

# Secure Communication Over Fading Channels

Yingbin Liang, *Member, IEEE*, H. Vincent Poor, *Fellow, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

**Abstract**—The fading broadcast channel with confidential messages (BCC) is investigated, where a source node has common information for two receivers (receivers 1 and 2), and has confidential information intended only for receiver 1. The confidential information needs to be kept as secret as possible from receiver 2. The broadcast channel from the source node to receivers 1 and 2 is corrupted by multiplicative fading gain coefficients in addition to additive Gaussian noise terms. The channel state information (CSI) is assumed to be known at both the transmitter and the receivers. The parallel BCC with independent subchannels is first studied, which serves as an information-theoretic model for the fading BCC. The secrecy capacity region of the parallel BCC is established, which gives the secrecy capacity region of the parallel BCC with degraded subchannels. The secrecy capacity region is then established for the parallel Gaussian BCC, and the optimal source power allocations that achieve the boundary of the secrecy capacity region are derived. In particular, the secrecy capacity region is established for the basic Gaussian BCC. The secrecy capacity results are then applied to study the fading BCC. The ergodic performance is first studied. The ergodic secrecy capacity region and the optimal power allocations that achieve the boundary of this region are derived. The outage performance is then studied, where a long-term power constraint is assumed. The power allocation is derived that minimizes the outage probability where either the target rate of the common message or the target rate of the confidential message is not achieved. The power allocation is also derived that minimizes the outage probability where the target rate of the confidential message is not achieved subject to the constraint that the target rate of the common message must be achieved for all channel states.

**Index Terms**—Confidential message, ergodic capacity, fading broadcast channel, Gaussian broadcast channel, outage probability, parallel broadcast channel, power allocation, secrecy capacity.

## I. INTRODUCTION

WIRELESS communication has an inherent broadcast nature, for which security issues are captured by a basic wiretap channel introduced by Wyner in [1]. In this model, a

Manuscript received November 21, 2006; revised October 5, 2007. This work was supported by the National Science Foundation under Grants ANI-03-38807, CNS-06-25637, and CCF-07-28208. This work was also supported by European Commission in the framework of the FP7 Network of Excellence in Wireless Communications NEWCOM ++. The material in this paper was presented in part at the 44th Annual Allerton Conference on Communication, Control and Computing, Monticello, IL, September 2006, in part at the Information Theory and Applications Workshop, San Diego, CA, January 2007, and in part at the IEEE International Symposium on Information Theory, Nice, France, June 2007.

Y. Liang is with the Department of Electrical Engineering, University of Hawaii, Holmes Hall, Honolulu, HI 96822 USA (e-mail: yingbinl@hawaii.edu).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Engineering Quadrangle, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

S. Shamai (Shitz) is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel (e-mail: sshlomo@ee.technion.ac.il).

Communicated by Y. Zheng, Guest Editor for Special Issue on Information Theoretic Security.

Color versions of Figures. 1, 3, and 5–11 in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2008.921678

source node wishes to transmit confidential information to a destination node and wishes to keep a wiretapper as ignorant of this information as possible. The performance measure of interest is the secrecy capacity, which is the largest communication rate achievable from the source node to the destination node with the wiretapper obtaining no information. The secrecy capacity was given in [1] for the discrete memoryless wiretap channel and in [2] for the Gaussian wiretap channel. Fading wiretap channels were studied recently in, e.g., [3]–[7], and multiple-antenna wiretap channels were studied in, e.g., [8]–[11]. A more general model of the wiretap channel was studied by Csiszár and Körner in [12], where the source node also has a common message for both receivers in addition to the confidential message for only one receiver. This channel is regarded as the broadcast channel with confidential messages (BCC). The capacity-equivocation region and the secrecy capacity region of the BCC were characterized in [12]. The BCC was further studied recently in [13]–[15], in which the source node transmits two confidential messages to two receivers, respectively.

In this paper, we investigate the fading BCC, which is based on the BCC studied in [12] with the channels from the source node to receivers 1 and 2 corrupted by multiplicative fading gain coefficients in addition to additive Gaussian noise terms. The fading BCC model captures the basic time-varying property of wireless channels, and hence understanding this channel plays an important role in solving security issues in wireless applications. For the fading BCC, we assume that the fading gain coefficients are stationary and ergodic over time. We further assume that the channel state information (CSI) is known at both the transmitter and the receivers. The CSI at the source node can be realized by a reliable feedback from the two receivers, who are legitimate members of the broadcast network and thus are supposed to receive information from the source node. We note that the cases in which the CSI is not known at the transmitter were considered in [7].

The fading BCC we study in this paper relates to or generalizes a few channels that have been previously studied in the literature. Compared to the fading broadcast channel that was studied in [16]–[20], the fading BCC requires an additional secrecy constraint that the confidential information for one receiver must be perfectly secret from the other receiver. Compared to the fading wiretap channel studied in [5] (the conference version of this paper), [6] and [7] (the full CSI case), the fading BCC we study in this paper assumes that the source node has a common message for both receivers in addition to the confidential message for receiver 1. Hence, the fading BCC includes the fading wiretap channel as a special case. The fading BCC also includes the parallel Gaussian wiretap channel studied in [21] (the case in which wiretappers cooperate) as a special case for the same reason as above and also because a power constraint is assumed for each subchannel in [21].

Before studying the fading BCC, we first study a more general model of the parallel BCC with  $L$  independent subchannels, where the source node communicates with receivers 1 and 2 over  $L$  parallel links. This model serves as a general information-theoretic model that includes the fading BCC as a special case. We establish the secrecy capacity region of the parallel BCC. In particular, we provide a converse proof to show that having independent inputs for each subchannel is optimal to achieve the secrecy capacity region. This fact does not follow directly from the single-letter characterization of the secrecy capacity region of the BCC given in [12]. The secrecy capacity region of the parallel BCC further gives the secrecy capacity region of the parallel BCC with degraded subchannels.

We further study the parallel Gaussian BCC, which is an example parallel BCC with degraded subchannels. We show that the secrecy capacity region of the parallel Gaussian BCC is a union over the rate regions achieved by all source power allocations (among the parallel subchannels). Moreover, we derive the optimal power allocations that achieve the boundary of the secrecy capacity region and hence completely characterize this region. The secrecy capacity region of the parallel Gaussian BCC also establishes the secrecy capacity region for the basic Gaussian BCC. This result complements the secrecy capacity region of the discrete memoryless BCC given by Csiszár and Körner in [12].

We then apply our results to investigate the fading BCC. We first study the ergodic performance, where no delay constraints on message transmission are assumed and the secrecy capacity region is averaged over all channel states. Now the fading BCC can be viewed as the parallel Gaussian BCC with each fading state corresponding to one subchannel. Thus, the secrecy capacity region of the parallel Gaussian BCC applies to the fading BCC. In particular, since the source node knows the CSI, it can dynamically change its transmission power with the channel state realization to achieve the best performance. We obtain the optimal power allocations that achieve the boundary of the secrecy capacity region for the fading BCC.

We further study the outage performance of the fading BCC, where messages must be transmitted over a certain time (one block) to satisfy the delay constraint. We adopt the block-fading model, where the fading coefficients remain constant over one block and change to another realization in the next block. The block length is assumed to be large enough to guarantee decoding in one block. We assume the power constraint at the source node applies over a large number of blocks (i.e., it is a long-term power constraint as in [22]). As in the analysis of the ergodic performance, we assume that the CSI is known both at the transmitter and at the receivers, and hence the source node can allocate its transmission power to achieve the best outage performance. We first obtain the power allocation that minimizes the outage probability where either the target rate of the common message or the target rate of the confidential message is not achieved. We then obtain the power allocation that minimizes the outage probability where the target rate of the confidential message is not achieved subject to the constraint that the target rate of the common message must be achieved for all channel states.

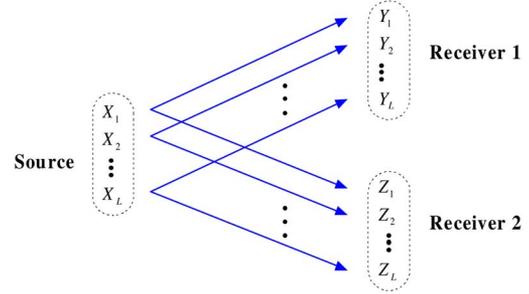


Fig. 1. The parallel BCC.

In this paper, we use  $X_{[1,L]}$  to indicate a group of variables  $(X_1, X_2, \dots, X_L)$ , and use  $X_{[1,L]}^n$  to indicate a group of vectors  $(X_1^n, X_2^n, \dots, X_L^n)$ , where  $X_l^n$  indicates the vector  $(X_{l1}, X_{l2}, \dots, X_{ln})$ . Throughout the paper, the logarithmic function is to the base 2.

The paper is organized as follows. We first study the parallel BCC with independent subchannels, and its special case of the parallel BCC with degraded subchannels. We next study the parallel Gaussian BCC. We then study the ergodic and outage performances of the fading BCC and demonstrate our results with numerical examples. We conclude the paper with a few remarks.

We note that the wiretap channel has also been studied in [23]–[35] and references therein. The topic of common randomness and secret key capacity in communication systems has been studied in [36]–[38]. Such communication system may be viewed as a wiretap channel with side information (which might be common randomness, i.e., a key at a certain rate). Other related work on this topic can be found in [39]–[47] and references therein. We also remark that the secrecy rate/capacity has also been studied for the multiple-access channel in, e.g., [48]–[51], the relay channel in, e.g., [52]–[56], and the interference channel in [13] and [57].

## II. PARALLEL BCCS

### A. Channel Model

We consider the parallel BCC with  $L$  independent subchannels (see Fig. 1), where there are one source node and two receivers. Each subchannel is assumed to be a general broadcast channel from the source node to the two receivers. As in the BCC, the source node wants to transmit common information to both receivers and confidential information to receiver 1. Moreover, the source node wishes to keep the confidential information as secret as possible from receiver 2.

More formally, the parallel BCC consists of  $L$  finite input alphabets  $\mathcal{X}_{[1,L]}$ , and  $2L$  finite output alphabets  $\mathcal{Y}_{[1,L]}$  and  $\mathcal{Z}_{[1,L]}$ . The transition probability distribution is given by

$$p(y_{[1,L]}, z_{[1,L]} | x_{[1,L]}) = \prod_{l=1}^L p_l(y_l, z_l | x_l) \quad (1)$$

where  $x_l \in \mathcal{X}_l$ ,  $y_l \in \mathcal{Y}_l$ , and  $z_l \in \mathcal{Z}_l$  for  $l = 1, \dots, L$ .

If the parallel BCC has only one subchannel, i.e.,  $L = 1$ , this channel becomes the BCC studied in [12]. Moreover, each subchannel is assumed to be a general broadcast channel as in [12] and is not necessarily degraded as assumed in [1].

A  $(2^{nR_0}, 2^{nR_1}, n)$  code consists of the following elements.

- Two message sets:  $\mathcal{W}_0 = \{1, 2, \dots, 2^{nR_0}\}$  and  $\mathcal{W}_1 = \{1, 2, \dots, 2^{nR_1}\}$  with the messages  $W_0$  and  $W_1$  uniformly distributed over the sets  $\mathcal{W}_0$  and  $\mathcal{W}_1$ , respectively;
- One (stochastic) encoder at the source node that maps each message pair  $(w_0, w_1) \in (\mathcal{W}_0, \mathcal{W}_1)$  to a codeword  $x_{[1,L]}^n$ ;
- Two decoders: one at receiver 1 that maps a received sequence  $y_{[1,L]}^n$  to a message pair  $(\hat{w}_0^{(1)}, \hat{w}_1) \in (\mathcal{W}_0, \mathcal{W}_1)$ ; the other at receiver 2 that maps a received sequence  $z_{[1,L]}^n$  to a message  $\hat{w}_0^{(2)} \in \mathcal{W}_0$ .

The secrecy level of the confidential message  $W_1$  achieved at receiver 2 is measured by the *equivocation rate* defined as follows:

$$\frac{1}{n} H(W_1 | Z_{[1,L]}^n). \quad (2)$$

The higher the equivocation rate, the less information that receiver 2 obtains about the confidential message  $W_1$ .

The average error probability is

$$P_e = \frac{1}{2^{nR_0} 2^{nR_1}} \times \sum_{w_0=1}^{2^{nR_0}} \sum_{w_1=1}^{2^{nR_1}} \Pr \left\{ (\hat{w}_0^{(1)}, \hat{w}_1) \neq (w_0, w_1) \text{ or } \hat{w}_0^{(2)} \neq w_0 \right\}. \quad (3)$$

A rate-equivocation triple  $(R_0, R_1, R_e)$  is *achievable* if there exists a sequence of  $(2^{nR_0}, 2^{nR_1}, n)$  codes with the average error probability  $P_e \rightarrow 0$  as  $n$  goes to infinity and with the equivocation rate  $R_e$  satisfying

$$R_e \leq \lim_{n \rightarrow \infty} \frac{1}{n} H(W_1 | Z_{[1,L]}^n). \quad (4)$$

In this paper, we focus on the case in which perfect secrecy is achieved, i.e., receiver 2 does not obtain any information about the message  $W_1$ . This happens if  $R_e = R_1$ . In this case, we define the secrecy capacity region in the following.

*Definition 1:* The *secrecy capacity region*  $\mathcal{C}_s$  is defined to be the set that includes all  $(R_0, R_1)$  such that  $(R_0, R_1, R_e = R_1)$  is achievable, i.e.,

$$\mathcal{C}_s = \{(R_0, R_1) : (R_0, R_1, R_e = R_1) \text{ is achievable}\}. \quad (5)$$

We note that the definition of secrecy capacity region was given in [1] and [12] as the performance measure in the case of perfect secrecy. In this paper, we study this classical notion in the context of fading broadcast channels.

We also note that the notion of secrecy studied in this paper is the same as that in cryptography. However, in this paper, we focus on secrecy achieved by exploiting the underlying physical channel, which is different from cryptographic methods based on shared keys. Coding schemes for accomplishing secrecy by physical layer approaches (which involve stochastic encoding) can be found in [1] and [12]. Here, we apply the information-theoretic approaches to study wireless fading broadcast channels.

## B. Secrecy Capacity Region of Parallel BCCs

For the parallel BCC, we obtain the following secrecy capacity region.

*Theorem 1:* The secrecy capacity region of the parallel BCC is given by

$$\mathcal{C}_s = \bigcup_{\prod_l [p(q_l)p(u_l | q_l)p(x_l | u_l)p(y_l, z_l | x_l)]} \left\{ \begin{array}{l} (R_0, R_1) : \\ R_0 \leq \min \left\{ \sum_{l=1}^L I(Q_l; Y_l), \sum_{l=1}^L I(Q_l; Z_l) \right\} \\ R_1 \leq \sum_{l=1}^L [I(U_l; Y_l | Q_l) - I(U_l; Z_l | Q_l)] \end{array} \right\} \quad (6)$$

where  $Q_l$  can be chosen as a deterministic function of  $U_l$  for  $l = 1, \dots, L$ .

*Proof:* See Appendix I.  $\square$

If the source node transmits only confidential information to receiver 1, i.e.,  $R_0 = 0$ , the parallel BCC becomes the parallel wiretap channel. The secrecy capacity of this channel is given in the following corollary.

*Corollary 1:* The secrecy capacity of the parallel wiretap channel is

$$C_s = \sum_{l=1}^L C_s^{(l)} \quad (7)$$

where  $C_s^{(l)}$  is the secrecy capacity of subchannel  $l$  and is given by

$$C_s^{(l)} = \max [I(U_l; Y_l) - I(U_l; Z_l)] \quad (8)$$

The maximum in the preceding equation is over the distributions  $p(u_l, x_l)p(y_l, z_l | x_l)$ , which satisfies the Markov chain condition  $U_l \rightarrow X_l \rightarrow (Y_l, Z_l)$ .

*Proof:* Corollary 1 follows from Theorem 1 by setting  $R_0 = 0$  and noticing that  $I(U_l; Y_l | Q_l) - I(U_l; Z_l | Q_l)$  is maximized by a constant  $Q_l$ .  $\square$

*Remark 1:* Theorem 1 implies an important property that having independent inputs for each subchannel is optimal. This fact does not follow directly from the single-letter result on the secrecy capacity of the BCC given in [12], although the parallel BCC can be viewed as a special case of the BCC. Hence, a converse proof is needed, which is provided in Appendix I.

We note that the secrecy capacity region  $\mathcal{C}_s^{(l)}$  of subchannel  $l$  is given in [12, Corollary 1], i.e.,

$$\mathcal{C}_s^{(l)} = \bigcup_{p(q_l)p(u_l | q_l)p(x_l | u_l)p(y_l, z_l | x_l)} \left\{ \begin{array}{l} (R_0, R_1) : \\ R_0 \leq \min \{I(Q_l; Y_l), I(Q_l; Z_l)\} \\ R_1 \leq I(U_l; Y_l | Q_l) - I(U_l; Z_l | Q_l) \end{array} \right\}. \quad (9)$$

We now define the sum of the secrecy capacity regions of the subchannels to be

$$\mathcal{C}_{\text{sum}} := \left\{ \begin{array}{l} (R_0, R_1) : \\ R_0 = \sum_l R_{l0}, \quad R_1 = \sum_l R_{l1}, \\ \text{with } (R_{l0}, R_{l1}) \in \mathcal{C}_s^{(l)} \text{ for } l = 1, \dots, L. \end{array} \right\} \quad (10)$$

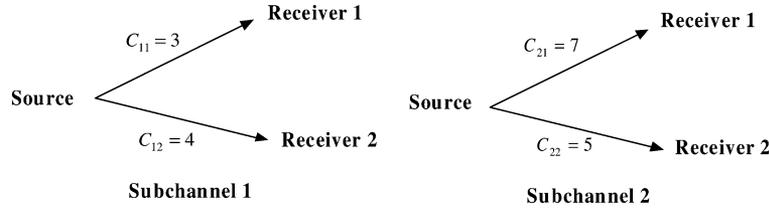


Fig. 2. A parallel BCC example.

*Remark 2:* The secrecy capacity region  $\mathcal{C}_s$  in Theorem 1 may be larger than the sum  $\mathcal{C}_{\text{sum}}$  of the secrecy capacity regions of the subchannels. Hence, the secrecy capacity region of the parallel BCC is achieved by coding over all parallel subchannels.

