Private Broadcasting over Independent Parallel Channels

Ashish Khisti *Member, IEEE,* and Tie Liu *Member, IEEE*

***Abstract*—We study broadcasting of two confidential messages to two groups of receivers over independent parallel sub-channels. One group consists of an arbitrary number of receivers, interested in a common message, whereas the other group has only one receiver. Each message must be confidential from the receiver(s) in the other group. Each of the sub-channels is assumed to be degraded in a certain fashion. While corner points of the capacity region of this setup were characterized in earlier works, we establish the complete capacity region, and show the optimality of a superposition coding technique. For Gaussian channels we establish the optimality of a Gaussian input distribution by applying an extremal information inequality. By extending our coding scheme to block-fading channels we demonstrate significant performance gains over a baseline time-sharing scheme.**

I. Introduction

There has been a considerable interest in the study of secure communication over wireless channels. An information theoretic model of secure communication is the wiretap channel [1], [2]. While the setting of the Gaussian wiretap channel requires that the the legitimate receiver's channel be stronger than the eavesdropper's channel [3] for the capacity to be non-zero, the fading channel does not impose such a condition [4]–[9]. By adapting power and rate of the codebooks based on the fading channel states, positive secrecy rates are achievable even when the legitimate receiver is weaker on average than the eavesdropper. In the present work, we study a new coding scheme for broadcasting confidential messages to two groups of receivers over fading channels. We will refer to this setup as *private broadcasting*. We will focus on the special case when there are an arbitrary number of receivers, say K, in one group, but only a single receiver in the second group.

We first consider private broadcasting over M independent, degraded parallel channels with two groups of receivers as above. We establish the capacity region and show that it reduces to previously known results at the corner points. When the group 2 message has zero rate, it reduces to broadcasting of a common confidential message to K receivers, in the presence of a single eavesdropper studied in [7]. When the group 1 message has zero rate it reduces to transmitting a confidential message to a single receiver in the presence of K eavesdroppers studied in [10]. In this work we generalize [7], [10] and establish the entire capacity region. A natural motivation for considering the setup of independent, degraded parallel channels is its application to Gaussian broadcast channels. Accordingly we extend our result to parallel Gaussian channels, under a power constraint, and establish the optimality of a Gaussian input distribution using an extremal information inequality. The proof involves obtaining a Lagrangian dual for every boundary point of the capacity region and then using an extremal inequality [11] to show that the expression is maximized using Gaussian inputs. We also present an extension to block-fading channels by suitably quantizing the continuous valued channel gains and demonstrate numerically that our proposed scheme can significantly improve upon a baseline time-sharing scheme.

Our coding scheme is based on a superposition technique. We use a secure-multicast codebook [7] for the message for group 1, and a secure-product codebook [10] for the message for group 2. Each codeword, in both these codebooks, consists of M sub-codewords, one corresponding to each parallel channel. In our superposition construction, the codeword from the secure-product codebook constitutes the base codeword whereas the codeword from the secure-multicast codebook constitutes the satellite codeword. We show that this way of combining the secure-product and secure-multicast codebooks achieves the entire capacity region, and discuss the intuition for the optimality of this structure.

In related work, references [12]–[14] study private broadcasting when there is one receiver in each group. References [15], [16] study private broadcasting with feedback over erasure and MIMO broadcast channels. Reference [17] studies interference alignment techniques for private broadcasting when there is a multi-antenna transmitter and arbitrary number of single-antenna receivers in each group. In contrast to [17], the present setup considers a parallel channel model but in the special case when there is a single receiver in group 2, but an arbitrary number of receivers in group 1. The general case when there are multiple receivers in both groups appears to be considerably more difficult. To our knowledge even the corner points of the capacity region (corresponding to one of the messages having zero rate) are not known. Furthermore even in the special setup treated in this paper, when the mutual secrecy constraint is not imposed, the capacity region is not known.

In the remainder of this paper, we present the problem setup

A. Khisti is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: akhisti@comm.utoronto.ca). T. Liu is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA (e-mail: tieliu@tamu.edu).

A. Khisti's work was supported by an NSERC (Natural Sciences Engineering Research Council) Discovery Grant and an Ontario Early Researcher Award. T. Liu was supported by the National Science Foundation under Grants CCF-08-45848 and CCF-09-16867. A portion of this work was presented at ISIT 2012.

TABLE I: Summary of Notation

User Parameters	Number of Users (Group 1)	K
	User Index (Group 1)	k
Parallel Channel Model	Number of Channels	M
	Sub-channel Index	i
	Channel Input Symbol	x_i
	Channel Output Symbol (Group 1)	$y_{k,i}$
	Channel Output Symbol (Group 2)	z_i
	Time Index	t
Gaussian Channel Model	Noise Power (Group 1 User)	$\sigma_{k,i}^2$
	Noise Power (Group 2 User)	δ_i^2
	Input Power Contraint	P_i
Fading Channel Model	Number of Coherence Blocks	M
	Coherence Block Index	i
	Length of Coherence Block	n
	Time Index within Coherence Block	t
	Channel Gain (Group 1 User)	$h_{k,i}$
	Channel Gain (Group 2 User)	g_i

and the summary of the main results in Section II and present a review of the secure-product code and secure-multicast code in Section III. We present our superposition construction in Section IV and the corresponding converse in Section V. Extensions to Gaussian channels and fading channels appear in Sections VI and VII respectively and conclusions appear in Section VIII.

Throughout in this paper, we use the sans serif font e.g., x to denote a random variable and the standard font e.g., x to denote its realization. We use bold font to denote vectors $\mathbf{x} = (x_1, \ldots, x_M)$. The notation $x_i(t)$ denotes the input symbol over sub-channel i at time t. The output at receiver k over sub-channel i at time t will be denoted using $y_{k,i}(t)$. Table I provides a summary of the notation used in the paper.

II. Problem Statement and Main Results

In this section we present the capacity result for three models — independent parallel channels in Subsection II-A, additive Gaussian noise channels with a power constraint in Subsection II-B, and fading channels in Subsection II-C respectively.

A. Independent Parallel Channels

Our setup involves two groups of receivers, with K receivers in group 1 but only one receiver in group 2. We assume that the channel can be decomposed into M independent and parallel channels. The output symbols at receiver k in group 1 across the M sub-channels is denoted by

$$\mathbf{y}_k = (y_{k,1}, y_{k,2}, \ldots, y_{k,M}), \quad k = 1, 2 \ldots, K, \tag{1}$$

whereas the output symbols of the group 2 receiver across the M sub-channels are denoted by

$$\mathbf{z} = (z_1, z_2, \ldots, z_M). \tag{2}$$

The channel input symbols are denoted by $\mathbf{x} = (x_1, \ldots, x_M)$. Each sub-channel is an independent discrete memoryless broadcast channel with the following degradation order:

$$\begin{aligned} &x_i \to y_{k,i} \to z_i, \quad \forall k \in \mathcal{Y}_i \\ &x_i \to z_i \to y_{k,i}, \quad \forall k \in \mathcal{Z}_i. \end{aligned} \tag{3}$$

Note that on each channel i, the receivers in group 1 can be partitioned into two groups. Those in set $\mathcal{Y}_i$ that are "stronger" than the group 2 receiver and those in set $\mathcal{Z}_i$ that are "weaker" than the group 2 receiver. No additional degradation across the users in in group 1 is required in our setup, although such an ordering naturally exists in Gaussian broadcast channels.

We intend to transmit message m_1 to all receivers in group 1, while the message m_2 must be transmitted to the receiver in group 2. A length-n private broadcast code encodes a message pair $(m_1, m_2) \in [1, 2^{nR_1}] \times [1, 2^{nR_2}]$ into a sequence $\mathbf{x}^n \triangleq (\mathbf{x}(1), \ldots, \mathbf{x}(n))$, where $\mathbf{x}(t) \triangleq (x_1(t), \ldots, x_M(t))$ denotes[1] the input across the M sub-channels at time t. The receiver k in group 1 decodes $\hat{m}_{1,k} = g_{1,k}(\mathbf{y}_k^n)$, and the receiver in group 2 decodes $\hat{m}_2 = g_2(\mathbf{z}^n)$. A rate pair (R_1, R_2) is achievable if there exists a private broadcast code such that $\Pr(m_1 \neq \hat{m}_{1,k}) \leq \varepsilon_n$, and $\Pr(m_2 \neq \hat{m}_2) \leq \varepsilon_n$, and furthermore the secrecy constraints

$$\frac{1}{n} I(m_1; \mathbf{z}^n) \leq \varepsilon_n, \quad \frac{1}{n} I(m_2; \mathbf{y}_k^n) \leq \varepsilon_n, \quad k = 1, 2, \ldots, K, \tag{4}$$

are also satisfied. Here $\{\varepsilon_n\}$ is a sequence, indexed by n that approaches zero as $n \to \infty$. The capacity region consists of the closure of all rate pairs (R_1, R_2) achieved by some private broadcast code.

Theorem 1: Let auxiliary variables $\{u_i\}_{1 \leq i \leq M}$ satisfy the Markov condition $u_i \to x_i \to (y_{1,i}, \ldots, y_{K,i}, z_i)$. The capacity region is given by the closure of all rate pairs (R_1, R_2) that satisfy the following constraints:

$$R_1 \leq \min_{1 \leq k \leq K} \left\{ \sum_{i=1}^{M} I(x_i; y_{k,i} | u_i, z_i) \right\} \tag{5}$$

$$R_2 \leq \min_{1 \leq k \leq K} \left\{ \sum_{i=1}^{M} I(u_i; z_i | y_{k,i}) \right\} \tag{6}$$

for some distributions $\{p_{u_i, x_i}\}_{1 \leq i \leq M}$. The alphabet of u_i satisfies the cardinality constraint $|\mathcal{U}_i| \leq |\mathcal{X}_i| + 2K - 1$. □

The coding theorem and converse for Theorem 1 are presented in section IV and V respectively.

B. Gaussian Channels

Consider the discrete-time real Gaussian model where the channel output over sub-channel i at time t is given by:

$$y_{k,i}(t) = x_i(t) + n_{k,i}(t) \tag{7}$$

$$z_i(t) = x_i(t) + w_i(t) \tag{8}$$

The additive noise symbols $n_{k,i}(t)$ are sampled i.i.d. $\mathcal{N}(0, \sigma_{k,i}^2)$ for each $t = 1, 2, \ldots, n$ and are independent

[1] The sequences $\mathbf{y}_k^n$ and $\mathbf{z}^n$ are defined similarly.

across the sub-channels. Similarly $w_i(t)$ is also sampled i.i.d. $\mathcal{N}(0, \delta_i^2)$ and is independent across the sub-channels. Since the capacity region of the channel only depends on the marginals of the additive noise $(n_{1,i}(t), \dots, n_{K,i}(t), w_i(t))$ and that Gaussian variables are infinitely divisible, without loss of generality we may assume that for each sub-channel i the receivers are degraded as in (3). We shall consider both the per sub-channel power constraint (almost surely [18, pp. 552, Ex. 22.2])

$$\frac{1}{n}\sum_{t=1}^{n} x_i^2(t) \le P_i, \quad \forall i = 1, \dots, M \tag{9}$$

and the sum-power constraint

$$\sum_{i=1}^{M} P_i \le P. \tag{10}$$

Theorem 2: The capacity region under the per sub-channel average power constraint (9) is given by the union of all rate pairs (R_1, R_2) that satisfy the following constraints:

$$R_1 \le \min_{1 \le k \le K} \left\{ \sum_{i=1}^{M} A_{k,i}^{(1)}(\mathbf{Q}) \right\} \tag{11}$$

$$R_2 \le \min_{1 \le k \le K} \left\{ \sum_{i=1}^{M} A_{k,i}^{(2)}(\mathbf{Q}) \right\} \tag{12}$$

for some power vector $\mathbf{Q} = (Q_1, \dots, Q_M)$, where $0 \le Q_i \le P_i$ for all $i = 1, \dots, M$,

$$A_{k,i}^{(1)}(\mathbf{Q}) := \left[\frac{1}{2}\log\left(\frac{Q_i + \sigma_{k,i}^2}{\sigma_{k,i}^2}\right) - \frac{1}{2}\log\left(\frac{Q_i + \delta_i^2}{\delta_i^2}\right) \right]^+ \tag{13}$$

$$A_{k,i}^{(2)}(\mathbf{Q}) := \left[\frac{1}{2}\log\left(\frac{P_i + \delta_i^2}{Q_i + \delta_i^2}\right) - \frac{1}{2}\log\left(\frac{P_i + \sigma_{k,i}^2}{Q_i + \sigma_{k,i}^2}\right) \right]^+ \tag{14}$$

and $x^+ := \max\{x, 0\}$. □

A proof of Theorem 2 is provided in section VI.

Corollary 1: The capacity region under the total average power constraint (10) is given by the union of all rate pairs (R_1, R_2) that satisfy the constraints (11) and (12) for some power vectors $\mathbf{P} = (P_1, \dots, P_M)$ and $\mathbf{Q} = (Q_1, \dots, Q_M)$, where $0 \le Q_i \le P_i$ for all $i = 1, \dots, M$ and $\sum_{i=1}^{M} P_i \le P$. □

The above Corollary follows directly from Theorem 2 and the well-known connection between the per sub-channel and the total average power constraints. We will not provide a proof of Corollary 1.

