diagonal covariance matrices $\{\varLambda_j^1\}_{j=1}^{K_1}$, $\{\varLambda_k^2\}_{k=1}^{K_2}$, $\{\varLambda_t^Z\}_{t=1}^{K_Z}$, respectively. The channel input $\mathbf{X}$ is subject to a trace constraint as

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

In this paper, we will be interested in Gaussian parallel DCMRWCs which means that the covariance matrices satisfy the following order:

$$\varLambda_j^1 \preceq \varLambda_k^2 \preceq \varLambda_t^Z \quad \forall (j, k, t). \qquad (15)$$

Since noise covariance matrices are diagonal, the order in (15) implies

$$\Lambda_{j,\ell\ell}^1 \le \Lambda_{k,\ell\ell}^2 \le \Lambda_{t,\ell\ell}^Z \quad \forall (j, k, t, \ell) \qquad (16)$$

where $\Lambda_{j,\ell\ell}^1, \Lambda_{k,\ell\ell}^2, \Lambda_{t,\ell\ell}^Z$ denote the $\ell$th diagonal element of $\varLambda_j^1$, $\varLambda_k^2$, $\varLambda_t^Z$, respectively.

The diagonality of noise covariance matrices also ensures the existence of diagonal matrices $\Lambda_Y^*$ and $\Lambda_Z^*$ such that

$$\Lambda_j^1 \preceq \Lambda_Y^* \preceq \Lambda_k^2 \preceq \Lambda_Z^* \preceq \Lambda_t^Z \quad \forall(k,j,t). \tag{17}$$

For example, we can select $\Lambda_Y^*$ as $\Lambda_{Y,\ell\ell} = \max_{j=1,\ldots,K_1} \Lambda_{j,\ell\ell}^1$ which already satisfies (17) because of $\max_{j=1,\ldots,K_1} \Lambda_{j,\ell\ell}^1 \leq \min_{k=1,\ldots,K_2} \Lambda_{k,\ell\ell}^2$ which is due to (16). Similarly, we can select $\Lambda_Z^*$. Thus, for Gaussian parallel compound multi-receiver channels, the two possible ways of defining degradedness, i.e., (2) and (3), are equivalent due to the equivalence of (15) and (17).

## C. Gaussian MIMO DCMRWC

The Gaussian MIMO DCMRWC is defined by

$$\mathbf{Y}_j^1 = \mathbf{X} + \mathbf{N}_j^1, \quad j = 1,\ldots,K_1 \tag{18}$$

$$\mathbf{Y}_k^2 = \mathbf{X} + \mathbf{N}_k^2, \quad k = 1,\ldots,K_2 \tag{19}$$

$$\mathbf{Z}_t = \mathbf{X} + \mathbf{N}_t^Z, \quad t = 1,\ldots,K_Z \tag{20}$$

where all column vectors $\{\mathbf{Y}_j^1\}_{j=1}^{K_1}$, $\{\mathbf{Y}_k^2\}_{k=1}^{K_2}$, $\{\mathbf{Z}_t\}_{t=1}^{K_Z}$, $\mathbf{X}$, $\{\mathbf{N}_j^1\}_{j=1}^{K_1}$, $\{\mathbf{N}_k^2\}_{k=1}^{K_2}$, $\{\mathbf{N}_t^Z\}_{t=1}^{K_Z}$ are of dimensions $M \times 1$. $\{\mathbf{N}_j^1\}_{j=1}^{K_1}$, $\{\mathbf{N}_k^2\}_{k=1}^{K_2}$, $\{\mathbf{N}_t^Z\}_{t=1}^{K_Z}$ are Gaussian random vectors with covariance matrices $\{\boldsymbol{\Sigma}_j^1\}_{j=1}^{K_1}$, $\{\boldsymbol{\Sigma}_k^2\}_{k=1}^{K_2}$, $\{\boldsymbol{\Sigma}_t^Z\}_{t=1}^{K_Z}$, respectively. Unlike in the case of Gaussian parallel channels, these covariance matrices are not necessarily diagonal. The channel input $\mathbf{X}$ is subject to a covariance constraint

$$E\left[\mathbf{X}\mathbf{X}^\top\right] \preceq \mathbf{S} \tag{21}$$

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

In this paper, we study Gaussian MIMO DCMRWCs for which there exist covariance matrices $\boldsymbol{\Sigma}_Y^*$ and $\boldsymbol{\Sigma}_Z^*$ such that

$$\boldsymbol{\Sigma}_j^1 \preceq \boldsymbol{\Sigma}_Y^* \preceq \boldsymbol{\Sigma}_k^2 \preceq \boldsymbol{\Sigma}_Z^* \preceq \boldsymbol{\Sigma}_t^Z \quad \forall(j,k,t). \tag{22}$$

We note that the order in (22), by which we define the degradedness, is more restrictive than the other possible order that can be used to define the degradedness, i.e.,

$$\boldsymbol{\Sigma}_j^1 \preceq \boldsymbol{\Sigma}_k^2 \preceq \boldsymbol{\Sigma}_t^Z \quad \forall(j,k,t). \tag{23}$$

In [16], a specific numerical example is provided to show that the order in (23) strictly subsumes the one in (22).

## D. Comments on Gaussian MIMO DCMRWC

We provide some comments about the way we define the Gaussian MIMO DCMRWC. The first one is about the covariance constraint in (21). Though it is more common to define capacity regions under a total power constraint, i.e., $\mathrm{tr}\left(E\left[\mathbf{X}\mathbf{X}^\top\right]\right) \leq P$, the covariance constraint in (21) is more general and it subsumes the total power constraint as a special case [19]. In particular, if we denote the secrecy capacity region under the constraint in (21) by $C(\mathbf{S})$, then the secrecy capacity region under the trace constraint, $\mathrm{tr}\left(E\left[\mathbf{X}\mathbf{X}^\top\right]\right) \leq P$, can be written as [19]

$$C^{\mathrm{trace}}(P) = \bigcup_{\mathbf{S}:\mathrm{tr}(\mathbf{S})\leq P} C(\mathbf{S}). \tag{24}$$

The second comment is about our assumption that $\mathbf{S}$ is strictly positive definite. This assumption does not lead to any loss of generality because for any Gaussian MIMO compound multi-receiver wiretap channel with a positive semidefinite covariance constraint, i.e., $\mathbf{S} \succeq \mathbf{0}$ and $|\mathbf{S}| = 0$, we can always construct an equivalent channel with the constraint $E\left[\mathbf{X}\mathbf{X}^\top\right] \preceq \mathbf{S}'$ where $\mathbf{S}' \succ \mathbf{0}$ (see [19, Lemma 2]), which has the same secrecy capacity region.

The last comment is about the assumption that the transmitter and all receivers have the same number of antennas. This assumption is implicit in the channel definition [see (18) – (20)] and also in the definition of degradedness [see (22)]. However, we can extend the definition of the Gaussian MIMO DCMRWC to include the cases where the number of transmit antennas and the number of receive antennas at each receiver are not necessarily the same. To this end, we first introduce the following channel model:

$$\mathbf{Y}_j^1 = \mathbf{H}_j^1 \mathbf{X} + \mathbf{N}_j^1, \quad j = 1,\ldots,K_1 \tag{25}$$

$$\mathbf{Y}_k^2 = \mathbf{H}_k^2 \mathbf{X} + \mathbf{N}_k^2, \quad k = 1,\ldots,K_2 \tag{26}$$

$$\mathbf{Z}_t = \mathbf{H}_t^Z \mathbf{X} + \mathbf{N}_t^Z, \quad t = 1,\ldots,K_Z \tag{27}$$

where $\mathbf{H}_j^1, \mathbf{H}_k^2, \mathbf{H}_t^Z$ are the channel matrices of sizes $r_j^1 \times t, r_k^2 \times t, r_t^Z \times t$, respectively, and $\mathbf{X}$ is of size $t \times 1$. The channel outputs $\mathbf{Y}_j^1, \mathbf{Y}_k^2, \mathbf{Z}_t$ are of sizes $r_j^1 \times 1, r_k^2 \times 1, r_t^Z \times 1$, respectively. The Gaussian noise vectors $\mathbf{N}_j^1, \mathbf{N}_k^2, \mathbf{N}_t^Z$ are assumed to have identity covariance matrices.

To define degradedness for the channel model given in (25) – (27), we need the following definition from [16]: A receive vector $\mathbf{Y}_a = \mathbf{H}_a \mathbf{X} + \mathbf{N}_a$ of size $r_a \times 1$ is said to be degraded with respect to $\mathbf{Y}_b = \mathbf{H}_b \mathbf{X} + \mathbf{N}_b$ of size $r_b \times 1$, if there exists a matrix $\mathbf{D}$ of size $r_a \times r_b$ such that $\mathbf{D}\mathbf{H}_b = \mathbf{H}_a$ and $\mathbf{D}\mathbf{D}^\top \preceq \mathbf{I}$. Using this equivalent definition of degradedness, we now give the equivalent definition of degradedness for the channel model in (25) – (27). To this end, we first introduce two fictitious users with observations $\mathbf{Y}^*$ and $\mathbf{Z}^*$, which are given by

$$\mathbf{Y}^* = \mathbf{H}_Y^* \mathbf{X} + \mathbf{N}_Y^* \tag{28}$$

$$\mathbf{Z}^* = \mathbf{H}_Z^* \mathbf{X} + \mathbf{N}_Z^*. \tag{29}$$

The Gaussian MIMO compound multi-receiver wiretap channel in (25) – (27) is said to be degraded if the following two conditions hold: 1) $\mathbf{Y}^*$ is degraded with respect to any user from the first group, and any user from the second group is degraded with respect to $\mathbf{Y}^*$, and 2) $\mathbf{Z}^*$ is degraded with respect to any user from the second group, and any eavesdropper is degraded with respect to $\mathbf{Z}^*$, where degradedness here is with respect to the definition given above.

In the rest of the paper, we consider the channel model given in (18) – (20) instead of the channel model given in (25) – (27), which is more general. However, if we establish the secrecy capacity region for the Gaussian MIMO DCMRWC defined by (18) – (20), we can also obtain the secrecy capacity region for the general Gaussian MIMO DCMRWC defined by (25) – (27) using the analysis carried out in [3, Sec. 7.1] and [16, Sec. V]. Since this analysis is quite standard and can be found in other works cited above, whenever we have a capacity result for the Gaussian MIMO DCMRWC defined by (18) – (20), we provide the extension of this capacity result to the general Gaussian MIMO DCMRWC defined by (25) – (27) without a proof.

## III. PROBLEM STATEMENT AND MAIN RESULTS

In this paper, we consider two different communication scenarios for the DCMRWC.

### A. First Scenario: External Eavesdroppers

In the first scenario, the transmitter wants to send a confidential message to users in the first group and a different confidential message to users in the second group, where both messages need to be kept confidential from the eavesdroppers. In this case, we assume that there is only one eavesdropper, i.e., $K_Z = 1$. The graphical illustration of the first scenario is given in Fig. 2.

An $(n, 2^{nR_1}, 2^{nR_2})$ code for the first scenario consists of two message sets $\mathcal{W}_1 = \{1, \ldots, 2^{nR_1}\}$, $\mathcal{W}_2 = \{1, \ldots, 2^{nR_2}\}$, an encoder $f : \mathcal{W}_1 \times \mathcal{W}_2 \to \mathcal{X}^n$, one decoder for each legitimate user in the first group $g_j^1 : \mathcal{Y}_j^{1,n} \to \mathcal{W}_1$, $j = 1, \ldots, K_1$, and one decoder for each legitimate user in the second group $g_k^2 : \mathcal{Y}_k^{2,n} \to \mathcal{W}_2$, $k = 1, \ldots, K_2$. The probability of error is defined as $P_e^n = \max\{P_e^{1,n}, P_e^{2,n}\}$, where $P_e^{1,n}$ and $P_e^{2,n}$ are given by

$$P_e^{1,n} = \max_{j \in \{1, \ldots, K_1\}} \Pr\left[g_j^1\left(Y_j^{1,n}\right) \neq W_1\right] \quad (30)$$

$$P_e^{2,n} = \max_{k \in \{1, \ldots, K_2\}} \Pr\left[g_k^2\left(Y_k^{2,n}\right) \neq W_2\right]. \quad (31)$$

A secrecy rate pair $(R_1, R_2)$ is said to be achievable if there exists an $(n, 2^{nR_1}, 2^{nR_2})$ code which has $\lim_{n \to \infty} P_e^n = 0$ and

$$\lim_{n \to \infty} \frac{1}{n} I(W_1, W_2; Z^n) = 0 \quad (32)$$

where we dropped the subscript of $Z_t$ since $K_Z = 1$. We note that (32) implies

$$\lim_{n \to \infty} \frac{1}{n} I(W_1; Z^n) = 0 \quad \text{and} \quad \lim_{n \to \infty} \frac{1}{n} I(W_2; Z^n) = 0. \quad (33)$$

From these definitions, it is clear that we are only interested in perfect secrecy rates of the channel. The secrecy capacity region is defined as the closure of all achievable secrecy rate pairs. A single-letter characterization of the secrecy capacity region is given as follows.

*Theorem 1:* The secrecy capacity region of the DCMRWC is given by the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{j=1, \ldots, K_1} I(X; Y_j^1 | U, Z) \quad (34)$$

$$R_2 \leq \min_{k=1, \ldots, K_2} I(U; Y_k^2 | Z) \quad (35)$$

where the union is over all $(U, X)$ such that

$$U \to X \to Y_j^1 \to Y^* \to Y_k^2 \to Z \quad (36)$$

for any $(j, k)$ pair.

Showing the achievability of this region is rather standard, thus is omitted here. We provide the converse proof in Appendix I. The presence of the fictitious user with observation $Y^*$ proves to be crucial in the converse proof. Essentially, it brings a conditional independence structure to the channel, which enables us to define the auxiliary random variable $U$, which, in turn, provides the converse proof.

As a side note, if we disable the eavesdropper by setting $Z = \phi$, the region in Theorem 1 reduces to the capacity region of the underlying degraded compound broadcast channel which was established in [16].

*1) Parallel DCMRWC:* In the upcoming section, we will consider the Gaussian parallel DCMRWC. For that purpose, here, we provide the secrecy capacity region of the parallel DCMRWC in a single-letter form.