This observation was also made in [58] for the broadcast channel with common messages. This fact follows because the common rate has the following property

$$R_0 = \min \left\{ \sum_{l=1}^L I(Q_l; Y_l), \sum_{l=1}^L I(Q_l; Z_l) \right\} \geq \sum_{l=1}^L \min\{I(Q_l; Y_l), I(Q_l; Z_l)\} = \sum_{l=1}^L R_{l0}. \quad (11)$$

This can also be seen from the following simple example. For simplicity, we consider the case in which the source node has only the common message for both receivers. We further assume  $L = 2$ , and each subchannel is a deterministic broadcast channel for  $l = 1, 2$  (see Fig. 2). For subchannel 1, the link capacities to receivers 1 and 2 are  $C_{11} = 3$  and  $C_{12} = 4$ , respectively. For subchannel 2, the link capacities to receivers 1 and 2 are  $C_{21} = 7$  and  $C_{22} = 5$ , respectively. The capacity of this parallel channel is given by

$$C = \min\{C_{11} + C_{21}, C_{12} + C_{22}\} = \min\{3 + 7, 4 + 5\} = 9. \quad (12)$$

However, the sum of the capacities of the two subchannels is

$$\sum_{l=1}^2 \min\{C_{l1}, C_{l2}\} = \min\{3, 4\} + \min\{7, 5\} = 8 \quad (13)$$

which is clearly smaller than the capacity given in (12).

Similarly to Theorem 14.6.1 in [59], we obtain the following lemma for the BCC studied in [12], which also applies to the parallel BCC we study in this paper.

*Lemma 1:* The secrecy capacity region of the BCC studied in [12] depends only on the marginal transition probability distributions  $p(y|x)$  of the channel from the source node to receiver 1 and  $p(z|x)$  of the channel from the source node to receiver 2.

*Proof:* The proof follows from the reasoning in [59, p. 454, Problem 10] and the fact that the equivocation rate  $\frac{1}{n}H(W_1 | Z_{[1,L]}^n)$  depends only on the marginal distribution of  $p(z|x)$ .  $\square$

One application of Lemma 1 is to obtain the following generalization of the result in [2] for the Gaussian wiretap channel, which is a special case of the BCC.

*Corollary 2:* The secrecy capacity of the Gaussian wiretap channel given in [2, Theorem 1] holds for the case, where the noise variables at the destination node and the wiretapper have a general correlation structure.

Lemma 1 will be useful in establishing the secrecy capacity region of the parallel Gaussian BCC in Section III.

### C. Parallel BCCs With Degraded Subchannels

We consider the parallel BCC with degraded subchannels (see Fig. 3), where each subchannel is either degraded such that the output at receiver 2 is a degraded version of the output at receiver 1, or degraded such that the output at receiver 1 is a degraded version of the output at receiver 2. Note that although each subchannel is degraded, the entire channel may not be degraded because the subchannels may not be degraded in the same fashion.

We define  $A$  to be the index set that includes all indices of subchannels, where the output at receiver 2 is a degraded version of the output at receiver 1, i.e.,

$$p_l(y_l, z_l | x_l) = p_l(y_l | x_l)p_l(z_l | y_l), \quad \text{for } l \in A. \quad (14)$$

Hence, the Markov chain condition  $X_l \rightarrow Y_l \rightarrow Z_l$  is satisfied for  $l \in A$ . We define  $A^c$  to be the complement of the set  $A$ , and  $A^c$  includes all indices of subchannels, where the output at receiver 1 is a degraded version of the output at receiver 2, i.e.,

$$p_l(y_l, z_l | x_l) = p_l(z_l | x_l)p_l(y_l | z_l), \quad \text{for } l \in A^c. \quad (15)$$

Hence, the Markov chain condition  $X_l \rightarrow Z_l \rightarrow Y_l$  is satisfied for  $l \in A^c$ . The channel transition probability distribution is given by

$$p(y_{[1,L]}, z_{[1,L]} | x_{[1,L]}) = \prod_{l \in A} [p_l(y_l | x_l)p_l(z_l | y_l)] \prod_{l \in A^c} [p_l(z_l | x_l)p_l(y_l | z_l)]. \quad (16)$$

For the parallel BCC with degraded subchannels, we apply Theorem 1 and obtain the following secrecy capacity region.

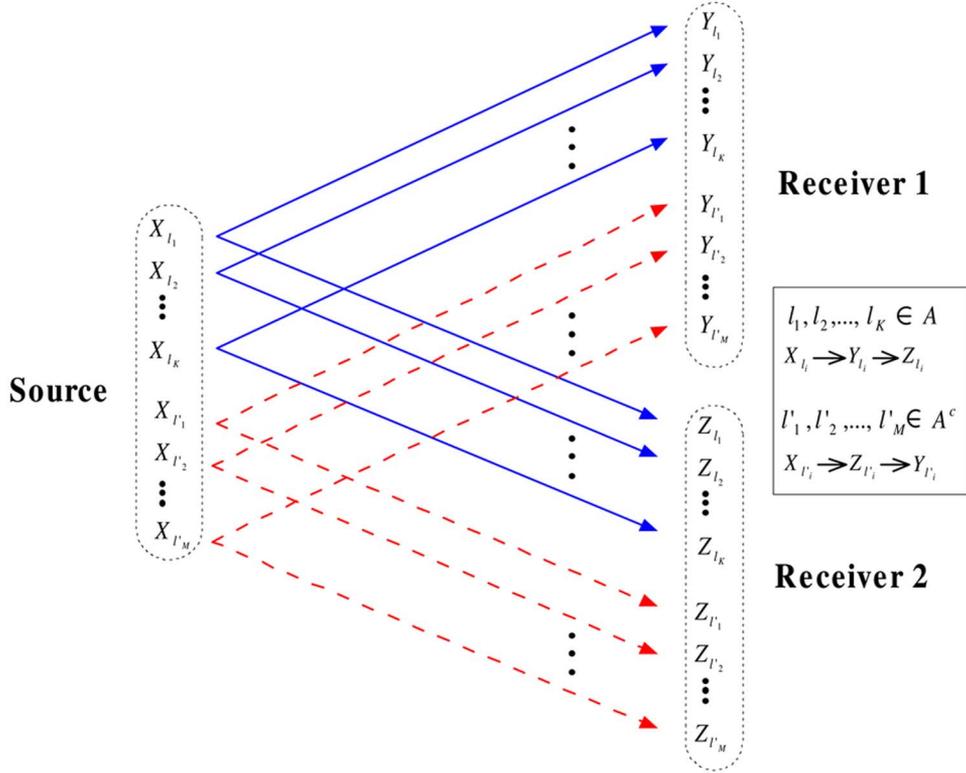


Fig. 3. The parallel BCC with degraded subchannels.

*Corollary 3:* The secrecy capacity region of the parallel BCC with degraded subchannels is

$$\mathcal{C}_s = \bigcup_{\substack{(R_0, R_1) : \\ R_0 \leq \min \left\{ \sum_{l \in A} I(Q_l; Y_l) + \sum_{l \in A^c} I(X_l; Y_l), \right. \\ \left. \sum_{l \in A} I(Q_l; Z_l) + \sum_{l \in A^c} I(X_l; Z_l) \right\} \\ R_1 \leq \sum_{l \in A} [I(X_l; Y_l | Q_l) - I(X_l; Z_l | Q_l)]}} \left\{ \begin{array}{l} (R_0, R_1) : \\ R_0 \leq \min \left\{ \sum_{l \in A} I(Q_l; Y_l) + \sum_{l \in A^c} I(X_l; Y_l), \right. \\ \left. \sum_{l \in A} I(Q_l; Z_l) + \sum_{l \in A^c} I(X_l; Z_l) \right\} \\ R_1 \leq \sum_{l \in A} [I(X_l; Y_l | Q_l) - I(X_l; Z_l | Q_l)] \end{array} \right\} \quad (17)$$

*Remark 3:* It can be seen that the common message  $W_0$  is sent over all subchannels, and the confidential message  $W_1$  for receiver 1 is sent only over the subchannels for which the output at receiver 2 is a degraded version of the output at receiver 1, i.e.,  $l \in A$ . Furthermore, over these subchannels, the messages  $W_0$  and  $W_1$  are sent by using the superposition encoding scheme.

*Proof:* The achievability follows from Theorem 1 by setting  $U_l = X_l$  for  $l \in A$  and setting  $Q_l = U_l = X_l$  for  $l \in A^c$ .

To show the converse, we first note that for  $l \in A^c$

$$I(Q_l; Y_l) \leq I(X_l; Y_l) \quad \text{and} \quad I(Q_l; Z_l) \leq I(X_l; Z_l) \quad (18)$$

which follow from the Markov condition  $Q_l \rightarrow X_l \rightarrow (Y_l Z_l)$ . We apply the bounds in (18) to the bound on  $R_0$  given in (6) and obtain the bound on  $R_0$  given in (17).

For  $l \in A^c$ , we also obtain

$$\begin{aligned} I(U_l; Y_l | Q_l) - I(U_l; Z_l | Q_l) &\leq I(U_l; Y_l Z_l | Q_l) - I(U_l; Z_l | Q_l) \\ &= I(U_l; Z_l | Q_l) + I(U_l; Y_l | Q_l Z_l) - I(U_l; Z_l | Q_l) \\ &= 0 \end{aligned} \quad (19)$$

where the last equality follows because  $I(U_l; Y_l | Q_l Z_l) = 0$  due to the degradedness condition (15).

For  $l \in A$ , we obtain the following bound:

$$\begin{aligned} I(U_l; Y_l | Q_l) - I(U_l; Z_l | Q_l) &= I(X_l U_l; Y_l | Q_l) - I(X_l; Y_l | Q_l U_l) \\ &\quad - I(X_l U_l; Z_l | Q_l) + I(X_l; Z_l | Q_l U_l) \\ &= I(X_l; Y_l | Q_l) + I(U_l; Y_l | Q_l X_l) \\ &\quad - I(X_l; Y_l | U_l) - I(X_l; Z_l | Q_l) \\ &\quad - I(U_l; Z_l | Q_l X_l) + I(X_l; Z_l | U_l) \\ &\stackrel{(a)}{=} I(X_l; Y_l | Q_l) - I(X_l; Y_l | U_l) - I(X_l; Z_l | Q_l) \\ &\quad + I(X_l; Z_l | U_l) \\ &\leq I(X_l; Y_l | Q_l) - I(X_l; Y_l | U_l) - I(X_l; Z_l | Q_l) \\ &\quad + I(X_l; Y_l Z_l | U_l) \\ &= I(X_l; Y_l | Q_l) - I(X_l; Y_l | U_l) - I(X_l; Z_l | Q_l) \\ &\quad + I(X_l; Y_l | U_l) + I(X_l; Z_l | U_l Y_l) \\ &\stackrel{(b)}{\leq} I(X_l; Y_l | Q_l) - I(X_l; Z_l | Q_l) \end{aligned} \quad (20)$$

where (a) follows because  $I(U_l; Y_l | Q_l X_l) = 0$  and  $I(U_l; Z_l | Q_l X_l) = 0$  due to the Markov chain condition  $Q_l \rightarrow U_l \rightarrow X_l \rightarrow (Y_l Z_l)$ , and (b) follows because  $I(X_l; Z_l | U_l Y_l) = 0$  due to the degradedness condition (14).

By applying the bounds (19) and (20) to the bound on  $R_1$  given in (17), we obtain the bound on  $R_1$  given in (6). This concludes the proof of the converse.  $\square$

### III. PARALLEL GAUSSIAN BCCS

In this section, we study the parallel Gaussian BCCs, where the channel outputs at receivers 1 and 2 are corrupted by additive

Gaussian noise terms. The channel input–output relationship is given by

$$Y_{li} = X_{li} + W_{li}, \quad Z_{li} = X_{li} + V_{li}, \quad \text{for } l = 1, \dots, L \quad (21)$$

where  $i$  is the time index. For  $l = 1, \dots, L$ , the noise processes  $\{W_{li}\}$  and  $\{V_{li}\}$  are independent and identically distributed (i.i.d.) with the components being zero-mean Gaussian random variables with variances  $\mu_l^2$  and  $\nu_l^2$ , respectively. We assume  $\mu_l^2 < \nu_l^2$  for  $l \in A$  and  $\mu_l^2 \geq \nu_l^2$  for  $l \in A^c$ . The channel input sequence  $X_{[1,L]}^n$  is subject to the average power constraints  $P$ , i.e.,

$$\frac{1}{n} \sum_{i=1}^n \sum_{l=1}^L \mathbb{E}[X_{li}^2] \leq P. \quad (22)$$

We now apply Lemma 1 to obtain the secrecy capacity region of the parallel Gaussian BCC. It can be seen from (21) that the subchannels of the parallel Gaussian BCC are not physically degraded. We consider the following subchannels:

$$\begin{aligned} Y_{li} &= X_{li} + W_{li}, & Z_{li} &= X_{li} + W_{li} + V'_{li}, & \text{for } l \in A; \\ Y_{li} &= X_{li} + V_{li} + W'_{li}, & Z_{li} &= X_{li} + V_{li}, & \text{for } l \in A^c \end{aligned} \quad (23)$$

where  $\{W'_{li}\}$  and  $\{V'_{li}\}$  are i.i.d. random processes with components being zero-mean Gaussian random variables with variances  $\mu_l^2 - \nu_l^2$  (for  $l \in A^c$ ) and  $\nu_l^2 - \mu_l^2$  (for  $l \in A$ ), respectively. Moreover,  $\{V'_{li}\}$  is independent of  $\{W_{li}\}$  and  $\{W'_{li}\}$  is independent of  $\{V_{li}\}$ . It can be seen that the channel defined in (23) has physically degraded subchannels. This channel has the same marginal distributions  $p(y|x)$  and  $p(z|x)$  as the parallel Gaussian BCC defined in (21). Hence, by Lemma 1, the two channels have the same secrecy capacity region.

For the channel defined in (23), we can apply Corollary 3 to obtain the secrecy capacity region. In particular, the degradedness of the subchannels allows the use of the entropy power inequality in the proof of the converse. The secrecy capacity region obtained for this channel also applies to the parallel Gaussian BCC defined in (21), and is presented in the following theorem.

**Theorem 2:** The secrecy capacity region of the parallel Gaussian BCC is

$$\mathcal{C}_s^g = \bigcup_{\underline{p} \in \mathcal{P}} \left\{ \begin{aligned} &(R_0, R_1) : \\ &R_0 \leq \min \left\{ \sum_{l \in A} \frac{1}{2} \log \left( 1 + \frac{p_{l0}}{\mu_l^2 + p_{l1}} \right) + \sum_{l \in A^c} \frac{1}{2} \log \left( 1 + \frac{p_{l0}}{\mu_l^2} \right), \right. \\ &\left. \sum_{l \in A} \frac{1}{2} \log \left( 1 + \frac{p_{l0}}{\nu_l^2 + p_{l1}} \right) + \sum_{l \in A^c} \frac{1}{2} \log \left( 1 + \frac{p_{l0}}{\nu_l^2} \right) \right\} \\ &R_1 \leq \sum_{l \in A} \left[ \frac{1}{2} \log \left( 1 + \frac{p_{l1}}{\mu_l^2} \right) - \frac{1}{2} \log \left( 1 + \frac{p_{l1}}{\nu_l^2} \right) \right] \end{aligned} \right\} \quad (24)$$

where  $\underline{p}$  is the power allocation vector, which consists of  $(p_{l0}, p_{l1})$  for  $l \in A$  and  $p_{l0}$  for  $l \in A^c$  as components, and the

set  $\mathcal{P}$  includes all power allocation vectors  $\underline{p}$  that satisfy the power constraint (22), i.e.,

$$\mathcal{P} := \left\{ \underline{p} : \sum_{l \in A} [p_{l0} + p_{l1}] + \sum_{l \in A^c} p_{l0} \leq P \right\}. \quad (25)$$

*Proof:* See Appendix II.  $\square$

Note that  $\underline{p}$  indicates the power allocation among all subchannels. For  $l \in A$ , since the source node transmits both common and confidential messages,  $p_{l0}$  and  $p_{l1}$  indicate the powers allocated to transmit the common and confidential messages, respectively. For  $l \in A^c$ , the source transmits only the common message, and  $p_{l0}$  indicates the power allocated to transmit the common message.

If  $L = 1$ , the parallel Gaussian BCC becomes the Gaussian BCC. The following secrecy capacity region of the Gaussian BCC follows directly from Theorem 2.

**Corollary 4:** The secrecy capacity region of the Gaussian BCC is

$$\mathcal{C}_s = \bigcup_{0 \leq \beta \leq 1} \left\{ \begin{aligned} &(R_0, R_1) : \\ &R_0 \leq \min \left\{ \frac{1}{2} \log \left( 1 + \frac{(1-\beta)P}{\mu^2 + \beta P} \right), \right. \\ &\quad \left. \frac{1}{2} \log \left( 1 + \frac{(1-\beta)P}{\nu^2 + \beta P} \right) \right\} \\ &R_1 \leq \left[ \frac{1}{2} \log \left( 1 + \frac{\beta P}{\mu^2} \right) - \frac{1}{2} \log \left( 1 + \frac{\beta P}{\nu^2} \right) \right]^+ \end{aligned} \right\} \quad (26)$$

where  $(x)^+ = x$  if  $x > 0$  and  $(x)^+ = 0$  if  $x \leq 0$ .

To characterize the secrecy capacity region of the parallel Gaussian BCC given in (24), we need to characterize every boundary point and the corresponding power allocation vector that achieves this boundary point. It is clear that the secrecy capacity region given in (24) is convex due to the converse proof in Appendix II. Hence, the boundary of the secrecy capacity region can be characterized as follows. For every point  $(R_0^*, R_1^*)$  on the boundary, there exist  $\gamma_0 > 0$  and  $\gamma_1 > 0$  such that  $(R_0^*, R_1^*)$  is the solution to the following optimization problem:

$$\max_{(R_0, R_1) \in \mathcal{C}_s^g} [\gamma_0 R_0 + \gamma_1 R_1]. \quad (27)$$

Therefore, the power allocation  $\underline{p}^*$  that achieves the boundary point  $(R_0^*, R_1^*)$  is the solution to the following optimization problem:

$$\begin{aligned} &\max_{\underline{p} \in \mathcal{P}} [\gamma_0 R_0(\underline{p}) + \gamma_1 R_1(\underline{p})] \\ &= \max_{\underline{p} \in \mathcal{P}} [\gamma_0 \min\{R_{01}(\underline{p}), R_{02}(\underline{p})\} + \gamma_1 R_1(\underline{p})] \end{aligned} \quad (28)$$

where  $R_0(\underline{p})$  and  $R_1(\underline{p})$  indicate the bounds on  $R_0$  and  $R_1$  in (24). We further define  $R_{01}(\underline{p})$  and  $R_{02}(\underline{p})$  to be the two terms over which the minimization in  $R_0(\underline{p})$  is taken, i.e.,  $R_0(\underline{p}) = \min\{R_{01}(\underline{p}), R_{02}(\underline{p})\}$ . The optimization (28) serves as a complete characterization of the boundary of the secrecy capacity

region of the parallel Gaussian BCC. The solution to (28) provides the power allocations that achieve the boundary of the secrecy capacity region. Our goal now is to solve the optimization problem (28).