C. Fading Channels

We consider a block-fading channel model with a coherence period of n complex symbols. The channel output in coherence block i is given by

$$\begin{aligned} y_{k,i}(t) &= h_{k,i} \cdot x_i(t) + n_{k,i}(t) \\ z_{k,i}(t) &= g_i \cdot x_i(t) + w_{k,i}(t) \end{aligned} \tag{15}$$

where the channel gains $h_{k,i}$ of the K receivers in group 1, and the channel gain g_i of the group 2 receiver are sampled independently in each coherence block $i \in \{1, 2, \dots, M\}$, and stay constant throughout the block. The coherence period will be taken to be sufficiently large, so that random coding arguments can be invoked in each coherence block. We impose a long-term power constraint (almost surely)

$$\frac{1}{Mn}\sum_{i=1}^{M}\sum_{t=1}^{n} |x_i(t)|^2 \le P. \tag{16}$$

We assume that all the additive noise variables in (15) are sampled i.i.d. $\mathcal{CN}(0,1)$. We are interested in the ergodic communication scenario where the number of blocks M used for communication can be arbitrarily large. Furthermore we assume that the channel gains in each coherence block are revealed to all terminals including the transmitter at the beginning of each coherence block.

Theorem 3: The private broadcasting capacity region for the fading channel model consists of all rate pairs (R_1, R_2) that satisfy the following constraints:

$$R_1 \le \min_{1 \le k \le K} E\left[\left\{ \log\left(\frac{1 + Q(\mathbf{h}, g)|h_k|^2}{1 + Q(\mathbf{h}, g)|g|^2}\right) \right\}^+ \right], \tag{17}$$

$$\begin{aligned} R_2 \le \min_{1 \le k \le K} E\Bigg[\Bigg\{ &\log\left(\frac{1 + P(\mathbf{h}, g)|g|^2}{1 + Q(\mathbf{h}, g)|g|^2}\right) \\ &- \log\left(\frac{1 + P(\mathbf{h}, g)|h_k|^2}{1 + Q(\mathbf{h}, g)|h_k|^2}\right) \Bigg\}^+ \Bigg], \end{aligned} \tag{18}$$

for some power allocation functions $P(\mathbf{h}, g)$ and $Q(\mathbf{h}, g)$ that satisfy $0 \le Q(\mathbf{h}, g) \le P(\mathbf{h}, g)$ for all $(\mathbf{h}, g) \in \mathbb{C}^{K+1}$, and $E[P(\mathbf{h}, g)] \le P$, where $\mathbf{h} := (h_1, \dots, h_K)$ denotes the channel gains of the receivers in group 1. □

A proof of Theorem 3 is provided in Section VII.

We note that in our discussion of fading channels we assume that perfect channel state information of the fading gains is available. Thus our setup considers honest but curious participants in the secrecy analysis. A scenario where such an assumption can be valid is when the average signal-to-noise-ratio of each user is known, e.g., when the users are stationary. In such a setting a receiver cannot consistently report a lower channel gain, since the average SNR will be lower than the true value known to the transmitter and thus be detected. We note however that an exhaustive treatment of malicious users is outside the scope of this paper.

Theorems 1, 2 and 3 constitute the main results in this paper.

III. Background

In this section we review two coding techniques from earlier works that will be utilized in our code construction. The first technique is the secure-product codebook construction, which achieves the corner point of the capacity region when $R_1 = 0$. The second technique is the secure-multicast codebook construction, which achieves the corner point of the capacity region when $R_2 = 0$. The review of these techniques is essential in our superposition construction that achieves the entire capacity region.

A. Secure-Product Codebook Construction

The secure-product codebook construction, introduced in [10], treats the case of independent parallel channels, with one legitimate receiver and multiple, say K, eavesdroppers. This corresponds to the corner point when $R_1 = 0$ in our setup in Section II-A. We first state the capacity associated with this corner point which was established in [10].

Proposition 1: The secrecy capacity associated with communicating a confidential message to the group 2 receiver, and treating all the group 1 receivers as eavesdroppers, in the parallel channel model in Section II-A is given by:

$$C_2 = \max_{p_{\mathsf{x}_1}(\cdot),\dots,p_{\mathsf{x}_M}(\cdot)} \min_{1\le k\le K} \sum_{i=1}^{M} I(\mathsf{x}_i;\mathsf{z}_i|\mathsf{y}_{k,i}) \tag{19}$$

□

We note that the converse essentially follows from a "pair-wise bound". For each eavesdropper considered separately we can show that the capacity is upper bounded by the right hand side in (19), and then take the worst eavesdropper. The achievability is based on a product codebook construction as discussed next. We note that a straightforward vector extension of the wiretap codebook [1], [2] to the parallel channel model results in the following rate:

$$R^{-} = \max_{p_{\mathsf{x}_1}(\cdot),\dots,p_{\mathsf{x}_M}(\cdot)} \min_{1\le k\le K} \sum_{i=1}^{M} \{I(\mathsf{x}_i;\mathsf{z}_i) - I(\mathsf{x}_i;\mathsf{y}_{k,i})\} \tag{20}$$

which is smaller than the capacity (19).

The product-codebook construction involves a capacity-achieving codebook for each parallel channel. For each sub-channel $i \in \{1,\dots,M\}$ we consider all binary sequences of length $N_i = n[I(\mathsf{x}_i;\mathsf{z}_i) - 2\varepsilon]$ i.e., we let $\mathcal{M}_i = \{0,1\}^{N_i}$. On channel i we generate a codebook $\mathcal{C}_i : \mathcal{M}_i \to \mathcal{X}^n$ consisting of 2^{N_i} codewords, i.e.,

$$\mathcal{C}_i = \{x_i^n(\bar{m}_i) : \bar{m}_i \in \mathcal{M}_i\}, \tag{21}$$

where we assume that each sequence x_i^n is sampled i.i.d. from a distribution $p_{\mathsf{x}_i}(\cdot)$ and the codebooks are revealed to all the terminals. We consider the cartesian product of such binary sequences on all sub-channels:

$$\begin{aligned}\mathcal{M} &= \mathcal{M}_1 \times \mathcal{M}_2 \dots \times \mathcal{M}_M &(22)\\ &= \{(\bar{m}_1,\dots,\bar{m}_M) : \bar{m}_i \in \mathcal{M}_i, i = 1,2,\dots,M\} &(23)\end{aligned}$$

and partition $\mathcal{M}$ into 2^{nR} bins, such that each bin has $\frac{|\mathcal{M}|}{2^{nR}}$ elements. Fig. 1 illustrates the secure-product codebook construction for the case when $M = 2$.

Given a message m, the encoder selects an element $(\bar{m}_1,\dots,\bar{m}_M) \in \mathcal{M}$ uniformly at random from the bin associated with m. It transmits the associated codeword sequences $(x_1^n(\bar{m}_1),\dots,x_M^n(\bar{m}_M))$ on the M corresponding sub-channels. We note that in this construction all the sub-messages $\bar{m}_j$ are selected to be independent, i.e.,

$$\Pr(\bar{m}_1 = \bar{m}_1,\dots,\bar{m}_M = \bar{m}_M) = \prod_{j=1}^{M} \Pr(\bar{m}_j = \bar{m}_j) = \frac{1}{|\mathcal{M}_1| \times |\mathcal{M}_2| \dots, |\mathcal{M}_M|}. \tag{24}$$

At the receiver, each codeword $x_i^n(\bar{m}_i)$ can be decoded with high probability from z_i^n, since the rate of each $\mathcal{C}_i$ in (21) is selected to be smaller than $I(\mathsf{x}_i;\mathsf{z}_i)$. In turn the receiver decodes the message bin with high probability.

In the secrecy analysis we exploit the independence of sub-messages in (24) to argue that the information leakage rate to the receiver k of group 1 on sub-channel i equals $\min\{I(\mathsf{x}_i;\mathsf{z}_i), I(\mathsf{x}_i;\mathsf{y}_{k,i})\}$. Since the rate of each sub-message equals $I(\mathsf{x}_i;\mathsf{z}_i)$, the total equivocation rate is given by (19). We omit the formal secrecy analysis in this paper.

B. Secure-Multicast Codebook Construction

The secure-multicast codebook construction, introduced in [7], treats the case of independent parallel channels, with K legitimate receivers and one eavesdropper. This corresponds to the corner point when $R_2 = 0$ in our setup in Section II-A. We first state the capacity associated with this corner point which was established in [7].

Proposition 2: The secrecy capacity associated with communicating a common confidential message to the group 1 receivers, and treating the group 2 receiver as an eavesdropper, in the parallel channel model in Section II-A is given by:

$$C_1 = \max_{p_{\mathsf{x}_1}(\cdot),\dots,p_{\mathsf{x}_M}(\cdot)} \min_{1\le k\le K} \sum_{i=1}^{M} I(\mathsf{x}_i;\mathsf{y}_{k,i}|\mathsf{z}_i). \tag{25}$$

□

In establishing (25) we note that the upper bound is again a "pair-wise bound". We consider each receiver separately and can show that the capacity is upper bounded by the right hand side in (25). The achievability is based on a multicast codebook construction as discussed below. We note that a straightforward extension of the wiretap codebook [1], [2] to the parallel channel setup results in the following rate:

$$R^{-} = \max_{p_{\mathsf{x}_1}(\cdot),\dots,p_{\mathsf{x}_M}(\cdot)} \min_{1\le k\le K} \sum_{i=1}^{M} \{I(\mathsf{x}_i;\mathsf{y}_{k,i}) - I(\mathsf{x}_i;\mathsf{z}_i)\} \tag{26}$$

which is sub-optimal. In the secure-multicast coding scheme, for each message, we sample a total of $L_i = 2^{n[I(\mathsf{x}_i;\mathsf{z}_i)+\varepsilon]}$ codewords on sub-channel i:

$$\mathcal{C}_i(m) = \{x_i^n(m,l_i), l_i \in [1,L_i]\}, \quad m \in \{1,2,\dots,2^{nR}\}, i \in \{1,\dots,M\}. \tag{27}$$

The construction of a secure-multicast codebook for the case of $M = 2$ parallel channels is shown in Fig. 2.

Each codeword in $\mathcal{C}_i(m)$ is sampled i.i.d. from $p_{\mathsf{x}_i}(\cdot)$ and the codebooks are revealed to all the terminals. When encoding message m, the encoder selects an index l_i, uniformly at random from the set $\{1,\dots,L_i\}$ and transmits the associated sequence $x_i^n(m,l_i)$ from $\mathcal{C}_i(m)$ on sub-channel i. Each

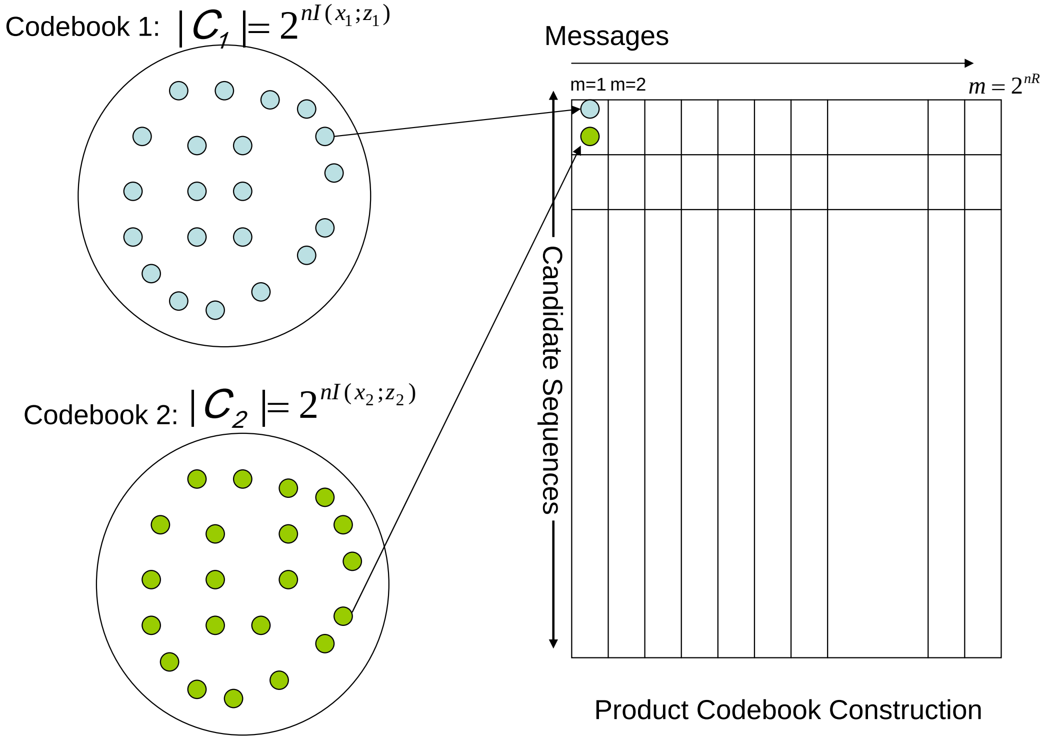
Codebook 1: $|C_1| = 2^{nI(x_1;z_1)}$
Messages
m=1 m=2
$m = 2^{nR}$
Candidate Sequences
Codebook 2: $|C_2| = 2^{nI(x_2;z_2)}$
Product Codebook Construction

Fig. 1: Product Codebook Construction for $M = 2$ parallel channels. We generate two codebooks $\mathcal{C}_1$ and $\mathcal{C}_2$ consisting of $\approx 2^{nI(x_1;z_1)}$ and $\approx 2^{nI(x_2;z_2)}$ codewords respectively and partition the set $\{\mathcal{C}_1 \times \mathcal{C}_2\}$ into 2^{nR} bins as shown. Each message corresponds to one bin index. There are approximately $2^{nI(x_1;z_1)+nI(x_2;z_2)-nR}$ sequence pairs per bin. One of these is selected uniformly at random and transmitted over the corresponding sub-channel.