*Theorem 2:* The secrecy capacity region of the parallel DCMRWC is given by the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{j=1, \ldots, K_1} \sum_{\ell=1}^{L} I(X_\ell; Y_{j\ell}^1 | U_\ell, Z_\ell) \quad (37)$$

$$R_2 \leq \min_{k=1, \ldots, K_2} \sum_{\ell=1}^{L} I(U_\ell; Y_{k\ell}^2 | Z_\ell) \quad (38)$$

where the union is over all distributions of the form $\prod_{\ell=1}^{L} p(u_\ell, x_\ell)$ such that

$$U_\ell \to X_\ell \to Y_{j\ell}^1 \to Y_\ell^* \to Y_{k\ell}^2 \to Z_\ell \quad (39)$$

for any $(j, k, \ell)$ triple.

Though Theorem 1 provides the secrecy capacity region for a rather general channel model including the parallel DCMRWC as a special case, we still need a converse proof to show that the region in Theorem 1 reduces to the region in Theorem 2 for parallel channels. In other words, we still need to show the optimality of independent signalling on each subchannel. This proof is provided in Appendix II.

*2) Gaussian Parallel DCMRWC:* We now obtain the secrecy capacity region of the Gaussian parallel DCMRWC. To that end, we need to evaluate the region given in Theorem 2, i.e., we need to find the optimal joint distribution $\prod_{\ell=1}^{L} p(u_\ell, x_\ell)$. We first introduce the following theorem which will be instrumental in evaluating the region in Theorem 2 for Gaussian parallel channels.

*Theorem 3:* Let $N_1$, $N^*$, $N_2$, $N_Z$ be zero-mean Gaussian random variables with variances $\sigma_1^2$, $\sigma_*^2$, $\sigma_2^2$, $\sigma_Z^2$, respectively, where

$$\sigma_1^2 \leq \sigma_*^2 \leq \sigma_2^2 \leq \sigma_Z^2. \quad (40)$$

Let $(U, X)$ be an arbitrarily dependent random variable pair, which is independent of $(N_1, N^*, N_2, N_Z)$, and the second-moment of $X$ be constrained as $E[X^2] \leq P$. Then, for any feasible $(U, X)$, we can find a $P^* \leq P$ such that

$$h(X + N_Z | U) - h(X + N^* | U) = \frac{1}{2} \log \frac{P^* + \sigma_Z^2}{P^* + \sigma_*^2} \quad (41)$$

and

$$h(X + N_Z | U) - h(X + N_1 | U) \geq \frac{1}{2} \log \frac{P^* + \sigma_Z^2}{P^* + \sigma_1^2} \quad (42)$$

$$h(X + N_Z | U) - h(X + N_2 | U) \leq \frac{1}{2} \log \frac{P^* + \sigma_Z^2}{P^* + \sigma_2^2} \quad (43)$$

for any $(\sigma_1^2, \sigma_2^2)$ satisfying the order in (40).

The proof of this theorem is provided in Appendix III. In this proof, Costa's entropy power inequality [17] plays a key role.

We now establish the secrecy capacity region of the Gaussian parallel DCMRWC.

*Theorem 4:* The secrecy capacity region of the Gaussian parallel DCMRWC is given by the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{j=1,\ldots,K_1} \sum_{\ell=1}^{L} \frac{1}{2} \log\left(1 + \frac{\beta_\ell P_\ell}{\Lambda_{j,\ell\ell}^1}\right) - \frac{1}{2}\log\left(1 + \frac{\beta_\ell P_\ell}{\Lambda_{Z,\ell\ell}}\right) \tag{44}$$

$$R_2 \leq \min_{k=1,\ldots,K_2} \sum_{\ell=1}^{L} \frac{1}{2}\log\left(1 + \frac{\bar{\beta}_\ell P_\ell}{\beta_\ell P_\ell + \Lambda_{k,\ell\ell}^2}\right) - \frac{1}{2}\log\left(1 + \frac{\bar{\beta}_\ell P_\ell}{\beta_\ell P_\ell + \Lambda_{Z,\ell\ell}}\right) \tag{45}$$

where the union is over all $\{P_\ell\}_{\ell=1}^{L}$ such that $\sum_{\ell=1}^{L} P_\ell = P$ and $\bar{\beta}_\ell = 1 - \beta_\ell \in [0,1]$, $\ell = 1,\ldots,L$.

The proof of this theorem is provided in Appendix IV. Here, $P_\ell$ denotes the part of the total available power $P$ which is devoted to the transmission in the $\ell$th subchannel. Furthermore, $\beta_\ell$ denotes the fraction of the power $P_\ell$ of the $\ell$th subchannel spent for the transmission to users in the first group.

*3) Gaussian MIMO DCMRWC:* In this section, we first obtain the secrecy capacity region of the Gaussian MIMO DCMRWC when $K_2 = 1$. To that end, we need to evaluate the region given in Theorem 1. In other words, we need to find the optimal random variable pair $(U, \mathbf{X})$. We are able to do this when there is only one user in the second group, i.e., $K_2 = 1$. For this, we need the following theorem.

*Theorem 5:* Let $(\mathbf{N}_1, \mathbf{N}^*, \mathbf{N}_Z)$ be zero-mean Gaussian random vectors with covariance matrices $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}^*, \boldsymbol{\Sigma}_Z$, respectively, where

$$\boldsymbol{\Sigma}_1 \preceq \boldsymbol{\Sigma}^* \preceq \boldsymbol{\Sigma}_Z. \tag{46}$$

Let $(U, \mathbf{X})$ be arbitrarily dependent random vector, which is independent of $(\mathbf{N}_1, \mathbf{N}^*, \mathbf{N}_Z)$, and let the second moment of $\mathbf{X}$ be constrained as $E\left[\mathbf{XX}^\top\right] \preceq \mathbf{S}$. Then, for any feasible $(U, \mathbf{X})$, we can find a positive semidefinite matrix $\mathbf{K}^*$ such that $\mathbf{K}^* \preceq \mathbf{S}$, and it satisfies

$$h(\mathbf{X} + \mathbf{N}_Z|U) - h(\mathbf{X} + \mathbf{N}^*|U) = \frac{1}{2}\log\frac{|\mathbf{K}^* + \boldsymbol{\Sigma}_Z|}{|\mathbf{K}^* + \boldsymbol{\Sigma}^*|} \tag{47}$$

and

$$h(\mathbf{X} + \mathbf{N}_Z|U) - h(\mathbf{X} + \mathbf{N}_1|U) \geq \frac{1}{2}\log\frac{|\mathbf{K}^* + \boldsymbol{\Sigma}_Z|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_1|} \tag{48}$$

for any $\boldsymbol{\Sigma}_1$ satisfying the order in (46).

The proof of this theorem can be found in [3]. Using this theorem, we can establish the secrecy capacity region of the Gaussian MIMO DCMRWC when $K_2 = 1$ as follows.

*Theorem 6:* The secrecy capacity region of the Gaussian MIMO DCMRWC when $K_2 = 1$ is given by the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{j=1,\ldots,K_1} \frac{1}{2}\log\frac{|\mathbf{K} + \boldsymbol{\Sigma}_j^1|}{|\boldsymbol{\Sigma}_j^1|} - \frac{1}{2}\log\frac{|\mathbf{K} + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|} \tag{49}$$

$$R_2 \leq \frac{1}{2}\log\frac{|\mathbf{S} + \boldsymbol{\Sigma}^2|}{|\mathbf{K} + \boldsymbol{\Sigma}^2|} - \frac{1}{2}\log\frac{|\mathbf{S} + \boldsymbol{\Sigma}_Z|}{|\mathbf{K} + \boldsymbol{\Sigma}_Z|} \tag{50}$$

where we dropped the subscript of $\boldsymbol{\Sigma}_k^2$ since $K_2 = 1$, and the union is over all positive semidefinite matrices $\mathbf{K}$ such that $\mathbf{K} \preceq \mathbf{S}$.

The proof of this theorem is given in Appendix V. Now, we would like to comment about why we can obtain the secrecy capacity region of the Gaussian MIMO DCMRWC only when $K_2 = 1$. The reason is that we can extend Theorem 3, which was used to obtain the secrecy capacity region of the Gaussian parallel DCMRWC, to vector case in Theorem 5 partially, i.e., not completely. In particular, we could not show that the matrix $\mathbf{K}^*$ in Theorem 5 also satisfies

$$h(\mathbf{X} + \mathbf{N}_Z|U) - h(\mathbf{X} + \mathbf{N}_2|U) \leq \frac{1}{2}\log\frac{|\mathbf{K}^* + \boldsymbol{\Sigma}_Z|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_2|} \tag{51}$$

for any Gaussian random vector $\mathbf{N}_2$ with covariance matrix satisfying $\boldsymbol{\Sigma}^* \preceq \boldsymbol{\Sigma}_2 \preceq \boldsymbol{\Sigma}_Z$. If (51) can be shown, the secrecy capacity region of the Gaussian MIMO DCMRWC can be obtained as the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{j=1,\ldots,K_1} \frac{1}{2}\log\frac{|\mathbf{K} + \boldsymbol{\Sigma}_j^1|}{|\boldsymbol{\Sigma}_j^1|} - \frac{1}{2}\log\frac{|\mathbf{K} + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|} \tag{52}$$

$$R_2 \leq \min_{k=1,\ldots,K_2} \frac{1}{2}\log\frac{|\mathbf{S} + \boldsymbol{\Sigma}_k^2|}{|\mathbf{K} + \boldsymbol{\Sigma}_k^2|} - \frac{1}{2}\log\frac{|\mathbf{S} + \boldsymbol{\Sigma}_Z|}{|\mathbf{K} + \boldsymbol{\Sigma}_Z|} \tag{53}$$

where the union is over all positive semidefinite matrices $\mathbf{K}$ such that $\mathbf{K} \preceq \mathbf{S}$.

Finally we note that using the analysis carried out in [3, Sec. 7.1] and [16, Sec. V], the capacity result given in Theorem 6 can be extended to the general Gaussian MIMO DCMRWC defined by (25) – (27) as follows.

*Corollary 1:* The secrecy capacity region of the general Gaussian MIMO DCMRWC, which is defined by (25) – (27), when $K_2 = 1$, is given by the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{j=1,\ldots,K_1} \frac{1}{2}\log\frac{|\mathbf{H}_j^1\mathbf{K}(\mathbf{H}_j^1)^\top + \boldsymbol{\Sigma}_j^1|}{|\boldsymbol{\Sigma}_j^1|} - \frac{1}{2}\log\frac{|\mathbf{H}_Z\mathbf{K}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|} \tag{54}$$

$$R_2 \leq \frac{1}{2}\log\frac{|\mathbf{H}^2\mathbf{S}(\mathbf{H}^2)^\top + \boldsymbol{\Sigma}^2|}{|\mathbf{H}^2\mathbf{K}(\mathbf{H}^2)^\top + \boldsymbol{\Sigma}^2|} - \frac{1}{2}\log\frac{|\mathbf{H}_Z\mathbf{S}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|}{|\mathbf{H}_Z\mathbf{K}\mathbf{H}_Z^\top + \boldsymbol{\Sigma}_Z|} \tag{55}$$

where we dropped the subscripts of $\boldsymbol{\Sigma}_k^2$, $\mathbf{H}_k^2$ since $K_2 = 1$, and the union is over all positive semidefinite matrices $\mathbf{K}$ such that $\mathbf{K} \preceq \mathbf{S}$.

### B. Second Scenario: Layered Confidential Messages

In the second scenario, the transmitter wants to send a confidential message to users in the first group which needs to be kept confidential from the second group of users and eavesdroppers. The transmitter also wants to send a different confidential message to users in the second group, which needs to be kept confidential from the eavesdroppers. As opposed to the first scenario, in this case, we do not put any restriction on the number of eavesdroppers. The graphical illustration of the second scenario is given in Fig. 3. The situation where there is only one user

in each group and one eavesdropper was investigated in [18]. Hence, this second scenario can be seen as a generalization of the model in [18] to a compound channel setting. Following the terminology of [18], we call this channel model DCMRWC with *layered messages*.

An $(n, 2^{nR_1}, 2^{nR_2})$ code for DCMRWC with *layered messages* consists of two message sets $\mathcal{W}_1 = \{1, \ldots, 2^{nR_1}\}, \mathcal{W}_2 = \{1, \ldots, 2^{nR_2}\}$ and an encoder $f : \mathcal{W}_1 \times \mathcal{W}_2 \to \mathcal{X}^n$, one decoder for each legitimate user in the first group $g_j^1 : \mathcal{Y}_j^{1,n} \to \mathcal{W}_1, \ j = 1, \ldots, K_1$, and one decoder for each legitimate user in the second group $g_k^2 : \mathcal{Y}_k^{2,n} \to \mathcal{W}_2, \ k = 1, \ldots, K_2$. The probability of error is defined as $P_e^n = \max\{P_e^{1,n}, P_e^{2,n}\}$ where $P_e^{1,n}$ and $P_e^{2,n}$ are given by

$$P_e^{1,n} = \max_{j \in \{1,\ldots,K_1\}} \Pr\left[g_j^1(Y_j^{1,n}) \neq W_1\right] \quad (56)$$

$$P_e^{2,n} = \max_{k \in \{1,\ldots,K_2\}} \Pr\left[g_k^2(Y_k^{2,n}) \neq W_2\right]. \quad (57)$$

A secrecy rate pair is said to be achievable if there exists an $(n, 2^{nR_1}, 2^{nR_2})$ code which has $\lim_{n \to \infty} P_e^n = 0$

$$\lim_{n \to \infty} \frac{1}{n} I(W_2; Z_t^n) = 0, \quad t = 1, \ldots, K_Z \quad (58)$$

and

$$\lim_{n \to \infty} \frac{1}{n} I(W_1; Y_k^{2,n} | W_2) = 0, \quad k = 1, \ldots, K_2. \quad (59)$$

We note that these two secrecy conditions imply

$$\lim_{n \to \infty} \frac{1}{n} I(W_1, W_2; Z_t^n) = 0, \quad t = 1, \ldots, K_Z. \quad (60)$$

Furthermore, it is clear that we are only interested in perfect secrecy rates of the channel. The secrecy capacity region is defined as the closure of all achievable secrecy rate pairs. A single-letter characterization of the secrecy capacity region is given as follows.