The optimization problem (28) is a max-min optimization, and can be solved by the approach used in [60]. The main idea is contained in Proposition 1 in [60], which is stated in the following lemma.

*Lemma 2:* The optimal  $\underline{p}^*$  that solves (28) falls into one of the following three cases:

- Case 1:  $\underline{p}^*$  maximizes  $\gamma_0 R_{01}(\underline{p}) + \gamma_1 R_1(\underline{p})$ ,  
and  $R_{01}(\underline{p}^*) < R_{02}(\underline{p}^*)$ ;
- Case 2:  $\underline{p}^*$  maximizes  $\gamma_0 R_{02}(\underline{p}) + \gamma_1 R_1(\underline{p})$ ,  
and  $R_{01}(\underline{p}^*) > R_{02}(\underline{p}^*)$ ;
- Case 3:  $\underline{p}^*$  maximizes  $\gamma_0(\alpha R_{01}(\underline{p}) + \bar{\alpha} R_{02}(\underline{p})) + \gamma_1 R_1(\underline{p})$ ,  
where  $0 \leq \alpha \leq 1$  is such that  $R_{01}(\underline{p}^*) = R_{02}(\underline{p}^*)$ ,  
and  $\bar{\alpha} = 1 - \alpha$ . (29)

By applying Lemma 2, we obtain the optimal power allocation  $\underline{p}^*$  that solves (28).

*Theorem 3:* The optimal power allocation vector  $\underline{p}^*$  that solves (28) and hence achieves the boundary of the secrecy capacity region of the parallel Gaussian BCC has one of the following three forms.

Case 1:  $\underline{p}^* = \underline{p}^{(1)}$  if the following  $\underline{p}^{(1)}$  satisfies  $R_{01}(\underline{p}^{(1)}) < R_{02}(\underline{p}^{(1)})$ .

For  $l \in A$ , if  $\frac{\gamma_1}{\gamma_0} > \frac{\nu_l^2}{\nu_l^2 - \mu_l^2}$ , then

$$p_{l0}^{(1)} = \left( \frac{\gamma_0}{2\lambda \ln 2} - \left( \frac{\gamma_1}{\gamma_0} - 1 \right) (\nu_l^2 - \mu_l^2) \right)^+ \\ \text{and} \\ p_{l1}^{(1)} = \left( \min \left\{ \frac{1}{2} \sqrt{(\nu_l^2 - \mu_l^2) \left( \nu_l^2 - \mu_l^2 + \frac{2\gamma_1}{\lambda \ln 2} \right)} \right. \right. \\ \left. \left. - \frac{1}{2} (\mu_l^2 + \nu_l^2), \frac{\gamma_1}{\gamma_0} (\nu_l^2 - \mu_l^2) - \nu_l^2 \right\} \right)^+.$$

Alternatively, if  $\frac{\gamma_1}{\gamma_0} \leq \frac{\nu_l^2}{\nu_l^2 - \mu_l^2}$ , then

$$p_{l0}^{(1)} = \left( \frac{\gamma_0}{2\lambda \ln 2} - \mu_l^2 \right)^+ \quad \text{and} \quad p_{l1}^{(1)} = 0$$

For  $l \in A^c$

$$p_{l0}^{(1)} = \left( \frac{\gamma_0}{2\lambda \ln 2} - \mu_l^2 \right)^+ \quad (30)$$

where  $\lambda$  is chosen to satisfy the power constraint

$$\sum_{l \in A} [p_{l0} + p_{l1}] + \sum_{l \in A^c} p_{l0} \leq P. \quad (31)$$

Case 2:  $\underline{p}^* = \underline{p}^{(2)}$  if the following  $\underline{p}^{(2)}$  satisfies  $R_{01}(\underline{p}^{(2)}) > R_{02}(\underline{p}^{(2)})$ .

For  $l \in A$ , if  $\frac{\gamma_1}{\gamma_0} > \frac{\mu_l^2}{\nu_l^2 - \mu_l^2}$ , then

$$p_{l0}^{(2)} = \left( \frac{\gamma_0}{2\lambda \ln 2} - \left( \frac{\gamma_1}{\gamma_0} + 1 \right) (\nu_l^2 - \mu_l^2) \right)^+$$

and

$$p_{l1}^{(2)} = \left( \min \left\{ \frac{1}{2} \sqrt{(\nu_l^2 - \mu_l^2) \left( \nu_l^2 - \mu_l^2 + \frac{2\gamma_1}{\lambda \ln 2} \right)} \right. \right. \\ \left. \left. - \frac{1}{2} (\mu_l^2 + \nu_l^2), \frac{\gamma_1}{\gamma_0} (\nu_l^2 - \mu_l^2) - \mu_l^2 \right\} \right)^+.$$

Alternatively, if  $\frac{\gamma_1}{\gamma_0} \leq \frac{\mu_l^2}{\nu_l^2 - \mu_l^2}$ , then

$$p_{l0}^{(2)} = \left( \frac{\gamma_0}{2\lambda \ln 2} - \nu_l^2 \right)^+ \quad \text{and} \quad p_{l1}^{(2)} = 0.$$

For  $l \in A^c$

$$p_{l0}^{(2)} = \left( \frac{\gamma_0}{2\lambda \ln 2} - \nu_l^2 \right)^+ \quad (32)$$

where  $\lambda$  is chosen to satisfy the power constraint defined in (31).

Case 3:  $\underline{p}^* = \underline{p}^{(\alpha)}$  if there exists  $0 \leq \alpha \leq 1$  such that the following  $\underline{p}^{(\alpha)}$  satisfies  $R_{01}(\underline{p}^{(\alpha)}) = R_{02}(\underline{p}^{(\alpha)})$ .

For  $l \in A$ , if  $\frac{\gamma_1}{\gamma_0} > \frac{\alpha \nu_l^2 + \bar{\alpha} \mu_l^2}{\nu_l^2 - \mu_l^2}$ , then

$$p_{l0}^{(\alpha)} = \left( \frac{1}{2} \sqrt{(\nu_l^2 - \mu_l^2 - \frac{\gamma_0}{2 \ln 2 \lambda})^2 + \frac{2\alpha\gamma_0}{\lambda \ln 2} (\nu_l^2 - \mu_l^2)} \right. \\ \left. + \frac{\gamma_0}{4 \ln 2 \lambda} - \left( \frac{\gamma_1}{\gamma_0} - \alpha + \frac{1}{2} \right) (\nu_l^2 - \mu_l^2) \right)^+$$

and

$$p_{l1}^{(\alpha)} = \left( \min \left\{ \frac{1}{2} \sqrt{(\nu_l^2 - \mu_l^2) \left( \nu_l^2 - \mu_l^2 + \frac{2\gamma_1}{\lambda \ln 2} \right)} \right. \right. \\ \left. \left. - \frac{1}{2} (\mu_l^2 + \nu_l^2), \frac{\gamma_1}{\gamma_0} (\nu_l^2 - \mu_l^2) - (\alpha \nu_l^2 + \bar{\alpha} \mu_l^2) \right\} \right)^+.$$

Alternatively, if  $\frac{\gamma_1}{\gamma_0} \leq \frac{\alpha \nu_l^2 + \bar{\alpha} \mu_l^2}{\nu_l^2 - \mu_l^2}$ , then

$$p_{l0}^{(\alpha)} = \left( \frac{1}{2} \sqrt{(\nu_l^2 - \mu_l^2 - \frac{\gamma_0}{2 \ln 2 \lambda})^2 + \frac{2\alpha\gamma_0}{\lambda \ln 2} (\nu_l^2 - \mu_l^2)} \right. \\ \left. - \frac{1}{2} \left( \mu_l^2 + \nu_l^2 - \frac{\gamma_0}{2 \ln 2 \lambda} \right) \right)^+$$

and

$$p_{l1}^{(\alpha)} = 0.$$

For  $l \in A^c$

$$p_{l0}^{(\alpha)} = \left( \frac{1}{2} \sqrt{(\nu_l^2 - \mu_l^2 - \frac{\gamma_0}{2 \ln 2 \lambda})^2 + \frac{2\alpha\gamma_0}{\lambda \ln 2} (\nu_l^2 - \mu_l^2)} \right. \\ \left. - \frac{1}{2} \left( \mu_l^2 + \nu_l^2 - \frac{\gamma_0}{2 \ln 2 \lambda} \right) \right)^+ \quad (33)$$

where  $\lambda$  is chosen to satisfy the power constraint defined in (31).

*Proof:* See Appendix III.  $\square$

Based on Theorem 3, we provide the following algorithm to find the optimal  $\underline{p}^*$ .

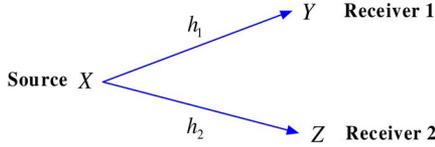


Fig. 4. The fading BCC.

---

 Algorithm to find  $\underline{p}^*$  that solves (28)
 

---

- Step 1. Find  $\underline{p}^{(1)}$  given in (30).  
 If  $R_{01}(\underline{p}^{(1)}) < R_{02}(\underline{p}^{(1)})$ , then  $\underline{p}^* = \underline{p}^{(1)}$  and finish.  
 Otherwise, go to Step 2.
- Step 2. Find  $\underline{p}^{(2)}$  given in (32).  
 If  $R_{01}(\underline{p}^{(2)}) > R_{02}(\underline{p}^{(2)})$ , then  $\underline{p}^* = \underline{p}^{(2)}$  and finish.  
 Otherwise, go to Step 3.
- Step 3. For a given  $\alpha$ , find  $\underline{p}^{(\alpha)}$  given in (33).  
 Search over  $0 \leq \alpha \leq 1$  to find  $\alpha$  that satisfies
- $$R_{01}(\underline{p}^{(\alpha)}) = R_{02}(\underline{p}^{(\alpha)}).$$
- Then  $\underline{p}^* = \underline{p}^{(\alpha)}$  and finish.
- 

#### IV. FADING BCCS: ERGODIC SECRECY CAPACITY REGION

In this section, we study the fading BCC (see Fig. 4), where the channels from the source node to receivers 1 and 2 are corrupted by multiplicative fading gain processes in addition to additive white Gaussian processes. The channel input–output relationship is given by

$$Y_i = h_{1i}X_i + W_i \quad \text{and} \quad Z_i = h_{2i}X_i + V_i \quad (34)$$

where  $i$  is the time index,  $X_i$  is the channel input at the time instant  $i$ , and  $Y_i$  and  $Z_i$  are channel outputs at the time instant  $i$  at receivers 1 and 2, respectively. The channel gain coefficients  $h_{1i}$  and  $h_{2i}$  are proper complex random variables. We define  $\underline{h}_i := (h_{1i}, h_{2i})$ , and assume  $\{\underline{h}_i\}$  is a stationary and ergodic vector random process. The noise processes  $\{W_i\}$  and  $\{V_i\}$  are zero-mean i.i.d. proper complex Gaussian with  $W_i$  and  $V_i$  having variances  $\mu^2$  and  $\nu^2$ , respectively. The input sequence  $\{X_i\}$  is subject to the average power constraint  $P$ , i.e.,  $\frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i^2] \leq P$ .

We assume that the CSI (i.e., the realization of  $\underline{h}_i$ ) is known at both the transmitter and the receivers instantaneously. The CSI at the source node can be realized by a reliable feedback from the two receivers, who are legitimate participants in the network and are thus supposed to receive information from the source node. Depending on the CSI, the source node can dynamically change its transmission power to achieve better performance. In this section, we assume that there are no delay constraints on the transmitted messages, and the performance criterion we study, i.e., the secrecy capacity region, is averaged over all channel states and is referred to as the *ergodic* secrecy capacity region.

It can be seen that for a given fading state, i.e., a realization of  $\underline{h}_i$ , the fading BCC is a Gaussian BCC. Hence, the fading BCC can be viewed as a parallel Gaussian BCC with each fading state corresponding to one subchannel. Thus, the following secrecy capacity region of the fading BCC follows from Theorem 2. In the following, for each channel state  $\underline{h}$ , we use  $p_0(\underline{h})$  and  $p_1(\underline{h})$  to denote the source powers allocated to transmit the common and confidential messages, respectively. We define  $p(\underline{h}) = (p_0(\underline{h}), p_1(\underline{h}))$ . We further define the following set  $\mathcal{P}$  that includes all power allocations that satisfy the power constraint:

$$\mathcal{P} = \{p(\underline{h}) = (p_0(\underline{h}), p_1(\underline{h})) : \mathbb{E}_A[p_0(\underline{h}) + p_1(\underline{h})] + \mathbb{E}_{A^c}[p_0(\underline{h})] \leq P\}. \quad (35)$$

Note that  $p_1(\underline{h}) = 0$  for  $\underline{h} \in A^c$ .

*Corollary 5:* The secrecy capacity region of the fading BCC is

$$C_s = \bigcup_{p(\underline{h}) \in \mathcal{P}} \left\{ \begin{array}{l} (R_0, R_1) : \\ R_0 \leq \min \left\{ \mathbb{E}_{\underline{h} \in A} \log \left( 1 + \frac{p_0(\underline{h})|h_1|^2}{\mu^2 + p_1(\underline{h})|h_1|^2} \right) \right. \\ \quad \left. + \mathbb{E}_{\underline{h} \in A^c} \log \left( 1 + \frac{p_0(\underline{h})|h_1|^2}{\mu^2} \right), \right. \\ \quad \left. \mathbb{E}_{\underline{h} \in A} \log \left( 1 + \frac{p_0(\underline{h})|h_2|^2}{\nu^2 + p_1(\underline{h})|h_2|^2} \right) \right. \\ \quad \left. + \mathbb{E}_{\underline{h} \in A^c} \log \left( 1 + \frac{p_0(\underline{h})|h_2|^2}{\nu^2} \right) \right\} \\ R_1 \leq \mathbb{E}_{\underline{h} \in A} \left[ \log \left( 1 + \frac{p_1(\underline{h})|h_1|^2}{\mu^2} \right) - \log \left( 1 + \frac{p_1(\underline{h})|h_2|^2}{\nu^2} \right) \right] \end{array} \right\} \quad (36)$$

where

$$A := \left\{ \underline{h} : \frac{|h_1|^2}{\mu^2} > \frac{|h_2|^2}{\nu^2} \right\}$$

where the random vector  $\underline{h} = (h_1, h_2)$  has the same distribution as the marginal distribution of the process  $\{\underline{h}_i\}$  at one time instant.

*Remark 4:* The secrecy capacity region given in Corollary 5 is established for fading processes  $\{\underline{h}_i\}$  where only ergodic and stationary conditions are assumed. The fading process  $\{\underline{h}_i\}$  can be correlated across time, and is not necessarily Gaussian.

*Remark 5:* The secrecy capacity region given in Corollary 5 also applies to the case in which the two component processes  $\{h_{1i}\}$  and  $\{h_{2i}\}$  are correlated. However, the secrecy capacity region does depend on the correlation between the two processes. In fact, the average  $E_h$  in (36) needs to be taken over the joint distributions of  $h_{1i}$  and  $h_{2i}$  to derive the correct secrecy capacity region.

This fact can be seen from the following example. We assume that only the confidential message is transmitted, i.e.,  $R_0 = 0$ . We also assume that both  $h_{1i}$  and  $h_{2i}$  take the values 0 and 1 with equal probabilities of 1/2. We first consider case 1 in which  $h_{1i} = h_{2i}$ . From (36) it is clear that  $R_1 = 0$ , i.e., no

secrecy capacity is possible, because the channels to the two receivers are the same for all channel realizations. However, for case 2 in which  $h_{1i} = 1 - h_{2i}$ , from (36) the secrecy capacity equals  $\frac{1}{2} \log(1 + \frac{2P}{\mu^2})$ , where the power is allocated only to the channel states where  $h_{1i} = 1$  and  $h_{2i} = 0$ . From this example, one can see that the correlation between  $h_{1i}$  and  $h_{2i}$  affects the secrecy capacity region, although the marginal distributions of  $h_{1i}$  and  $h_{2i}$  are the same for the two cases. We further note that this fact is consistent with Lemma 1 because  $h_{1i}$  and  $h_{2i}$  are now considered as channel parameters (with each realization corresponding to one subchannel) due to CSI availability at the transmitter. The channel statistics come only from the additive Gaussian noise terms for each subchannel with one realization of  $h_{1i}$  and  $h_{2i}$ .

*Remark 6:* The secrecy capacity region in Corollary 5 is established for the case with general correlation between the noise variables  $W_i$  and  $V_i$ .

From the bound on  $R_1$  in (36), it can be seen that as long as  $A$  is not a zero probability event, positive secrecy rate can be achieved. Since fading introduces more randomness to the channel, it is more likely that the channel from the source node to receiver 1 is better than the channel from the source node to receiver 2 for some channel states, and hence positive secrecy capacity can be achieved by exploiting these channel states.

Since the source node is assumed to know the CSI, it can allocate its power according to the instantaneous channel realization to achieve the best performance, i.e., the boundary of the secrecy capacity region. The optimal power allocation that achieves the boundary of the secrecy capacity region for the fading BCC can be derived from Theorem 3 and is given in the following.

*Corollary 6:* The optimal power allocation  $p^*(\underline{h})$  that achieves the boundary of the secrecy capacity region of the fading BCC falls into one of the following three cases.

Case 1:  $p^*(\underline{h}) = p^{(1)}(\underline{h})$  if the following  $p^{(1)}(\underline{h})$  satisfies

$$R_{01}(p^{(1)}(\underline{h})) < R_{02}(p^{(1)}(\underline{h})).$$

For  $\underline{h} \in A$ , if  $\frac{\gamma_1}{\gamma_0} > \frac{\nu^2|h_1|^2}{\nu^2|h_1|^2 - \mu^2|h_2|^2}$ , then we have the first set of equations at the bottom of the page. Alternatively, if  $\frac{\gamma_1}{\gamma_0} \leq \frac{\nu^2|h_1|^2}{\nu^2|h_1|^2 - \mu^2|h_2|^2}$ , then

$$p_0^{(1)}(\underline{h}) = \left( \frac{\gamma_0}{\lambda \ln 2} - \frac{\mu^2}{|h_1|^2} \right)^+ \quad \text{and} \quad p_1^{(1)}(\underline{h}) = 0.$$

For  $\underline{h} \in A^c$

$$p_0^{(1)}(\underline{h}) = \left( \frac{\gamma_0}{\lambda \ln 2} - \frac{\mu^2}{|h_1|^2} \right)^+ \quad (37)$$

where the parameter  $\lambda$  is chosen to satisfy the following power constraint:

$$E_A[p_0(\underline{h}) + p_1(\underline{h})] + E_{A^c}[p_0(\underline{h})] \leq P. \quad (38)$$

Case 2:  $p^*(\underline{h}) = p^{(2)}(\underline{h})$  if the following  $p^{(2)}(\underline{h})$  satisfies  $R_{01}(p^{(2)}(\underline{h})) > R_{02}(p^{(2)}(\underline{h}))$ .