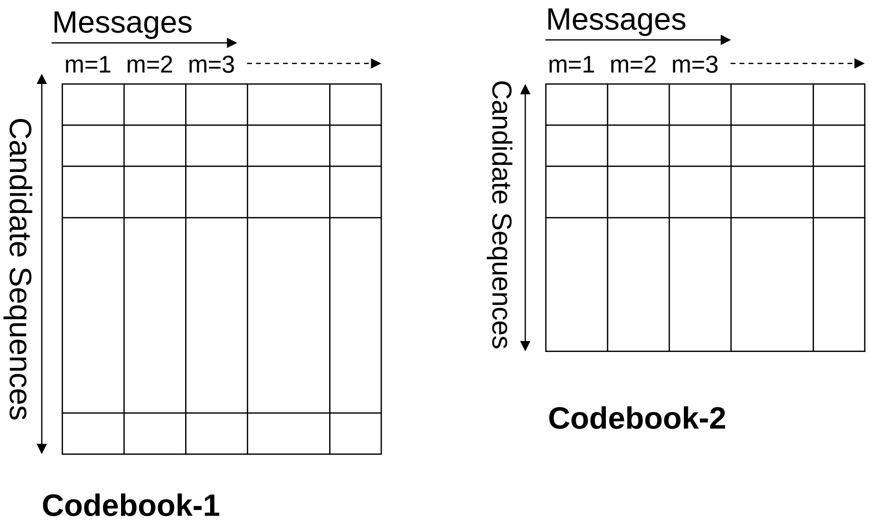

Fig. 2: Secure-Multicast Codebook Construction for $M = 2$ parallel channels. We generate two codebooks $\mathcal{C}_1$ and $\mathcal{C}_2$ one for each sub-channel. Each codebook consists of 2^{nR} messages. For each message, we generate $\approx 2^{nI(x_1;z_1)}$ sequences in codebook $\mathcal{C}_1$ and $\approx 2^{nI(x_2;z_2)}$ sequences in codebook $\mathcal{C}_2$.

legitimate receiver searches for a message m and indices $(l_1, \dots, l_M)$ such that $(x_i^n(m, l_i), y_{k,i}^n) \in T_\varepsilon^n(x_i, y_{k,i})$ are jointly typical. It can be shown that for any rate that is below the capacity in (25), the error probability at each receiver goes to zero.

For the secrecy analysis, we note that for the proposed L_i, we satisfy the secrecy condition

$$H(m|z_i^n) \geq H(m) - n\varepsilon_n$$

for a suitable sequence ε_n that goes to zero as $n \to \infty$. Furthermore since the input sequences

$$\{x_1^n(m, l_1), \dots, x_M^n(m, l_M)\},$$

are conditionally independent given m, it can be shown [7] that even when the eavesdropper combines all its channel outputs $(z_1^n, \dots, z_M^n)$, the secrecy condition remains satisfied i.e.,

$$H(m|z_1^n, \dots, z_M^n) \geq H(m) - n\varepsilon_n'$$

for a suitable sequence ε_n'. We again omit the formal secrecy analysis as it will be considered in the more general superposition coding scheme.

IV. Coding Theorem

Our coding scheme is based upon a superposition approach. The base layer consists of codewords of the secure-product codebook discussed in Section III-A which is used to encode message m_2. The satellite codewords correspond to a secure-multicast codebook, discussed in Section III-B, which is used to encode message m_1.

In our discussion we fix auxiliary variables $(u_1, \dots u_M)$ and the distributions $p_{x_i|u_i}(\cdot)$. The message m_2 for the group 2 receiver is encoded using a secure-product codebook. Let $\mathcal{M}_{2,i}$ be the set of all binary sequences of length $N_{2,i} = n(I(u_i; z_i) - 2\varepsilon)$ i.e., $\mathcal{M}_{2,i} := \{0,1\}^{N_{2,i}}$. On channel i, we generate a codebook $\mathcal{C}_{2,i} : \mathcal{M}_{2,i} \to \mathcal{U}_i^n$ consisting of $|\mathcal{M}_{2,i}|$ codewords, i.e.,

$$\mathcal{C}_{2,i} := \{u_i^n(\bar{m}_{2,i}) : \bar{m}_{2,i} \in \mathcal{M}_{2,i}\}, \tag{28}$$

where each sequence u_i^n is sampled i.i.d. from the distribution $p_{u_i}(\cdot)$. Let

$$\begin{aligned} \mathcal{M}_2 &:= \mathcal{M}_{2,1} \times \mathcal{M}_{2,2} \times \dots \times \mathcal{M}_{2,M} & (29)\\ &= \{(\bar{m}_{2,1}, \dots, \bar{m}_{2,M}) : \bar{m}_{2,i} \in \mathcal{M}_{2,i}, i = 1, \dots, M\}. & (30) \end{aligned}$$

We define $\mathcal{C}_2 = \mathcal{C}_{2,1} \times \mathcal{C}_{2,2} \dots \times \mathcal{C}_{2,M}$ as the overall product codebook associated with $\mathcal{M}_2$. We partition the set $\mathcal{M}_2$ into 2^{nR_2} bins such that there are $L_2 = 2^{n\left\{\sum_{i=1}^M I(u_i;z_i) - R_2 - M\varepsilon\right\}}$ sequences in each bin. Each bin corresponds to one message $m_2 \in [1, 2^{nR_2}]$. Thus given a message m_2 the encoder selects one sequence $(\bar{m}_{2,1}, \dots, \bar{m}_{2,M}) \in \mathcal{M}_2$ uniformly at random from the corresponding bin. On channel i we select the codeword $u_i^n \in \mathcal{C}_{2,i}$ associated with $\bar{m}_{2,i}$. We note that from our construction, each sequence in $\mathcal{M}_2$ is equally likely i.e.,

$$\begin{aligned} \Pr(\bar{m}_{2,1} = \bar{m}_{2,1}, \dots, \bar{m}_{2,M} = \bar{m}_{2,M}) &= \prod_{j=1}^M \Pr(\bar{m}_{2,j} = \bar{m}_{2,j}) \\ &= \frac{1}{|\mathcal{M}_{2,1}| \times |\mathcal{M}_{2,2}| \dots, |\mathcal{M}_{2,M}|}. \end{aligned} \tag{31}$$

The codebook associated with m_1 is a secure-multicast codebook. For each $u_i^n \in \mathcal{C}_{2,i}$, and each $m_1 \in [1, 2^{nR_1}]$ we construct a codebook $\mathcal{C}_{1,i}(u_i^n, m_1)$ consisting of a total of $L_{1,i} = 2^{n(I(x_i;z_i|u_i)+\varepsilon)}$ codeword sequences of length n, each sampled i.i.d. from the distribution $\prod_{j=1}^n p_{x_i|u_i}(x_{ij}|u_{ij})$ and revealed to all the terminals.

Given a message $m_1 \in [1, 2^{nR_1}]$ and codewords $(u_1^n, \dots, u_M^n)$, selected in the base layer, we select the sequence x_i^n from the codebook $\mathcal{C}_{1,i}(u_i^n, m_1)$ corresponding to a randomly and uniformly generated index $l_{1,i}$. The sequence x_i^n is transmitted on sub-channel i.

We let $\mathcal{C}$ to be the overall codebook consisting of the base layers and the refinement layers, which is revealed to all the terminals in the network before the start of the communication. The following property will be useful in our subsequent analysis.

Lemma 1: The sequences $(x_1^n, x_2^n, \dots, x_M^n)$ are conditionally independent given m_1 and the codebook $\mathcal{C}$, i.e.,

$$p\left(x_1^n, x_2^n, \dots, x_M^n | m_1, \mathcal{C}\right) = \prod_{i=1}^M p(x_i^n | m_1, \mathcal{C}). \tag{32}$$

Proof: See Appendix A. ■

A. Decoding and Error Analysis

1) Decoding of Message m_1: Receiver k in group 1 selects those sub-channels $\mathcal{J}_k$ where he is stronger than the group 2 receiver:

$$\mathcal{J}_k = \left\{i \in [1, M] : x_i \to y_{k,i} \to z_i\right\} \tag{33}$$

- For each $i \in \mathcal{J}_k$, receiver k selects a sequence $\hat{u}_i^n \in \mathcal{C}_{2,i}$ such that[2] $(\hat{u}_i^n, y_{k,i}^n) \in T_\varepsilon^n(u_i, y_{k,i})$. We define $\mathcal{E}_k$ as the event that there exists some $i \in \mathcal{J}_k$ such that $\{\hat{u}_i^n \neq u_i^n\}$.
- Receiver k then searches for a message $\hat{m}_1 \in [1, 2^{nR_1}]$ with the following property: for each $i \in \mathcal{J}_k$ there exists a codeword $x_i^n \in \mathcal{C}_{1,i}(m_1, \hat{u}_i^n)$ such that $(x_i^n, y_{k,i}^n) \in T_\varepsilon^n(x_i, y_{k,i}|u_i)$. An error is declared if $\hat{m}_1 \neq m_1$.

Now observe that

$$\Pr(\hat{m}_1 \neq m_1) \leq \Pr(\mathcal{E}_k) + \Pr(\hat{m}_1 \neq m_1 | \mathcal{E}_k^c). \tag{34}$$

Since $|\mathcal{C}_{2,i}| \leq 2^{n(I(u_i;z_i)-\varepsilon)}$ and $I(u_i; y_{k,i}) \geq I(u_i; z_i)$ for each $i \in \mathcal{J}_k$, it follows that $\Pr(\mathcal{E}_k) \leq M\varepsilon$.

[2] We will use the notion of strong typicality. The set $T_\varepsilon^n(x, y)$ denotes the ε-strongly typical set.

To bound the second term in (34) we use the union bound and analysis of typical events.

$$\Pr(\hat{m}_1 \neq m_1 | \mathcal{E}_k^c) \leq 2^{nR_1} \prod_{i \in \mathcal{J}_k} \left\{ |\mathcal{C}_{1,i}| \, 2^{-n(I(x_i;y_{k,i}|u_i)-\varepsilon)} \right\} \tag{35}$$

$$\leq 2^{nR_1} 2^{-n\sum_{i\in\mathcal{J}_k}(I(x_i;y_{k,i}|u_i)-I(x_i;z_i|u_i)-2\varepsilon)} \tag{36}$$

$$= 2^{nR_1} 2^{-n\sum_{i\in\mathcal{J}_k}(I(x_i;y_{k,i}|u_i,z_i)-2\varepsilon)} \tag{37}$$

which goes to zero provided that $R_1 \leq \sum_{i\in\mathcal{J}_k} I(x_i;y_{k,i}|u_i,z_i) - (2M+1)\varepsilon$. Since $\varepsilon > 0$ is arbitrary, our choice of R_1 in (5) thus guarantees that the error probability associated with message m_1 vanishes to zero.

2) Decoding of message m_2: The receiver in group 2 decodes message $\bar{m}_{2,i}$ on sub-channel i by searching for a sequence $u_i^n \in \mathcal{C}_{2,i}$ that is jointly typical with z_i^n. Since the number of codewords in $\mathcal{C}_{2,i}$ does not exceed $2^{n(I(u_i;z_i)-2\varepsilon)}$, this event succeeds with high probability. Hence the receiver correctly decodes $(\bar{m}_{2,1},\ldots,\bar{m}_{2,M})$ and in turn message m_2 with high probability.

B. Secrecy Analysis

In order to establish the secrecy of message m_1 we need to show that

$$\frac{1}{n} I(m_1;\mathbf{z}^n|\mathcal{C}) \leq \varepsilon_n \tag{38}$$

where recall that $\mathcal{C}$ denotes the overall codebook $\{\mathcal{C}_{1,i},\mathcal{C}_{2,i}\}_{1\leq i\leq M}$ in our construction.

Using Lemma 1 and the fact that the channels are independent, we have that $z_1^n,\ldots,z_M^n$ are conditionally independent given m_1 and $\mathcal{C}$, it can be shown that

$$I(m_1;\mathbf{z}^n|\mathcal{C}) \leq \sum_{i=1}^{M} I(m_1;z_i^n|\mathcal{C}). \tag{39}$$

Since in our conditional codebook construction, there are $2^{n(I(x_i;z_i|u_i)+\varepsilon)}$ sequences in each codebook $\mathcal{C}_{1,i}(u_i^n,m_1)$, it follows from standard arguments that $\frac{1}{n}I(m_1;z_i^n|\mathcal{C}) \leq \varepsilon_n$. The secrecy constraint (38) now follows.