*Theorem 7:* The secrecy capacity region of the DCMRWC with layered messages is given by the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{\substack{j=1,\ldots,K_1 \\ k=1,\ldots,K_2}} I(X; Y_j^1 | U, Y_k^2) \quad (61)$$

$$R_2 \leq \min_{\substack{k=1,\ldots,K_2 \\ t=1,\ldots,K_Z}} I(U; Y_k^2 | Z_t) \quad (62)$$

where the union is over all random variable pairs $(U, X)$ such that

$$U \to X \to Y_j^1 \to Y^* \to Y_k^2 \to Z^* \to Z_t \quad (63)$$

for any triple $(j, k, t)$.

The proof of this theorem is given in Appendix VI. Similar to the converse proof of Theorem 1, the presence of the fictitious users $Y^*$ and $Z^*$ plays an important role here as well. In particular, these two random variables introduce a conditional independence structure to the channel which enables us to define the auxiliary random variable $U$ that yields a tight outer

bound. Despite this similarity in the role of fictitious users in converse proofs, there is a significant difference between Theorems 1 and 7; in particular, it does not seem to be possible to extend Theorem 1 to an arbitrary number of eavesdroppers, while Theorem 7 holds for any number of eavesdroppers. This is due to the difference of two communication scenarios. In the second scenario, since we assume that users in the second group as well as the eavesdroppers wiretap users in the first group, we are able to provide a converse proof for the general situation of arbitrary number of eavesdroppers.

As an aside, if we set $K_1 = K_2 = K_Z = 1$, then DCMRWC with layered messages reduces to the degraded multi-receiver wiretap channel with layered messages of [18], the secrecy capacity region in Theorem 7 reduces to the secrecy capacity region of the channel model in [18].

*1) Parallel DCMRWC With Layered Messages:* In the next section, we investigate the Gaussian parallel DCMRWC with layered messages. To that end, here we obtain the secrecy capacity region of the parallel DCMRWC with layered messages in a single-letter form as follows.

*Theorem 8:* The secrecy capacity region of the parallel DCMRWC with layered messages is given by the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{\substack{j=1,\ldots,K_1 \\ k=1,\ldots,K_2}} \sum_{\ell=1}^{L} I(X_\ell; Y_{j\ell}^1 | U_\ell, Y_{k\ell}^2) \quad (64)$$

$$R_2 \leq \min_{\substack{k=1,\ldots,K_2 \\ t=1,\ldots,K_Z}} \sum_{\ell=1}^{L} I(U_\ell; Y_{k\ell}^2 | Z_{t\ell}) \quad (65)$$

where the union is over all $\prod_{\ell=1}^{L} p(u_\ell, x_\ell)$ such that

$$U_\ell \to X_\ell \to Y_{j\ell}^1 \to Y_\ell^* \to Y_{k\ell}^2 \to Z_\ell^* \to Z_{t\ell} \quad (66)$$

for any $(\ell, j, k, t)$.

Since the parallel DCMRWC with layered messages is a special case of the DCMRWC with layered messages, Theorem 7 implicitly gives the secrecy capacity region of the parallel DCMRWC with layered messages. However, we still need to show that the region in Theorem 7 is equivalent to the region in Theorem 8. That is, we need to prove the optimality of independent signalling in each subchannel. The proof of Theorem 8 is provided in Appendix VII.

*2) Gaussian Parallel DCMRWC With Layered Messages:* We now obtain the secrecy capacity region of the Gaussian parallel DCMRWC with layered messages. To that end, we need to evaluate the region given in Theorem 8, i.e., we need to find the optimal distribution $\prod_{\ell=1}^{L} p(u_\ell, x_\ell)$. We first introduce the following theorem, which is an extension of Theorem 3.

*Theorem 9:* Let $N_1, N^*, N_2, \tilde{N}, N_Z$ be zero-mean Gaussian random variables with variances $\sigma_1^2, \sigma_*^2, \sigma_2^2, \tilde{\sigma}^2, \sigma_Z^2$, respectively, where

$$\sigma_1^2 \leq \sigma_*^2 \leq \sigma_2^2 \leq \tilde{\sigma}^2 \leq \sigma_Z^2. \quad (67)$$

Let $(U, X)$ be an arbitrarily dependent random variable pair, which is independent of $(N_1, N^*, N_2, \tilde{N}, N_Z)$, and the second

moment of $X$ be constrained as $E\left[X^2\right] \leq P$. Then, for any feasible $(U, X)$, we can find a $P^* \leq P$ such that

$$h(X + \tilde{N}|U) - h(X + N^*|U) = \frac{1}{2}\log\frac{P^* + \tilde{\sigma}^2}{P^* + \sigma_*^2} \quad (68)$$

and

$$h(X + N_Z|U) - h(X + N_2|U) \leq \frac{1}{2}\log\frac{P^* + \sigma_Z^2}{P^* + \sigma_2^2} \quad (69)$$

$$h(X + N_2|U) - h(X + N_1|U) \geq \frac{1}{2}\log\frac{P^* + \sigma_2^2}{P^* + \sigma_1^2} \quad (70)$$

for any $(\sigma_1^2, \sigma_2^2, \sigma_Z^2)$ satisfying the order in (67).

The proof of this theorem is given in Appendix VIII. The proof of this theorem basically relies on Theorem 3 and Costa's entropy power inequality [17].

Using this theorem, we can establish the secrecy capacity region of the Gaussian parallel DCMRWC with layered messages as follows.

*Theorem 10:* The secrecy capacity region of the Gaussian parallel DCMRWC with layered messages is given by the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{\substack{j=1,\dots,K_1 \\ k=1,\dots,K_2}} \sum_{\ell=1}^{L} \frac{1}{2}\log\left(1 + \frac{\beta_\ell P_\ell}{\Lambda_{j,\ell\ell}^1}\right) - \frac{1}{2}\log\left(1 + \frac{\beta_\ell P_\ell}{\Lambda_{k,\ell\ell}^2}\right) \quad (71)$$

$$R_2 \leq \min_{\substack{k=1,\dots,K_2 \\ t=1,\dots,K_Z}} \sum_{\ell=1}^{L} \frac{1}{2}\log\left(1 + \frac{\bar{\beta}_\ell P_\ell}{\beta_\ell P_\ell + \Lambda_{k,\ell\ell}^2}\right)$$
$$- \frac{1}{2}\log\left(1 + \frac{\bar{\beta}_\ell P_\ell}{\beta_\ell P_\ell + \Lambda_{t,\ell\ell}^Z}\right) \quad (72)$$

where $\bar{\beta}_\ell = 1 - \beta_\ell \in [0, 1]$, $\ell = 1, \dots, L$, and the union is over all $\{P_\ell\}_{\ell=1}^{L}$ such that $\sum_{\ell=1}^{L} P_\ell = P$.

The proof of this theorem is given in Appendix IX. Similar to Theorem 4, here also, $P_\ell$ denotes the amount of power $P$ devoted to the transmission in the $\ell$th subchannel. Similarly, $\beta_\ell$ is the fraction of the power $P_\ell$ of the $\ell$th subchannel spent for the transmission to users in the first group.

*3) Gaussian MIMO DCMRWC With Layered Messages:* We now obtain the secrecy capacity region of the Gaussian MIMO DCMRWC with layered messages. To that end, we need to evaluate the region given in Theorem 7, i.e., find the optimal random vector pair $(U, \mathbf{X})$. We are able to find the optimal random vector pair $(U, \mathbf{X})$ when there is only one user in the second group, i.e., $K_2 = 1$. To obtain that result, we first need the following generalization of Theorem 5.

*Theorem 11:* Let $(\mathbf{N}_1, \mathbf{N}_2, \mathbf{N}^*, \mathbf{N}_Z)$ be Gaussian random vectors with covariance matrices $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$, $\boldsymbol{\Sigma}^*$, $\boldsymbol{\Sigma}_Z$, respectively, where

$$\boldsymbol{\Sigma}_1 \preceq \boldsymbol{\Sigma}_2 \preceq \boldsymbol{\Sigma}^* \preceq \boldsymbol{\Sigma}_Z. \quad (73)$$

Let $(U, \mathbf{X})$ be an arbitrarily dependent random vector pair, which is independent of $(\mathbf{N}_1, \mathbf{N}_2, \mathbf{N}^*, \mathbf{N}_Z)$, and the second moment of $\mathbf{X}$ be constrained as $E\left[\mathbf{X}\mathbf{X}^\top\right] \preceq \mathbf{S}$. Then, for any

feasible $(U, \mathbf{X})$, there exists a positive semidefinite matrix $\mathbf{K}^*$ such that $\mathbf{K}^* \preceq \mathbf{S}$, and it satisfies

$$h(\mathbf{X} + \mathbf{N}^*|U) - h(\mathbf{X} + \mathbf{N}_2|U) = \frac{1}{2}\log\frac{|\mathbf{K}^* + \boldsymbol{\Sigma}^*|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_2|} \quad (74)$$

and

$$h(\mathbf{X} + \mathbf{N}_Z|U) - h(\mathbf{X} + \mathbf{N}_2|U) \leq \frac{1}{2}\log\frac{|\mathbf{K}^* + \boldsymbol{\Sigma}_Z|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_2|} \quad (75)$$

$$h(\mathbf{X} + \mathbf{N}_2|U) - h(\mathbf{X} + \mathbf{N}_1|U) \geq \frac{1}{2}\log\frac{|\mathbf{K}^* + \boldsymbol{\Sigma}_2|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_1|} \quad (76)$$

for any $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_Z)$ satisfying the order in (73).

The proof of this theorem is given in Appendix X. Using this theorem, we can find the secrecy capacity region of the Gaussian MIMO DCMRWC with layered messages when $K_2 = 1$ as follows.

*Theorem 12:* The secrecy capacity region of the Gaussian MIMO DCMRWC with layered messages when $K_2 = 1$ is given by the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{j=1,\dots,K_1} \frac{1}{2}\log\frac{|\mathbf{K} + \boldsymbol{\Sigma}_j^1|}{|\boldsymbol{\Sigma}_j^1|} - \frac{1}{2}\log\frac{|\mathbf{K} + \boldsymbol{\Sigma}^2|}{|\boldsymbol{\Sigma}^2|} \quad (77)$$

$$R_2 \leq \min_{t=1,\dots,K_Z} \frac{1}{2}\log\frac{|\mathbf{S} + \boldsymbol{\Sigma}^2|}{|\mathbf{K} + \boldsymbol{\Sigma}^2|} - \frac{1}{2}\log\frac{|\mathbf{S} + \boldsymbol{\Sigma}_t^Z|}{|\mathbf{K} + \boldsymbol{\Sigma}_t^Z|} \quad (78)$$

where the union is over all positive semidefinite matrices $\mathbf{K}$ such that $\mathbf{K} \preceq \mathbf{S}$.

The proof of this theorem is given in Appendix XI. As an aside, if we set $K_1 = K_Z = 1$ in this theorem, we can recover the secrecy capacity region of the degraded multi-receiver wiretap channel with layered messages that was established in [18].

Finally we note that using the analysis carried out in [3, Sec. 7.1] and [16, Sec. V], the capacity result given in Theorem 12 can be extended to the general Gaussian MIMO DCMRWC defined by (25) – (27) as follows.

*Corollary 2:* The secrecy capacity region of the general Gaussian MIMO DCMRWC, defined by (25) – (27), with layered messages when $K_2 = 1$ is given by the union of rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min_{j=1,\dots,K_1} \frac{1}{2}\log\frac{|\mathbf{H}_j^1\mathbf{K}(\mathbf{H}_j^1)^\top + \boldsymbol{\Sigma}_j^1|}{|\boldsymbol{\Sigma}_j^1|}$$
$$- \frac{1}{2}\log\frac{|\mathbf{H}^2\mathbf{K}(\mathbf{H}^2)^\top + \boldsymbol{\Sigma}^2|}{|\boldsymbol{\Sigma}^2|} \quad (79)$$

$$R_2 \leq \min_{t=1,\dots,K_Z} \frac{1}{2}\log\frac{|\mathbf{H}^2\mathbf{S}(\mathbf{H}^2)^\top + \boldsymbol{\Sigma}^2|}{|\mathbf{H}^2\mathbf{K}(\mathbf{H}^2)^\top + \boldsymbol{\Sigma}^2|}$$
$$- \frac{1}{2}\log\frac{|\mathbf{H}_t^Z\mathbf{S}(\mathbf{H}_t^Z)^\top + \boldsymbol{\Sigma}_t^Z|}{|\mathbf{H}_t^Z\mathbf{K}(\mathbf{H}_t^Z)^\top + \boldsymbol{\Sigma}_t^Z|} \quad (80)$$

where the union is over all positive semidefinite matrices $\mathbf{K}$ such that $\mathbf{K} \preceq \mathbf{S}$.

## IV. CONCLUSION

In this paper, we studied two different communication scenarios for the DCMRWC. In the first scenario, the transmitter wants to send a confidential message to users in the first group, and a different confidential message to users in the second

group, where both messages are to be kept confidential from an eavesdropper. We establish the secrecy capacity region of the general discrete memoryless channel model, the parallel channel model, and the Gaussian parallel channel model. For the Gaussian MIMO channel model, we obtain the secrecy capacity region when there is only one user in the second group.

In the second scenario we study, the transmitter sends a confidential message to users in the first group which is wiretapped by both users in the second group and eavesdroppers. In addition to this message sent to the first group of users, the transmitter sends a different message to users in the second group which needs to be kept confidential only from the eavesdroppers. In this case, we do not put any restriction on the number of eavesdroppers. As in the first scenario, we establish the secrecy capacity region for the general discrete memoryless channel model, the parallel channel model, and the Gaussian parallel channel model. For the Gaussian MIMO channel model, we obtain the secrecy capacity region when there is only one user in the second group.

## APPENDIX I
### PROOF OF THEOREM 1

Achievability is clear. We provide the converse proof. For an arbitrary code achieving the secrecy rates $(R_1, R_2)$, there exist $(\epsilon_{1,n}, \epsilon_{2,n})$ and $\gamma_n$ which vanish as $n \to \infty$ such that

$$H(W_1|Y_j^{1,n}) \leq n\epsilon_{1,n}, \quad j = 1, \ldots, K_1 \tag{81}$$

$$H(W_2|Y_k^{2,n}) \leq n\epsilon_{2,n}, \quad k = 1, \ldots, K_2 \tag{82}$$

$$I(W_1, W_2; Z^n) \leq n\gamma_n \tag{83}$$

where (81) and (82) are due to Fano's lemma, and (83) is due to the perfect secrecy requirement stated in (32).