For  $\underline{h} \in A$ , if  $\frac{\gamma_1}{\gamma_0} > \frac{\mu^2|h_2|^2}{\nu^2|h_1|^2 - \mu^2|h_2|^2}$ , then we get the second set of equations at the bottom of the page. Alternatively, if  $\frac{\gamma_1}{\gamma_0} \leq \frac{\mu^2|h_2|^2}{\nu^2|h_1|^2 - \mu^2|h_2|^2}$ , then

$$p_0^{(2)}(\underline{h}) = \left( \frac{\gamma_0}{\lambda \ln 2} - \frac{\nu^2}{|h_2|^2} \right)^+ \quad \text{and} \quad p_1^{(1)}(\underline{h}) = 0.$$

For  $\underline{h} \in A^c$

$$p_0^{(2)}(\underline{h}) = \left( \frac{\gamma_0}{\lambda \ln 2} - \frac{\nu^2}{|h_2|^2} \right)^+ \quad (39)$$

where  $\lambda$  is chosen to satisfy the power constraint defined in (38).

Case 3:  $p^*(\underline{h}) = p^{(\alpha)}(\underline{h})$  if there exists  $0 \leq \alpha \leq 1$  such that the following  $p^{(\alpha)}(\underline{h})$  satisfies  $R_{01}(p^{(\alpha)}(\underline{h})) = R_{02}(p^{(\alpha)}(\underline{h}))$ .

For  $\underline{h} \in A$ , if

$$\frac{\gamma_1}{\gamma_0} > \frac{\alpha\nu^2|h_1|^2 + \bar{\alpha}\mu^2|h_2|^2}{\nu^2|h_1|^2 - \mu^2|h_2|^2}$$

then we get the first set of equations at the bottom of the following page. Alternatively, if

$$\frac{\gamma_1}{\gamma_0} \leq \frac{\alpha\nu^2|h_1|^2 + \bar{\alpha}\mu^2|h_2|^2}{\nu^2|h_1|^2 - \mu^2|h_2|^2}$$

$$p_0^{(1)}(\underline{h}) = \left( \frac{\gamma_0}{\lambda \ln 2} - \left( \frac{\gamma_1}{\gamma_0} - 1 \right) \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right) \right)^+$$

and

$$p_1^{(1)}(\underline{h}) = \left( \min \left\{ \frac{1}{2} \sqrt{\left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right) \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} + \frac{4\gamma_1}{\lambda \ln 2} \right)} - \frac{1}{2} \left( \frac{\mu^2}{|h_1|^2} + \frac{\nu^2}{|h_2|^2} \right), \frac{\gamma_1}{\gamma_0} \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right) - \frac{\nu^2}{|h_2|^2} \right\} \right)^+.$$

$$p_0^{(2)}(\underline{h}) = \left( \frac{\gamma_0}{\lambda \ln 2} - \left( \frac{\gamma_1}{\gamma_0} + 1 \right) \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right) \right)^+$$

and

$$p_1^{(2)}(\underline{h}) = \left( \min \left\{ \frac{1}{2} \sqrt{\left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right) \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} + \frac{4\gamma_1}{\lambda \ln 2} \right)} - \frac{1}{2} \left( \frac{\mu^2}{|h_1|^2} + \frac{\nu^2}{|h_2|^2} \right), \frac{\gamma_1}{\gamma_0} \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right) - \frac{\mu^2}{|h_1|^2} \right\} \right)^+.$$

then we get the second set of equations at the bottom of the page. For  $l \in A^c$ , we get (40) at the bottom of the page, where  $\lambda$  is chosen to satisfy the power constraint defined in (38).

If the source node does not have common messages for both receivers, and only has confidential messages for receiver 1, the fading BCC becomes the fading wiretap channel. For this channel, the secrecy capacity is readily obtained from Corollaries 5 and 6.

*Corollary 7:* The secrecy capacity of the fading wiretap channel is

$$C_s = \max_{E_A[p(\underline{h})] \leq P} E_A \left[ \log \left( 1 + \frac{p(\underline{h})|h_1|^2}{\mu^2} \right) - \log \left( 1 + \frac{p(\underline{h})|h_2|^2}{\nu^2} \right) \right]. \quad (41)$$

The optimal power allocation that achieves the secrecy capacity in (41) is given by the last expression at the bottom of the page, where  $\lambda$  is chosen to satisfy the power constraint  $E_A[p(\underline{h})] = P$ .

### V. FADING BCCs: OUTAGE PERFORMANCE

In Section IV, we considered the ergodic secrecy capacity region for the fading BCC. In this case, messages can be coded

over long block lengths and hence over all channel realizations. This applies to wireless systems in which transmission delay can be tolerated. In this section, we consider wireless systems in which there is a stringent delay constraint, and messages must be transmitted within a certain time.

We adopt the channel model described in (34). However, we now make the block fading assumption, in which the fading coefficients  $h_{1i}$  and  $h_{2i}$  remain constant over one block and change to another realization in the next block in an ergodic and stationary manner. Moreover, we assume that the block length is large enough such that coding over one block can achieve small probability of error. We assume that the delay constraint is within the block length. Coding over multiple blocks and, hence, over multiple channel state realizations is not allowed. We also assume that both the transmitter and the receivers know the channel state information.

We use  $(\check{R}_0, \check{R}_1)$  to indicate a target rate pair, i.e., the common and confidential messages need to be transmitted to the two receivers at the rates  $\check{R}_0$  and  $\check{R}_1$ , respectively, in each block (each fading state realization). If the target rate pair is not achieved for one block, an outage is claimed. We define the outage probability to be

$$P_{\text{out}} = \Pr\{(\check{R}_0, \check{R}_1) \notin C_s(\underline{h}, p(\underline{h}))\} \quad (43)$$

$$p_0^{(\alpha)}(\underline{h}) = \left( \frac{1}{2} \sqrt{\left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} - \frac{\gamma_0}{2 \ln \lambda} \right)^2 + \frac{4\alpha\gamma_0}{\lambda \ln 2} \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right)} + \frac{\gamma_0}{2 \ln 2\lambda} - \left( \frac{\gamma_1}{\gamma_0} - \alpha + \frac{1}{2} \right) \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right) \right)^+$$

and

$$p_1^{(\alpha)}(\underline{h}) = \left( \min \left\{ \frac{1}{2} \sqrt{\left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right) \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} + \frac{4\gamma_1}{\lambda \ln 2} \right)} - \frac{1}{2} \left( \frac{\mu^2}{|h_1|^2} + \frac{\nu^2}{|h_2|^2} \right), \frac{\gamma_1}{\gamma_0} \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right) - \left( \alpha \frac{\nu^2}{|h_2|^2} + \bar{\alpha} \frac{\mu^2}{|h_1|^2} \right) \right\} \right)^+.$$

$$p_0^{(\alpha)}(\underline{h}) = \left( \frac{1}{2} \sqrt{\left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} - \frac{\gamma_0}{2 \ln \lambda} \right)^2 + \frac{4\alpha\gamma_0}{\lambda \ln 2} \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right)} - \frac{1}{2} \left( \frac{\mu^2}{|h_1|^2} + \frac{\nu^2}{|h_2|^2} - \frac{\gamma_0}{\lambda \ln 2} \right) \right)^+$$

and

$$p_1^{(\alpha)}(\underline{h}) = 0.$$

$$p_0^{(\alpha)}(\underline{h}) = \left( \frac{1}{2} \sqrt{\left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} - \frac{\gamma_0}{\lambda \ln 2} \right)^2 + \frac{4\alpha\gamma_0}{\lambda \ln 2} \left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right)} - \frac{1}{2} \left( \frac{\mu^2}{|h_1|^2} + \frac{\nu^2}{|h_2|^2} - \frac{\gamma_0}{\lambda \ln 2} \right) \right)^+ \quad (40)$$

$$p^*(\underline{h}) = \begin{cases} \left( \frac{1}{\lambda \ln 2} - \frac{\mu^2}{|h_1|^2} \right)^+, & \text{if } |h_2|^2 = 0 \\ \left( \frac{1}{2} \sqrt{\left( \frac{\nu^2}{|h_2|^2} - \frac{\mu^2}{|h_1|^2} \right) \left( \frac{4}{\lambda \ln 2} - \frac{\mu^2}{|h_1|^2} + \frac{\nu^2}{|h_2|^2} \right)} - \frac{1}{2} \left( \frac{\mu^2}{|h_1|^2} + \frac{\nu^2}{|h_2|^2} \right) \right)^+, & \text{if } |h_2|^2 > 0, \underline{h} \in A \\ 0, & \text{otherwise} \end{cases} \quad (42)$$

where  $\mathcal{C}_s(\underline{h}, p(\underline{h}))$  is the secrecy capacity region for the channel with fading state realization  $\underline{h}$ , and  $p(\underline{h})$  indicates the transmission power used by the source node for this fading state. The source node is able to adapt its transmission power to the instantaneous channel state realization, i.e.,  $p(\underline{h})$  is a function of  $\underline{h}$ , because CSI is assumed to be known at the transmitter. We assume that the power constraint applies over a large number of blocks and hence over all fading state realizations (i.e., it is a long-term power constraint as in [22]); that is, we assume

$$\mathbb{E}[p(\underline{h})] \leq P. \quad (44)$$

We define the set  $\mathcal{P} := \{p(\underline{h}) : \mathbb{E}[p(\underline{h})] \leq P\}$ .

It is clear from (43) that the outage probability depends on the power allocation function  $p(\underline{h})$ . Our goal is to study the power allocation  $p^*(\underline{h})$  that minimizes the outage probability, i.e.,

$$p^*(\underline{h}) = \arg \min_{p(\underline{h}) \in \mathcal{P}} P_{\text{out}}. \quad (45)$$

To understand this problem, we note that for each channel state, the source node knows how much power it needs to use to support the target rate pair. If the power needed is too large, the source node may decide not to transmit and claim outage in order to save power for other channel states that need lower power to support the target rate pair. Hence, the source node first allocates power to those channel states that need lower power to support the target rate, and then to those channel states that need higher power to support the target rate until all power is utilized. This suggests that the power allocation is a threshold decision.

To formally solve the optimization problem (45), we apply the approach in [22], where the minimum outage probability of the fading channel without a secrecy constraint was studied. This approach was also applied in [19] to study the minimum outage probability of the fading broadcast channel without a secrecy constraint. In the following, we describe the power allocation that minimizes the outage probability of the fading BCC, i.e., the solution to (45).

For a given block with the fading state realization  $\underline{h} = (h_1, h_2)$ , the channel we consider is a Gaussian BCC. From Corollary 4, the secrecy capacity region is given as follows.

If  $\underline{h} \in A$ , then

$$\mathcal{C}_s(\underline{h}, p(\underline{h})) = \bigcup_{0 \leq \beta \leq 1} \left\{ \begin{array}{l} (R_0, R_1) : \\ R_0 \leq \log \left( 1 + \frac{(1-\beta)p(\underline{h})|h_2|^2}{\nu^2 + \beta p(\underline{h})|h_2|^2} \right), \\ R_1 \leq \log \left( 1 + \frac{\beta p(\underline{h})|h_1|^2}{\mu^2} \right) - \log \left( 1 + \frac{\beta p(\underline{h})|h_2|^2}{\nu^2} \right) \end{array} \right\}$$

Alternatively, if  $\underline{h} \in A^c$ , then

$$\mathcal{C}_s(\underline{h}, p(\underline{h})) = \left\{ \begin{array}{l} (R_0, R_1) : \\ R_0 \leq \log \left( 1 + \frac{p(\underline{h})|h_1|^2}{\mu^2} \right), \\ R_1 = 0 \end{array} \right\}. \quad (46)$$

We now use (46) to compute the minimum power that is needed to achieve the target rate pair  $(\check{R}_0, \check{R}_1)$ . It is clear from

(46) that, only when  $\underline{h} \in A$  can we possibly achieve a positive  $\check{R}_1$ . The minimum power needed to achieve  $\check{R}_1$  is

$$\beta p(\underline{h}) = \begin{cases} \frac{2^{\check{R}_1} - 1}{\frac{|h_1|^2}{\mu^2} - 2^{\check{R}_1} \frac{|h_2|^2}{\nu^2}}, & \text{if } \check{R}_1 < \log \frac{|h_1|^2 \nu^2}{|h_2|^2 \mu^2} \\ \infty, & \text{otherwise.} \end{cases} \quad (47)$$

The minimum power to support  $\check{R}_0$  is then given by

$$(1 - \beta)p(\underline{h}) = \begin{cases} \frac{(2^{\check{R}_0} - 1) \left( \frac{|h_1|^2 \nu^2}{|h_2|^2 \mu^2} - 1 \right)}{\frac{|h_1|^2}{\mu^2} - 2^{\check{R}_1} \frac{|h_2|^2}{\nu^2}}, & \text{if } \check{R}_1 < \log \frac{|h_1|^2 \nu^2}{|h_2|^2 \mu^2} \\ \infty, & \text{otherwise.} \end{cases} \quad (48)$$

Hence, the minimum power needed to support  $(\check{R}_0, \check{R}_1)$  is

$$p^{\min}(\underline{h}) = \begin{cases} \frac{2^{\check{R}_0} \left( \frac{|h_1|^2 \nu^2}{|h_2|^2 \mu^2} - 1 \right) + 2^{\check{R}_1} - \frac{|h_1|^2 \nu^2}{|h_2|^2 \mu^2}}{\frac{|h_1|^2}{\mu^2} - 2^{\check{R}_1} \frac{|h_2|^2}{\nu^2}}, & \text{if } \check{R}_1 < \log \frac{|h_1|^2 \nu^2}{|h_2|^2 \mu^2} \\ \infty, & \text{otherwise.} \end{cases} \quad (49)$$

For  $s > 0$ , we define

$$\begin{aligned} \mathcal{R}(s) &= \{\underline{h} : p^{\min}(\underline{h}) < s\} \\ \bar{\mathcal{R}}(s) &= \{\underline{h} : p^{\min}(\underline{h}) \leq s\} \end{aligned} \quad (50)$$

where  $p^{\min}(\underline{h})$  is given in (49).

The average powers that are needed to support the rate pair  $(\check{R}_0, \check{R}_1)$  for the channel states in  $\mathcal{R}(s)$  and  $\bar{\mathcal{R}}(s)$  are

$$p(s) = \mathbb{E}_{\underline{h} \in \mathcal{R}(s)} [p^{\min}(\underline{h})]$$

and

$$\bar{p}(s) = \mathbb{E}_{\underline{h} \in \bar{\mathcal{R}}(s)} [p^{\min}(\underline{h})]. \quad (51)$$

For the given power constraint  $P$ , define

$$s^* = \sup\{s : p(s) < P\}$$

and

$$w^* = \frac{P - p(s^*)}{\bar{p}(s^*) - p(s^*)}. \quad (52)$$

We then obtain the following optimal power allocation  $p^*(\underline{h})$ .

*Proposition 1:* The power allocation  $p^*(\underline{h})$  that solves (45), and hence minimizes the outage probability for a given target rate pair  $(\check{R}_0, \check{R}_1)$ , is given by

$$p^*(\underline{h}) = \begin{cases} p^{\min}(\underline{h}), & \text{if } \underline{h} \in \mathcal{R}(s^*) \\ p^{\min}(\underline{h}), & \text{with prob. } w^* \text{ if } \underline{h} \in \bar{\mathcal{R}}(s^*) \setminus \mathcal{R}(s^*) \\ 0, & \text{if } \underline{h} \notin \bar{\mathcal{R}}(s^*) \end{cases} \quad (53)$$

where  $p^{\min}(\underline{h})$  is given in (49).

*Proof:* Proposition 1 follows by using (49) and applying Lemma 3 and Proposition 4 in [22].  $\square$

It can be seen that the optimal power allocation  $p^*(\underline{h})$  is a threshold solution. The power is first allocated to the fading

states that need smaller amounts of power to achieve the target rate pair, and is then allocated to the fading states that need larger amounts of power to achieve the target rate pair. When the total power  $P$  is used up by these fading states, no further power is allocated to other states.

In the above problem setting, an outage is claimed if either  $\check{R}_0$  or  $\check{R}_1$  is not achieved. This results in an optimal solution  $p^*(\underline{h})$  that allocates power only to those states for which both  $\check{R}_0$  and  $\check{R}_1$  can be achieved by relatively small power consumption. However, some channel states may support one target rate with small power consumption, but need a large amount of power to support both rates. These states are unlikely to be allocated power. For example, even if the channels from the source node to both receivers are good to transmit common messages, it may happen that no power is allocated to this fading state if the channel from the source node to receiver 1 is worse than the channel from the source node to receiver 2 so that  $\check{R}_1$  cannot be achieved. Sometimes this is not reasonable, because the two messages are independent, and one message should be transmitted whenever the channel is good to transmit it. It should not depend on whether the other message is transmitted or not. Nevertheless, the solution to the problem (45) we have considered is useful if we consider the following two more reasonable problems.

It is clear that the solution to the problem (45) immediately implies the optimal power allocation for the case where only the confidential message is transmitted and only the target rate  $\check{R}_1$  is assumed. Now the minimum power to achieve  $\check{R}_1$  is given by

$$P^{\min}(\underline{h}) = \begin{cases} \frac{2^{\check{R}_1} - 1}{\frac{|h_1|^2}{\mu^2} - 2^{\check{R}_1} \frac{|h_2|^2}{\nu^2}}, & \text{if } \check{R}_1 < \log \frac{|h_1|^2 \nu^2}{|h_2|^2 \mu^2} \\ \infty, & \text{otherwise.} \end{cases} \quad (54)$$

The power allocation that minimizes the outage probability follows from Proposition 1 by using (54) to replace (49) in (50)–(53).

We next consider a scenario in which the source node has both common and confidential messages to transmit. We assume that the common message is required to be transmitted at a constant rate  $\check{R}_0$  for all channel states, i.e., no outage is allowed for the common rate. This scenario applies to wireless systems in which a constant common rate must be satisfied. Since  $\check{R}_0$  must be achieved for all channel states, the total power must be large enough to support this rate, i.e.,

$$P \geq P_0 := \mathbb{E}[p_0(\underline{h})] \quad (55)$$

where  $p_0(\underline{h})$  is the power that is needed to support the rate  $\check{R}_0$  for the channel state  $\underline{h}$  and is given by

$$p_0(\underline{h}) = \frac{2^{\check{R}_0} - 1}{\min \left\{ \frac{|h_1|^2}{\mu^2}, \frac{|h_2|^2}{\nu^2} \right\}}. \quad (56)$$

In addition to the common message, the source node wishes to transmit confidential information to receiver 1 at a target rate  $\check{R}_1$  and with as small an outage probability as possible. We note that a similar problem was studied in [19] for the broadcast channel with separate messages for two receivers and the rate to one receiver must be constant for all channel states.