To establish secrecy of message m_2 with respect to user 1 in group 1, we show that

$$\frac{1}{n} H(m_2|\mathbf{y}_1^n, m_1, \mathcal{C}) \geq R_2 - \varepsilon_n. \tag{40}$$

where for simplicity we drop the subscript associated with user 1 in the sequence $\mathbf{y}_1^n$. Without loss of generality, we assume that sub-channels $i = 1,2,\ldots,L$ satisfy $x_i \to z_i \to y_i$ while sub-channels $i = L+1,\ldots,M$ satisfy $x_i \to y_i \to z_i$. Now consider

$$H(m_2|\mathbf{y}_1^n, m_1, \mathcal{C}) \tag{41}$$

$$= H(m_2|y_1^n,\ldots,y_M^n,m_1,\mathcal{C}) \tag{42}$$

$$\begin{aligned} &= H(\bar{m}_{2,1}^M|y_1^n,\ldots,y_M^n,m_1,\mathcal{C}) \\ &\qquad - H(\bar{m}_{2,1}^M|m_2,m_1,y_1^n,\ldots,y_M^n,\mathcal{C}) \end{aligned} \tag{43}$$

$$\begin{aligned} &= \sum_{j=1}^{M} H(\bar{m}_{2,j}|y_j^n,m_1,\mathcal{C}) \\ &\qquad - H(\bar{m}_{2,1}^M|m_2,m_1,y_1^n,\ldots,y_M^n,\mathcal{C}) \end{aligned} \tag{44}$$

$$\begin{aligned} &\geq \sum_{j=1}^{L} H(\bar{m}_{2,j}|y_j^n,m_1,\mathcal{C}) \\ &\qquad - H(\bar{m}_{2,1}^M|m_2,m_1,y_1^n,\ldots,y_M^n,\mathcal{C}) \end{aligned} \tag{45}$$

where we have introduced $\bar{m}_{2,1}^M \triangleq (\bar{m}_{2,1},\ldots,\bar{m}_{2,M})$ in (43), (44) follows by establishing that the collection of pairs $\{(m_{2,1},y_1^n),\ldots,(m_{2,M},y_M^n)\}$ is conditionally independent given m_1, which can be establishes in a manner similar to the proof of Lemma 1 and (45) follows from the fact that the entropy function is non-negative and therefore we can drop the terms $L+1,\ldots,M$ in the first summation.

We lower bound the first term in (45). Recall that $\bar{m}_{2,j}$ is uniformly distributed over $\mathcal{C}_{2,j}$ with $|\mathcal{C}_{2,j}| = 2^{n(I(u_j;z_j)-\varepsilon)}$. Furthermore, the corresponding codeword u_j^n is the base codeword in $\mathcal{C}_{1,j}(m_1,u_j^n)$ and

$$\left|\mathcal{C}_{1,j}(m_1,u_j^n)\right| = 2^{n(I(x_j;z_j|u_j)-\varepsilon)} \geq 2^{n(I(x_j;y_j|u_j)-\varepsilon)}, \tag{46}$$

since the channel satisfies the relation $x_j \to z_j \to y_j$ for $j = 1,\ldots,L$. Since the satellite codeword x_j^n is uniformly selected from $\mathcal{C}_{1,j}$ it follows that [18, Remark 22.2, pp. 554-555]

$$\frac{1}{n} H(\bar{m}_{2,j}|y_j^n,m_1,\mathcal{C}) \geq I(u_j;z_j) - I(u_j;y_j) - \varepsilon. \tag{47}$$

and therefore using the fact that $u_j \to z_j \to y_j$, we have

$$\frac{1}{n}\sum_{j=1}^{L} H(\bar{m}_{2,j}|y_j^n,m_1,\mathcal{C}) \geq \sum_{j=1}^{L} I(u_j;z_j|y_j) - L\varepsilon. \tag{48}$$

We next upper bound the second term in (45) and show that

$$\frac{1}{n} H(\bar{m}_{2,1}^M|m_2,m_1,y_1^n,\ldots,y_M^n,\mathcal{C}) \leq \sum_{i=1}^{L} I(u_i;z_i|y_i) - R_2 + \varepsilon. \tag{49}$$

Note that:

$$\begin{aligned} &H(\bar{m}_{2,1}^M|m_2,m_1,y_1^n,\ldots,y_M^n,\mathcal{C}) \leq \\ &\qquad H(\bar{m}_{2,1}^M|m_2,m_1,y_1^n,\ldots,y_L^n,z_{L+1}^n,\ldots z_M^n,\mathcal{C}) \end{aligned} \tag{50}$$

since z_j^n is a degraded version of y_j^n on channels $j \in \{L+1,\ldots,M\}$. Note that for each message m_2, there are a total of $2^{n(\tilde{R}-R_2)}$ different message sequences $\bar{m}_{2,1}^M$ in its associated bin, where:

$$\tilde{R} = \frac{1}{n} H(\bar{m}_{2,1},\ldots,\bar{m}_{2,M}) \tag{51}$$

$$= \sum_{i=1}^{M} \left\{ I(u_i;z_i) - 2\varepsilon \right\}, \tag{52}$$

and furthermore, $R_2 = \frac{1}{n}H(m_2) \leq \sum_{i=1}^{L} I(u_i;z_i|y_i) - (2M+1)\varepsilon$.

Furthermore for each value of $\bar{m}_{2,j}$ and m_1, we select a codeword x_j^n uniformly at random from the codebook

$\mathcal{C}_1(u_j^n, m_1)$. Further using (46), we can conclude that (c.f. [18, Lemma 22.1, Remark 22.2, pp. 554-555], [19, Lemma 1])

$$\begin{aligned}
&\frac{1}{n}H(\bar{m}_{2,1}^M | m_2, m_1, y_1^n, \ldots, y_L^n, z_{L+1}^n, \ldots z_M^n, \mathcal{C}) \\
&\quad \leq \tilde{R} - R_2 - I(u_1, \ldots, u_M; y_1, \ldots, y_L, z_{L+1}, \ldots, z_M) + \varepsilon && (53) \\
&\quad \leq \tilde{R} - R_2 - \sum_{i=1}^{L} I(u_i; y_i) - \sum_{i=L+1}^{M} I(u_i; z_i) + \varepsilon && (54) \\
&\quad = \sum_{i=1}^{L} I(u_i; z_i | y_i) - R_2 - \varepsilon && (55)
\end{aligned}$$

where we use the independence of $(u_1, \ldots, u_M)$ in (54) and substitute (52) for $\tilde{R}$.

Substituting (48) and (55) into (45) we have that

$$\frac{1}{n}H(m_2 | \mathbf{y}_1^n, m_1, \mathcal{C}) \geq R_2 - (L+1)\varepsilon, \tag{56}$$

Since $\varepsilon > 0$ can be arbitrarily small, this establishes the secrecy of message m_2 with respect to user 1 in group 1. The secrecy with respect to every other user can be established in a similar fashion. Finally we note that since the average equivocation over the codebooks $\mathcal{C}$ satisfies the required constraint, there must exist at least one codebook in this ensemble with this property.

We conclude this section with an intuitive explanation behind the optimality of the superposition approach. Note that our approach uses the codewords for the group 2 user as cloud centers and the codewords of the group 1 user as satellite codewords. To explain this, note that on any given channel, say channel i, there is an ordering of receivers as in (3). Receivers in group 1 in the set $\mathcal{Z}_i$ are weaker than the group 2 user. It can be seen that these receivers do not learn any information on channel i. Thus among all the set of active users on any given channel, the group 2 user is the weakest user. Therefore the associated codeword of the group 2 user constitutes the cloud centre.

V. Converse

We first show that there exists a choice of auxiliary variables $u_i(j)$ that satisfies the Markov chain condition in Theorem 1 i.e.,

$$u_i(j) \to x_i(j) \to (y_{1,i}(j), \ldots, y_{K,i}(j), z_i(j)),$$

such that the rates R_1 and R_2 are upper bounded by

$$nR_1 \leq \sum_{i=1}^{M}\sum_{j=1}^{n} I(x_i(j); y_{k,i}(j) | u_i(j), z_i(j)) + 2n\varepsilon_n \tag{57}$$

$$nR_2 \leq \sum_{i=1}^{M}\sum_{j=1}^{n} I(u_i(j); z_i(j) | y_{k,i}(j)) + 2n\varepsilon_n \tag{58}$$

for each $k \in \{1, \ldots, K\}$ and for some sequence $\{\varepsilon_n\}$ that goes to zero as $n \to \infty$.

A. Special case of $K = 2$

We first treat the special case when $K = 2$ and $M = 2$ channels. The study of this special case first will make the notation used in the general case clearer. We assume that the output symbols at receivers 1 and 2 in group 1 are given by $\mathbf{y}_1 \triangleq (y_{1,1}, y_{1,2})$ and $\mathbf{y}_2 \triangleq (y_{2,1}, y_{2,2})$ respectively. Furthermore we assume the sub-channels 1 and 2 satisfy the following degradation order:

$$x_1 \to y_{1,1} \to z_1 \to y_{2,1}, \tag{59}$$

$$x_2 \to y_{2,2} \to z_2 \to y_{1,2} \tag{60}$$

Thus user 1 in group 1 is the strongest user on channel 1, and weakest on channel 2. Likewise user 2 in group 1 is the strongest user on channel 2, and weakest on channel 1. The group 2 user is in the middle in both channels.

For convenience, we define:

$$\bar{\mathbf{y}}_1^n \triangleq (\bar{y}_{1,1}^n, \bar{y}_{1,2}^n) = (z_1^n, y_{1,2}^n) \tag{61}$$

$$\bar{\mathbf{y}}_2^n \triangleq (\bar{y}_{2,1}^n, \bar{y}_{2,2}^n) = (y_{2,1}^n, z_2^n) \tag{62}$$

i.e., $\bar{\mathbf{y}}_1^n$ is obtained by degrading user 1 on channel 1 to z_1^n, and $\bar{\mathbf{y}}_2^n$ is obtained similarly. Clearly from the secrecy constraint of m_2, it also follows that $\frac{1}{n}I(m_2; \bar{\mathbf{y}}_k^n) \leq \varepsilon_n$ for $k = 1, 2$. By combining Fano's inequality and the secrecy constraint, we obtain the following upper bound on R_2:

$$\begin{aligned}
&nR_2 \leq I(m_2; \mathbf{z}^n) - I(m_2; \bar{\mathbf{y}}_1^n) + 2n\varepsilon_n && (63) \\
&\leq I(m_2; z_2^n | y_{1,2}^n, z_1^n) + 2n\varepsilon_n && (64) \\
&\leq \sum_{j=1}^{n} I(m_2; z_2(j) \mid z_2^{j-1}, y_{1,2}^n, z_1^n) + 2n\varepsilon_n && (65) \\
&= \sum_{j=1}^{n} I(m_2; z_2(j) \mid z_2^{j-1}, y_{1,2}^{j-1}, y_{1,2}(j), y_{1,2,j+1}^n, z_1^n) + 2n\varepsilon_n && (66) \\
&\leq \sum_{j=1}^{n} I(m_2, z_2^{j-1}, y_{1,2}^{j-1}, y_{1,2,j+1}^n, z_1^n; z_2(j) \mid y_{1,2}(j)) + 2n\varepsilon_n && (67) \\
&\leq \sum_{j=1}^{n} I(m_2, \bar{\mathbf{z}}_{\{2\}}^{j-1}, \bar{\mathbf{z}}_{\{2\},j+1}^n, \bar{\mathbf{z}}_{\{1\}}^n; z_2(j) \mid y_{1,2}(j)) + 2n\varepsilon_n && (68)
\end{aligned}$$

where the justification of the steps is as follows. Eq. (63) follows by combining the Fano's Inequality and secrecy constraints. Eq. (64) follows by substituting in (61), (65) follows from the chain rule of mutual information, while (67) follows from the non-negativity of the mutual information expression. In (68) we have introduced the following notation:

$$\bar{\mathbf{z}}_{\{1\}}^n \triangleq (z_1^n, y_{2,1}^n) \tag{69}$$

$$\bar{\mathbf{z}}_{\{2\}}^{j-1} \triangleq \left(z_2^{j-1}, y_{1,2}^{j-1}\right) \tag{70}$$

$$\bar{\mathbf{z}}_{\{2\},j+1}^n \triangleq \left(z_{2,j+1}^n, y_{1,2,j+1}^n\right). \tag{71}$$

Here $\bar{\mathbf{z}}_{\{1\}}^n$ denotes the channel output of the group 2 receiver and the output of weaker receiver in group 1, and the other notations are defined similarly. We let

$$u_2(j) \triangleq \left(m_2, \bar{\mathbf{z}}_{\{2\}}^{j-1}, \bar{\mathbf{z}}_{\{2\},j+1}^{n}, \bar{\mathbf{z}}_{\{1\}}^{n}\right) \tag{72}$$

and note that $u_2(j) \rightarrow x_2(j) \rightarrow (y_{2,2}(j), z_2(j), y_{1,2}(j))$ holds. Thus (68) reduces to the following:

$$nR_2 \leq \sum_{j=1}^{n} I(u_2(j); z_2(j)|y_{1,2}(j)) + 2n\varepsilon_n. \tag{73}$$

In a similar fashion we can show that

$$nR_2 \leq \sum_{j=1}^{n} I(u_1(j); z_1(j)|y_{2,1}(j)) + 2n\varepsilon_n \tag{74}$$

where

$$u_1(j) \triangleq \left(m_2, \bar{\mathbf{z}}_{\{1\}}^{j-1}, \bar{\mathbf{z}}_{\{1\},j+1}^{n}, \bar{\mathbf{z}}_{\{2\}}^{n}\right) \tag{75}$$

satisfies the Markov chain $u_1(j) \rightarrow x_1(j) \rightarrow (y_{1,1}(j), z_1(j), y_{2,1}(j))$.

