# Fading Cognitive Multiple-Access Channels With Confidential Messages

Ruoheng Liu, *Member, IEEE*, Yingbin Liang, *Member, IEEE*, and H. Vincent Poor, *Fellow, IEEE*

***Abstract*—The fading cognitive multiple-access channel with confidential messages (CMAC-CM) is investigated, in which two users (users 1 and 2) wish to transmit a common message to a destination and user 1 also has a confidential message intended for the destination. The two users transmit to the destination via a multiple access channel, and user 2 also receives noisy channel outputs. Such channel outputs potentially help user 2 to learn user 1's confidential information (although they are not exploited by user 2 for channel transmission). Hence, user 1 views user 2 as an eavesdropper and wishes to keep its confidential message as secret as possible from user 2. A parallel CMAC-CM with independent subchannels is first studied. The secrecy capacity region of the parallel CMAC-CM is established, which yields the secrecy capacity regions of the parallel CMAC-CM with degraded subchannels and the parallel Gaussian CMAC-CM. These results are then applied to study the fading CMAC-CM, in which both the user-to-user channel and the user-to-destination channel are corrupted by multiplicative fading gain coefficients in addition to additive white Gaussian noise. The channel state information (CSI) is assumed to be known at both the users and the destination. With the CSI, users can dynamically change their transmission powers with the channel realization to achieve the optimal performance. The closed-form power allocation function that achieves every boundary point of the secrecy capacity region is derived.**

***Index Terms*—Equivocation, fading channel, multiple-access channel, parallel channel, power allocation, secrecy capacity, secure communication.**

## I. Introduction

WIRELESS transmissions lack physical boundaries and so any adversary within range can receive them. Thus, security is one of the most important issues in wireless communications. One approach to security involves applying encryption algorithms to make messages unintelligible to adversaries. Unfortunately, these security methods are often designed without consideration of the specific properties of wireless networks. More specifically, encryption methods tend to be layer-specific and ignore the most fundamental communication layer, i.e., the physical-layer, whereby devices communicate through the encoding and modulation of information into waveforms.

The first study of secure communication via physical layer approaches was captured by a basic wiretap channel introduced by Wyner in [1]. In this model, a single source-destination communication link is eavesdropped upon by an eavesdropper via a degraded channel. The source node wishes to send confidential information to the destination node in a reliable manner as well as to keep the eavesdropper as ignorant of this information as possible. The performance measure of interest is the secrecy capacity which characterizes the largest possible reliable communication rate from the source node to the destination node with the eavesdropper obtaining no source information. Wyner's formulation was generalized by Csiszár and Körner in [2], in which the secrecy capacity region is established for a more general model referred to as the broadcast channel with confidential messages (BCC) [2].

More recently, multi-terminal communication with confidential messages has been studied intensively (see [3] for a recent survey of progress in this area). Among these studies, a generalization of both the wiretap channel and the classical multiple-access channel (MAC) was studied in [4], in which each user also receives channel outputs, and hence may obtain the confidential information sent by the other user from the channel output it receives. In this communication scenario, each user views the other user as an eavesdropper, and wishes to keep its confidential information as secret as possible from the other user. The authors of [4] investigated the rate-equivocation region and secrecy capacity region for this channel. Other related studies on secure communication over multiple-access channels can be found in [5]–[11].

Fading has traditionally been considered to be an obstacle to providing reliable wireless communication. However, over the past fifteen years, it has been demonstrated that fading can help improve capacity, reliability, and confidentiality of wireless networks. The impact of fading on secure communication was studied in, e.g., [12]–[14]. More specifically, [12] studied the secrecy capacity of ergodic fading BCCs when the channel state information (CSI) is known at all communicating nodes; [13] considered the ergodic scenario of fading wiretap channel in which the transmitter has no CSI about the eavesdropper channel, and the coding scheme relies on the ability to code within each coherence interval to guarantee secrecy; and [14] studied the outage preference of secure communication over wireless channels, in which the transmitter has no CSI about

Manuscript received December 29, 2009; revised November 21, 2010; accepted January 31, 2011. Date of current version July 29, 2011. R. Liu and H. V. Poor were supported in part by the National Science Foundation under Grant CNS-09-05398, and in part by the Air Force Office of Scientific Research under Grant FA9550-08-1-0480. Y. Liang was supported by a National Science Foundation CAREER Award under Grant CCF-10-26565 and Grant CCF-10-26566. The material in this paper was presented in part at the 47th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, September 30–October 2, 2009.

R. Liu is with Alcatel-Lucent, Murray Hill, NJ 07974 USA (e-mail: ruoheng.liu@alcatel-lucent.com).

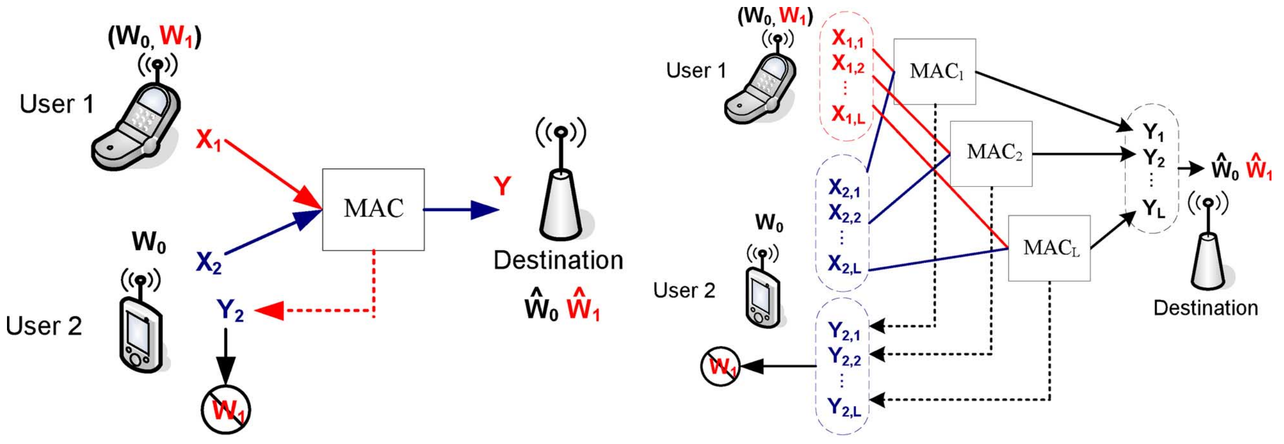
Y. Liang is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: yliang06@syr.edu).