We need to find the power allocation that minimizes the outage probability that the target rate  $\check{R}_1$  is not achieved. The optimization problem is summarized as follows:

$$\begin{aligned} & \text{Minimize} && \Pr\{(\check{R}_0, \check{R}_1) \notin \mathcal{C}(\underline{h}, p(\underline{h}))\} \\ & \text{Subject to} && \mathbb{E}[p(\underline{h})] \leq P \\ & && \check{R}_0 \text{ is achieved for all } \underline{h}, \text{ i.e., } p(\underline{h}) \geq p_0(\underline{h}). \end{aligned} \quad (57)$$

We can change the problem (57) to the following equivalent problem:

$$\begin{aligned} & \text{Minimize} && \Pr\{(\check{R}_0, \check{R}_1) \notin \mathcal{C}(\underline{h}, p(\underline{h}))\} \\ & \text{Subject to} && \mathbb{E}[\Delta p(\underline{h})] \leq P - P_0 \end{aligned} \quad (58)$$

where  $\Delta p(\underline{h}) = p(\underline{h}) - p_0(\underline{h})$  is the difference between the power needed to support  $(\check{R}_0, \check{R}_1)$  and the power needed to support  $\check{R}_0$  only.

It can be seen that the problem (58) is the same as the problem (45) with  $p(\underline{h})$  in (45) being replaced by  $\Delta p(\underline{h})$ . Thus, the optimal  $\Delta p^*(\underline{h})$  can be derived from Proposition 1 with  $p^{\min}(\underline{h})$  in (50)–(53) being replaced by  $\Delta p^{\min}(\underline{h})$ , which is the minimum difference between the power needed to support  $(\check{R}_0, \check{R}_1)$  and the power needed to support  $\check{R}_0$  and is given by

$$\Delta p^{\min}(\underline{h}) = \begin{cases} \frac{2^{\check{R}_0} (2^{\check{R}_1} - 1)}{\frac{|h_1|^2}{\mu^2} - 2^{\check{R}_1} \frac{|h_2|^2}{\nu^2}}, & \text{if } \check{R}_1 < \log \frac{|h_1|^2 \nu^2}{|h_2|^2 \mu^2} \\ \infty, & \text{otherwise.} \end{cases} \quad (59)$$

## VI. NUMERICAL RESULTS

In this section, we provide numerical results to demonstrate the ergodic and outage performance for the fading BCC.

We first consider the Rayleigh-fading BCC, where  $h_1$  and  $h_2$  are zero-mean proper complex Gaussian random variables. Hence  $|h_1|^2$  and  $|h_2|^2$  are exponentially distributed with parameters  $\sigma_1$  and  $\sigma_2$ . We assume the source power  $P = 5$  dB, and fix  $\sigma_1 = 1$ . In Fig. 5, we plot the boundaries of the secrecy capacity regions corresponding to  $\sigma_2 = 0.4, 0.7$  and 1, respectively. It can be seen that as  $\sigma_2$  decreases, the secrecy rate  $\check{R}_1$  of the confidential message improves, but the rate  $\check{R}_0$  of the common message decreases. This fact follows because smaller  $\sigma_2$  implies a worse channel from the source node to receiver 2. Thus, confidential information can be forwarded to receiver 1 at a higher rate. However, the rate of the common information is limited by the worse channel from the source node to receiver 2.

For the Rayleigh-fading BCC with  $\sigma_1 = 1$  and  $\sigma_2 = 0.4$ , we plot the boundary of the secrecy capacity region in Fig. 6. The three cases (see Corollary 6) to derive the boundary achieving power allocations are also indicated with the corresponding boundary points. It can be seen that the boundary points with large  $\check{R}_1$  are achieved by the power allocations derived from Case 1, and are indicated by the line with circles on the graph. The boundary points with large  $\check{R}_0$  are achieved by the optimal power allocations derived from Case 2, and are indicated by the line with squares. Between the boundary points achieved by Case 1 and Case 2, the boundary points are achieved by the

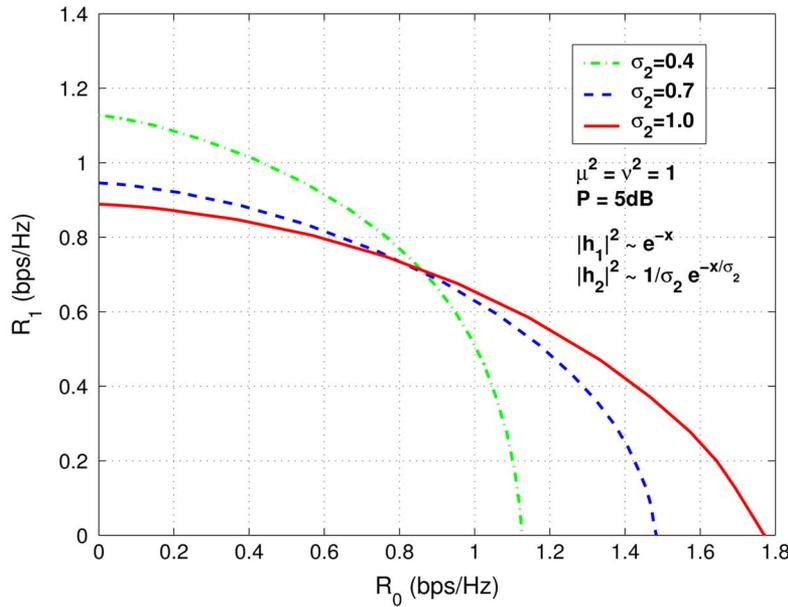


Fig. 5. Secrecy capacity regions for a Rayleigh-fading BCC.

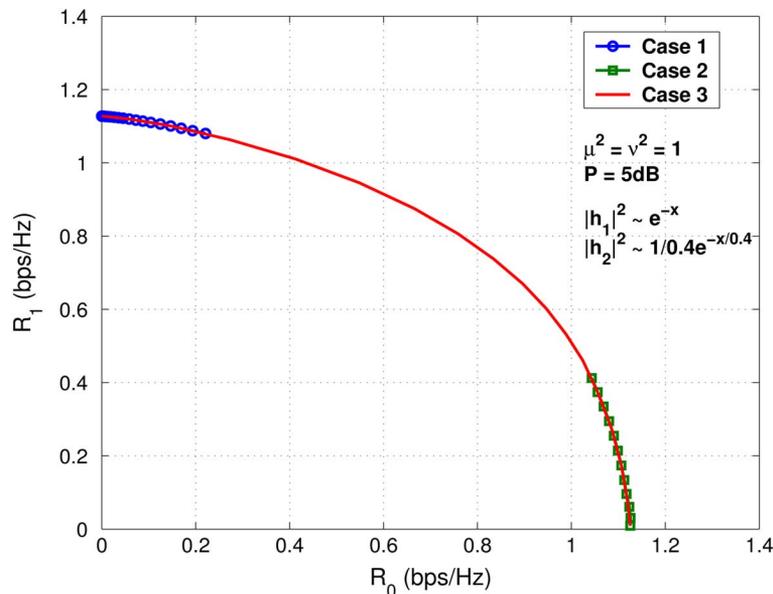


Fig. 6. Three cases for power allocation optimization to achieve the boundary of the secrecy capacity region for a Rayleigh-fading BCC.

power allocations derived from Case 3, and are indicated by the plain solid line.

The intuition as to how the three cases associate with the boundary points is given as follows. To achieve large secrecy rate  $R_1$ , most channel states in the set  $A$  where receiver 1 has a stronger channel than receiver 2 are used to transmit the confidential message. The common message is hence transmitted mostly over the channel states in the set  $A^c$ , over which the common rate is limited by the channel from the source node to receiver 1. Thus, power allocation needs to optimize the rate of this channel, and hence the optimal power allocation follows from Case 1. To achieve large  $R_0$ , the common message is forwarded over the channel states both in  $A$  and  $A^c$ . It can be seen that on average the source node has a much worse channel to

receiver 2 than to receiver 1, and hence, the channel from the source node to receiver 2 limits the common rate. Power allocation now needs to optimize the rate to receiver 2, and the optimal power allocation follows from Case 2. Between these two cases, power allocation needs to balance the rates to receivers 1 and 2 and hence follows from Case 3.

We next consider the case in which  $R_0 = 0$ , i.e., only the confidential message is transmitted from the source node to receiver 1. We assume  $\sigma_1 = \sigma_2 = 1$ . In Fig. 7(a), we plot the optimal power allocation  $p_1^*(\underline{h})$  as a function of  $\underline{h}$ . It can be seen from the graph that most of the source power is allocated to the channel states with small  $|h_2|^2$ . This behavior is shown more clearly in Fig. 7(b), which plots  $p_1^*(\underline{h})$  as a function of  $|h_1|^2$  for different values of  $|h_2|^2$ , and in Fig. 7(c), which plots  $p_1^*(\underline{h})$  as a

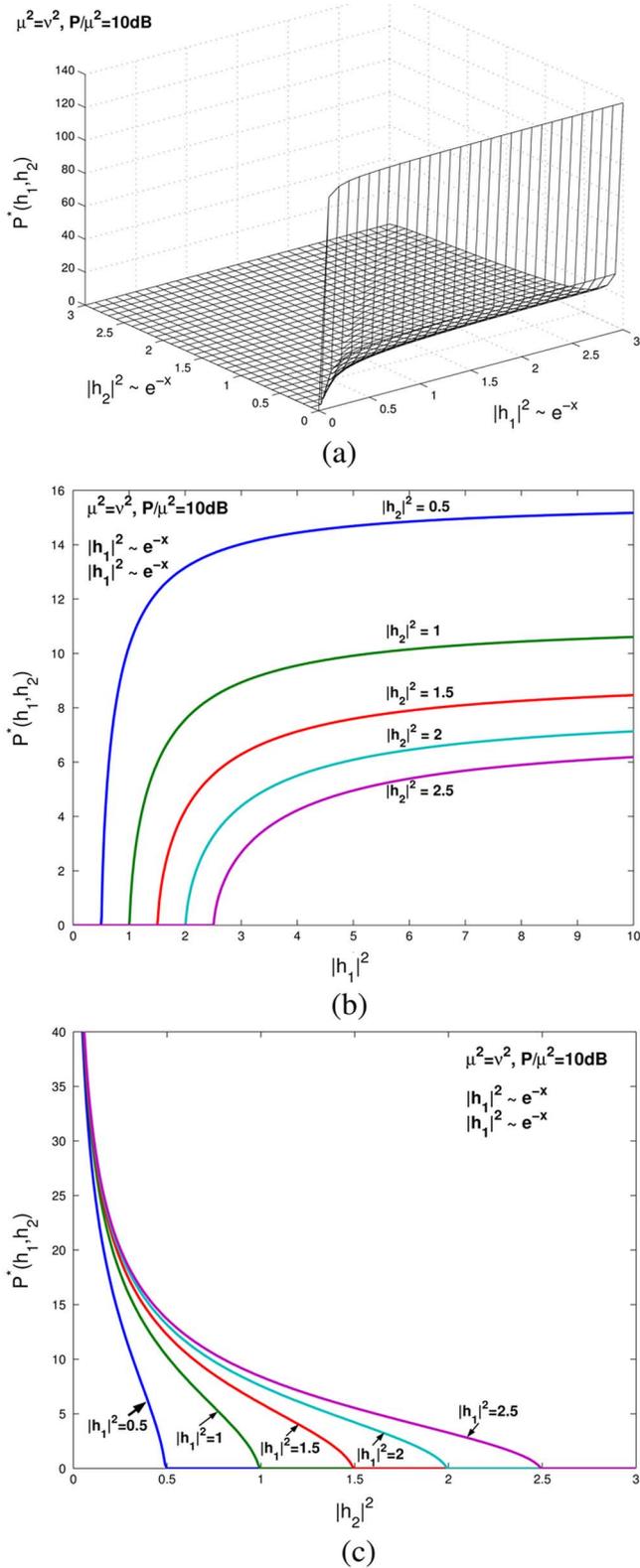


Fig. 7. Optimal power allocation  $p_1^*(\underline{h})$  for a Rayleigh-fading BCC with  $R_0 = 0$ . (a)  $p_1^*(\underline{h})$  as a function of  $(h_1, h_2)$ . (b)  $p_1^*(\underline{h})$  as a function of  $|h_1|^2$ . (c)  $p_1^*(\underline{h})$  as a function of  $|h_2|^2$ .

function of  $|h_2|^2$  for different values of  $|h_1|^2$ . The source node allocates more power to the channel states with larger  $|h_1|^2$  to forward more confidential information to the destination node, and allocates less power for the channel states with larger  $|h_2|^2$  to prevent receiver 2 for obtaining the confidential information.

It can also be seen from Fig. 7(b) and Fig. 7(c) that the source node transmits only when the channel from the source node to receiver 1 is better than the channel from the source node to receiver 2.

For the case in which  $R_0 = 0$ , Fig. 8 plots the secrecy capacity achieved by the optimal power allocation, and compares it with the secrecy rate achieved by a uniform power allocation, i.e., allocating the same power for all channel states  $\underline{h} \in A$ . It can be seen that the uniform power allocation does not provide performance close to the secrecy capacity for the signal-to-noise ratio (SNR) values of interest. This is in contrast to the Rayleigh-fading channel without the secrecy constraint, where the uniform power allocation can be close to optimal even for moderate SNRs. This also demonstrates that the exact channel state information is important to achieve higher secrecy rate.

We now consider the outage performance of the Rayleigh-fading BCC. We assume  $\sigma_1 = 10$  and  $\sigma_2 = 0.5$ . We first assume the target rate  $\hat{R}_0 = 0$ , i.e., only the confidential message is transmitted. In Fig. 9, we plot the outage probabilities corresponding to different values of the target rate  $\hat{R}_1$ . The outage probability decreases as the target rate  $\hat{R}_1$  decreases. For a fixed  $\hat{R}_1$ , the outage probability is bounded below by a certain threshold. This fact follows because the outage probability cannot be prevented for those channel states where the channel from the source node to receiver 1 is worse than the channel from the source node to receiver 2. In Fig. 10, we compare the outage probability when  $R_0 = 0$  with the outage probability when  $R_0 = 0.2$  b/s/Hz. It can be seen that even a small positive common rate  $R_0$  can cause a large increase in outage probability. In Fig. 11, we compare the outage probability minimized by the power allocation and that achieved by the equal power allocation for all channel states. It can be seen that optimizing the power allocation significantly reduces the outage probability.

## VII. CONCLUSION

We have established the secrecy capacity region of the parallel BCC, where a converse proof has shown that having independent inputs to each subchannel is optimal. We have also established the secrecy capacity region for the parallel Gaussian BCC, and have characterized the optimal power allocations that achieve the boundary of this region. One fundamental result we have established is the secrecy capacity region of the Gaussian BCC, which complements the secrecy capacity region of the discrete memoryless BCC given by Csiszár and Körner in [12].

We have further applied our results to obtain the ergodic secrecy capacity region for the fading BCC and the optimal power allocations that achieve the boundary of the secrecy capacity region. Our results generalize the secrecy capacity results that have been recently obtained in [5], [6], and [7] (full CSI case). We have also studied the outage performance of the fading BCC, and have obtained the power allocation that minimizes the outage probability that certain target rates are not achieved.

## APPENDIX I PROOF OF THEOREM 1

The achievability follows from [12, Corollary 1] by setting  $Q = (Q_1, \dots, Q_L)$  ( $Q$  is indicated by  $U$  in

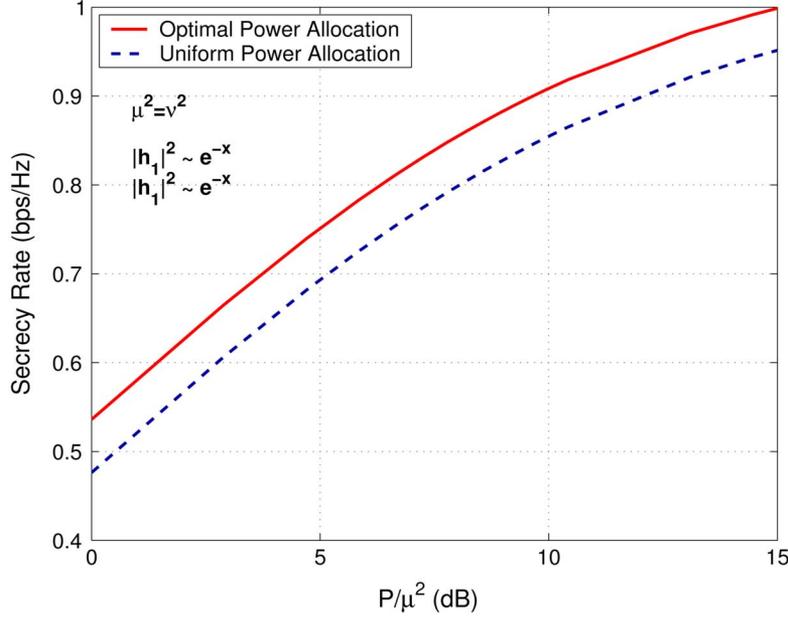


Fig. 8. Comparison of the secrecy capacity achieved by optimal power allocation with the secrecy rate achieved by uniform power allocation for a Rayleigh-fading BCC with  $R_0 = 0$ .

[12]),  $U = (U_1, \dots, U_L)$  ( $U$  is indicated by  $V$  in [12]),  $X = (X_1, \dots, X_L)$ ,  $Y = (Y_1, \dots, Y_L)$ , and  $Z = (Z_1, \dots, Z_L)$  with  $Q, U$ , and  $X$  having independent components. Furthermore, we choose the components of these random vectors to satisfy the Markov chain conditions:  $Q_l \rightarrow U_l \rightarrow X_l \rightarrow (Y_l, Z_l)$  for  $l = 1, 2, \dots, L$ .

To show the converse, we consider a code  $(2^{nR_0}, 2^{nR_1}, n)$  with average error probability  $P_e$ . The probability distribution on  $W_0 \times W_1 \times \mathcal{X}_{[1,L]}^n \times \mathcal{Y}_{[1,L]}^n \times \mathcal{Z}_{[1,L]}^n$  is given by

$$p(w_0, w_1, x_{[1,L]}^n, y_{[1,L]}^n, z_{[1,L]}^n) = p(w_0)p(w_1)p(x_{[1,L]}^n | w_0, w_1) \prod_{i=1}^n \prod_{l=1}^L p_l(y_{li}, z_{li} | x_{li}) \quad (60)$$

By Fano's inequality [59, Sec. 2.11], we have

$$\begin{aligned} H(W_0, W_1 | Y_{[1,L]}^n) &\leq n(R_0 + R_1)P_e + 1 := n\delta_1 \\ H(W_0 | Z_{[1,L]}^n) &\leq nR_0P_e + 1 := n\delta_2 \end{aligned} \quad (61)$$

where  $\delta_1, \delta_2 \rightarrow 0$  if  $P_e \rightarrow 0$ .