We let q_1 and q_2 be independent and uniformly distributed over the interval $\{1, 2, \ldots, n\}$. Let $u_1 \triangleq (u_1(q_1), q_1)$ and let $u_2 \triangleq (u_2(q_2), q_2)$. We have that

$$R_2 \leq \min\{I(u_1; z_1|y_{2,1}), I(u_2; z_2|y_{1,2})\} + 2\varepsilon_n \tag{76}$$

To obtain an upper bound on R_1 we consider the secrecy constraint with respect to the group 2 receiver and apply Fano's inequality for user 2 in group 1.

$$\begin{aligned}
nR_1 &\leq I(m_1; \mathbf{y}_2^n) - I(m_1; \mathbf{z}^n, m_2) + 2n\varepsilon_n && (77)\\
&\leq I(m_1; y_{2,2}^n|z_1^n, z_2^n, m_2) + 2n\varepsilon_n && (78)\\
&= \sum_{j=1}^{n} I(m_1; y_{2,2}(j)|y_{2,2}^{j-1}, z_1^n, z_2^n, m_2) + 2n\varepsilon_n && (79)\\
&= \sum_{j=1}^{n} H(y_{2,2}(j)|y_{2,2}^{j-1}, z_1^n, z_2^n, m_2)\\
&\quad - H(y_{2,2}(j)|y_{2,2}^{j-1}, z_1^n, z_2^n, m_1, m_2) && (80)\\
&= \sum_{j=1}^{n} H(y_{2,2}(j)|y_{2,2}^{j-1}, z_1^n, z_2^n, m_2)\\
&\quad - H(y_{2,2}(j)|y_{2,2}^{j-1}, z_1^n, z_2^n, m_1, m_2, x_2(j)) && (81)\\
&= \sum_{j=1}^{n} H(y_{2,2}(j)|y_{2,2}^{j-1}, z_1^n, z_2^n, m_2)\\
&\quad - H(y_{2,2}(j)|z_2(j), x_2(j)) && (82)\\
&\leq \sum_{j=1}^{n} H(y_{2,2}(j)|z_1^n, z_2^n, m_2)\\
&\quad - H(y_{2,2}(j)|z_2(j), x_2(j)) && (83)\\
&= \sum_{j=1}^{n} H(y_{2,2}(j)|\bar{\mathbf{z}}_{\{1\}}^{n}, \bar{\mathbf{z}}_{\{2\}}^{j-1}, \bar{\mathbf{z}}_{\{2\},j+1}^{n}, m_2, z_2(j))\\
&\quad - H(y_{2,2}(j)|z_2(j), x_2(j)) && (84)
\end{aligned}$$

where (77) follows upon applying Fano's inequality and secrecy constraints respectively, (78) follows from the degraded structure of the channels. Eq. (82) follows from the fact that the channels are memoryless and independent and hence we have that

$$(y_{2,2}(j), z_2(j)) \rightarrow x_2(j) \rightarrow (y_{2,2}^{j-1}, z_1^n, z_2^{j-1}, z_{2,j+1}^n, m_1) \tag{85}$$

holds. Eq. (83) follows from the fact that conditioning reduces entropy. Finally in (84) we use the property that

$$\left(\bar{\mathbf{z}}_{\{1\}}^{n}, \bar{\mathbf{z}}_{\{2\}}^{j-1}, \bar{\mathbf{z}}_{\{2\},j+1}^{n}\right) \rightarrow (z_1^n, z_2^{j-1}, z_{2,j+1}^n, z_2(j), m_2) \rightarrow y_{2,2}(j) \tag{86}$$

since the additional components in $\left(\bar{\mathbf{z}}_{\{1\}}^{n}, \bar{\mathbf{z}}_{\{2\}}^{j-1}, \bar{\mathbf{z}}_{\{2\},j+1}^{n}\right)$ are degraded versions of $(z_1^n, z_2^{j-1}, z_{2,j+1}^n)$ and the noise across the channels is independent.

Upon substituting (72) in (84) and using the associated Markov condition we have that

$$\begin{aligned}
nR_1 &\leq \sum_{j=1}^{n} I(x_2(j); y_{2,2}(j)|z_2(j), u_2(j)) + 2n\varepsilon_n && (87)\\
&= n\left(I(x_2; y_{2,2}|z_2, u_2) + 2\varepsilon_n\right) && (88)
\end{aligned}$$

In a similar fashion, we can show that

$$\begin{aligned}
nR_1 &\leq \sum_{j=1}^{n} I(x_1(j); y_{1,1}(j)|z_1(j), u_1(j)) + 2n\varepsilon_n && (89)\\
&= n\left(I(x_1; y_{1,1}|z_1, u_1) + 2\varepsilon_n\right) && (90)
\end{aligned}$$

and thus upon combining (88) and (90) we have that

$$R_1 \leq \min\{I(x_1; y_{1,1}|z_1, u_1), I(x_2; y_{2,2}|z_2, u_2)\} + 2\varepsilon_n \tag{91}$$

Upon using the structure of the channel (59) and (60) it can be easily seen that the upper bound in (57) and (58) reduces to (91) and (76) respectively.

Remark 1: The upper bounds in in (91) and (76) can be explained intuitively as follows. The upper bounds

$$\begin{aligned}
R_1 &\leq I(x_1; y_{1,1}|z_1, u_1) && (92)\\
R_2 &\leq I(u_1; z_1|y_{1,2}) && (93)
\end{aligned}$$

correspond to a setup where the message m_1 only needs to be decoded by user 1 in group 1 with the group 2 user as an eavesdropper, whereas the message m_2 only needs to be secure against user 2 in group 1. We relax the decoding of message m_1 at user 2 in group 1 and relax the secrecy constraint for m_2 associated with user 1. In this case one can show that the channel (60) should not be used and the above constraints on R_1 and R_2 can be established. The constraints on R_1 and R_2 involving u_2 can be established by a similar approach. □

B. General case

In establishing the converse for the general case we essentially follow similar steps, however the definition of auxiliary variables $u_1(j)$ and $u_2(j)$ is more involved. In particular by extending the definitions in (72) and (75), the choice of $u_i(j)$ is given by the following:

$$u_i(j) = \left\{m_2, \bar{\mathbf{z}}_{\{i\},j+1}^{n}, \bar{\mathbf{z}}_{\{i\}}^{j-1}, \bar{\mathbf{Z}}_{\{\backslash i\}}^{n}\right\} \tag{94}$$

where we introduce

$$\bar{\mathbf{z}}^n_{\{i\}} := \left(z^n_i, y^n_{k,i}, \forall k \in \mathcal{Z}_i\right), \tag{95}$$

$$\bar{\mathbf{z}}^{j-1}_{\{i\}} := \left(z^{j-1}_i, y^{j-1}_{k,i}, \forall k \in \mathcal{Z}_i\right), \tag{96}$$

$$\bar{\mathbf{z}}^n_{\{i\},j+1} := \left(z^n_{i,j+1}, y^n_{k,i,j+1}, \forall k \in \mathcal{Z}_i\right), \tag{97}$$

$$\bar{\mathbf{Z}}^n_{\{\backslash i\}} := \left(\bar{\mathbf{z}}^n_{\{1\}}, \ldots, \bar{\mathbf{z}}^n_{\{i-1\}}, \bar{\mathbf{z}}^n_{\{i+1\}}, \ldots, \bar{\mathbf{z}}^n_{\{M\}}\right). \tag{98}$$

where the set $\mathcal{Z}_i$ denotes users in group 1 that are weaker than the group 2 user on channel i (c.f. (3)), and observe our choice of $u_i(j)$ in (94) indeed satisfies the Markov condition

$$u_i(j) \to x_i(j) \to (y_{i,1}(j), \ldots, y_{i,K}(j), z_i(j)) \tag{99}$$

since the sub-channels are independent and memoryless. Note that $\bar{\mathbf{z}}^n_{\{i\}}$ is the collection of the group 2 receiver's channel output as well as the output of all the receivers in group 1 that are degraded with respect to the group 2 receiver on channel i. We begin with the secrecy constraint associated with message m_2 with respect to user k in group 1. Let us define the following:

$$\bar{y}^n_{k,i} := \begin{cases} y^n_{k,i}, & x_i \to z_i \to y_{k,i} \\ z^n_i, & x_i \to y_{k,i} \to z_i, \end{cases} \tag{100}$$

$$\bar{\mathbf{y}}^n_k := (\bar{y}^n_{k,1}, \ldots, \bar{y}^n_{k,M}), \quad \mathbf{z}^n := (z^n_1, \ldots, z^n_M), \tag{101}$$

$$\bar{\mathbf{y}}^n_{k,[i]} := (\bar{y}^n_{k,1}, \ldots, \bar{y}^n_{k,i}), \quad \mathbf{z}^n_{[i]} := (z^n_1, \ldots, z^n_i), \tag{102}$$

$$\bar{\mathbf{y}}^n_{k,\{\backslash i\}} := (\bar{y}^n_{k,1}, \cdots, \bar{y}^n_{k,i-1}, \bar{y}^n_{k,i+1}, \cdots, \bar{y}^n_{k,M}) \tag{103}$$

Thus $\bar{\mathbf{y}}^n_k$ corresponds to a weaker receiver, whose output on channel i is degraded to z^n_i, if user k is stronger than the group 2 user on this sub-channel. Clearly we have that $\frac{1}{n}I(m_2; \bar{\mathbf{y}}^n_k) \leq \varepsilon_n$ whenever $\frac{1}{n}I(m_2; \mathbf{y}^n_k) \leq \varepsilon_n$. We thus have

$$n(R_2 - 2\varepsilon_n) \leq I(m_2; \mathbf{z}^n) - I(m_2; \bar{\mathbf{y}}^n_k) \tag{104}$$

$$\leq I(m_2; \mathbf{z}^n | \bar{\mathbf{y}}^n_k) \tag{105}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{n} I(m_2; z_i(j) | z^{j-1}_i, \mathbf{z}^n_{[i-1]}, \bar{\mathbf{y}}^n_k) \tag{106}$$

$$\leq \sum_{i=1}^{M} \sum_{j=1}^{n} \Bigg\{ \qquad I(m_2, z^{j-1}_i, z^n_{i,j+1}, \mathbf{z}^n_{[i-1]}, \bar{\mathbf{y}}^n_{k,\{\backslash i\}}, \bar{y}^{j-1}_{k,i}, \bar{y}^n_{k,i,j+1}; z_i(j) | \bar{y}_{k,i}(j)) \Bigg\} \tag{107}$$

$$\leq \sum_{i=1}^{M} \sum_{j=1}^{n} I(m_2, \bar{\mathbf{Z}}^n_{\{\backslash i\}}, \bar{\mathbf{z}}^n_{\{i\},j+1}, \bar{\mathbf{z}}^{j-1}_{\{i\}}; z_i(j) | \bar{y}_{k,i}(j)) \tag{108}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{n} I(u_i(j); z_i(j) | \bar{y}_{k,i}(j)) \tag{109}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{n} I(u_i(j); z_i(j) | y_{k,i}(j)) \tag{110}$$

where (108) follows from the fact that

$$(\mathbf{z}^n_{[i-1]}, \bar{\mathbf{y}}^n_{k,\{\backslash i\}}) \subseteq \bar{\mathbf{Z}}^n_{\{\backslash i\}}, \quad (z^{j-1}_i, \bar{y}^{j-1}_{k,i}) \subseteq \bar{\mathbf{z}}^{j-1}_{\{i\}}, (z^n_{i,j+1}, \bar{y}^n_{k,i,j+1}) \subseteq \bar{\mathbf{z}}^n_{\{i\},j+1}, \tag{111}$$

and (110) follows from the fact whenever $y_{k,i}(j) \neq \bar{y}_{k,i}(j)$ then $z_i(j)$ is a degraded version of $y_{k,i}(j)$ and from (100), we have that

$$I(u_i(j); z_i(j) | y_{k,i}(j)) = I(u_i(j); z_i(j) | \bar{y}_{k,i}(j)) = 0. \tag{112}$$

This establishes (58).