We define the following auxiliary random variables:

$$U_i = W_2 Y^{*,i-1} Z_{i+1}^n, \quad i = 1, \ldots, n \tag{84}$$

which satisfy the following Markov chain:

$$U_i \to X_i \to Y_{j,i}^1 \to Y_i^* \to Y_{k,i}^2 \to Z_i, \quad i = 1, \ldots, n \tag{85}$$

for any $(j, k)$ pair. The Markov chain in (85) is a consequence of the fact that the channel is memoryless and degraded.

We first bound the rate of the second message

$$nR_2 = H(W_2) \tag{86}$$

$$\leq I(W_2; Y_k^{2,n}) + n\epsilon_{2,n} \tag{87}$$

$$\leq I(W_2; Y_k^{2,n}) - I(W_2; Z^n) + n(\epsilon_{2,n} + \gamma_n) \tag{88}$$

$$= I(W_2; Y_k^{2,n}|Z^n) + n(\epsilon_{2,n} + \gamma_n) \tag{89}$$

$$= \sum_{i=1}^n I(W_2; Y_{k,i}^2|Y_k^{2,i-1}, Z^n) + n(\epsilon_{2,n} + \gamma_n) \tag{90}$$

$$= \sum_{i=1}^n I(W_2; Y_{k,i}^2|Y_k^{2,i-1}, Z_{i+1}^n, Z_i) + n(\epsilon_{2,n} + \gamma_n) \tag{91}$$

$$\leq \sum_{i=1}^n I(Y_k^{2,i-1}, Z_{i+1}^n, W_2; Y_{k,i}^2|Z_i) + n(\epsilon_{2,n} + \gamma_n) \tag{92}$$

$$\leq \sum_{i=1}^n I(Y^{*,i-1}, Y_k^{2,i-1}, Z_{i+1}^n, W_2; Y_{k,i}^2|Z_i) + n(\epsilon_{2,n} + \gamma_n) \tag{93}$$

$$= \sum_{i=1}^n I(Y^{*,i-1}, Z_{i+1}^n, W_2; Y_{k,i}^2|Z_i) + n(\epsilon_{2,n} + \gamma_n) \tag{94}$$

$$= \sum_{i=1}^n I(U_i; Y_{k,i}^2|Z_i) + n(\epsilon_{2,n} + \gamma_n) \tag{95}$$

where (87) is due to (82), (88) is a consequence of (83), (89) comes from the Markov chain

$$W_2 \to Y_k^{2,n} \to Z^n, \quad k = 1, \ldots, K_2 \tag{96}$$

which is a consequence of the fact that the channel is degraded, (91) comes from the Markov chain

$$Z^{i-1} \to Y_k^{2,i-1} \to (Y_{k,i}^2, Z_i^n, W_2), \quad k = 1, \ldots, K_2 \tag{97}$$

which is due to the fact that the channel is degraded and memoryless, and (94) is a consequence of the Markov chain

$$Y_k^{2,i-1} \to Y^{*,i-1} \to (W_2, Z_i^n, Y_{k,i}^2), \quad k = 1, \ldots, K_2 \tag{98}$$

which is due to the Markov chain in (2) and the fact that the channel is memoryless.

Next we bound the rate of the first message

$$nR_1 = H(W_1) \tag{99}$$

$$= H(W_1|W_2) \tag{100}$$

$$\leq I(W_1; Y_j^{1,n}|W_2) + n\epsilon_{1,n} \tag{101}$$

$$\leq I(W_1; Y_j^{1,n}|W_2) - I(W_1; Z^n|W_2) + n(\epsilon_{1,n} + \gamma_n) \tag{102}$$

$$= I(W_1; Y_j^{1,n}|W_2, Z^n) + n(\epsilon_{1,n} + \gamma_n) \tag{103}$$

$$= \sum_{i=1}^n I(W_1; Y_{j,i}^1|W_2, Z^n, Y_j^{1,i-1}) + n(\epsilon_{1,n} + \gamma_n) \tag{104}$$

$$= \sum_{i=1}^n I(W_1; Y_{j,i}^1|W_2, Z_{i+1}^n, Y_j^{1,i-1}, Z_i) + n(\epsilon_{1,n} + \gamma_n) \tag{105}$$

$$= \sum_{i=1}^n I(W_1; Y_{j,i}^1|W_2, Z_{i+1}^n, Y_j^{1,i-1}, Y^{*,i-1}, Z_i) + n(\epsilon_{1,n} + \gamma_n) \tag{106}$$

$$\leq \sum_{i=1}^n I(X_i, W_1; Y_{j,i}^1|W_2, Z_{i+1}^n, Y_j^{1,i-1}, Y^{*,i-1}, Z_i) + n(\epsilon_{1,n} + \gamma_n) \tag{107}$$

$$= \sum_{i=1}^n I(X_i; Y_{j,i}^1|W_2, Z_{i+1}^n, Y_j^{1,i-1}, Y^{*,i-1}, Z_i) + n(\epsilon_{1,n} + \gamma_n) \tag{108}$$

$$= \sum_{i=1}^n H(Y_{j,i}^1|W_2, Z_{i+1}^n, Y_j^{1,i-1}, Y^{*,i-1}, Z_i) - H(Y_{j,i}^1|W_2, Z_{i+1}^n, Y_j^{1,i-1}, Y^{*,i-1}, Z_i, X_i) + n(\epsilon_{1,n} + \gamma_n) \tag{109}$$

$$\leq \sum_{i=1}^{n} H(Y_{j,i}^1 | W_2, Z_{i+1}^n, Y^{*,i-1}, Z_i)$$
$$- H(Y_{j,i}^1 | W_2, Z_{i+1}^n, Y_j^{1,i-1}, Y^{*,i-1}, Z_i, X_i)$$
$$+ n(\epsilon_{1,n} + \gamma_n) \tag{110}$$

$$= \sum_{i=1}^{n} H(Y_{j,i}^1 | W_2, Z_{i+1}^n, Y^{*,i-1}, Z_i)$$
$$- H(Y_{j,i}^1 | W_2, Z_{i+1}^n, Y^{*,i-1}, Z_i, X_i) + n(\epsilon_{1,n} + \gamma_n) \tag{111}$$

$$= \sum_{i=1}^{n} I(X_i; Y_{j,i}^1 | W_2, Z_{i+1}^n, Y^{*,i-1}, Z_i) + n(\epsilon_{1,n} + \gamma_n) \tag{112}$$

$$= \sum_{i=1}^{n} I(X_i; Y_{j,i}^1 | U_i, Z_i) + n(\epsilon_{1,n} + \gamma_n) \tag{113}$$

where (101) is due to (81), (102) is a consequence of (83), (103) comes from the Markov chain

$$(W_2, W_1) \to Y_j^{1,n} \to Z^n, \quad j = 1, \dots, K_1 \tag{114}$$

which is due to the fact that the channel is degraded, (105) comes from the Markov chain

$$Z^{i-1} \to Y_j^{1,i-1} \to (W_1, W_2, Y_{j,i}^1, Z_i^n), \quad j = 1, \dots, K_1 \tag{115}$$

which is a consequence of the fact that the channel is degraded and memoryless, (106) follows from the Markov chain

$$Y^{*,i-1} \to Y_j^{1,i-1} \to (W_1, W_2, Y_{j,i}^1, Z_i^n) \tag{116}$$

for all $j = 1, \dots, K_1$, which results from the Markov chain in (2) and the fact that the channel is memoryless, (108) is a consequence of the Markov chain

$$(Y_{j,i}^1, Z_i) \to X_i \to (Y^{*,i-1}, Y_j^{1,i-1}, Z_{i+1}^n, W_1, W_2) \tag{117}$$

for all $j = 1, \dots, K_1$, which is due to the fact that the channel is memoryless, (110) comes from the fact that conditioning cannot increase entropy, and (111) is again due to the Markov chain in (117).

Next, we define a uniformly distributed random variable $Q \in \{1, \dots, n\}$, and $U = (Q, U_Q)$, $X = X_Q$, $Y_j^1 = Y_{j,Q}^1$, $Y_k^2 = Y_{k,Q}^2$, and $Z = Z_Q$. Using these definitions in (95) and (113), we obtain the single-letter expressions in Theorem 1.

## APPENDIX II
## PROOF OF THEOREM 2

The achievability of this region follows from Theorem 1 by selecting $(U, X) = (U_1, X_1, \dots, U_L, X_L)$ with a joint distribution of the product form $p(u, x) = \prod_{\ell=1}^{L} p(u_\ell, x_\ell)$. We next provide the converse proof. To that end, we define the following auxiliary random variables:

$$U_\ell = U Y_{[1:\ell-1]}^* Z_{[\ell+1:L]}, \quad \ell = 1, \dots, L \tag{118}$$

which satisfy the Markov chain

$$U_\ell \to X_\ell \to (Y_{j\ell}^1, Y_{k\ell}^2, Z_\ell) \tag{119}$$

for any $(j, k, \ell)$ triple because of the facts that the channel is memoryless and subchannels are independent.

We bound the rate of the second message as follows:

$$R_2 \leq \min_{k=1,\dots,K_2} I(U; Y_{k[1:L]}^2 | Z_{[1:L]}) \tag{120}$$

$$= \min_{k=1,\dots,K_2} \sum_{\ell=1}^{L} I(U; Y_{k\ell}^2 | Y_{k[1:\ell-1]}^2, Z_{[1:L]}) \tag{121}$$

$$= \min_{k=1,\dots,K_2} \sum_{\ell=1}^{L} I(U; Y_{k\ell}^2 | Y_{k[1:\ell-1]}^2, Z_{[\ell:L]}) \tag{122}$$

$$\leq \min_{k=1,\dots,K_2} \sum_{\ell=1}^{L} I(U, Y_{k[1:\ell-1]}^2, Z_{[\ell+1:L]}; Y_{k\ell}^2 | Z_\ell) \tag{123}$$

$$\leq \min_{k=1,\dots,K_2} \sum_{\ell=1}^{L} I(U, Y_{[1:\ell-1]}^*, Y_{k[1:\ell-1]}^2, Z_{[\ell+1:L]}; Y_{k\ell}^2 | Z_\ell) \tag{124}$$

$$= \min_{k=1,\dots,K_2} \sum_{\ell=1}^{L} I(U, Y_{[1:\ell-1]}^*, Z_{[\ell+1:L]}; Y_{k\ell}^2 | Z_\ell) \tag{125}$$

$$= \min_{k=1,\dots,K_2} \sum_{\ell=1}^{L} I(U_\ell; Y_{k\ell}^2 | Z_\ell) \tag{126}$$

where (122) follows from the Markov chain

$$Z_{[1:\ell-1]} \to Y_{k[1:\ell-1]}^2 \to (U, Y_{k\ell}^2, Z_{[\ell:L]}) \tag{127}$$

which is a consequence of the facts that the channel is degraded and memoryless, and subchannels are independent, and (125) is due to the Markov chain

$$Y_{k[1:\ell-1]}^2 \to Y_{[1:\ell-1]}^* \to U, Y_{k\ell}^2, Z_{[\ell:L]} \tag{128}$$

which is a consequence of the Markov chain in (10) and the facts that the channel is memoryless and subchannels are independent.

We next bound the rate of the first message as follows:

$$R_1 \leq \min_{j=1,\dots,K_1} I(X_{[1:L]}; Y_{j[1:L]}^1 | U, Z_{[1:L]}) \tag{129}$$

$$= \min_{j=1,\dots,K_1} \sum_{\ell=1}^{L} I(X_{[1:L]}; Y_{j\ell}^1 | U, Y_{j[1:\ell-1]}^1, Z_{[1:L]}) \tag{130}$$

$$= \min_{j=1,\dots,K_1} \sum_{\ell=1}^{L} I(X_{[1:L]}; Y_{j\ell}^1 | U, Y_{j[1:\ell-1]}^1, Z_{[\ell:L]}) \tag{131}$$

$$= \min_{j=1,\dots,K_1} \sum_{\ell=1}^{L} I(X_{[1:L]}; Y_{j\ell}^1 | U, Y_{j[1:\ell-1]}^1, Y_{[1:\ell-1]}^*, Z_{[\ell:L]}) \tag{132}$$

$$= \min_{j=1,\dots,K_1} \sum_{\ell=1}^{L} I(X_{[1:L]}; Y_{j\ell}^1 | U_\ell, Y_{j[1:\ell-1]}^1, Z_\ell) \tag{133}$$

$$= \min_{j=1,\dots,K_1} \sum_{\ell=1}^{L} H(Y_{j\ell}^1 | U_\ell, Y_{j[1:\ell-1]}^1, Z_\ell)$$
$$- H(Y_{j\ell}^1 | U_\ell, Y_{j[1:\ell-1]}^1, Z_\ell, X_{[1:L]}) \tag{134}$$

$$\leq \min_{j=1,\ldots,K_1} \sum_{\ell=1}^{L} H(Y^1_{j\ell}|U_\ell, Z_\ell)$$
$$- H(Y^1_{j\ell}|U_\ell, Y^1_{j[1:\ell-1]}, Z_\ell, X_{[1:L]}) \quad (135)$$

$$= \min_{j=1,\ldots,K_1} \sum_{\ell=1}^{L} H(Y^1_{j\ell}|U_\ell, Z_\ell) - H(Y^1_{j\ell}|U_\ell, Z_\ell, X_\ell) (136)$$

$$= \min_{j=1,\ldots,K_1} \sum_{\ell=1}^{L} I(X_\ell; Y^1_{j\ell}|U_\ell, Z_\ell) \quad (137)$$

where (131) and (132) follow from the Markov chain

$$Z_{[1:\ell-1]} \rightarrow Y^*_{[1:\ell-1]} \rightarrow Y^1_{j[1:\ell-1]} \rightarrow (U, Y^1_{j\ell}, Z_{[\ell:L]}, X_{[1:L]})$$
$$(138)$$

which is due to the facts that the channel is degraded and memoryless, subchannels are independent, and the Markov chain in (10), (135) results from the fact that conditioning cannot increase entropy, (136) comes from the Markov chain

$$(Y^1_{j\ell}, Z_\ell) \rightarrow X_\ell \rightarrow (U_\ell, Y^1_{j[1:\ell-1]}, X_{[1:\ell-1]}, X_{[\ell+1:L]}) \quad (139)$$

which is a consequence of the facts that the channel is memoryless, and subchannels are independent.