For  $l = 1, 2, \dots, L$ , we define the following auxiliary random variables:

$$\begin{aligned} Q_{li} &:= (W_0, Y_{[1,l-1]}^n, Y_l^{i-1}, Z_{[l,i+1]}^n, Z_{[l+1,L]}^n) \\ U_{li} &:= (W_1, Q_{li}). \end{aligned} \quad (62)$$

We note that  $(Q_{li}, U_{li}, X_{li}, Y_{li}, Z_{li})$  satisfies the following Markov chain condition:

$$Q_{li} \rightarrow U_{li} \rightarrow X_{li} \rightarrow (Y_{li}, Z_{li}). \quad (63)$$

We first bound the common rate  $R_0$  as follows:

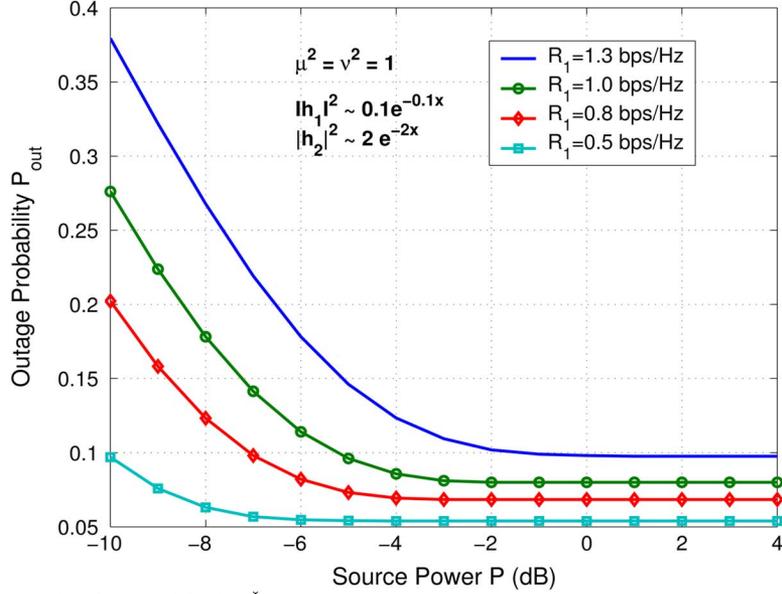
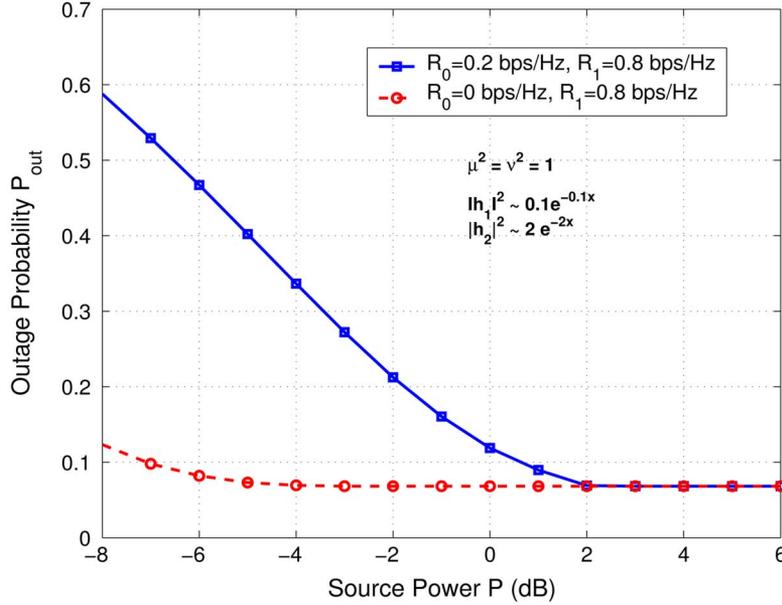
$$\begin{aligned} nR_0 &= H(W_0) \\ &= I(W_0; Y_{[1,L]}^n) + H(W_0 | Y_{[1,L]}^n) \end{aligned}$$

$$\begin{aligned} &\stackrel{(a)}{\leq} \sum_{l=1}^L I(W_0; Y_l^n | Y_{[1,l-1]}^n) + n\delta_1 \\ &\stackrel{(b)}{=} \sum_{l=1}^L \sum_{i=1}^n I(W_0; Y_{li} | Y_{[1,l-1]}^n Y_l^{i-1}) + n\delta_1 \\ &\stackrel{(c)}{\leq} \sum_{l=1}^L \sum_{i=1}^n I(W_0 Y_{[1,l-1]}^n Y_l^{i-1} Z_{[l,i+1]}^n Z_{[l+1,L]}^n; Y_{li}) \\ &\quad + n\delta_1 \\ &\stackrel{(d)}{=} \sum_{l=1}^L \sum_{i=1}^n I(Q_{li}; Y_{li}) + n\delta_1 \end{aligned} \quad (64)$$

where (a) follows from the chain rule and Fano's inequality, (b) follows from the chain rule, (c) follows because  $I(A; B | C) \leq I(A, C; B)$ , and (d) follows from the definition (62).

We can also bound the common rate  $R_0$  as follows:

$$\begin{aligned} nR_0 &= H(W_0) \\ &= I(W_0; Z_{[1,L]}^n) + H(W_0 | Z_{[1,L]}^n) \\ &\leq \sum_{l=1}^L I(W_0; Z_l^n | Z_{[l+1,L]}^n) + n\delta_2 \\ &= \sum_{l=1}^L \sum_{i=1}^n I(W_0; Z_{li} | Z_{[l,i+1]}^n Z_{[l+1,L]}^n) + n\delta_2 \\ &\leq \sum_{l=1}^L \sum_{i=1}^n I(W_0 Y_{[1,l-1]}^n Y_l^{i-1} Z_{[l,i+1]}^n Z_{[l+1,L]}^n; Z_{li}) \\ &\quad + n\delta_2 \\ &= \sum_{l=1}^L \sum_{i=1}^n I(Q_{li}; Z_{li}) + n\delta_2. \end{aligned} \quad (65)$$


 Fig. 9. Outage probabilities for a Rayleigh-fading BCC with  $\tilde{R}_0 = 0$ .

 Fig. 10. Comparison of outage probabilities when  $R_0 = 0$  with those when  $R_0 = 0.2$  b/s/Hz for a Rayleigh-fading BCC.

We now bound the rate  $R_1$  and obtain

$$\begin{aligned}
 nR_1 &\stackrel{(a)}{=} nR_e \leq H(W_1 | Z_{[1,L]}^n) \\
 &= H(W_1 | W_0 Z_{[1,L]}^n) + I(W_1; W_0 | Z_{[1,L]}^n) \\
 &\stackrel{(b)}{\leq} I(W_1; Y_{[1,L]}^n | W_0) - I(W_1; Z_{[1,L]}^n | W_0) \\
 &\quad + H(W_1 | W_0 Y_{[1,L]}^n) + H(W_0 | Z_{[1,L]}^n) \\
 &\stackrel{(c)}{\leq} I(W_1; Y_{[1,L]}^n | W_0) - I(W_1; Z_{[1,L]}^n | W_0) \\
 &\quad + n(\delta_1 + \delta_2) \\
 &= \sum_{l=1}^L \left[ I(W_1; Y_l^n | W_0 Y_{[1,l-1]}^n) \right. \\
 &\quad \left. - I(W_1; Z_l^n | W_0 Z_{[l+1,L]}^n) \right] + n(\delta_1 + \delta_2) \\
 &= \sum_{l=1}^L \sum_{i=1}^n \left[ I(W_1; Y_{li} | W_0 Y_{[1,l-1]}^n Y_l^{i-1}) \right. \\
 &\quad \left. - I(W_1; Z_{li} | W_0 Z_{[i+1,L]}^n Z_{[l+1,L]}^n) \right] + n(\delta_1 + \delta_2)
 \end{aligned}$$

$$\begin{aligned}
 &- I(W_1; Z_{li} | W_0 Z_{[i+1,L]}^n Z_{[l+1,L]}^n) \Big] + n(\delta_1 + \delta_2) \\
 &\stackrel{(d)}{=} \sum_{l=1}^L \sum_{i=1}^n \left[ I(W_1 Z_{[i+1,L]}^n Z_{[l+1,L]}^n; Y_{li} | W_0 Y_{[1,l-1]}^n Y_l^{i-1}) \right. \\
 &\quad - I(Z_{[i+1,L]}^n Z_{[l+1,L]}^n; Y_{li} | W_0 W_1 Y_{[1,l-1]}^n Y_l^{i-1}) \\
 &\quad - I(W_1 Y_{[1,l-1]}^n Y_l^{i-1}; Z_{li} | W_0 Z_{[i+1,L]}^n Z_{[l+1,L]}^n) \\
 &\quad \left. + I(Y_{[1,l-1]}^n Y_l^{i-1}; Z_{li} | W_0 W_1 Z_{[i+1,L]}^n Z_{[l+1,L]}^n) \right] \\
 &\quad + n(\delta_1 + \delta_2) \\
 &\stackrel{(e)}{=} \sum_{l=1}^L \sum_{i=1}^n \left[ I(W_1 Z_{[i+1,L]}^n Z_{[l+1,L]}^n; Y_{li} | W_0 Y_{[1,l-1]}^n Y_l^{i-1}) \right. \\
 &\quad \left. - I(W_1 Y_{[1,l-1]}^n Y_l^{i-1}; Z_{li} | W_0 Z_{[i+1,L]}^n Z_{[l+1,L]}^n) \right] \\
 &\quad + n(\delta_1 + \delta_2)
 \end{aligned}$$

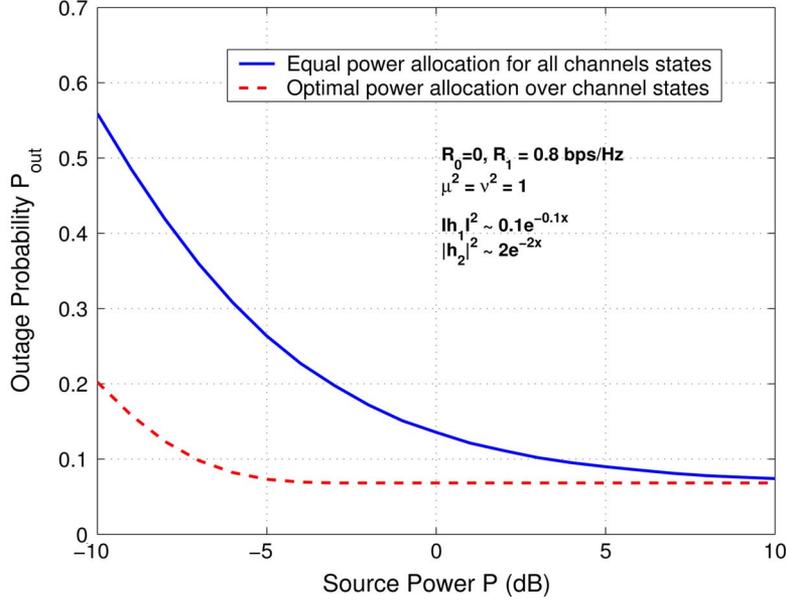


Fig. 11. Comparison of outage probabilities achieved by optimal power allocation with those achieved by equal power allocation for a Rayleigh-fading BCC with  $R_0 = 0$ .

$$\begin{aligned}
&= \sum_{l=1}^L \sum_{i=1}^n \left[ I \left( Z_{l[i+1]}^n Z_{[l+1,L]}^n; Y_{li} \mid W_0 Y_{[1,l-1]}^n Y_l^{i-1} \right) \right. \\
&\quad + I \left( W_1; Y_{li} \mid W_0 Y_{[1,l-1]}^n Y_l^{i-1} Z_{l[i+1]}^n Z_{[l+1,L]}^n \right) \\
&\quad - I \left( Y_{[1,l-1]}^n Y_l^{i-1}; Z_{li} \mid W_0 Z_{l[i+1]}^n Z_{[l+1,L]}^n \right) \\
&\quad \left. - I \left( W_1; Z_{li} \mid W_0 Y_{[1,l-1]}^n Y_l^{i-1} Z_{l[i+1]}^n Z_{[l+1,L]}^n \right) \right] \\
&\quad + n(\delta_1 + \delta_2) \\
&\stackrel{(f)}{=} \sum_{l=1}^L \sum_{i=1}^n \left[ I \left( W_1; Y_{li} \mid W_0 Y_{[1,l-1]}^n Y_l^{i-1} Z_{l[i+1]}^n Z_{[l+1,L]}^n \right) \right. \\
&\quad \left. - I \left( W_1; Z_{li} \mid W_0 Y_{[1,l-1]}^n Y_l^{i-1} Z_{l[i+1]}^n Z_{[l+1,L]}^n \right) \right] \\
&\quad + n(\delta_1 + \delta_2) \\
&\stackrel{(g)}{=} \sum_{l=1}^L \sum_{i=1}^n \left[ I(U_{li}; Y_{li} \mid Q_{li}) - I(U_{li}; Z_{li} \mid Q_{li}) \right] \\
&\quad + n(\delta_1 + \delta_2) \tag{66}
\end{aligned}$$

where (a) follows from the perfect secrecy condition, (b) follows because  $I(W_1; W_0 \mid Z_{[1,L]}^n) \leq H(W_0 \mid Z_{[1,L]}^n)$ , (c) follows from Fano's inequality, (d) follows from the chain rule, (e) and (f) follow from Lemma 7 in [12], and (g) follows from the definition (62).

We introduce a random variable  $G$  that is independent of all other random variables, and is uniformly distributed over  $\{1, 2, \dots, n\}$ . Define  $Q_l = (G, Q_{lG})$ ,  $U_l = (G, U_{lG})$ ,  $X_l = X_{lG}$ ,  $Y_l = Y_{lG}$ , and  $Z_l = Z_{lG}$  for  $l = 1, \dots, L$ . Note that  $(Q_l, U_l, X_l, Y_l, Z_l)$  satisfies the following Markov chain condition:

$$Q_l \rightarrow U_l \rightarrow X_l \rightarrow (Y_l, Z_l), \quad \text{for } l = 1, \dots, L. \tag{67}$$

Using the above definitions, (64), (65), and (66) become

$$R_0 \leq \sum_{l=1}^L I(Q_{lG}; Y_{lG} \mid G) + \delta_1$$

$$\begin{aligned}
&\leq \sum_{l=1}^L I(Q_l; Y_l) + \delta_1 \\
R_0 &\leq \sum_{l=1}^L I(Q_{lG}; Z_{lG} \mid G) + \delta_2 \\
&\leq \sum_{l=1}^L I(Q_l; Z_l) + \delta_2 \\
R_1 &\leq \sum_{l=1}^L [I(U_l; Y_l \mid Q_l) - I(U_l; Z_l \mid Q_l)] + \delta_1 + \delta_2. \tag{68}
\end{aligned}$$

Therefore, an outer bound on the secrecy capacity region  $\mathcal{C}_s$  is given by the following set:

$$\bigcup \{ (R_0, R_1) \text{ that satisfy (68)} \} \tag{69}$$

where the union is over all probability distributions  $p(q_{[1,L]}, u_{[1,L]}, x_{[1,L]}, y_{[1,L]}, z_{[1,L]})$ . Finally, we note that each term in (68) depends only on the distribution  $p(q_l, u_l, x_l, y_l, z_l)$ . Hence, there is no loss of optimality to consider only those distributions that have the form  $\prod_{l=1}^L p(q_l, u_l, x_l) p(y_l, z_l \mid x_l)$ . This concludes the converse proof.

## APPENDIX II PROOF OF THEOREM 2

By Lemma 1, we need to prove Theorem 2 only for the channel defined by (23).

The achievability follows by applying Corollary 3 and choosing the following input distribution:

$$\begin{aligned}
&\text{for } l \in A, \quad Q_l \sim \mathcal{N}(0, p_{l0}), \quad X_l' \sim \mathcal{N}(0, p_{l1}) \\
&\quad \text{with } X_l' \text{ independent of } Q_l, \\
&\quad X_l = Q_l + X_l'; \\
&\text{for } l \in A^c; \quad X_l \sim \mathcal{N}(0, p_{l0}).
\end{aligned} \tag{70}$$

To show the converse, we first apply (64) and obtain

$$\begin{aligned}
nR_0 &\leq \sum_{l=1}^L \sum_{i=1}^n I(Q_{li}; Y_{li}) + n\delta_1 \\
&\leq \sum_{l \in A} \sum_{i=1}^n I(Q_{li}; Y_{li}) + \sum_{l \in A^c} \sum_{i=1}^n I(X_{li}; Y_{li}) + n\delta_1 \\
&= \sum_{l \in A} \sum_{i=1}^n [h(Y_{li}) - h(Y_{li}|Q_{li})] \\
&\quad + \sum_{l \in A^c} \sum_{i=1}^n [h(Y_{li}) - h(Y_{li}|X_{li})] + n\delta_1 \\
&\leq \sum_{l \in A} \sum_{i=1}^n \frac{1}{2} \log 2\pi e (EX_{li}^2 + \mu_l^2) - \sum_{l \in A} \sum_{i=1}^n h(Y_{li}|Q_{li}) \\
&\quad + \sum_{l \in A^c} \sum_{i=1}^n \frac{1}{2} \log 2\pi e (EX_{li}^2 + \mu_l^2) \\
&\quad - \sum_{l \in A^c} \sum_{i=1}^n \frac{n}{2} \log 2\pi e \mu_l^2 + n\delta_1 \\
&\leq \sum_{l \in A} \frac{n}{2} \log 2\pi e \left( \frac{1}{n} \sum_{i=1}^n EX_{li}^2 + \mu_l^2 \right) - \sum_{l \in A} \sum_{i=1}^n h(Y_{li}|Q_{li}) \\
&\quad + \sum_{l \in A^c} \frac{n}{2} \log 2\pi e \left( \frac{1}{n} \sum_{i=1}^n EX_{li}^2 + \mu_l^2 \right) \\
&\quad - \sum_{l \in A^c} \frac{n}{2} \log 2\pi e \mu_l^2 + n\delta_1 \\
&\leq \sum_{l \in A} \frac{n}{2} \log 2\pi e (p_l + \mu_l^2) - \sum_{l \in A} \sum_{i=1}^n h(Y_{li}|Q_{li}) \\
&\quad + \sum_{l \in A^c} \frac{n}{2} \log \left( 1 + \frac{p_l}{\mu_l^2} \right) + n\delta_1 \tag{71}
\end{aligned}$$

where we define  $p_l = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_{li}^2]$ .