Next, we upper bound R_1 as follows:

$$n(R_1 - 2\varepsilon_n) \leq I(m_1; \mathbf{y}^n_k) - I(m_1; \mathbf{z}^n, m_2) \tag{113}$$

$$\leq I(m_1; \mathbf{y}^n_k | \mathbf{z}^n, m_2) \tag{114}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{n} I(m_1; y_{k,i}(j) | y^{j-1}_{k,i}, \mathbf{y}^n_{k,[i-1]}, \mathbf{z}^n, m_2) \tag{115}$$

$$\leq \sum_{i=1}^{M} \sum_{j=1}^{n} H(y_{k,i}(j) | y^{j-1}_{k,i}, \mathbf{y}^n_{k,[i-1]}, \mathbf{z}^n, m_2) - H(y_{k,i}(j) | y^{j-1}_{k,i}, \mathbf{y}^n_{k,i-1}, \mathbf{z}^n, m_1, m_2, x_i(j)) \tag{116}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{n} H(y_{k,i}(j) | y^{j-1}_{k,i}, \mathbf{y}^n_{k,[i-1]}, \mathbf{z}^n, m_2) - H(y_{k,i}(j) | x_i(j), z_i(j)) \tag{117}$$

$$\leq \sum_{i=1}^{M} \sum_{j=1}^{n} H(y_{k,i}(j) | \mathbf{z}^n, m_2) - H(y_{k,i}(j) | x_i(j), z_i(j)) \tag{118}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{n} H(y_{k,i}(j) | \bar{\mathbf{Z}}^n_{\{\backslash i\}}, \bar{\mathbf{z}}^{j-1}_{\{i\}}, \bar{\mathbf{z}}^n_{\{i\},j+1}, z_i(j), m_2) - H(y_{k,i}(j) | x_i(j), z_i(j)) \tag{119}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{n} H(y_{k,i}(j) | u_i(j), z_i(j)) - H(y_{k,i}(j) | x_i(j), z_i(j), u_i(j)) \tag{120}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{n} I(x_i(j); y_{k,i}(j) | u_i(j), z_i(j)), \tag{121}$$

where we use the notation $\mathbf{y}^n_{k,[i-1]} \triangleq (y^n_{k,1}, \ldots, y^n_{k,i-1})$ in (115), (117) follows from the fact that for our channel model $(y_{k,i}(j), z_i(j))$ are independent of all other random variables given $x_i(j)$ whereas (119) follows from the fact that even though $\mathbf{z}^n \subseteq \{\bar{\mathbf{Z}}^n_{\{\backslash i\}}, \bar{\mathbf{z}}^{j-1}_{\{i\}}, \bar{\mathbf{z}}^n_{\{i\},j+1}, z_i(j)\}$ holds, the additional elements in the latter are only a degraded version of $\mathbf{z}^n$. Since the channels are independent and memoryless, these additional terms in the conditioning do not reduce the entropy term. This establishes (57).

To complete the converse, let q_i to be a random variable uniformly distributed over the set $\{1, 2, \ldots, n\}$ and furthermore we let $u_i = (u_i(q_i), q_i)$, $x_i = x_i(q_i)$ etc. Then (57) and (58) can be reduced to

$$R_1 - 2\varepsilon_n \le \sum_{i=1}^{M} I(x_i; y_{k,i}|u_i, z_i, q_i) = \sum_{i=1}^{M} I(x_i; y_{k,i}|u_i, z_i) \tag{122}$$

$$R_2 - 2\varepsilon_n \le \sum_{i=1}^{M} I(u_i; z_i|y_{k,i}, q_i) \le \sum_{i=1}^{M} I(u_i; z_i|y_{k,i}). \tag{123}$$

The upper bound on the cardinality of $\mathcal{U}_i$ follows by a straightforward application of Caratheodory's theorem and the proof is omitted.

VI. Gaussian Channels

In this section we provide a proof for Theorem 2. Note that the achievability of the rate pairs (R_1, R_2) constrained by (11) and (12) follows that of those constrained by (5) and (6) by setting $x_i = u_i + v_i$, where u_i and v_i are independent $\mathcal{N}(0, P_i - Q_i)$ and $\mathcal{N}(0, Q_i)$ respectively for some $0 \le Q_i \le P_i$ and $i = 1, \ldots, M$. For the rest of the section, we shall focus on proving the converse result.

Considering proof by contradiction, let us assume that (R_1^o, R_2^o) is an achievable rate pair that lies *outside* the rate region constrained by (11) and (12). Note that the maximum rate for message m_1 is given by the right-hand side of (11) by setting $Q_i = P_i$ for all $i = 1, \ldots, M$ [7], and the maximum rate for message m_2 is given by the right-hand side of (12) by setting $Q_i = 0$ for all $i = 1, \ldots, M$ [4], [10]. Thus, without loss of generality we may assume that $R_2^0 = R_2^* + \delta$ for some $\delta > 0$ where R_2^* is given by

$$\begin{aligned}
&\max_{(\mathbf{Q}, R_2)} && R_2 \\
&\text{subject to} && R_1^o \le \sum_{i=1}^{M} A_{k,i}^{(1)}(\mathbf{Q}), \qquad \forall k = 1, \ldots, K && (124) \\
& && R_2 \le \sum_{i=1}^{M} A_{k,i}^{(2)}(\mathbf{Q}), \qquad \forall k = 1, \ldots, K && (125) \\
& && Q_i \ge 0, \qquad \forall i = 1, \ldots, M && (126) \\
& && Q_i \le P_i, \qquad \forall i = 1, \ldots, M. && (127)
\end{aligned}$$

For each $k = 1, \ldots, K$ and $i = 1, \ldots, M$ let α_k, β_k, $M_{1,i}$ and $M_{2,i}$ be the Lagrangians that correspond to the constrains (124)–(127) respectively, and let

$$\begin{aligned}
L := R_2 + \sum_{k=1}^{K} \alpha_k \left[\sum_{i=1}^{M} A_{k,i}^{(1)}(\mathbf{Q}) - R_1^o \right] \\
+ \sum_{k=1}^{K} \beta_k \left[\sum_{i=1}^{M} A_{k,i}^{(2)}(\mathbf{Q}) - R_2 \right] + \sum_{i=1}^{M} M_{1,i} Q_i + \sum_{i=1}^{M} M_{2,i}(P_i - Q_i).
\end{aligned} \tag{128}$$

It is straightforward to verify that the above optimization program that determines R_2^* is a convex program. Therefore, taking partial derivatives of L over Q_i, $i = 1 \ldots, M$ and R_2 respectively gives the following set of Karush-Kuhn-Tucker (KKT) conditions, which must be satisfied by any *optimal* solution $(\mathbf{Q}^*, R_2^*)$:

$$\begin{aligned}
&\sum_{k \in \mathcal{Y}_i} \alpha_k (Q_i^* + \sigma_{k,i}^2)^{-1} + \sum_{k \in \mathcal{Z}_i} \beta_k (Q_i^* + \sigma_{k,i}^2)^{-1} + M_{1,i} \\
&\quad = \left(\sum_{k \in \mathcal{Y}_i} \alpha_k + \sum_{k \in \mathcal{Z}_i} \beta_k \right) (Q_i^* + \delta_i^2)^{-1} + M_{2,i} && (129) \\
&\sum_{k=1}^{K} \beta_k = 1 && (130) \\
&\alpha_k \left[\sum_{i=1}^{M} A_{k,i}^{(1)}(\mathbf{Q}^*) - R_1^o \right] = 0, \ \forall k = 1, \ldots, K && (131) \\
&\beta_k \left[\sum_{i=1}^{M} A_{k,i}^{(2)}(\mathbf{Q}^*) - R_2^* \right] = 0, \ \forall k = 1, \ldots, K && (132) \\
&M_{1,i} Q_i^* = 0, \ \forall i = 1, \ldots, M && (133) \\
&M_{2,i}(P_i - Q_i^*) = 0, \ \forall i = 1, \ldots, M && (134) \\
&\alpha_k, \beta_k \ge 0, \ \forall k = 1, \ldots, K && (135) \\
&M_{1,i}, M_{2,i} \ge 0, \ \forall i = 1, \ldots, M && (136)
\end{aligned}$$

where recall that

$$\mathcal{Y}_i := \{k : \sigma_{k,i}^2 < \delta_i^2\} \quad \text{and} \quad \mathcal{Z}_i := \{k : \sigma_{k,i}^2 > \delta_i^2\}. \tag{137}$$

Note that $\delta > 0$, so we have

$$\begin{aligned}
&\left(\sum_{k=1}^{K} \alpha_k \right) R_1^o + R_2^o > \left(\sum_{k=1}^{K} \alpha_k \right) R_1^o + R_2^* && (138) \\
&= \sum_{k=1}^{K} (\alpha_k R_1^o + \beta_k R_2^*) && (139) \\
&= \sum_{k=1}^{K} \left[\alpha_k \sum_{i=1}^{M} A_{k,i}^{(1)}(\mathbf{Q}^*) + \beta_k \sum_{i=1}^{M} A_{k,i}^{(2)}(\mathbf{Q}^*) \right] && (140) \\
&= \sum_{i=1}^{M} \sum_{k=1}^{K} \left[\alpha_k A_{k,i}^{(1)}(\mathbf{Q}^*) + \beta_k A_{k,i}^{(2)}(\mathbf{Q}^*) \right], && (141)
\end{aligned}$$

where (139) follows from the KKT condition (130), and (140) follows from the KKT conditions (131) and (132).

Next, we shall show that by assumption (R_1^o, R_2^o) is achievable, so we have

$$\left(\sum_{k=1}^{K} \alpha_k \right) R_1^o + R_2^o \le \sum_{i=1}^{M} \sum_{k=1}^{K} \left[\alpha_k A_{k,i}^{(1)}(\mathbf{Q}^*) + \beta_k A_{k,i}^{(2)}(\mathbf{Q}^*) \right] \tag{142}$$

which is an apparent contradiction to (141) and hence will help to complete the proof of the theorem. To prove (142), let us apply the converse part of Theorem 1 on (R_1^o, R_2^o) and write

$$\begin{aligned}
\left(\sum_{k=1}^{K} \alpha_k \right) R_1^o + R_2^o \le \left(\sum_{k=1}^{K} \alpha_k \right) \min_{1 \le k \le K} \left\{ \sum_{i=1}^{M} I(x_i; y_{k,i}|u_i, z_i) \right\} \\
+ \min_{1 \le k \le K} \left\{ \sum_{i=1}^{M} I(u_i; z_i|y_{k,i}) \right\}
\end{aligned} \tag{143}$$

$$\le \sum_{k=1}^{K}\left[\alpha_k \sum_{i=1}^{M} I(x_i;y_{k,i}|u_i,z_i)\right]+\sum_{k=1}^{K}\left[\beta_k\sum_{i=1}^{M} I(u_i;z_i|y_{k,i})\right] \tag{144}$$

$$=\sum_{i=1}^{M}\sum_{k=1}^{K}\left[\alpha_k I(x_i;y_{k,i}|u_i,z_i)+\beta_k I(u_i;z_i|y_{k,i})\right], \tag{145}$$

where (144) follows from the well-known fact that minimum is no more than any weighted mean. By the degradedness assumption (3), we have

$$I(x_i;y_{k,i}|u_i,z_i)=I(x_i;y_{k,i}|u_i)-I(x_i;z_i|u_i) \tag{146}$$

$$=h(y_{k,i}|u_i)-h(z_i|u_i)-h(n_{k,i})+h(w_i) \tag{147}$$

$$=h(y_{k,i}|u_i)-h(z_i|u_i)-\frac{1}{2}\log\left(\frac{\sigma_{k,i}^2}{\delta_i^2}\right) \tag{148}$$

for any $k\in\mathcal{Y}_i$ and $I(x_i;y_{k,i}|u_i,z_i)=0$ for any $k\notin\mathcal{Y}_i$. Similarly,

$$I(u_i;z_i|y_{k,i})=I(u_i;z_i)-I(u_i;y_{k,i}) \tag{149}$$

$$=h(z_i)-h(y_{k,i})-h(z_i|u_i)+h(y_{k,i}|u_i) \tag{150}$$

$$\le\frac{1}{2}\log\left(\frac{P_i+\delta_i^2}{P_i+\sigma_{k,i}^2}\right)-h(z_i|u_i)+h(y_{k,i}|u_i) \tag{151}$$

for any $k\in\mathcal{Z}_i$, where (151) follows from the worst additive noise Lemma [20], and $I(u_i;z_i|y_{k,i})=0$ for any $k\notin\mathcal{Z}_i$. Thus, for each $i=1,\ldots,M$ we have

$$\begin{aligned}&\sum_{k=1}^{K}\left[\alpha_k I(x_i;y_{k,i}|u_i,z_i)+\beta_k I(u_i;z_i|y_{k,i})\right]\\&\le\sum_{k\in\mathcal{Y}_i}\alpha_k\left[h(y_{k,i}|u_i)-h(z_i|u_i)-\frac{1}{2}\log\left(\frac{\sigma_{k,i}^2}{\delta_i^2}\right)\right]+\\&\qquad\sum_{k\in\mathcal{Z}_i}\beta_k\left[\frac{1}{2}\log\left(\frac{P_i+\delta_i^2}{P_i+\sigma_{k,i}^2}\right)-h(z_i|u_i)+h(y_{k,i}|u_i)\right]\end{aligned} \tag{152}$$

$$\begin{aligned}&=\sum_{k\in\mathcal{Y}_i}\alpha_k h(y_{k,i}|u_i)+\sum_{k\in\mathcal{Z}_i}\beta_k h(y_{k,i}|u_i)\\&\quad-\left(\sum_{k\in\mathcal{Y}_i}\alpha_k+\sum_{k\in\mathcal{Z}_i}\beta_k\right)h(z_i|u_i)-\\&\qquad\sum_{k\in\mathcal{Y}_i}\frac{\alpha_k}{2}\log\left(\frac{\sigma_{k,i}^2}{\delta_i^2}\right)+\sum_{k\in\mathcal{Z}_i}\frac{\beta_k}{2}\log\left(\frac{P_i+\delta_i^2}{P_i+\sigma_{k,i}^2}\right).\end{aligned} \tag{153}$$

We have the following result [11, Lemma 1], which can be proved either directly from Costa's entropy-power inequality [21] or from the classical entropy-power inequality of Shannon [22] and Stam [23] via the "change-of-variable" technique of Watanabe and Oohama [24, Rem. 6].