In view of (126) and (137), we obtain the single-letter expressions in Theorem 2. Finally, we note that each expression in the bounds given by (126) and (137) depend on the joint distribution $p(u_{[1:L]}, x_{[1:L]})$ through its marginals $p(u_\ell, x_\ell)$. Thus, there is no loss of optimality to choose $p(u_{[1:L]}, x_{[1:L]}) = \prod_{\ell=1}^{L} p(u_\ell, x_\ell)$. This completes the converse proof.

## APPENDIX III
## PROOF OF THEOREM 3

We first note that

$$\frac{1}{2} \log \frac{\sigma^2_*}{\sigma^2_Z} \leq h(X + N^*|U) - h(X + N_Z|U)$$
$$\leq \frac{1}{2} \log \frac{P + \sigma^2_*}{P + \sigma^2_Z} \quad (140)$$

where the right-hand side can be shown via the entropy power inequality [20], [21]. To show the left-hand side, let us define a Gaussian random variable $\tilde{N}$ with variance $\sigma^2_Z - \sigma^2_*$, and independent of $(U, X, N^*)$. Thus, we can write down the difference of differential entropy terms in (140) as

$$h(X + N^*|U) - h(X + N_Z|U)$$
$$= h(X + N^*|U) - h(X + N^* + \tilde{N}|U) \quad (141)$$
$$= -I(\tilde{N}; X + N^* + \tilde{N}|U) \quad (142)$$
$$= -h(\tilde{N}|U) + h(\tilde{N}|U, X + N^* + \tilde{N}) \quad (143)$$
$$\geq -h(\tilde{N}|U) + h(\tilde{N}|U, X + N^* + \tilde{N}, X) \quad (144)$$
$$= -h(\tilde{N}) + h(\tilde{N}|N^* + \tilde{N}) \quad (145)$$
$$= \frac{1}{2} \log \frac{\sigma^2_*}{\sigma^2_Z} \quad (146)$$

where (144) is due to the fact that conditioning cannot increase entropy and (145) is a consequence of the fact that $(U, X)$ and $(N^*, \tilde{N})$ are independent.

Equation (140) implies that there exists $P^*$ such that $P^* \leq P$ and

$$h(X + N^*|U) - h(X + N_Z|U) = \frac{1}{2} \log \frac{P^* + \sigma^2_*}{P^* + \sigma^2_Z} \quad (147)$$

which will be used frequently hereafter.

We now state Costa's entropy power inequality [17] which will be used in the upcoming proof.[1]

*Lemma 1 ([17, Theorem 1]):* Let $(U, X)$ be an arbitrarily dependent random variable pair, which is independent of $N$, where $N$ is a Gaussian random variable. Then, for $0 \leq t \leq 1$, we have

$$e^{2h(X+\sqrt{t}N|U)} \geq (1-t)e^{2h(X|U)} + te^{2h(X+N|U)}. \quad (148)$$

We now consider (42). We first note that we can write $N^*$ as

$$N^* = N_1 + \sqrt{t_1}\tilde{N}_1 \quad (149)$$

where $\tilde{N}_1$ is a Gaussian random variable with variance $\sigma^2_Z - \sigma^2_1$, which is independent of $(U, X, N_1)$. $t_1$ in (149) is given by

$$t_1 = \frac{\sigma^2_* - \sigma^2_1}{\sigma^2_Z - \sigma^2_1} \quad (150)$$

where it is clear that $t_1 \in [0, 1]$. Using (149) and Costa's entropy power inequality [17], we get

$$e^{2h(X+N^*|U)} = e^{2h(X+N_1+\sqrt{t_1}\tilde{N}_1|U)} \quad (151)$$
$$\geq (1-t_1)e^{2h(X+N_1|U)} + t_1 e^{2h(X+N_Z|U)}$$
$$(152)$$

which is equivalent to

$$(1-t_1)e^{2[h(X+N_1|U)-h(X+N_Z|U)]} + t_1$$
$$\leq e^{2[h(X+N^*|U)-h(X+N_Z|U)]} \quad (153)$$
$$= \frac{P^* + \sigma^2_*}{P^* + \sigma^2_Z} \quad (154)$$

where (154) is obtained by using (147). Equation (154) is equivalent to

$$h(X + N_1|U) - h(X + N_Z|U)$$
$$\leq \frac{1}{2} \log \frac{1}{1-t_1} \left( \frac{P^* + \sigma^2_*}{P^* + \sigma^2_Z} - t_1 \right) \quad (155)$$
$$= \frac{1}{2} \log \left( \frac{P^*}{P^* + \sigma^2_Z} + \frac{1}{1-t_1} \frac{\sigma^2_* - t_1\sigma^2_Z}{P^* + \sigma^2_Z} \right) \quad (156)$$
$$= \frac{1}{2} \log \frac{P^* + \sigma^2_1}{P^* + \sigma^2_Z} \quad (157)$$

[1]Although, Theorem 1 in [17] states the inequality for a constant $U$, using Jensen's inequality, the current form of the inequality for an arbitrary $U$ can be shown.

where we used the definition of $t_1$ given in (150) to obtain (157). Equation (157) proves (42).

We now consider (43). First, we note that we can write $N_2$

$$N_2 = N^* + \sqrt{t_2}\tilde{N}_Z \tag{158}$$

where $\tilde{N}_Z$ is a Gaussian random variable with variance $\sigma_Z^2 - \sigma_*^2$, which is independent of $(U, X, N^*)$. $t_2$ in (158) is given by

$$t_2 = \frac{\sigma_2^2 - \sigma_*^2}{\sigma_Z^2 - \sigma_*^2} \tag{159}$$

where it is clear that $t_2 \in [0, 1]$. Using (158) and Costa's entropy power inequality [17], we get

$$e^{2h(X+N_2|U)} = e^{2h(X+N^*+\sqrt{t_2}\tilde{N}_Z|U)} \tag{160}$$

$$\geq (1-t_2)e^{2h(X+N^*|U)} + t_2 e^{2h(X+N_Z|U)} \tag{161}$$

which is equivalent to

$$e^{2[h(X+N_2|U)-h(X+N_Z|U)]}$$
$$\geq (1-t_2)e^{2[h(X+N^*|U)-h(X+N_Z|U)]} + t_2 \tag{162}$$
$$= (1-t_2)\frac{P^* + \sigma_*^2}{P^* + \sigma_Z^2} + t_2 \tag{163}$$
$$= \frac{P^* + \sigma_2^2}{P^* + \sigma_Z^2} \tag{164}$$

where (164) is obtained by using the definition of $t_2$ given in (159). Equation (164) is equivalent to

$$h(X + N_Z|U) - h(X + N_2|U) \leq \frac{1}{2}\log\frac{P^* + \sigma_Z^2}{P^* + \sigma_2^2} \tag{165}$$

which is (43). This completes the proof of Theorem 3.

## APPENDIX IV
## PROOF OF THEOREM 4

Achievability is clear. We provide the converse proof. To this end, let us fix the distribution $\prod_{\ell=1}^{L} p(u_\ell, x_\ell)$ such that

$$E\left[X_\ell^2\right] = P_\ell, \quad \ell = 1, \dots, L \tag{166}$$

and $\sum_{\ell=1}^{L} P_\ell \leq P$. We first establish the bound on $R_2$ given in (45). To this end, we start with (38). Using the Markov chain $U_\ell \to Y_{k\ell}^2 \to Z_\ell$, we have

$$R_2 \leq \min_{k=1,\dots,K_2} \sum_{\ell=1}^{L} I(U_\ell; Y_{k\ell}^2) - I(U_\ell; Z_\ell) \tag{167}$$

$$= \min_{k=1,\dots,K_2} \sum_{\ell=1}^{L} \left[h(Y_{k\ell}^2) - h(Z_\ell)\right] + \left[h(Z_\ell|U) - h(Y_{k\ell}^2|U)\right] \tag{168}$$

$$\leq \min_{k=1,\dots,K_2} \sum_{\ell=1}^{L} \frac{1}{2}\log\frac{P_\ell + \Lambda_{k,\ell\ell}^2}{P_\ell + \Lambda_{Z,\ell\ell}} + \left[h(Z_\ell|U) - h(Y_{k\ell}^2|U)\right] \tag{169}$$

where (169) comes from the fact that Gaussian $X_\ell$ maximizes

$$h(Y_{k\ell}^2) - h(Z_\ell) \tag{170}$$

which can be shown via the entropy power inequality [20], [21]. We now use Theorem 3. For that purpose, we introduce the diagonal covariance matrix $\Lambda^*$ which satisfies

$$\Lambda_j^1 \preceq \Lambda^* \preceq \Lambda_k^2 \tag{171}$$

for any $(j, k)$ pair, and in particular, for the diagonal elements of these matrices, we have

$$\Lambda_{j,\ell\ell}^1 \leq \Lambda_{\ell\ell}^* \leq \Lambda_{k,\ell\ell}^2 \tag{172}$$

for any triple $(j, k, \ell)$. Thus, due to Theorem 3, for any selection of $\{(U_\ell, X_\ell)\}_{\ell=1}^{L}$, there exists a $P_\ell^*$ such that

$$P_\ell^* \leq P_\ell \tag{173}$$

$$h(Z_\ell|U_\ell) - h(Y_{j\ell}^1|U_\ell) \geq \frac{1}{2}\log\frac{P_\ell^* + \Lambda_{Z,\ell\ell}}{P_\ell^* + \Lambda_{j,\ell\ell}^1} \tag{174}$$

$$h(Z_\ell|U_\ell) - h(Y_{k\ell}^2|U_\ell) \leq \frac{1}{2}\log\frac{P_\ell^* + \Lambda_{Z,\ell\ell}}{P_\ell^* + \Lambda_{k,\ell\ell}^2} \tag{175}$$

for any triple $(j, k, \ell)$. Using (175) in (169), we get

$$R_2 \leq \min_{k=1,\dots,K_2} \sum_{\ell=1}^{L} \frac{1}{2}\log\frac{P_\ell + \Lambda_{k,\ell\ell}^2}{P_\ell^* + \Lambda_{k,\ell\ell}^2} - \frac{1}{2}\log\frac{P_\ell + \Lambda_{Z,\ell\ell}}{P_\ell^* + \Lambda_{Z,\ell\ell}}. \tag{176}$$

We define $P_\ell^* = \beta_\ell P_\ell$ and $\bar{\beta}_\ell = 1 - \beta_\ell$, $\ell = 1, \dots, L$, where $\beta_\ell \in [0, 1]$ due to (173). Thus, we have established the desired bound on $R_2$ given in (45). We now bound $R_1$. We start with (37). Using the Markov chain $(U_\ell, X_\ell) \to Y_{j\ell}^1 \to Z_\ell$, we have

$$R_1 \leq \min_{j=1,\dots,K_1} \sum_{\ell=1}^{L} I(X_\ell; Y_{j\ell}^1|U_\ell) - I(X_\ell; Z_\ell|U_\ell) \tag{177}$$

$$= \min_{j=1,\dots,K_1} \sum_{\ell=1}^{L} h(Y_{j\ell}^1|U_\ell) - h(Z_\ell|U_\ell) - \frac{1}{2}\log\frac{\Lambda_{j,\ell\ell}^1}{\Lambda_{Z,\ell\ell}} \tag{178}$$

$$\leq \min_{j=1,\dots,K_1} \sum_{\ell=1}^{L} \frac{1}{2}\log\frac{P_\ell^* + \Lambda_{j,\ell\ell}^1}{P_\ell^* + \Lambda_{Z,\ell\ell}} - \frac{1}{2}\log\frac{\Lambda_{j,\ell\ell}^1}{\Lambda_{Z,\ell\ell}} \tag{179}$$

where (179) comes from (174). Since we defined $P_\ell^* = \beta_\ell P_\ell$, (179) is the desired bound on $R_1$ given in (44), completing the proof.

## APPENDIX V
## PROOF OF THEOREM 6

The main tools for the proof of Theorem 6 are Theorem 5, and the following so-called worst additive noise lemma [22], [23].

*Lemma 2 ([22], [23]):* Let $\mathbf{N}$ be a Gaussian random vector with covariance matrix $\Sigma$, and $\mathbf{K}_X$ be a positive semidefinite matrix. Consider the following optimization problem:

$$\min_{p(\mathbf{x})} \quad I(\mathbf{N}; \mathbf{N} + \mathbf{X}) \quad \text{s.t.} \quad \text{Cov}(\mathbf{X}) = \mathbf{K}_X \tag{180}$$

where $\mathbf{X}$ and $\mathbf{N}$ are independent. A Gaussian $\mathbf{X}$ is the minimizer of this optimization problem.

We first bound $R_2$. Assume we fixed the distribution of $(U, \mathbf{X})$ such that $\mathrm{Cov}(\mathbf{X}) = \mathbf{K}_X$. Then, we have

$$R_2 \leq I(U; \mathbf{Y}^2) - I(U; \mathbf{Z}) \tag{181}$$

$$= h(\mathbf{Y}^2) - h(\mathbf{Z}) + [h(\mathbf{Z}|U) - h(\mathbf{Y}^2|U)] \tag{182}$$

$$\leq \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}^2|}{|\mathbf{S} + \boldsymbol{\Sigma}_Z|} + [h(\mathbf{Z}|U) - h(\mathbf{Y}^2|U)]. \tag{183}$$

To show (183), consider $\tilde{\mathbf{N}}$ which is a Gaussian random vector with covariance matrix $\boldsymbol{\Sigma}_Z - \boldsymbol{\Sigma}^2$, and is independent of $(U, \mathbf{X}, \mathbf{N}^2)$. Thus, we can write

$$h(\mathbf{Y}^2) - h(\mathbf{Z}) = h(\mathbf{Z}|\tilde{\mathbf{N}}) - h(\mathbf{Z}) \tag{184}$$

$$= -I(\tilde{\mathbf{N}}; \mathbf{X} + \mathbf{N}^2 + \tilde{\mathbf{N}}) \tag{185}$$

$$\leq \frac{1}{2} \log \frac{|\mathbf{K}_X + \boldsymbol{\Sigma}^2|}{|\mathbf{K}_X + \boldsymbol{\Sigma}_Z|} \tag{186}$$

$$\leq \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}^2|}{|\mathbf{S} + \boldsymbol{\Sigma}_Z|} \tag{187}$$

where (186) is due to Lemma 2, and (187) follows from the fact that

$$\frac{|\mathbf{A}|}{|\mathbf{A} + \mathbf{B}|} \leq \frac{|\mathbf{A} + \boldsymbol{\Delta}|}{|\mathbf{A} + \mathbf{B} + \boldsymbol{\Delta}|} \tag{188}$$

for $\mathbf{A} \succeq \mathbf{0}$, $\mathbf{B} \succ \mathbf{0}$, $\boldsymbol{\Delta} \succeq \mathbf{0}$ [3], [19].