It is easy to see that for  $l \in A$

$$\sum_{i=1}^n h(Y_{li}|Q_{li}) \leq \sum_{i=1}^n h(Y_{li}) \leq \frac{n}{2} \log 2\pi e (p_l + \mu_l^2) \tag{72}$$

and

$$\sum_{i=1}^n h(Y_{li}|Q_{li}) \geq \sum_{i=1}^n h(Y_{li}|X_{li}) = \frac{n}{2} \log 2\pi e \mu_l^2. \tag{73}$$

Hence, there exists  $0 \leq \beta_l \leq 1$  such that

$$\sum_{i=1}^n h(Y_{li}|Q_{li}) = \frac{n}{2} \log 2\pi e (\beta_l p_l + \mu_l^2). \tag{74}$$

Applying (74) to (71), we obtain

$$\begin{aligned}
nR_0 &\leq \sum_{l \in A} \frac{n}{2} \log \left( 1 + \frac{(1-\beta_l)p_l}{\beta_l p_l + \mu_l^2} \right) \\
&\quad + \sum_{l \in A^c} \frac{n}{2} \log \left( 1 + \frac{p_l}{\mu_l^2} \right) + n\delta_1. \tag{75}
\end{aligned}$$

We apply (65), follow the steps that are similar to those in (71), and obtain the following bound:

$$\begin{aligned}
nR_0 &\leq \sum_{l \in A} \frac{n}{2} \log 2\pi e (p_l + \nu_l^2) - \sum_{l \in A} \sum_{i=1}^n h(Z_{li}|Q_{li}) \\
&\quad + \sum_{l \in A^c} \frac{n}{2} \log \left( 1 + \frac{p_l}{\nu_l^2} \right) + n\delta_2. \tag{76}
\end{aligned}$$

For the second term in the preceding equation, we apply the entropy power inequality and obtain

$$\begin{aligned}
h(Z_{li}|Q_{li} = q_{li}) &= h(Y_{li} + V'_{li}|Q_{li} = q_{li}) \\
&\geq \frac{1}{2} \log \left( 2^{2h(Y_{li}|Q_{li}=q_{li})} + 2^{2h(V'_{li}|Q_{li}=q_{li})} \right) \\
&= \frac{1}{2} \log \left( 2^{2h(Y_{li}|Q_{li}=q_{li})} + 2\pi e (\nu_l^2 - \mu_l^2) \right). \tag{77}
\end{aligned}$$

Hence

$$\begin{aligned}
&\sum_{i=1}^n h(Z_{li}|Q_{li}) \\
&\geq \frac{1}{2} \sum_{i=1}^n \mathbb{E} \log \left( 2^{2h(Y_{li}|Q_{li}=q_{li})} + 2\pi e (\nu_l^2 - \mu_l^2) \right) \\
&\stackrel{(a)}{\geq} \frac{1}{2} \sum_{i=1}^n \log \left( 2^{2\mathbb{E}h(Y_{li}|Q_{li}=q_{li})} + 2\pi e (\nu_l^2 - \mu_l^2) \right) \\
&= \frac{1}{2} \sum_{i=1}^n \log \left( 2^{2h(Y_{li}|Q_{li})} + 2\pi e (\nu_l^2 - \mu_l^2) \right) \\
&\stackrel{(b)}{\geq} \frac{n}{2} \log \left( 2^{2\frac{1}{n} \sum_{i=1}^n h(Y_{li}|Q_{li})} + 2\pi e (\nu_l^2 - \mu_l^2) \right) \\
&\stackrel{(c)}{=} \frac{n}{2} \log \left( 2\pi e (\beta_l p_l + \mu_l^2) + 2\pi e (\nu_l^2 - \mu_l^2) \right) \\
&= \frac{n}{2} \log \left( 2\pi e (\beta_l p_l + \nu_l^2) \right) \tag{78}
\end{aligned}$$

where (a) and (b) follow from Jensen's inequality and the fact that  $\log(2^x + c)$  is a convex function of  $x$ , and (c) follows from (74).

By applying (78) to (76), we obtain

$$\begin{aligned}
nR_0 &\leq \sum_{l \in A} \frac{n}{2} \log \left( 1 + \frac{(1-\beta_l)p_l}{\beta_l p_l + \nu_l^2} \right) \\
&\quad + \sum_{l \in A^c} \frac{n}{2} \log \left( 1 + \frac{p_l}{\nu_l^2} \right) + n\delta_2. \tag{79}
\end{aligned}$$

We apply (66), follow steps similar to those in (20), and obtain the following bound:

$$\begin{aligned}
nR_1 &\leq \sum_{l \in A} \sum_{i=1}^n [I(X_{li}; Y_{li}|Q_{li}) - I(X_{li}; Z_{li}|Q_{li})] + n(\delta_1 + \delta_2) \\
&= \sum_{l \in A} \sum_{i=1}^n [h(Y_{li}|Q_{li}) - h(Y_{li}|X_{li}; Q_{li}) - h(Z_{li}|Q_{li}) \\
&\quad + h(Z_{li}|X_{li}; Q_{li})] + n(\delta_1 + \delta_2) \\
&\leq \sum_{l \in A} \left[ \frac{n}{2} \log \left( 1 + \frac{\beta_l p_l}{\mu_l^2} \right) - \frac{n}{2} \log \left( 1 + \frac{\beta_l p_l}{\nu_l^2} \right) \right] + n(\delta_1 + \delta_2) \tag{80}
\end{aligned}$$

where the last equality follows from (74) and (78).

For  $l \in A$ , we define  $p_{l0} = (1 - \beta_l)p_l$  and  $p_{l1} = \beta_l p_l$ . For  $l \in A^c$ , we define  $p_{l0} = p_l$ . It is clear from (22) that

$$\sum_{l \in A} [p_{l0} + p_{l1}] + \sum_{l \in A^c} p_{l0} \leq P. \quad (81)$$

This concludes the proof of the converse.

### APPENDIX III PROOF OF THEOREM 3

We apply Lemma 2 and consider the following three cases. For each case, we apply the technique in [17] to solve the optimization problem.

**Case 1:** We need to find  $\underline{p}^{(1)} \in \mathcal{P}$  that maximizes

$$\gamma_0 R_{01}(\underline{p}) + \gamma_1 R_{11}(\underline{p}).$$

If  $\underline{p}^{(1)}$  satisfies  $R_{01}(\underline{p}^{(1)}) < R_{02}(\underline{p}^{(1)})$ , then the optimal  $\underline{p}^* = \underline{p}^{(1)}$ .

The Lagrangian is given by

$$\begin{aligned} \mathcal{L} = & \sum_{l \in A} \left[ \frac{\gamma_0}{2} \log \left( 1 + \frac{p_{l0}}{\mu_l^2 + p_{l1}} \right) + \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l1}}{\mu_l^2} \right) \right. \\ & \left. - \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l1}}{\nu_l^2} \right) \right] + \sum_{l \in A^c} \frac{\gamma_0}{2} \log \left( 1 + \frac{p_{l0}}{\mu_l^2} \right) \\ & - \lambda \left[ \sum_{l \in A} [p_{l0} + p_{l1}] + \sum_{l \in A^c} p_{l0} \right] \end{aligned} \quad (82)$$

where  $\lambda$  is a Lagrange multiplier.

For  $l \in A^c$ ,  $p_{l0}^{(1)}$  needs to maximize the following  $\mathcal{L}_l$ :

$$\begin{aligned} \mathcal{L}_l = & \frac{\gamma_0}{2} \log \left( 1 + \frac{p_{l0}}{\mu_l^2} \right) - \lambda p_{l0} \\ = & \int_0^{p_{l0}} \left( \frac{\gamma_0}{2 \ln 2} \frac{1}{\mu_l^2 + x} - \lambda \right) dx. \end{aligned} \quad (83)$$

It is clear that  $p_{l0}^{(1)}$  that optimizes  $\mathcal{L}_l$  is either the root of the following equation:

$$\frac{\gamma_0}{2 \ln 2} \frac{1}{\mu_l^2 + x} - \lambda = 0 \quad (84)$$

if the root is positive or zero, i.e.,

$$p_{l0}^{(1)} = \left( \frac{\gamma_0}{2 \lambda \ln 2} - \mu_l^2 \right)^+. \quad (85)$$

For  $l \in A$ ,  $p_{l0}^{(1)}$  and  $p_{l1}^{(1)}$  need to maximize the following  $\mathcal{L}_l$ :

$$\begin{aligned} \mathcal{L}_l = & \frac{\gamma_0}{2} \log \left( 1 + \frac{p_{l0}}{\mu_l^2 + p_{l1}} \right) + \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l1}}{\mu_l^2} \right) \\ & - \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l1}}{\nu_l^2} \right) - \lambda (p_{l0} + p_{l1}) \\ = & \int_{p_{l1}}^{p_{l1} + p_{l0}} u_{l0}^{(1)}(x) dx + \int_0^{p_{l1}} u_{l1}^{(1)}(x) dx \\ \leq & \int_0^\infty \left( \max \{ u_{l0}^{(1)}(x), u_{l1}^{(1)}(x) \} \right)^+ dx \end{aligned} \quad (86)$$

where

$$\begin{aligned} u_{l0}^{(1)}(x) &= \frac{\gamma_0}{2 \ln 2} \frac{1}{\mu_l^2 + x} - \lambda \\ u_{l1}^{(1)}(x) &= \frac{\gamma_1}{2 \ln 2} \left( \frac{1}{\mu_l^2 + x} - \frac{1}{\nu_l^2 + x} \right) - \lambda. \end{aligned} \quad (87)$$

We next derive  $p_{l0}^{(1)}$  and  $p_{l1}^{(1)}$  that achieve the upper bound on  $\mathcal{L}_l$  in (86) and hence maximize  $\mathcal{L}_l$ .

We define  $x_{l0}^{(1)}$  to be the root of  $u_{l0}^{(1)}(x) = 0$  and  $x_{l1}^{(1)}$  to be the largest root of  $u_{l1}^{(1)}(x) = 0$ , i.e.,

$$\begin{aligned} x_{l0}^{(1)} &= \frac{\gamma_0}{2 \lambda \ln 2} - \mu_l^2 \\ x_{l1}^{(1)} &= \frac{1}{2} \sqrt{(\nu_l^2 - \mu_l^2) \left( \nu_l^2 - \mu_l^2 + \frac{2 \gamma_1}{\lambda \ln 2} \right)} - \frac{1}{2} (\mu_l^2 + \nu_l^2). \end{aligned} \quad (88)$$

It can be seen that  $u_{l0}^{(1)}(x)$  and  $u_{l1}^{(1)}(x)$  intersect only once at

$$x_{lr}^{(1)} = \frac{\gamma_1}{\gamma_0} (\nu_l^2 - \mu_l^2) - \nu_l^2. \quad (89)$$

In the following, we consider two cases.

1)  $\frac{\gamma_1}{\gamma_0} > \frac{\nu_l^2}{\nu_l^2 - \mu_l^2}$ , i.e.,  $x_{lr}^{(1)}$  is positive.

It is easy to see that  $u_{l1}^{(1)}(0) > u_{l0}^{(1)}(0)$ . The optimal  $p_{l0}^{(1)}$  and  $p_{l1}^{(1)}$  depend on the value of  $\lambda$  and fall into the following three possibilities.

(a) If  $u_{l1}^{(1)}(0) < 0$  (see Fig. 12(1)-(a)), i.e.,

$$\frac{\gamma_1 (\nu_l^2 - \mu_l^2)}{2 \mu_l^2 \nu_l^2 \ln 2} < \lambda$$

then both  $u_{l0}^{(1)}(x)$  and  $u_{l1}^{(1)}(x)$  are negative for  $x > 0$ . The upper bound on  $\mathcal{L}_l$  in (86) is achieved by  $p_{l0}^{(1)} = 0$  and  $p_{l1}^{(1)} = 0$ .

(b) If  $u_{l1}^{(1)}(0) \geq 0$  and  $x_{l0}^{(1)} < x_{lr}^{(1)}$  (see Fig. 12(1)-(b)), i.e.,

$$\frac{\gamma_0^2}{2 \ln 2 (\gamma_1 - \gamma_0) (\nu_l^2 - \mu_l^2)} < \lambda \leq \frac{\gamma_1 (\nu_l^2 - \mu_l^2)}{2 \mu_l^2 \nu_l^2 \ln 2}$$

then the upper bound on  $\mathcal{L}_l$  in (86) is achieved by  $p_{l0}^{(1)} = 0$  and  $p_{l1}^{(1)} = x_{l1}^{(1)}$ .

(c) If  $x_{l0}^{(1)} \geq x_{lr}^{(1)}$  (see Fig. 12(1)-(c)), i.e.,

$$\lambda \leq \frac{\gamma_0^2}{2 \ln 2 (\gamma_1 - \gamma_0) (\nu_l^2 - \mu_l^2)}$$

then the upper bound on  $\mathcal{L}_l$  in (86) is achieved by  $p_{l0}^{(1)} = x_{l0}^{(1)} - x_{lr}^{(1)}$  and  $p_{l1}^{(1)} = x_{lr}^{(1)}$ .

In summary, we obtain

$$p_{l0}^{(1)} = [x_{l0}^{(1)} - x_{lr}^{(1)}]^+ \quad \text{and} \quad p_{l1}^{(1)} = \left[ \min \{ x_{l1}^{(1)}, x_{lr}^{(1)} \} \right]^+. \quad (90)$$

2)  $\frac{\gamma_1}{\gamma_0} \leq \frac{\nu_l^2}{\nu_l^2 - \mu_l^2}$ , i.e.,  $x_{lr}^{(1)}$  is zero or negative.

It is easy to see that  $u_{l0}^{(1)}(0) \geq u_{l1}^{(1)}(0)$ .

(a) If  $u_{l0}^{(1)}(0) \leq 0$  (see Fig. 13(2)-(a)), i.e.,

$$\frac{\gamma_0}{2 \mu_l^2 \ln 2} \leq \lambda$$

then the upper bound on  $\mathcal{L}_l$  in (86) is achieved by  $p_{l0}^{(1)} = 0$  and  $p_{l1}^{(1)} = 0$ .

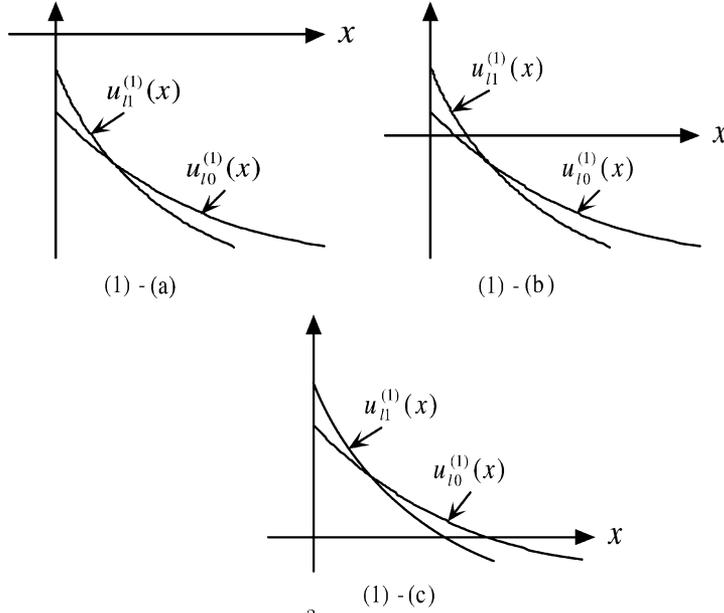


Fig. 12. Illustration of  $u_{l_0}^{(1)}(x)$  and  $u_{l_1}^{(1)}(x)$  for Case 1 with  $\frac{\gamma_1}{\gamma_0} > \frac{\nu_l^2}{\nu_l^2 - \mu_l^2}$ .

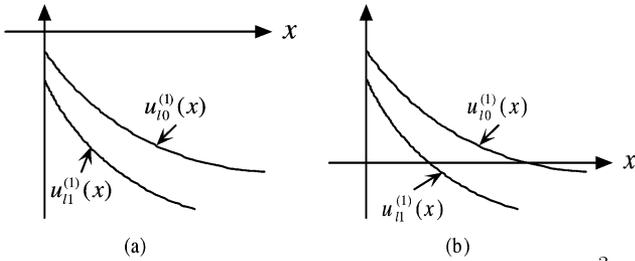


Fig. 13. Illustration of  $u_{l_0}^{(1)}(x)$  and  $u_{l_1}^{(1)}(x)$  for Case 1 with  $\frac{\gamma_1}{\gamma_0} \le \frac{\nu_l^2}{\nu_l^2 - \mu_l^2}$ .

(b) If  $u_{l_0}^{(1)}(0) > 0$  (see Fig. 13(b)), i.e.,

$$\lambda < \frac{\gamma_0}{2\mu_l^2 \ln 2}$$

then the upper bound on  $\mathcal{L}_l$  in (86) is achieved by  $p_{l_0}^{(1)} = x_{l_0}^{(1)}$  and  $p_{l_1}^{(1)} = 0$ .

In summary, we obtain

$$p_{l_0}^{(1)} = [x_{l_0}^{(1)}]^+, \quad p_{l_1}^{(1)} = 0. \quad (91)$$

The Lagrange parameter  $\lambda$  needs to be chosen to satisfy the power constraint

$$\sum_{l \in A} [p_{l_0} + p_{l_1}] + \sum_{l \in A^c} p_{l_0} \leq P. \quad (92)$$

According to Lemma 2, if the condition  $R_{01}(\underline{p}^{(1)}) < R_{02}(\underline{p}^{(1)})$  is satisfied, then the optimal  $\underline{p}^* = \underline{p}^{(1)}$ .

**Case 2:** We need to find  $\underline{p}^{(2)} \in \underline{\mathcal{P}}$  that maximizes

$$\gamma_0 R_{02}(\underline{p}) + \gamma_1 R_{11}(\underline{p}).$$

If  $\underline{p}^{(2)}$  satisfies  $R_{01}(\underline{p}^{(2)}) > R_{02}(\underline{p}^{(2)})$ , then the optimal  $\underline{p}^* = \underline{p}^{(2)}$ .

The Lagrangian is given by

$$\mathcal{L} = \sum_{l \in A} \left[ \frac{\gamma_0}{2} \log \left( 1 + \frac{p_{l_0}}{\nu_l^2 + p_{l_1}} \right) + \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l_1}}{\mu_l^2} \right) - \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l_1}}{\nu_l^2} \right) \right]$$

$$+ \sum_{l \in A^c} \frac{\gamma_0}{2} \log \left( 1 + \frac{p_{l_0}}{\nu_l^2} \right) - \lambda \left[ \sum_{l \in A} [p_{l_0} + p_{l_1}] + \sum_{l \in A^c} p_{l_0} \right] \quad (93)$$

where  $\lambda$  is a Lagrange multiplier.