Lemma 2: For any real scalars α_k, β_k, Q_i^*, $M_{1,i}$ and $M_{2,i}$ that satisfy KKT conditions (129) and (133)–(136), we have

$$\begin{aligned}&\sum_{k\in\mathcal{Y}_i}\alpha_k h(y_{k,i}|u_i)+\sum_{k\in\mathcal{Z}_i}\beta_k h(y_{k,i}|u_i)-\left(\sum_{k\in\mathcal{Y}_i}\alpha_k+\sum_{k\in\mathcal{Z}_i}\beta_k\right)h(z_i|u_i)\\&\le\sum_{k\in\mathcal{Y}_i}\frac{\alpha_k}{2}\log(Q_i^*+\sigma_{k,i}^2)+\sum_{k\in\mathcal{Z}_i}\frac{\beta_k}{2}\log(Q_i^*+\sigma_{k,i}^2)-\\&\quad\frac{\sum_{k\in\mathcal{Y}_i}\alpha_k+\sum_{k\in\mathcal{Z}_i}\beta_k}{2}\log(Q_i^*+\delta_i^2)\end{aligned} \tag{154}$$

for any (u_i,x_i) that is independent of the additive Gaussian noise $(n_{1,i},\ldots,n_{K,i},w_i)$ and such that $E[x_i^2]\le P_i$.

Substituting (154) into (153) gives

$$\begin{aligned}&\sum_{k=1}^{K}\left[\alpha_k I(x_i;y_{k,i}|u_i,z_i)+\beta_k I(u_i;z_i|y_{k,i})\right]\\&\le\sum_{k\in\mathcal{Y}_i}\frac{\alpha_k}{2}\log(Q_i^*+\sigma_{k,i}^2)+\sum_{k\in\mathcal{Z}_i}\frac{\beta_k}{2}\log(Q_i^*+\sigma_{k,i}^2)\\&\quad-\frac{\sum_{k\in\mathcal{Y}_i}\alpha_k+\sum_{k\in\mathcal{Z}_i}\beta_k}{2}\log(Q_i^*+\delta_i^2)-\\&\qquad\sum_{k\in\mathcal{Y}_i}\frac{\alpha_k}{2}\log\left(\frac{\sigma_{k,i}^2}{\delta_i^2}\right)+\sum_{k\in\mathcal{Z}_i}\frac{\beta_k}{2}\log\left(\frac{P_i+\delta_i^2}{P_i+\sigma_{k,i}^2}\right)\end{aligned} \tag{155}$$

$$\begin{aligned}&=\sum_{k\in\mathcal{Y}_i}\alpha_k\left[\frac{1}{2}\log\left(\frac{Q_i^*+\sigma_{k,i}^2}{\sigma_{k,i}^2}\right)-\frac{1}{2}\log\left(\frac{Q_i^*+\delta_i^2}{\delta_i^2}\right)\right]+\\&\qquad\sum_{k\in\mathcal{Z}_i}\beta_k\left[\frac{1}{2}\log\left(\frac{P_i+\delta_i^2}{Q_i^*+\delta_i^2}\right)-\frac{1}{2}\log\left(\frac{P_i+\sigma_{k,i}^2}{Q_i^*+\sigma_{k,i}^2}\right)\right]\end{aligned} \tag{156}$$

$$=\sum_{k=1}^{K}\left[\alpha_k A_{k,i}^{(1)}(\mathbf{Q}^*)+\beta_k A_{k,i}^{(2)}(\mathbf{Q}^*)\right]. \tag{157}$$

Further substituting (157) into (145) completes the proof of (142). We have thus completed the proof of Theorem 2.

VII. Fading Channels

To establish the connection to fading channels, first observe that Theorem 2 and Corollary 1 can be extended in the following way. Consider the following scalar Gaussian broadcast channel with $K+1$ users:

$$y_k(t)=x(t)+n_k(t) \tag{158}$$

$$z(t)=x(t)+w(t),\quad t=1,\ldots,n. \tag{159}$$

At each time sample t, the additive noise $(n_1(t),\ldots,n_K(t),w(t))$ are independent zero-mean Gaussian with the variances $(\sigma_1^2,\ldots,\sigma_K^2,\delta^2)$ selected at random as $(\sigma_{1,i}^2,\ldots,\sigma_{K,i}^2,\delta_i^2)$ with probability p_i, $i=1,\ldots,M$. Both the selection of the noise variances and the realization of the additive noise are assumed to be independent across the time index t and revealed to all the terminals. We are interested in the ergodic scenario where the duration of communication can be arbitrarily large. The following extension of Theorem 2 readily follows and its proof will be omitted.

Corollary 2: For the scalar Gaussian broadcast channel considered above, the capacity region consists of all rate pairs (R_1, R_2) that satisfy

$$R_1 \leq \min_{1\leq k\leq K} \sum_{i=1}^{M} p_i \left[\frac{1}{2}\log\left(\frac{Q_i+\sigma_{k,i}^2}{\sigma_{k,i}^2}\right) - \frac{1}{2}\log\left(\frac{Q_i+\delta_i^2}{\delta_i^2}\right)\right]^+ \tag{160}$$

$$R_2 \leq \min_{1\leq k\leq K} \sum_{i=1}^{M} p_i \left[\frac{1}{2}\log\left(\frac{P_i+\delta_i^2}{Q_i+\delta_i^2}\right) - \frac{1}{2}\log\left(\frac{P_i+\sigma_{k,i}^2}{Q_i+\sigma_{k,i}^2}\right)\right]^+ \tag{161}$$

for some $0 \leq Q_i \leq P_i$ and $i = 1, \ldots, M$. □

Clearly if the fading coefficients in (15) are all discrete-valued, then the result in Theorem 3 follows immediately from Corollary 2. When the fading coefficients are continuous valued, we can generalize Theorem 2 by suitably quantizing the channel gains as discussed below.

First without loss of generality, we assume that each fading coefficient is real-valued, since each receiver can cancel out the phase of the fading gain through a suitable multiplication at the receiver. Consider a discrete set

$$\mathcal{A} := \{A_1, A_2, \ldots, A_N, A_{N+1}\}$$

where $A_i \leq A_{i+1}$, $A_1 := 0$, $A_N := J$ and $A_{N+1} := \infty$ holds.

Given a set of channel gains $(\mathsf{h}_{1,i}, \ldots, \mathsf{h}_{K,i}, \mathsf{g}_i)$ in coherence block i, we discretize them to one of $(N+1)^{K+1}$ states as described below.

- Encoding message m_1: Suppose that the channel gain of receiver k satisfies $A_q \leq \mathsf{h}_{k,i} \leq A_{q+1}$, then we assume that the channel gain equals $\hat{\mathsf{h}}_{k,i} = A_q$. If the channel gain of the group 2 user satisfies $A_q \leq \mathsf{g}_i \leq A_{q+1}$ then we assume that its channel gain equals $\bar{\mathsf{g}}_i = A_{q+1}$.
- Encoding message m_2: Suppose that the channel gain of the group 2 receiver satisfies $A_q \leq \mathsf{g}_i \leq A_{q+1}$, then we assume that the channel gain equals $\hat{\mathsf{g}}_i = A_q$. If the channel gain of a group 1 receiver satisfies $A_q \leq \mathsf{h}_{k,i} \leq A_{q+1}$ then we assume it equals $\bar{\mathsf{h}}_{k,i} = A_{q+1}$.

Thus the set of discretized channel gains coherence block i, i.e., $(\hat{\mathsf{h}}_{1,i}, \ldots, \hat{\mathsf{h}}_{K,i}, \hat{\mathsf{g}}_i)$ (and equivalently $(\bar{\mathsf{h}}_{1,i}, \ldots, \bar{\mathsf{h}}_{K,i}, \bar{\mathsf{g}}_i)$) belongs to one of $L = (N+1)^{K+1}$ possible values. The above discretization maps the system to one of L possible sub-channels, indexed by state $\mathbf{s}_j \equiv (\mathsf{s}_{1,j}, \ldots, \mathsf{s}_{K,j}, \mathsf{s}_{K+1,j})$. Here $\mathsf{s}_{k,j}$ for $k = 1, \ldots, K$ denotes the discretized channel gains $\hat{\mathsf{h}}_k$ of the receiver in group 1 whereas $\mathsf{s}_{K+1,j}$ denotes the discretized channel gain $\hat{\mathsf{g}}$ of the group 2 receiver. Since there is a one-to-one relation between $\hat{\mathsf{h}}_k$ and $\bar{\mathsf{h}}_k$ and similarly between $\hat{\mathsf{g}}$ and $\bar{\mathsf{g}}$ we can also express $\mathbf{s}_j \equiv (\bar{\mathsf{s}}_{1,j}, \ldots, \bar{\mathsf{s}}_{K,j}, \bar{\mathsf{s}}_{K+1,j})$ where $\bar{\mathsf{s}}_{k,j} = \bar{\mathsf{h}}_k$ for $k = 1, \ldots, K$ and $\bar{\mathsf{s}}_{K+1,j} = \bar{\mathsf{g}}$.

With the above quantization procedure we can consider a coding scheme associated for $L = (N+1)^{K+1}$ parallel channels, where each parallel channel corresponds to one state realization $\mathbf{s}_j$. Using Corollary 2 the following rate pair (R_1, R_2) is achievable:

$$R_1 \leq \min_{1\leq k\leq K} \sum_{j=1}^{L} \Pr(\mathbf{s}_j) A_{k,j}^{(1)}(\mathbf{s}_j) \tag{162}$$

$$R_2 \leq \min_{1\leq k\leq K} \sum_{j=1}^{L} \Pr(\mathbf{s}_j) A_{k,j}^{(2)}(\mathbf{s}_j), \tag{163}$$

where

$$A_{k,j}^{(1)}(\mathbf{s}_j) := \left\{\log\frac{1+Q(\mathbf{s}_j)|\mathsf{s}_{k,j}|^2}{1+Q(\mathbf{s}_j)|\bar{\mathsf{s}}_{K+1,j}|^2}\right\}^+ \tag{164}$$

$$A_{k,j}^{(2)}(\mathbf{s}_j) := \left\{\log\frac{1+P(\mathbf{s}_j)|\mathsf{s}_{K+1,j}|^2}{1+Q(\mathbf{s}_j)|\mathsf{s}_{K+1,j}|^2} - \log\frac{1+P(\mathbf{s}_j)|\bar{\mathsf{s}}_{k,j}|^2}{1+Q(\mathbf{s}_j)|\bar{\mathsf{s}}_{k,j}|^2}\right\}^+ . \tag{165}$$

Recall that $A_{N+1} = J$, denotes the largest value of the discretized channel gain. For any fixed J upon taking the limit $N \to \infty$, we have that

$$\lim_{N\to\infty} \sum_{j=1}^{L} \Pr(\mathbf{s}_j) A_{k,j}^{(1)}(\mathbf{s}_j) \geq \oint_0^J \int_0^J A_k^{(1)}(\mathbf{h}, g) dF(\mathbf{h}) dF(g) \tag{166}$$

where $\mathbf{h} = (\mathsf{h}_1, \ldots, \mathsf{h}_K)$ and g denote the channel gains of the group 1 and group 2 users and $dF(\cdot)$ the corresponding distribution and

$$A_k^{(1)}(\mathbf{h}, \mathsf{g}) = \left\{\log\frac{1+Q(\mathbf{h},\mathsf{g})|\mathsf{h}_k|^2}{1+Q(\mathbf{h},\mathsf{g})|\mathsf{g}|^2}\right\}^+,$$

We only have a lower bound in (166) as we do not account for the contribution of the channel gains greater than J. Also since whenever $\mathsf{g} > J$, note that $A_k^{(1)}(\mathbf{h}, \mathsf{g}) = 0$. Thus, it follows that,

$$\oint_0^J \int_0^J A_k^{(1)}(\mathbf{h}, g) dF(\mathbf{h}) dF(g) = \oint_0^J \int_0^\infty A_k^{(1)}(\mathbf{h}, g) dF(\mathbf{h}) dF(g). \tag{167}$$

Finally, by taking J arbitrarily large, the following rate is achievable

$$R_1 = \min_{1\leq k\leq K} \oint_0^\infty \int_0^\infty A_k^{(1)}(\mathbf{h}, g) dF(g) dF(\mathbf{h}) \tag{168}$$

as required. In a similar fashion the achievability of R_2 can be established.