For the rest of the proof, we need Theorem 5. According to Theorem 5, for any $(U, \mathbf{X})$, there exists a $\mathbf{0} \preceq \mathbf{K} \preceq \mathrm{Cov}(\mathbf{X}|U)$ such that

$$h(\mathbf{Z}|U) - h(\mathbf{Y}^2|U) = \frac{1}{2} \log \frac{|\mathbf{K} + \boldsymbol{\Sigma}_Z|}{|\mathbf{K} + \boldsymbol{\Sigma}^2|} \tag{189}$$

$$h(\mathbf{Z}|U) - h(\mathbf{Y}_j^1|U) \geq \frac{1}{2} \log \frac{|\mathbf{K} + \boldsymbol{\Sigma}_Z|}{|\mathbf{K} + \boldsymbol{\Sigma}_j^1|}, \quad j = 1, \ldots, K_1 \tag{190}$$

because $\boldsymbol{\Sigma}_j^1 \preceq \boldsymbol{\Sigma}^2$, $j = 1, \ldots, K_1$. Using (189) in (183) yields

$$R_2 \leq \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}^2|}{|\mathbf{K} + \boldsymbol{\Sigma}^2|} - \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_Z|}{|\mathbf{K} + \boldsymbol{\Sigma}_Z|} \tag{191}$$

which is the desired bound on $R_2$.

The desired bound on $R_1$ can be obtained as follows:

$$R_1 \leq \min_{j=1,\ldots,K_1} I(\mathbf{X}; \mathbf{Y}_j^1|U) - I(\mathbf{X}; \mathbf{Z}|U) \tag{192}$$

$$= \min_{j=1,\ldots,K_1} h(\mathbf{Y}_j^1|U) - h(\mathbf{Z}|U) - \frac{1}{2} \log \frac{|\boldsymbol{\Sigma}_j^1|}{|\boldsymbol{\Sigma}_Z|} \tag{193}$$

$$\leq \min_{j=1,\ldots,K_1} \frac{1}{2} \log \frac{|\mathbf{K} + \boldsymbol{\Sigma}_j^1|}{|\mathbf{K} + \boldsymbol{\Sigma}_Z|} - \frac{1}{2} \log \frac{|\boldsymbol{\Sigma}_j^1|}{|\boldsymbol{\Sigma}_Z|} \tag{194}$$

$$= \min_{j=1,\ldots,K_1} \frac{1}{2} \log \frac{|\mathbf{K} + \boldsymbol{\Sigma}_j^1|}{|\boldsymbol{\Sigma}_j^1|} - \frac{1}{2} \log \frac{|\mathbf{K} + \boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_Z|} \tag{195}$$

where (194) is due to (190). This completes the proof of Theorem 6.

# APPENDIX VI
# PROOF OF THEOREM 7

We first show the achievability of the region given in Theorem 7, then provide the converse proof.

*A) Achievability:* First, we present the following lemma which simplifies the achievability proof.

*Lemma 3 ([11, Lemma A.1]):* Let $U$, $X$, $Z_1$, $Z_2$, $Y_1^2$, $Y_2^2$ be random variables such that they satisfy the following Markov chains:

$$U \to X \to Z_1, Z_2 \tag{196}$$

$$U \to X \to Y_1^2, Y_2^2. \tag{197}$$

If $I(U; Z_1) < I(U; Z_2)$, there exists a random variable $\tilde{Z}$ such that $I(U; Z_1, \tilde{Z}) = I(U; Z_2)$ and $\tilde{Z}$ satisfies the following Markov chain:

$$U \to X \to (Z_1, Z_2) \to \tilde{Z}. \tag{198}$$

Similarly, if $I(X; Y_1^2|U) < I(X; Y_2^2|U)$, there exists a random variable $\tilde{Y}^2$ such that $I(X; Y_1^2, \tilde{Y}^2|U) = I(X; Y_2^2|U)$ and $\tilde{Y}^2$ satisfies the following Markov chain:

$$U \to X \to (Y_1^2, Y_2^2) \to \tilde{Y}^2. \tag{199}$$

We now show the achievability of the region given in Theorem 7. First, we fix the distribution $p(u, x)$.

*Codebook generation:*

1) Generate $2^{n(R_2 + \tilde{R}_2)}$ length-$n$ $\mathbf{u}$ sequences through $p(\mathbf{u}) = \prod_{i=1}^n p(u_i)$ where $\tilde{R}_2 = \max_{t=1,\ldots,K_Z} I(U; Z_t)$. We index $\mathbf{u}$ sequences as $\mathbf{u}(w_2, \tilde{w}_2)$ where $w_2 \in \{1, \ldots, 2^{nR_2}\}$, and $\tilde{w}_2 \in \{1, \ldots, 2^{n\tilde{R}_2}\}$.
2) For each $\mathbf{u}$, generate $2^{n(R_1 + \tilde{R}_1)}$ length-$n$ $\mathbf{x}$ sequences through $p(\mathbf{x}|\mathbf{u}) = \prod_{i=1}^n p(x_i|u_i)$ where $\tilde{R}_1 = \max_{k=1,\ldots,K_2} I(X; Y_k^2|U)$. We index $\mathbf{x}$ sequences as $\mathbf{x}(w_1, \tilde{w}_1|\mathbf{w}_2)$ where $\mathbf{w}_2 = (w_2, \tilde{w}_2)$, $w_1 \in \{1, \ldots, 2^{nR_1}\}$, and $\tilde{w}_1 \in \{1, \ldots, 2^{n\tilde{R}_1}\}$.

*Encoding:*

If $(w_1, w_2)$ is the message to be transmitted, we pick $\tilde{w}_1, \tilde{w}_2$ independently and uniformly, and send the corresponding $\mathbf{x}$.

*Decoding:*

The legitimate users can decode the messages with vanishingly small probability of error, if the rates satisfy

$$R_1 \leq \min_{\substack{j=1,\ldots,K_1 \\ k=1,\ldots,K_2}} I(X; Y_j^1|U) - I(X; Y_k^2|U) \tag{200}$$

$$R_2 \leq \min_{\substack{k=1,\ldots,K_2 \\ t=1,\ldots,K_Z}} I(U; Y_k^2) - I(U; Z_t) \tag{201}$$

which is the same as the region given in Theorem 7 because of the degradedness of the channel.

*Equivocation computation:*

We now show that this coding scheme satisfies the secrecy requirements given in (58) and (59). To this end, we will take a shortcut by using Lemma 3, as it is done in [11]. To show (58), we consider the enhanced eavesdroppers with observations $(Z_t, \tilde{Z}_t)$ such that $I(U; Z_t, \tilde{Z}_t) = \max_{t=1,\ldots,K_Z} I(U; Z_t)$, where the existence of the random variable $\tilde{Z}_t$ is ensured by Lemma 3. Following the equivocation computation in [18, Appendix A], one can get

$$\lim_{n \to \infty} \frac{1}{n} I(W_2; Z_t^n, \tilde{Z}_t^n) = 0 \tag{202}$$

which implies that the secrecy requirement in (58) is satisfied.

Next we show that the proposed encoding scheme satisfies the secrecy requirement in (59) as well. Similar to what we did to show (58), we take a shortcut by using Lemma 3. In particular, we consider the enhanced second group of users with observations $(Y_k^2, \tilde{Y}_k^2)$ such that $I(X; Y_k^2, \tilde{Y}_k^2|U) = \max_{k=1,\ldots,K_2} I(X; Y_k^2, \tilde{Y}_k^2|U)$ where the existence of the random variable $\tilde{Y}_k^2$ is ensured by Lemma 3. Following the equivocation computation in [18, Appendix A], one can get

$$\lim_{n \to \infty} \frac{1}{n} I(W_1; Y_k^{2,n}, \tilde{Y}_k^{2,n}|W_2) = 0 \tag{203}$$

which implies that the secrecy requirement in (59) is satisfied. This completes the achievability proof of Theorem 7.

*B) Converse:* First, we note that for an arbitrary code achieving the secrecy rate pairs $(R_1, R_2)$, there exist $(\epsilon_{1,n}, \epsilon_{2,n})$ and $(\gamma_{1,n}, \gamma_{2,n})$ which vanish as $n \to \infty$ such that

$$H(W_1|Y_j^{1,n}) \leq n\epsilon_{1,n}, \quad j = 1, \ldots, K_1 \tag{204}$$

$$H(W_2|Y_k^{2,n}) \leq n\epsilon_{2,n}, \quad k = 1, \ldots, K_2 \tag{205}$$

$$I(W_2; Z_t^n) \leq n\gamma_{2,n}, \quad t = 1, \ldots, K_Z \tag{206}$$

$$I(W_1; Y_k^{2,n}|W_2) \leq n\gamma_{1,n}, \quad k = 1, \ldots, K_2 \tag{207}$$

where (204) and (205) are due to Fano's lemma, and (206) and (207) come from perfect secrecy requirements in (58) and (59).

We now define the following auxiliary random variables:

$$U_i = W_2 Y^{*,i-1} Z_{i+1}^{*,n} \tag{208}$$

for $i = 1, \ldots, n$, which satisfy the Markov chains

$$U_i \to X_i \to Y_{j,i}^1 \to Y_i^* \to Y_{k,i}^2 \to Z_i^* \to Z_{t,i} \tag{209}$$

for any $(j, k, t)$ triple and for $i = 1, \ldots, n$. The Markov chain in (209) is a consequence of the fact that the channel is memoryless and degraded.

We first establish the desired bound on $R_2$ as follows:

$$nR_2 \leq I(W_2; Y_{k,i}^{2,n}|Z_{t,i+1}^n, Y_k^{2,i-1}, Z_{t,i}) + n(\epsilon_{2,n} + \gamma_{2,n}) \tag{210}$$

$$\leq \sum_{i=1}^n I(Z_{t,i+1}^n, Y_k^{2,i-1}, W_2; Y_{k,i}^2|Z_{t,i}) + n(\epsilon_{2,n} + \gamma_{2,n}) \tag{211}$$

$$\leq \sum_{i=1}^n I(Z_{i+1}^{*,n}, Y^{*,i-1}, Z_{t,i+1}^n, Y_k^{2,i-1}, W_2; Y_{k,i}^2|Z_{t,i}) + n(\epsilon_{2,n} + \gamma_{2,n}) \tag{212}$$

$$\leq \sum_{i=1}^n I(Z_{i+1}^{*,n}, Y^{*,i-1}, W_2; Y_{k,i}^2|Z_{t,i}) + n(\epsilon_{2,n} + \gamma_{2,n}) \tag{213}$$

$$= \sum_{i=1}^n I(U_i; Y_{k,i}^2|Z_{t,i}) + n(\epsilon_{2,n} + \gamma_{2,n}) \tag{214}$$

where (210) can be obtained by following the steps similar to (86) – (91) in Appendix I and (213) is due to the Markov chain

$$(Z_{t,i+1}^n, Y_k^{2,i-1}) \to (Z_{i+1}^{*,n}, Y^{*,i-1}) \to (W_2, Y_{k,i}^2, Z_{t,i}) \tag{215}$$

which is a consequence of the Markov chain in (2).

We now establish the bound on $R_1$ as follows:

$$nR_1$$
$$\leq \sum_{i=1}^n I(W_1; Y_{j,i}^1|W_2, Y_{k,i+1}^{2,n}, Y_j^{1,i-1}, Y_{k,i}^2) + n(\epsilon_{1,n} + \gamma_{1,n}) \tag{216}$$

$$= \sum_{i=1}^n I(W_1; Y_{j,i}^1|W_2, Y_{k,i+1}^{2,n}, Y_j^{1,i-1}, Z_{i+1}^{*,n}, Y^{*,i-1}, Y_{k,i}^2)$$
$$+ n(\epsilon_{1,n} + \gamma_{1,n}) \tag{217}$$

$$= \sum_{i=1}^n I(W_1; Y_{j,i}^1|U_i, Y_{k,i+1}^{2,n}, Y_j^{1,i-1}, Y_{k,i}^2) + n(\epsilon_{1,n} + \gamma_{1,n}) \tag{218}$$

$$\leq \sum_{i=1}^n I(X_i, W_1; Y_{j,i}^1|U_i, Y_{k,i+1}^{2,n}, Y_j^{1,i-1}, Y_{k,i}^2)$$
$$+ n(\epsilon_{1,n} + \gamma_{1,n}) \tag{219}$$

$$= \sum_{i=1}^n I(X_i; Y_{j,i}^1|U_i, Y_{k,i+1}^{2,n}, Y_j^{1,i-1}, Y_{k,i}^2) + n(\epsilon_{1,n} + \gamma_{1,n}) \tag{220}$$

$$= \sum_{i=1}^n H(Y_{j,i}^1|U_i, Y_{k,i+1}^{2,n}, Y_j^{1,i-1}, Y_{k,i}^2)$$
$$- H(Y_{j,i}^1|U_i, Y_{k,i+1}^{2,n}, Y_j^{1,i-1}, Y_{k,i}^2, X_i)$$
$$+ n(\epsilon_{1,n} + \gamma_{1,n}) \tag{221}$$

$$= \sum_{i=1}^n H(Y_{j,i}^1|U_i, Y_{k,i+1}^{2,n}, Y_j^{1,i-1}, Y_{k,i}^2)$$
$$- H(Y_{j,i}^1|U_i, Y_{k,i}^2, X_i) + n(\epsilon_{1,n} + \gamma_{1,n}) \tag{222}$$

$$\leq \sum_{i=1}^n H(Y_{j,i}^1|U_i, Y_{k,i}^2) - H(Y_{j,i}^1|U_i, Y_{k,i}^2, X_i)$$
$$+ n(\epsilon_{1,n} + \gamma_{1,n}) \tag{223}$$

$$= \sum_{i=1}^n I(X_i; Y_{j,i}^1|U_i, Y_{k,i}^2) + n(\epsilon_{1,n} + \gamma_{1,n}) \tag{224}$$

where (216) can be obtained by following the steps similar to (99) – (105) in Appendix I, (217) is a consequence of the Markov chain

$$(Z_{i+1}^{*,n}, Y^{*,i-1}) \to (Y_{k,i+1}^{2,n}, Y_j^{1,i-1}) \to (W_2, W_1, Y_{j,i}^1, Y_{k,i}^2) \tag{225}$$

which results from the Markov chain in (2), (220) comes from the Markov chain

$$(Y_{k,i}^2, Y_{j,i}^1) \to X_i \to (W_1, W_2, U_i, Y_{k,i+1}^{2,n}, Y_j^{1,i-1}) \tag{226}$$

which is due to the fact that the channel is memoryless, (222) is also due to the Markov chain in (226), and (223) comes from the fact that conditioning cannot increase entropy.