For  $l \in A^c$ , it is easy to see that

$$p_{l_0}^{(2)} = \left( \frac{\gamma_0}{2\lambda \ln 2} - \nu_l^2 \right)^+. \quad (94)$$

For  $l \in A$ ,  $p_{l_0}^{(2)}$ , and  $p_{l_1}^{(2)}$  need to maximize the following  $\mathcal{L}_l$ :

$$\begin{aligned} \mathcal{L}_l &= \frac{\gamma_0}{2} \log \left( 1 + \frac{p_{l_0}}{\nu_l^2 + p_{l_1}} \right) + \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l_1}}{\mu_l^2} \right) \\ &\quad - \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l_1}}{\nu_l^2} \right) - \lambda (p_{l_0} + p_{l_1}) \\ &= \int_{p_{l_1}}^{p_{l_1} + p_{l_0}} u_{l_0}^{(2)}(x) dx + \int_0^{p_{l_1}} u_{l_1}^{(2)}(x) dx \\ &\leq \int_0^\infty \left( \max \{ u_{l_0}^{(2)}(x), u_{l_1}^{(2)}(x) \} \right)^+ dx \end{aligned} \quad (95)$$

where

$$\begin{aligned} u_{l_0}^{(2)}(x) &= \frac{\gamma_0}{2 \ln 2} \frac{1}{\nu_l^2 + x} - \lambda \\ u_{l_1}^{(2)}(x) &= \frac{\gamma_1}{2 \ln 2} \left( \frac{1}{\mu_l^2 + x} - \frac{1}{\nu_l^2 + x} \right) - \lambda. \end{aligned} \quad (96)$$

We define  $x_{l_0}^{(2)}$  to be the root of  $u_{l_0}^{(2)}(x) = 0$  and  $x_{l_1}^{(2)}$  to be the root of  $u_{l_1}^{(2)}(x) = 0$ , i.e.,

$$\begin{aligned} x_{l_0}^{(2)} &= \frac{\gamma_0}{2\lambda \ln 2} - \nu_l^2 \\ x_{l_1}^{(2)} &= \frac{1}{2} \sqrt{(\nu_l^2 - \mu_l^2) \left( \nu_l^2 - \mu_l^2 + \frac{2\gamma_1}{\lambda \ln 2} \right)} - \frac{1}{2} (\mu_l^2 + \nu_l^2). \end{aligned} \quad (97)$$

It is clear that  $u_{i0}^{(2)}(x)$  and  $u_{ir}^{(2)}(x)$  intersect only once at

$$x_{ir}^{(2)} = \frac{\gamma_1}{\gamma_0} (\nu_i^2 - \mu_i^2) - \mu_i^2. \quad (98)$$

Following steps similar to those in Case 1, we obtain the following  $p_{i0}^{(2)}$  and  $p_{ir}^{(2)}$  that achieve the upper bound on  $\mathcal{L}_i$  in (95) and hence maximize  $\mathcal{L}_i$ .

1) If  $\frac{\gamma_1}{\gamma_0} > \frac{\mu_i^2}{\nu_i^2 - \mu_i^2}$ , i.e.,  $x_{ir}^{(2)}$  is positive, then

$$p_{i0}^{(2)} = [x_{i0}^{(2)} - x_{ir}^{(2)}]^+ \quad \text{and} \quad p_{ir}^{(2)} = [\min \{x_{i1}^{(2)}, x_{ir}^{(2)}\}]^+. \quad (99)$$

2) If  $\frac{\gamma_1}{\gamma_0} \leq \frac{\mu_i^2}{\nu_i^2 - \mu_i^2}$ , i.e.,  $x_{ir}^{(2)}$  is zero or negative, then

$$p_{i0}^{(2)} = [x_{i0}^{(2)}]^+ \quad \text{and} \quad p_{ir}^{(2)} = 0. \quad (100)$$

The Lagrange multiplier  $\lambda$  needs to be chosen to satisfy the power constraint. According to Lemma 2, if the condition  $R_{01}(\underline{p}^{(2)}) > R_{02}(\underline{p}^{(2)})$  is satisfied, then the optimal  $\underline{p}^* = \underline{p}^{(2)}$ .

**Case 3:** We need to find  $\underline{p}^{(\alpha)} \in \mathcal{P}$  that maximizes

$$\gamma_0(\alpha R_{01}(\underline{p}) + \bar{\alpha} R_{02}(\underline{p})) + \gamma_1 R_1(\underline{p})$$

for a given  $0 \leq \alpha \leq 1$ . We then choose  $\alpha$  to satisfy  $R_{01}(\underline{p}^{(\alpha)}) = R_{02}(\underline{p}^{(\alpha)})$ . According to Lemma 2, if we have not found the optimal  $\underline{p}^*$  in Cases 1 and 2, such an  $\alpha$  must exist.

The Lagrangian is given by

$$\begin{aligned} \mathcal{L} = & \sum_{l \in A} \left[ \frac{\gamma_0 \alpha}{2} \log \left( 1 + \frac{p_{l0}}{\mu_l^2 + p_{l1}} \right) + \frac{\gamma_0 \bar{\alpha}}{2} \log \left( 1 + \frac{p_{l0}}{\nu_l^2 + p_{l1}} \right) \right. \\ & \left. + \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l1}}{\mu_l^2} \right) - \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l1}}{\nu_l^2} \right) \right] \\ & + \sum_{l \in A^c} \left[ \frac{\gamma_0 \alpha}{2} \log \left( 1 + \frac{p_{l0}}{\mu_l^2} \right) + \frac{\gamma_0 \bar{\alpha}}{2} \log \left( 1 + \frac{p_{l0}}{\nu_l^2} \right) \right] \\ & - \lambda \left[ \sum_{l \in A} [p_{l0} + p_{l1}] + \sum_{l \in A^c} p_{l0} \right] \end{aligned} \quad (101)$$

where  $\lambda$  is a Lagrange multiplier.

For  $l \in A^c$ , it is clear that  $p_{l0}^{(\alpha)}$  is either the root of the following equation:

$$\frac{\gamma_0 \alpha}{2 \ln 2} \frac{1}{\mu_l^2 + x} + \frac{\gamma_0 \bar{\alpha}}{2 \ln 2} \frac{1}{\nu_l^2 + x} = \lambda \quad (102)$$

if the root is positive, or zero. Hence,  $p_{l0}^{(\alpha)}$  is given by

$$\begin{aligned} p_{l0}^{(\alpha)} = & \left( \frac{1}{2} \sqrt{\left( \nu_l^2 - \mu_l^2 - \frac{\gamma_0}{2 \ln 2 \lambda} \right)^2 + \frac{2 \alpha \gamma_0}{\lambda \ln 2} (\nu_l^2 - \mu_l^2)} \right. \\ & \left. - \frac{1}{2} \left( \mu_l^2 + \nu_l^2 - \frac{\gamma_0}{2 \ln 2 \lambda} \right) \right)^+. \end{aligned} \quad (103)$$

For  $l \in A$ ,  $p_{l0}^{(\alpha)}$ , and  $p_{l1}^{(\alpha)}$  need to maximize the following  $\mathcal{L}_l$ :

$$\begin{aligned} \mathcal{L}_l = & \frac{\gamma_0 \alpha}{2} \log \left( 1 + \frac{p_{l0}}{\mu_l^2 + p_{l1}} \right) + \frac{\gamma_0 \bar{\alpha}}{2} \log \left( 1 + \frac{p_{l0}}{\nu_l^2 + p_{l1}} \right) \\ & + \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l1}}{\mu_l^2} \right) - \frac{\gamma_1}{2} \log \left( 1 + \frac{p_{l1}}{\nu_l^2} \right) - \lambda (p_{l0} + p_{l1}) \\ = & \int_{p_{l1}}^{p_{l1} + p_{l0}} u_{i0}^{(\alpha)}(x) dx + \int_0^{p_{l1}} u_{i1}^{(\alpha)}(x) dx \\ \leq & \int_0^\infty \left( \max \{u_{i0}^{(\alpha)}(x), u_{i1}^{(\alpha)}(x)\} \right)^+ dx \end{aligned} \quad (104)$$

where

$$\begin{aligned} u_{i0}^{(\alpha)}(x) &= \frac{\gamma_0}{2 \ln 2} \left( \frac{\alpha}{\mu_l^2 + x} + \frac{\bar{\alpha}}{\nu_l^2 + x} \right) - \lambda \\ u_{i1}^{(\alpha)}(x) &= \frac{\gamma_1}{2 \ln 2} \left( \frac{1}{\mu_l^2 + x} - \frac{1}{\nu_l^2 + x} \right) - \lambda. \end{aligned} \quad (105)$$

We define  $x_{i0}^{(\alpha)}$  to be the largest root of  $u_{i0}^{(\alpha)}(x) = 0$  and  $x_{i1}^{(\alpha)}$  to be the largest root of  $u_{i1}^{(\alpha)}(x) = 0$ , i.e.,

$$\begin{aligned} x_{i0}^{(\alpha)} &= \frac{1}{2} \sqrt{\left( \nu_l^2 - \mu_l^2 - \frac{\gamma_0}{2 \ln 2 \lambda} \right)^2 + \frac{2 \alpha \gamma_0}{\lambda \ln 2} (\nu_l^2 - \mu_l^2)} \\ & \quad - \frac{1}{2} \left( \mu_l^2 + \nu_l^2 - \frac{\gamma_0}{2 \ln 2 \lambda} \right) \\ x_{i1}^{(\alpha)} &= \frac{1}{2} \sqrt{(\nu_l^2 - \mu_l^2) \left( \nu_l^2 - \mu_l^2 + \frac{2 \gamma_1}{\lambda \ln 2} \right)} \\ & \quad - \frac{1}{2} (\mu_l^2 + \nu_l^2). \end{aligned} \quad (106)$$

It is clear that  $u_{i0}^{(\alpha)}(x)$  and  $u_{i1}^{(\alpha)}(x)$  intersect only once at

$$x_{ir}^{(\alpha)} = \frac{\gamma_1}{\gamma_0} (\nu_l^2 - \mu_l^2) - (\alpha \nu_l^2 + \bar{\alpha} \mu_l^2). \quad (107)$$

Following steps similar to those in Case 1, we obtain

1) If  $\frac{\gamma_1}{\gamma_0} > \frac{\alpha \nu_l^2 + \bar{\alpha} \mu_l^2}{\nu_l^2 - \mu_l^2}$ , i.e.,  $x_{ir}^{(\alpha)}$  is positive, then

$$p_{l0}^{(\alpha)} = [x_{l0}^{(\alpha)} - x_{lr}^{(\alpha)}]^+ \quad \text{and} \quad p_{l1}^{(\alpha)} = [\min \{x_{l1}^{(\alpha)}, x_{lr}^{(\alpha)}\}]^+. \quad (108)$$

2) If  $\frac{\gamma_1}{\gamma_0} \leq \frac{\alpha \nu_l^2 + \bar{\alpha} \mu_l^2}{\nu_l^2 - \mu_l^2}$ , i.e.,  $x_{ir}^{(\alpha)}$  is negative or zero, then

$$p_{l0}^{(\alpha)} = [x_{l0}^{(\alpha)}]^+ \quad \text{and} \quad p_{l1}^{(\alpha)} = 0. \quad (109)$$

The Lagrange multiplier  $\lambda$  needs to be chosen to satisfy the power constraint. We finally choose  $\alpha$  to satisfy  $R_{01}(\underline{p}^{(\alpha)}) = R_{02}(\underline{p}^{(\alpha)})$ . Then  $\underline{p}^* = \underline{p}^{(\alpha)}$ .

#### ACKNOWLEDGMENT

The authors would like to thank the reviewers and the Guest Editor for their comments and suggestions.

## REFERENCES

- [1] A. D. Wyner, "The wiretap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
- [2] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wiretap channel," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 451–456, Jul. 1978.
- [3] P. Parada and R. Blahut, "Secrecy capacity of SIMO and slow fading channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Adelaide, Australia, Sep. 2005, pp. 2152–2155.
- [4] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seattle, WA, Jul. 2006, pp. 356–360.
- [5] Y. Liang and H. V. Poor, "Secure communication over fading channels," in *Proc. 44th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2006, pp. 817–823.
- [6] Z. Li, R. Yates, and W. Trappe, "Secrecy capacity of independent parallel channels," in *Proc. 44th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2006, pp. 841–848.
- [7] P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, submitted for publication.
- [8] A. Khisti and G. Wornell, "The MIMOME channel," in *Proc. 45th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2007, pp. 625–632.
- [9] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," in *Proc. 45th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2007, pp. 848–855.
- [10] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multi-antenna wiretap channel," *IEEE Trans. Inf. Theory*, [Online]. Available: [http://arxiv.org/PS\\_cache/arxiv/pdf/0710/0710.4105v1.pdf](http://arxiv.org/PS_cache/arxiv/pdf/0710/0710.4105v1.pdf), submitted for publication
- [11] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wiretap channel: The 2-2-1 channel," *IEEE Trans. Inf. Theory*, submitted for publication.
- [12] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.
- [13] R. Liu, I. Maric, P. Spasojevic, and R. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate region," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.
- [14] R. Liu and H. V. Poor, "Secrecy capacity region of a multi-antenna Gaussian broadcast channel with confidential messages," *IEEE Trans. Inf. Theory*, submitted for publication.
- [15] J. Xu and B. Chen, "Broadcast confidential and public messages," in *Proc. 42nd Conf. Information Sciences and Systems (CISS)*, Princeton, NJ, Mar. 2008.
- [16] D. Hughes-Hartogs, "The Capacity of the Degraded Spectral Gaussian Broadcast Channel," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1975.
- [17] D. N. C. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Ulm, Germany, Jun. 1997, p. 27.
- [18] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels-Part I: Ergodic capacity," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1083–1102, Mar. 2001.
- [19] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels-Part II: Outage capacity," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1103–1127, Mar. 2001.
- [20] N. Jindal and A. Goldsmith, "Capacity and dirty paper coding for Gaussian broadcast channels with common information," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Chicago, IL, Jun./Jul. 2004, p. 215.
- [21] 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.
- [22] G. Caire, G. Taricco, and E. Biglieri, "Optimal power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, Jul. 1999.
- [23] A. Carleial and M. Hellman, "A note on Wyner's wiretap channel," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 3, pp. 387–390, May 1977.
- [24] S. K. Leung-Yan-Cheong, "On a special class of wiretap channels," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 5, pp. 625–627, Sep. 1977.
- [25] 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.
- [26] M. van Dijk, "On a special class of broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 712–714, Mar. 1997.
- [27] C. Mitrpant, A. J. H. Vinck, and Y. Luo, "An achievable region for the Gaussian wiretap channel with side information," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2181–2190, May 2006.
- [28] H. Koga and N. Sato, "On an upper bound of the secrecy capacity for a general wiretap channel," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Adelaide, Australia, Sep. 2005, pp. 1641–1645.
- [29] G. Cohen and G. Zémor, "Generalized coset schemes for the wiretap channel: Applications to biometrics," in *Proc. Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 46.
- [30] A. Thangaraj, S. Dihadar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Application of LDPC codes to the wiretap channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2933–2945, Aug. 2007.
- [31] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1562–1575, Apr. 2006.
- [32] A. Khisti, A. Tchamkerten, and G. Wornell, "Secure broadcasting," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453–2469, Jun. 2008.
- [33] 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.
- [34] R. Liu, Y. Liang, H. V. Poor, and P. Spasojevic, "Secure nested codes for type II wiretap channels," in *Proc. IEEE Information Theory Workshop (ITW)*, Lake Tahoe, CA, Sep. 2007, pp. 337–342.
- [35] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wiretap channels," in *Proc. 45th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2007, pp. 136–143.
- [36] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography-Part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.
- [37] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography-Part II: CR capacity," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 225–240, Jan. 1998.
- [38] U. M. Maurer, "Secret key agreement by public discussion based on common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
- [39] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 344–366, Mar. 2000.
- [40] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, Dec. 2004.
- [41] I. Csiszár and P. Narayan, "Secrecy capacities for multiterminal channel models," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Adelaide, Australia, Sep. 2005, pp. 2138–2141.
- [42] U. M. Maurer and S. Wolf, "From weak to strong information-theoretic key agreement," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Sorrento, Italy, Jun. 2000, p. 18.
- [43] U. M. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels-Part I. Definitions and a completeness result," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 822–831, Apr. 2003.
- [44] U. M. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels-Part II. Privacy amplification," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 832–838, Apr. 2003.
- [45] U. M. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels-Part III. Privacy amplification," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 839–851, Apr. 2003.
- [46] S. Venkatesan and V. Anantharam, "The common randomness capacity of a pair of independent discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 215–224, Jan. 1998.
- [47] S. Venkatesan and V. Anantharam, "The common randomness capacity of a network of discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 367–387, Mar. 2000.
- [48] Y. Liang and H. V. Poor, "Multiple access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.
- [49] R. Liu, I. Maric, R. Yates, and P. Spasojevic, "The discrete memoryless multiple access channel with confidential messages," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seattle, WA, Jul. 2006, pp. 957–961.
- [50] E. Tekin and A. Yener, "The Gaussian multiple access wiretap channel," *IEEE Trans. Inf. Theory*, submitted for publication.
- [51] E. Tekin and A. Yener, "Achievable rates for the general Gaussian multiple access wiretap channel with collective secrecy," in *Proc. 44th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2006, pp. 809–816.

- [52] Y. Oohama, "Coding for relay channels with confidential messages," in *Proc. IEEE Information Theory Workshop (ITW)*, Cairns, Australia, Sep. 2001, pp. 87–89.
- [53] Y. Oohama, "Relay channels with confidential messages," *IEEE Trans. Information Theory*, to be published.
- [54] L. Lai and H. El Gamal, "Cooperative secrecy: The relay-eavesdropper channel," in *Proc. IEEE Int. Symp. Information Theory*, Nice, France, Jun. 2007, pp. 931–935. The extended journal version "The relay-eavesdropper channel: Cooperation for secrecy," was submitted for publication in *IEEE Trans. Inf. Theory*.
- [55] X. He and A. Yener, "Secrecy when the relay is the eavesdropper," in *Proc. 3rd Information Theory and Applications Workshop*, San Diego, CA, Jan. 2008, to be published.
- [56] M. Yuksel and E. Erkip, "Secure communication with a relay helping the wiretapper," in *Proc. IEEE Information Theory Workshop (ITW)*, Lake Tahoe, CA, Sep. 2007, pp. 595–600.
- [57] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai (Shitz), and S. Verdú, "Cognitive interference channels with confidential messages," in *Proc. 45th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2007, pp. 1–6.
- [58] A. A. El Gamal, "Capacity of the product and sum of two unmatched broadcast channels," *Probl. Inf. Transm.*, vol. 16, no. 1, pp. 1–16, Jan.–Mar. 1980.
- [59] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [60] Y. Liang, V. V. Veeravalli, and H. V. Poor, "Resource allocation for wireless fading relay channels: Max-min solution," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3432–3453, Oct. 2007.