The converse follows by noticing that if the channel gains are revealed non-causally to the terminals, the system reduces to a parallel channel model and the result in Theorem 2 immediately applies.

Numerical Results

In order to evaluate the achievable rate region, we assume that the fading gains are all sampled from $\mathcal{CN}(0, 1)$. Furthermore instead of finding the optimal power allocation we assume a potentially sub-optimal power allocation[3]:

$$Q(\mathbf{h}, \mathsf{g}) = \begin{cases} P, & |\mathsf{g}|^2 \geq \theta \\ 0, & |\mathsf{g}|^2 < \theta. \end{cases} \tag{169}$$

[3] For any boundary point of the capacity region in Theorem 3, the associated expression $\lambda_1 R_1 + \lambda_2 R_2$ is a concave function of $Q(\cdot)$. Thus one can apply KKT conditions to characterize the optimal power allocation strategy. However the presence of a common message for group 1 receivers makes the optimality conditions quite involved. See [25] for a related problem in absence of secrecy constraints.

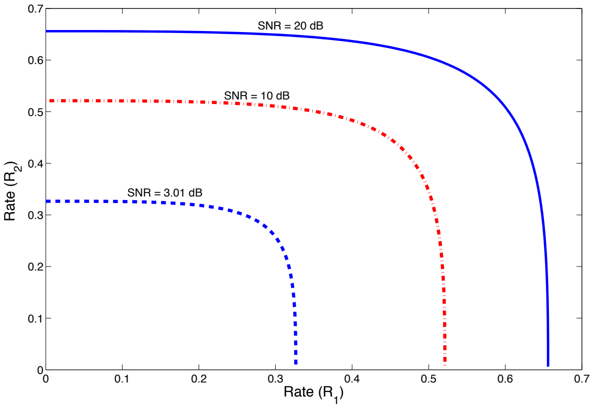

Fig. 3: Achievable rates (nats/symbol) for the two groups at different SNR values for i.i.d. Rayleigh Fading $h, g \sim \mathcal{N}(0, 1)$. The x-axis shows the rate R_1 for group 1 whereas the y-axis shows the rate R_2 for group 2.

where θ is a certain fixed parameter and assume that $P(\mathbf{h}, g) = P$ for all values of $(\mathbf{h}, g)$. Notice that our power allocation does not depend on the channel gains of the receivers in group 1. This is a reasonable simplification when K is large and the channel gains $(h_1, \ldots, h_K)$ are identically distributed. The achievable rate expressions (17) and (18) reduce to:

$$R_1 \leq \Pr(|g|^2 \leq \theta) E\left[\left\{\log \frac{1 + P|h|^2}{1 + P|g|^2}\right\}^+ \middle| |g|^2 \leq \theta\right] \quad (170)$$

$$R_2 \leq \Pr(|g|^2 \geq \theta) E\left[\left\{\log \frac{1 + P|g|^2}{1 + P|h|^2}\right\}^+ \middle| |g|^2 \geq \theta\right] \quad (171)$$

In Fig. 3, we plot the achievable rates for $P \in \{2, 10, 100\}$. We make the following observations:

- The corner points for R_1 and R_2 are obtained by setting $\theta = \infty$ and $\theta = 0$ respectively. By symmetry of the rate expressions in (170) and (171), it is clear that both the corner points evaluate to the same numerical constant.
- As we approach the corner point $(0, R_2)$ the boundary of the capacity region is nearly flat. Any coherence block, where $|g(i)| \leq \min_{1 \leq k \leq K} |h_k(i)|$ is clearly not useful to the group 2 receiver. By transmitting m_1 in these slots one can increase the rate R_1 without decreasing R_2.
- As we approach the corner point $(R_1, 0)$, the boundary of the capacity region is nearly vertical. The argument is very similar to the previous case. In any period where $|g(i)| \geq \max_{1 \leq k \leq K} |h_k(i)|$ one cannot transmit to group 1. By transmitting m_2 in these slots we increase R_2 without decreasing R_1.
- We observe that a natural alternative to the proposed scheme is time-sharing. The rate achieved by such a scheme corresponds to a straight line connecting the corner points. The rate-loss associated with such a scheme is significant compared to the proposed scheme.

VIII. Conclusions

We consider the problem of private broadcasting to two group of users under a mutual secrecy constraint. We focus on the special case when there can be an arbitrary number of receivers in one of the groups, but only a single receiver in the other group. Furthermore the channel can be decomposed into parallel, degraded independent channels. We establish the optimality of a superposition construction where the base layer is formed by the codewords of the group 2 user whereas the satellite codebook is formed by the codewords for group 1 users. Our capacity result is a generalization of previous works [7], [10] on compound wiretap channels, which correspond to the corner points of our region. We further treat the case of Gaussian channels with a power constraint and establish the optimality of a Gaussian input distribution by invoking a suitable extremal information inequality. We also extend our coding scheme to a class of block-fading channels and numerically demonstrate significant gains over a baseline time-sharing scheme in Rayleigh fading channels.

In future work it would be of interest to extend our result in a number of directions. The case when there are multiple receivers in both groups 1 and 2 is of interest. Similarly the case of MIMO channels is of interest. It should however be noted that these problems could be considerably more challenging as even the corner points of the capacity region are not known. Likewise it would be interesting to revisit the setup considered in this paper when the mutual secrecy constraint on the messages is not imposed. To our knowledge, the capacity region in this case also remains an open problem.

Appendix A
Proof of Lemma 1

We are required to show that $x_1^n, \ldots, x_M^n$ are conditionally independent given m_1. Note that

$$p(x_1^n, \ldots, x_M^n | m_1) = \sum_{\{\bar{m}_{2,i}\}} p(x_1^n, \ldots, x_M^n, \bar{m}_{2,1}, \ldots, \bar{m}_{2,M} | m_1, \mathcal{C}) \quad (172)$$

$$= \sum_{\{\bar{m}_{2,i}\}} p(x_1^n, \ldots, x_M^n | m_1, \bar{m}_{2,1}, \ldots, \bar{m}_{2,M}, , \mathcal{C}) p(\bar{m}_{2,1}, \ldots, \bar{m}_{2,M}) \quad (173)$$

$$= \sum_{\{\bar{m}_{2,i}\}} \Bigg\{ p(x_1^n, \ldots, x_M^n | m_1, \bar{m}_{2,1}, \ldots, \bar{m}_{2,M}, \mathcal{C}) \times \tag{174}$$

$$p(\bar{m}_{2,1}) \ldots p(\bar{m}_{2,M}) \Bigg\} \tag{175}$$

$$= \sum_{\{\bar{m}_{2,i}\}} \Bigg\{ p(x_1^n | m_1, \bar{m}_{2,1}, \mathcal{C}) \ldots p(x_M^n | m_1, \bar{m}_{2,M}, \mathcal{C}) \times$$

$$p(\bar{m}_{2,1}) \ldots p(\bar{m}_{2,M}) \Bigg\} \tag{176}$$

$$= \prod_{i=1}^{M} \sum_{\bar{m}_{2,i}} p(x_i^n | m_1, \bar{m}_{2,i}, \mathcal{C}) p(\bar{m}_{2,i}) \tag{177}$$

$$= \prod_{i=1}^{M} \sum_{\bar{m}_{2,i}} p(x_i^n, \bar{m}_{2,i} | m_1, \mathcal{C}) \tag{178}$$

$$= \prod_{i=1}^{M} p(x_i^n | m_1, \mathcal{C}) \tag{179}$$

where (173) follows from the fact that the messages $\bar{m}_{2,1}, \ldots, \bar{m}_{2,M}$ are independent of $(m_1, \mathcal{C})$; (175) follows from the fact that the messages satisfy (31); (176) follows from the fact that each $x_i^n \in \mathcal{C}_{1,i}(m_1, u_i^n)$ and u_i^n is the codeword associated with message $\bar{m}_{2,i}$ and furthermore x_i^n is selected independently for each i. Eq. (179) establishes the conditional independence of the messages and completes the proof.

References

[1] A. D. Wyner, "The wiretap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, 1975.

[2] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, pp. 339–348, 1978.

[3] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wiretap channel," *IEEE Trans. Inform. Theory*, vol. 24, pp. 451–456, 1978.

[4] P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[5] Z. Li, R. D. Yates, and W. Trappe, "Achieving secret communication for fast Rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 9, no. 9, pp. 2792–2799, 2010.

[6] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2470–2492, 2008.

[7] A. Khisti, A. Tchamkerten, and G. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2453—2469, 2008.

[8] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2515–2534, 2008.

[9] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai, "Compound wiretap channels," *EURASIP Journal on Wireless Communications and Networking - Special issue on wireless physical layer security*, Mar. 2009.

[10] T. Liu, V. Prabhakaran, and S. Vishwanath, "The secrecy capacity of a class of parallel Gaussian compound wiretap channels," in *Proc. Int. Symp. Inform. Theory*, 2008, pp. 116—120.

[11] J. Chen, "Rate region of Gaussian multiple description coding with individual and central distortion constraints," *IEEE Trans. Inform. Theory*, vol. 55, no. 9, pp. 3991–4005, 2009.

[12] N. Cai and K. Y. Lam, "How to broadcast privacy: Secret coding for deterministic broadcast channels," *Numbers, Information, and Complexity (Festschrift for Rudolf Ahlswede), eds: I. Althofer, N. Cai, G. Dueck, L. Khachatrian, M. Pinsker, A. Sarkozy, I. Wegener, and Z. Zhang*, pp. 353–368, 2000.

[13] R. Liu, T. Liu, H. Poor, and S. Shamai, "Multiple-input multiple-output gaussian broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, no. 9, pp. 4215 – 4227, 2010.

[14] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy capacity regions," *IEEE Trans. Inform. Theory*, June 2008.

[15] L. Czap, V. M. Prabhakaran, S. N. Diggavi, and C. Fragouli, "Broadcasting private messages securely," in *ISIT*, 2012, pp. 428–432.

[16] S. Yang, P. Piantanida, M. Kobayashi, and S. Shamai, "On the secrecy degrees of freedom of multi-antenna wiretap channels with delayed CSIT," in *ISIT*, 2011, pp. 2866–2870.

[17] A. Khisti, "Interference alignment for the multi-antenna compound wiretap channel," *IEEE Trans. Inform. Theory*, vol. 57, no. 5, pp. 2967—2993, 2011.

[18] A. E. Gamal and Y. H. Kim, *Network Information Theory*. Cambridge, UK: Cambridge University Press, 2011.

[19] Y. Chia and A. E. Gamal, "Three-receiver broadcast channels with common and confidential messages," *IEEE Trans. Inform. Theory*, vol. 58, no. 5, pp. 2748–2765, 2012.

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

[21] M. H. M. Costa, "A new entropy power inequality," *IEEE Trans. Inform. Theory*, vol. 31, no. 6, pp. 751–760, 1985.

[22] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.

[23] A. J. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon," *Inform. Contr.*, vol. 2, no. 2, pp. 102–112, 1959.

[24] S. Watanabe and Y. Oohama, "Secret key agreement from vector Gaussian sources by rate limited public communication," *IEEE Trans. Inform. Foren. Security*, vol. 6, no. 3, pp. 541–550, 2011.

[25] N. Jindal and A. Goldsmith, "Capacity and dirty paper coding for Gaussian broadcast channels with common information," in *ISIT*, 2004, p. 215.

Ashish Khisti (S02, M09) is an Assistant Professor in the Electrical and Computer Engineering (ECE) Department and a Canada Research Chair (Tier II) in Network Information Theory at the University of Toronto, Toronto, Ontario, Canada. He received his BASc degree in Engineering Sciences from University of Toronto in 2002 and his S.M. and Ph.D. degrees from the Massachusetts Institute of Tech- nology (MIT), Cambridge, MA, USA in 2004 and 2008, respectively. He has been with the University of Toronto since 2009. His research interests span the areas of information theory, wireless physical layer security and streaming communication systems. During his graduate studies, Professor Khisti was a recipient of the NSERC postgraduate fellowship, Harold H. Hazen Teaching award and the Morris Joseph Levin Masterworks award. At the University of Toronto he is a recipient of the Ontario Early Researcher Award (2012) and a Hewlett-Packard IRP award (2011, 2012). He is an associate editor of the IEEE TRANSACTIONS ON COMMUNICATIONS.

Tie Liu received his B.S. (1998) and M.S. (2000) degrees, both in Electrical Engineering, from Tsinghua University, Beijing, China and a second M.S. degree in Mathematics (2004) and Ph.D. degree in Electrical and Computer Engineering (2006) from the University of Illinois at Urbana-Champaign. Since August 2006 he has been with Texas A &M University, where he is currently an Associate Professor with the Department of Electrical and Computer Engineering. His primary research interest is in understanding the fundamental performance limits of communication and cryptographic systems via the lens of information theory. Dr. Liu is a recipient of the M. E. Van Valkenburg Graduate Research Award (2006) from the University of Illinois at Urbana-Champaign and the Faculty Early Career Development (CAREER) Award (2009) from the National Science Foundation.