Single-letterization can be accomplished as outlined in the proof of Theorem 1, completing the converse proof.

## APPENDIX VII
### PROOF OF THEOREM 8

The achievability of the region given in Theorem 8 can be shown by selecting $(U, X) = (U_1, X_1, \ldots, U_L, X_L)$ with a

joint distribution of the form $p(u,x) = \prod_{\ell=1}^{L} p(u_\ell, x_\ell)$. We next provide an outline of the converse proof. To that end, we define the following auxiliary random variables:

$$U_\ell = U Y^*_{[1:\ell-1]} Z^*_{[\ell+1:L]}, \quad \ell = 1, \ldots, L \qquad (227)$$

which satisfy the Markov chains

$$U_\ell \to X_\ell \to Y^1_{j\ell} \to Y^*_\ell \to Y^2_{k\ell} \to Z^*_\ell \to Z_{t\ell}, \quad \ell = 1, \ldots, L \qquad (228)$$

for any $(j, k, t)$ triple. These Markov chains are a consequence of the facts that the channel is memoryless and degraded, and subchannels are independent. Once the auxiliary random variables $\{U_\ell\}_{\ell=1}^L$ in (227) are identified, the rest of the converse proof is similar to the converse proof of Theorem 2 given in Appendix II. In particular, to obtain the desired bound on $R_2$, we start with

$$R_2 \leq \min_{\substack{k=1,\ldots,K_2 \\ t=1,\ldots,K_Z}} I(U; Y^2_{k[1:L]} | Z_{t[1:L]}) \qquad (229)$$

which is a direct consequence of Theorem 7. Next, following the steps similar to (121) – (126) in Appendix II, one can reach the desired bound:

$$R_2 \leq \min_{\substack{k=1,\ldots,K_2 \\ t=1,\ldots,K_Z}} \sum_{\ell=1}^{L} I(U_\ell; Y^2_{k\ell} | Z_{t\ell}). \qquad (230)$$

Similarly, to obtain the desired bound on $R_1$, we start with

$$R_1 \leq \min_{\substack{j=1,\ldots,K_1 \\ k=1,\ldots,K_2}} I(X; Y^1_{j[1:L]} | U, Y^2_{k[1:L]}) \qquad (231)$$

which is also a direct consequence of Theorem 7. Next, following the steps similar to (130) – (137) in Appendix II, one can each the desired bound:

$$R_1 \leq \min_{\substack{j=1,\ldots,K_1 \\ k=1,\ldots,K_2}} \sum_{\ell=1}^{L} I(X_\ell; Y^1_{j\ell} | U_\ell, Y^2_{k\ell}). \qquad (232)$$

To complete the converse proof, we note that each expression in the bounds given by (230) and (232) depend on the the joint distribution $p(u_{[1:L]}, x_{[1:L]})$ through its marginals $p(u_\ell, x_\ell)$. Thus, there is no loss of optimality to choose $p(u_{[1:L]}, x_{[1:L]}) = \prod_{\ell=1}^{L} p(u_\ell, x_\ell)$. This completes the converse proof.

## APPENDIX VIII
## PROOF OF THEOREM 9

According to Theorem 3, there exists a $P^* \leq P$ such that

$$h(X + \tilde{N} | U) - h(X + N^* | U) = \frac{1}{2} \log \frac{P^* + \tilde{\sigma}^2}{P^* + \sigma_*^2} \qquad (233)$$

$$h(X + \tilde{N} | U) - h(X + N_2 | U) \leq \frac{1}{2} \log \frac{P^* + \tilde{\sigma}^2}{P^* + \sigma_2^2} \qquad (234)$$

$$h(X + \tilde{N} | U) - h(X + N_1 | U) \geq \frac{1}{2} \log \frac{P^* + \tilde{\sigma}^2}{P^* + \sigma_1^2} \qquad (235)$$

for any $(\sigma_1^2, \sigma_2^2)$ as long as they satisfy

$$\sigma_1^2 \leq \sigma_*^2 \leq \sigma_2^2 \leq \tilde{\sigma}^2. \qquad (236)$$

We first show (70). To this end, we note that (233) and (234) imply

$$h(X + N_2 | U) - h(X + N^* | U) \geq \frac{1}{2} \log \frac{P^* + \sigma_2^2}{P^* + \sigma_*^2}. \qquad (237)$$

Furthermore, (233) and (235) imply

$$h(X + N^* | U) - h(X + N_1 | U) \geq \frac{1}{2} \log \frac{P^* + \sigma_*^2}{P^* + \sigma_1^2}. \qquad (238)$$

Combining (237) and (238) yields

$$h(X + N_2 | U) - h(X + N_1 | U) \geq \frac{1}{2} \log \frac{P^* + \sigma_2^2}{P^* + \sigma_1^2} \qquad (239)$$

which is the desired result in (70).

We now show (69). We first note that we can write $\tilde{N}$ as

$$\tilde{N} = N_2 + \sqrt{t} \tilde{N}_Z \qquad (240)$$

where $\tilde{N}_Z$ is a zero-mean Gaussian random variable with variance $\sigma_Z^2 - \sigma_2^2$, and independent of $(U, X, N_2)$. $t$ in (240) is given by

$$t = \frac{\tilde{\sigma}^2 - \sigma_2^2}{\sigma_Z^2 - \sigma_2^2} \qquad (241)$$

where it is clear that $t \in [0, 1]$. We now use Costa's entropy power inequality [17] to arrive at (69)

$$e^{2h(X+\tilde{N}|U)} = e^{2h(X+N_2+\sqrt{t}\tilde{N}_Z|U)} \qquad (242)$$
$$\geq (1-t) e^{2h(X+N_2|U)} + t e^{2h(X+N_Z|U)} \qquad (243)$$

which is equivalent to

$$e^{2[h(X+\tilde{N}|U) - h(X+N_2|U)]}$$
$$\geq (1-t) + t e^{2[h(X+N_Z|U) - h(X+N_2|U)]} \qquad (244)$$

which can be written as

$$h(X + N_Z | U) - h(X + N_2 | U)$$
$$\leq \frac{1}{2} \log \left[ \frac{1}{t} e^{2[h(X+\tilde{N}|U) - h(X+N_2|U)]} - \frac{1-t}{t} \right] \qquad (245)$$
$$\leq \frac{1}{2} \log \left[ \frac{1}{t} \frac{P^* + \tilde{\sigma}^2}{P^* + \sigma_2^2} - \frac{1-t}{t} \right] \qquad (246)$$
$$= \frac{1}{2} \log \left[ \frac{P^*}{P^* + \sigma_2^2} - \frac{1}{t} \frac{\tilde{\sigma}^2 - (1-t)\sigma_2^2}{P^* + \sigma_2^2} \right] \qquad (247)$$
$$= \frac{1}{2} \log \frac{P^* + \sigma_Z^2}{P^* + \sigma_2^2} \qquad (248)$$

where (246) is due to (234) and (248) comes from (241). Since (248) is the desired result in (69), this completes the proof.

## APPENDIX IX
### PROOF OF THEOREM 10

Achievability is clear. We provide the converse proof. We fix the distribution $\prod_{\ell=1}^{L} p(u_\ell, x_\ell)$ such that

$$E\left[X_\ell^2\right] = P_\ell, \quad \ell = 1, \ldots, L \qquad (249)$$

and $\sum_{\ell=1}^{L} P_\ell = P$. We first establish the bound on $R_2$ given in (72). To this end, we start with (65). Using the Markov chain $U_\ell \to Y_{k\ell}^2 \to Z_{t\ell}$, we have

$$R_2 \leq \min_{\substack{k=1,\ldots,K_2 \\ t=1,\ldots,K_Z}} \sum_{\ell=1}^{L} I(U_\ell; Y_{k\ell}^2) - I(U_\ell; Z_{t\ell}) \qquad (250)$$

$$= \min_{\substack{k=1,\ldots,K_2 \\ t=1,\ldots,K_Z}} \sum_{\ell=1}^{L} h(Y_{k\ell}^2) - h(Z_{t\ell}) + \left[h(Z_{t\ell}|U_\ell) - h(Y_{k\ell}^2|U_\ell)\right] \qquad (251)$$

$$\leq \min_{\substack{k=1,\ldots,K_2 \\ t=1,\ldots,K_Z}} \sum_{\ell=1}^{L} \frac{1}{2} \log \frac{P_\ell + \Lambda_{k,\ell\ell}^2}{P_\ell + \Lambda_{t,\ell\ell}^Z} + \left[h(Z_{t\ell}|U_\ell) - h(Y_{k\ell}^2|U_\ell)\right] \qquad (252)$$

where (252) comes from the fact that $h(Y_{k\ell}^2) - h(Z_{t\ell})$ is maximized by Gaussian distribution which can be shown by using the entropy power inequality [20], [21]. We now use Theorem 9. For that purpose, we introduce $\Lambda_Y^*$ and $\Lambda_Z^*$ which satisfy

$$\Lambda_j^1 \preceq \Lambda_Y^* \preceq \Lambda_k^2 \preceq \Lambda_Z^* \preceq \Lambda_t^Z \qquad (253)$$

for any $(j, k, t)$ triple, and in particular, for the diagonal, elements of these matrices, we have

$$\Lambda_{j,\ell\ell}^1 \leq \Lambda_{Y,\ell\ell}^* \leq \Lambda_{k,\ell\ell}^2 \leq \Lambda_{Z,\ell\ell}^* \leq \Lambda_{t,\ell\ell}^Z \qquad (254)$$

for any $(j, k, t, \ell)$. Thus, due to Theorem 9, for any selection of $\{(U_\ell, X_\ell)\}_{\ell=1}^{L}$, we have

$$P_\ell^* \leq P_\ell \qquad (255)$$

$$h(Z_{t\ell}|U_\ell) - h(Y_{k\ell}^2|U_\ell) \leq \frac{1}{2} \log \frac{P_\ell^* + \Lambda_{t,\ell\ell}^Z}{P_\ell^* + \Lambda_{k,\ell\ell}^2} \qquad (256)$$

$$h(Y_{k\ell}^2|U_\ell) - h(Y_{j\ell}^1|U_\ell) \geq \frac{1}{2} \log \frac{P_\ell^* + \Lambda_{k,\ell\ell}^2}{P_\ell^* + \Lambda_{j,\ell\ell}^1} \qquad (257)$$

for any $(k, j, t, \ell)$. Using (256) in (252) yields

$$R_2 \leq \min_{\substack{k=1,\ldots,K_2 \\ t=1,\ldots,K_Z}} \sum_{\ell=1}^{L} \frac{1}{2} \log \frac{P_\ell + \Lambda_{k,\ell\ell}^2}{P_\ell^* + \Lambda_{k,\ell\ell}^2} - \frac{1}{2} \log \frac{P_\ell + \Lambda_{t,\ell\ell}^Z}{P_\ell^* + \Lambda_{t,\ell\ell}^Z}. \qquad (258)$$

By defining $P_\ell^* = \beta_\ell P_\ell$ and $\bar{\beta}_\ell = 1 - \beta_\ell$, $\ell = 1, \ldots, L$, where $\beta_\ell \in [0, 1]$ due to (255), we get the desired bound on $R_2$ given in (72).

We now bound $R_1$. We start with (64). Using the Markov chain $U_\ell \to X_\ell \to Y_{j\ell}^1 \to Y_{k\ell}^2$, we have

$$R_1 \leq \min_{\substack{j=1,\ldots,K_1 \\ k=1,\ldots,K_2}} \sum_{\ell=1}^{L} I(X_\ell; Y_{j\ell}^1|U_\ell) - I(X_\ell; Y_{k\ell}^2|U_\ell) \qquad (259)$$

$$= \min_{\substack{j=1,\ldots,K_1 \\ k=1,\ldots,K_2}} \sum_{\ell=1}^{L} h(Y_{j\ell}^1|U_\ell) - h(Y_{k\ell}^2|U_\ell) - \frac{1}{2} \log \frac{\Lambda_{j,\ell\ell}^1}{\Lambda_{k,\ell\ell}^2} \qquad (260)$$

$$\leq \min_{\substack{j=1,\ldots,K_1 \\ k=1,\ldots,K_2}} \sum_{\ell=1}^{L} \frac{1}{2} \log \frac{P_\ell^* + \Lambda_{j,\ell\ell}^1}{P_\ell^* + \Lambda_{k,\ell\ell}^2} - \frac{1}{2} \log \frac{\Lambda_{j,\ell\ell}^1}{\Lambda_{k,\ell\ell}^2} \qquad (261)$$

$$= \min_{\substack{j=1,\ldots,K_1 \\ k=1,\ldots,K_2}} \sum_{\ell=1}^{L} \frac{1}{2} \log \left(1 + \frac{\beta_\ell P_\ell}{\Lambda_{j,\ell\ell}^1}\right) - \frac{1}{2} \log \left(1 + \frac{\beta_\ell P_\ell}{\Lambda_{k,\ell\ell}^2}\right) \qquad (262)$$

where (261) is due to (257). Since (262) is the desired bound on $R_1$ given in (71), this completes the proof.

## APPENDIX X
### BACKGROUND INFORMATION FOR APPENDIX XI

In Appendix XI, we need some properties of the Fisher information and the differential entropy, which are provided here.

*Definition 1 ([3, Definition 3]):* Let $(\mathbf{U}, \mathbf{X})$ be an arbitrarily correlated length-$n$ random vector pair with well-defined densities. The conditional Fisher information matrix of $\mathbf{X}$ given $\mathbf{U}$ is defined as

$$\mathbf{J}(\mathbf{X}|\mathbf{U}) = E\left[\rho(\mathbf{X}|\mathbf{U})\rho(\mathbf{X}|\mathbf{U})^\top\right] \qquad (263)$$

where the expectation is over the joint density $f(\mathbf{u}, \mathbf{x})$, and the conditional score function $\rho(\mathbf{x}|\mathbf{u})$ is

$$\rho(\mathbf{x}|\mathbf{u}) = \nabla \log f(\mathbf{x}|\mathbf{u}) \qquad (264)$$

$$= \left[\frac{\partial \log f(\mathbf{x}|\mathbf{u})}{\partial x_1} \cdots \frac{\partial \log f(\mathbf{x}|\mathbf{u})}{\partial x_n}\right]^\top. \qquad (265)$$

The following lemma will be used in the upcoming proof. In fact, an unconditional version of this lemma is proved in [3, Lemma 6].

*Lemma 4:* Let $\mathbf{T}$, $\mathbf{U}$, $\mathbf{V}_1$, $\mathbf{V}_2$ be random vectors such that $(\mathbf{T}, \mathbf{U})$ and $(\mathbf{V}_1, \mathbf{V}_2)$ are independent. Moreover, let $\mathbf{V}_1$, $\mathbf{V}_2$ be Gaussian random vectors with covariances matrices $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$ such that $\mathbf{0} \prec \mathbf{\Sigma}_1 \preceq \mathbf{\Sigma}_2$. Then, we have

$$\mathbf{J}^{-1}(\mathbf{U} + \mathbf{V}_2|\mathbf{T}) - \mathbf{\Sigma}_2 \succeq \mathbf{J}^{-1}(\mathbf{U} + \mathbf{V}_1|\mathbf{T}) - \mathbf{\Sigma}_1. \qquad (266)$$

The following lemma is also instrumental for the upcoming proof whose proof can be found in [3].

*Lemma 5 ([3, Lemma 8]):* Let $\mathbf{K}_1$, $\mathbf{K}_2$ be positive semidefinite matrices satisfying $\mathbf{0} \preceq \mathbf{K}_1 \preceq \mathbf{K}_2$, and $\mathbf{f}(\mathbf{K})$ be a matrix-valued function such that $\mathbf{f}(\mathbf{K}) \succeq \mathbf{0}$ for $\mathbf{K}_1 \preceq \mathbf{K} \preceq \mathbf{K}_2$. Moreover, we assume that $\mathbf{f}(\mathbf{K})$ is gradient of some scalar field. Then, we have

$$\int_{\mathbf{K}_1}^{\mathbf{K}_2} \mathbf{f}(\mathbf{K}) d\mathbf{K} \geq 0. \tag{267}$$

The following generalization of the de Bruin identity [20], [21] is due to [24]. In [24], the unconditional form of this identity, i.e., the case where $U = \phi$, is proved. However, its generalization to this conditional form for an arbitrary $U$ is rather straightforward, and given in [3, Lemma 16].

*Lemma 6 ([3, Lemma 16]):* Let $(\mathbf{U}, \mathbf{X})$ be an arbitrarily correlated random vector pair with finite second order moments, and be independent of the random vector $\mathbf{N}$ which is zero-mean Gaussian with covariance matrix $\boldsymbol{\Sigma}_N \succ \mathbf{0}$. Then, we have

$$\nabla_{\boldsymbol{\Sigma}_N} h(\mathbf{X} + \mathbf{N}|U) = \frac{1}{2}\mathbf{J}(\mathbf{X} + \mathbf{N}|U). \tag{268}$$

## APPENDIX XI
## PROOF OF THEOREM 11

According to Theorem 5, for any selection of $(U, \mathbf{X})$, there exists a $\mathbf{K}^* \preceq \mathbf{S}$ such that

$$h(\mathbf{X} + \mathbf{N}^*|U) - h(\mathbf{X} + \mathbf{N}_2|U) = \frac{1}{2} \log \frac{|\mathbf{K}^* + \boldsymbol{\Sigma}^*|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_2|} \tag{269}$$

$$h(\mathbf{X} + \mathbf{N}^*|U) - h(\mathbf{X} + \mathbf{N}_1|U) \geq \frac{1}{2} \log \frac{|\mathbf{K}^* + \boldsymbol{\Sigma}^*|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_1|} \tag{270}$$

for any $\boldsymbol{\Sigma}_1$ such that $\boldsymbol{\Sigma}_1 \preceq \boldsymbol{\Sigma}_2$. Furthermore, $\mathbf{K}^*$ satisfies [3]

$$\mathbf{K}^* \preceq \mathbf{J}^{-1}(\mathbf{X} + \mathbf{N}^*|U) - \boldsymbol{\Sigma}^* \tag{271}$$

Equations (269) and (270) already imply

$$h(\mathbf{X} + \mathbf{N}_2|U) - h(\mathbf{X} + \mathbf{N}_1|U) \geq \frac{1}{2} \log \frac{|\mathbf{K}^* + \boldsymbol{\Sigma}_2|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_1|} \tag{272}$$

for any $\boldsymbol{\Sigma}_1$ such that $\boldsymbol{\Sigma}_1 \preceq \boldsymbol{\Sigma}_2$, which is the desired inequality in (76).

We now prove (75). For that purpose, we note that (271) implies

$$\mathbf{K}^* \preceq \mathbf{J}^{-1}(\mathbf{X} + \mathbf{N}|U) - \boldsymbol{\Sigma}_N \tag{273}$$

for any Gaussian random vector $\mathbf{N}$, independent of $(U, \mathbf{X})$, with covariance matrix $\boldsymbol{\Sigma}_N$ such that $\boldsymbol{\Sigma}_N \succeq \boldsymbol{\Sigma}^*$ because of Lemma 4. The order in (273) is equivalent to

$$\mathbf{J}(\mathbf{X} + \mathbf{N}|U) \preceq (\mathbf{K}^* + \boldsymbol{\Sigma}_N)^{-1}, \quad \boldsymbol{\Sigma}^* \preceq \boldsymbol{\Sigma}_N. \tag{274}$$

Now, we can obtain (75) as follows:

$$\begin{aligned} &h(\mathbf{X} + \mathbf{N}_Z|U) - h(\mathbf{X} + \mathbf{N}_2|U) \\ &= h(\mathbf{X} + \mathbf{N}_Z|U) - h(\mathbf{X} + \mathbf{N}^*|U) \\ &\quad + h(\mathbf{X} + \mathbf{N}^*|U) - h(\mathbf{X} + \mathbf{N}_2|U) \end{aligned} \tag{275}$$

$$= h(\mathbf{X} + \mathbf{N}_Z|U) - h(\mathbf{X} + \mathbf{N}^*|U) + \frac{1}{2} \log \frac{|\mathbf{K}^* + \boldsymbol{\Sigma}^*|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_2|} \tag{276}$$

$$= \frac{1}{2} \int_{\boldsymbol{\Sigma}^*}^{\boldsymbol{\Sigma}_Z} \mathbf{J}(\mathbf{X} + \mathbf{N}|U) \, d\boldsymbol{\Sigma}_N + \frac{1}{2} \log \frac{|\mathbf{K}^* + \boldsymbol{\Sigma}^*|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_2|} \tag{277}$$

$$\leq \frac{1}{2} \int_{\boldsymbol{\Sigma}^*}^{\boldsymbol{\Sigma}_Z} (\mathbf{K}^* + \boldsymbol{\Sigma}_N)^{-1} d\boldsymbol{\Sigma}_N + \frac{1}{2} \log \frac{|\mathbf{K}^* + \boldsymbol{\Sigma}^*|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_2|} \tag{278}$$

$$\leq \frac{1}{2} \log \frac{|\mathbf{K}^* + \boldsymbol{\Sigma}_Z|}{|\mathbf{K}^* + \boldsymbol{\Sigma}_2|} \tag{279}$$

where (276) is due to (269), (277) is obtained by using Lemma 6, and (278) comes from Lemma 5 by noting (274). Since (279) is the desired inequality in (75), this completes the proof.

## APPENDIX XII
## PROOF OF THEOREM 12

We first establish the desired bound on $R_2$ given in (80) as follows:

$$R_2 \leq \min_{t=1,\ldots,K_Z} I(U; \mathbf{Y}^2) - I(U; \mathbf{Z}_t) \tag{280}$$

$$= \min_{t=1,\ldots,K_Z} h(\mathbf{Y}^2) - h(\mathbf{Z}_t) + \left[h(\mathbf{Z}_t|U) - h(\mathbf{Y}^2|U)\right] \tag{281}$$

$$\leq \min_{t=1,\ldots,K_Z} \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}^2|}{|\mathbf{S} + \boldsymbol{\Sigma}_t^Z|} + \left[h(\mathbf{Z}_t|U) - h(\mathbf{Y}^2|U)\right] \tag{282}$$

where (280) comes from Theorem 7 by noting the Markov chain $U \to \mathbf{Y}^2 \to \mathbf{Z}_t$, and (282) can be obtained by using the worst additive noise lemma, i.e., Lemma 2, as it is done in the proof of Theorem 6. We now use Theorem 11. According to Theorem 11, for any selection of $(U, \mathbf{X})$, there exists a positive semidefinite matrix $\mathbf{K}$ such that $\mathbf{K} \preceq \mathbf{S}$ and

$$h(\mathbf{Z}_t|U) - h(\mathbf{Y}^2|U) \leq \frac{1}{2} \log \frac{|\mathbf{K} + \boldsymbol{\Sigma}_t^Z|}{|\mathbf{K} + \boldsymbol{\Sigma}^2|} \tag{283}$$

$$h(\mathbf{Y}^2|U) - h(\mathbf{Y}_j^1|U) \geq \frac{1}{2} \log \frac{|\mathbf{K} + \boldsymbol{\Sigma}^2|}{|\mathbf{K} + \boldsymbol{\Sigma}_j^1|} \tag{284}$$

for any $(j, t)$ pair. Using (283) in (280) yields

$$R_2 \leq \min_{t=1,\ldots,K_Z} \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}^2|}{|\mathbf{K} + \boldsymbol{\Sigma}^2|} - \frac{1}{2} \log \frac{|\mathbf{S} + \boldsymbol{\Sigma}_t^Z|}{|\mathbf{K} + \boldsymbol{\Sigma}_t^Z|} \tag{285}$$

which is the desired bound on $R_2$ given in (80).

We now obtain the desired bound on $R_1$ given in (79) as follows:

$$R_1 \leq \min_{j=1,\ldots,K_1} I(\mathbf{X}; \mathbf{Y}_j^1|U) - I(\mathbf{X}; \mathbf{Y}^2|U) \qquad (286)$$

$$= \min_{j=1,\ldots,K_1} h(\mathbf{Y}_j^1|U) - h(\mathbf{Y}^2|U) - \frac{1}{2}\log\frac{|\boldsymbol{\Sigma}_j^1|}{|\boldsymbol{\Sigma}^2|} \quad(287)$$

$$\leq \min_{j=1,\ldots,K_1} \frac{1}{2}\log\frac{|\mathbf{K}+\boldsymbol{\Sigma}_j^1|}{|\boldsymbol{\Sigma}_j^1|} - \frac{1}{2}\log\frac{|\mathbf{K}+\boldsymbol{\Sigma}^2|}{|\boldsymbol{\Sigma}^2|} \quad(288)$$

where (286) comes from Theorem 7 by noting the Markov chain $U \to \mathbf{X} \to \mathbf{Y}_j^1 \to \mathbf{Y}^2$ and (288) is obtained by using (284). Since (288) is the desired bound on $R_1$ given in (79), this completes the proof.

## REFERENCES

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

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

[3] 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.

[4] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453–2469, Jun. 2008.

[5] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "The secrecy rate region of the broadcast channel," presented at the 46th Annu. Allerton Conf. Commun., Contr. Comput., Sep. 2008, (also available at [arXiv:0806.4200]).

[6] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "Secrecy rate region of the broadcast channel with an eavesdropper," *IEEE Trans. Inf. Theory*, Oct. 2009, (also available at [arXiv:0910.3658]), submitted for publication.

[7] E. Ekrem and S. Ulukus, "On secure broadcasting," presented at the 42nd Asilomar Conf. Signals, Syst. Comput., Oct. 2008.

[8] E. Ekrem and S. Ulukus, "Secrecy capacity of a class of broadcast channels with an eavesdropper," *EURASIP J. Wireless Commun. Netw.,*, vol. 2009, Oct. 2009.

[9] Y.-K. Chia and A. E. Gamal, "3-receiver broadcast channels with common and confidential messages," presented at the IEEE Int. Symp. Inf. Theory, Jul. 2009, (also available at [arXiv:0910.1407]).

[10] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "The secrecy capacity region of the Gaussian MIMO broadcast channel," *IEEE Trans. Inf. Theory*, Mar. 2009, (also available at [arXiv:0903.3261]), submitted for publication.

[11] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wire-tap channels," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, Oct. 2009.

[12] H. Yamamoto, "Coding theorem for secret sharing communication systems with two noisy channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 3, pp. 572–578, May 1989.

[13] H. Yamamoto, "A coding theorem for secret sharing communication systems with two Gaussian wiretap channels," *IEEE Trans. Inf. Theory*, vol. 37, no. 3, pp. 634–638, May 1991.

[14] P. Wang, G. Yu, and Z. Zhang, "On the secrecy capacity of fading wireless channel with multiple eavesdroppers," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 1301–1305.

[15] T. Liu, V. Prabhakaran, and S. Viswanath, "The secrecy capacity of a class of parallel Gaussian compound wiretap channels," in *IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 116–120.

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

[17] M. Costa, "A new entropy power inequality," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 6, pp. 751–760, Nov. 1985.

[18] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "A vector generalization of Costa's entropy-power inequality with applications," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1865–1879, Apr. 2010.

[19] 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.

[20] A. J. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon," *Inf. Control*, no. 2, pp. 101–112, Jun. 1959.

[21] N. M. Blachman, "The convolution inequality for entropy powers," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 2, pp. 267–271, Apr. 1965.

[22] 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.

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

[24] D. P. Palomar and S. Verdu, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 141–154, Jan. 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.