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

Communicated by S. Ulukus, Associate Editor for Communication Networks.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2011.2158471



(a)

(b)

Fig. 1. Cognitive multiple-access channel with confidential messages. (a) CMAC-CM. (b) Parallel CMAC-CM.

either the legitimate receiver's channel or the eavesdropper's channel. More studies on secure communication over fading channels can be found in [3] and the references therein.

In this paper, we investigate the fading cognitive multiple-access channel with both common and confidential messages. Our study is a natural extension of the study of secure communication over the MAC in [4]. More specifically, we consider the scenario in which two users (users 1 and 2) wish to transmit a common message to a destination and user 1 also has a confidential message intended for the destination. The two users transmit to the destination via a multiple-access channel, and user 2 also receives noisy channel outputs. Such channel outputs potentially help user 2 to learn user 1's confidential information. Hence, user 1 views user 2 as an eavesdropper and wishes to keep its confidential message as secret as possible from user 2. We refer to this model as the cognitive MAC with one confidential message (CMAC-CM) (see Fig. 1(a)); this terminology reflects the fact that this channel also models cognitive communication in which the secondary user (user 1) helps the primary user (user 2) to send a common message $W_0$, and also has a confidential message $W_1$ intended for the destination, needing to be kept secret from the primary user. We note that in this paper, we focus on the impact of user 2's outputs on secrecy of user 1's confidential message. The potential help of these outputs as feedback information for user 2 to adapt transmission schemes is not considered. We refer the reader to [8], [15]–[17] for recent studies that exploit feedback and cooperation to improve secrecy. We also note that although user 1 may receive the channel outputs as well, these outputs do not play a role in affecting secrecy (and hence are not considered) because user 2 does not have a confidential message. A more general scenario with both users having confidential messages was studied in [4], in which outputs at both users affect secrecy. This paper focuses on the case with only one confidential message for the sake of tractability of optimal solutions. Our approach can be extended to the more general scenario although without an optimality guarantee for the solutions.

Our focus in this paper is on the situation in which both the user-to-user and the user-to-destination channels are corrupted by multiplicative fading gain coefficients in addition to additive white Gaussian noise. The fading CMAC-CM model captures the basic time-varying and superposition properties of wireless channels, and thus understanding this channel plays an important role in addressing security issues in wireless application. For the fading CMAC-CM, we assume that the fading gain coefficients are stationary and ergodic over time and that the CSI is known at both users and the destination. Note that knowledge of the user-to-destination CSI is necessary in order to cooperatively transmit the common message, and thus should be provided through state feedback from the destination terminal to the user terminals. Knowledge of CSI between the user terminals can be obtained via the reciprocity property of those channels. Users are motivated to exchange the CSI in order to enable better cooperation for sending the common message.

To solve the fading CMAC-CM problem, we first consider a general information-theoretic model, i.e., the parallel MAC with $L$ independent subchannels. As shown in Fig. 1(b), the two users communicate with the destination over $L$ parallel links and each of the $L$ links is eavesdropped upon by user 2. We establish the secrecy capacity region for the parallel CMAC-CM (with $L > 1$). In particular, we provide a converse proof to show that having independent inputs for each subchannel is optimal for achieving the secrecy capacity region. We note that this result does not follow directly from the capacity result given in [4] (with $L = 1$). The secrecy capacity region of the parallel CMAC-CM further yields the secrecy capacity region of the parallel CMAC-CM with degraded subchannels. Next, we consider the parallel Gaussian CMAC-CM, which is an example parallel CMAC-CM with degraded subchannels. Based on the maximum-entropy theorem [18] and the extremal inequality [19], we show that the secrecy capacity region of the parallel Gaussian CMAC-CM is achievable by using jointly Gaussian inputs and optimizing power allocations among the parallel subchannels for the two users. We then apply this result to investigate the fading CMAC-CM. We study the ergodic performance, for which no delay constraint on message transmission is assumed and the secrecy capacity region is averaged over all channel states. In fact, the fading CMAC-CM can be viewed as the parallel Gaussian CMAC-CM with each fading state corresponding to one subchannel. Hence, the secrecy capacity region of the parallel Gaussian CMAC-CM applies to the fading CMAC-CM. Since both users know the CSI, users can dynamically change their transmission powers with the channel realization to achieve the optimal performance. The optimal power allocation that achieves every boundary point of the secrecy capacity region can be characterized as a solution to a nonconvex optimization problem. The Karush–Kuhn–Tucker (KKT) conditions (as necessary conditions) greatly facilitate exploitation of the specific structure of the problem, and enable us to obtain a closed-form solution for the optimal power allocation strategy for the two users.

The remainder of this paper is organized as follows. We first study the parallel CMAC-CM with independent subchannels and its special case of the parallel CMAC-CM with degraded subschannels in Section II. Next, we investigate the secrecy capacity region of the parallel Gaussian CMAC-CM in Section III and the ergodic performance of the fading CMAC-CM in Section IV. We then provide some numerical examples in Section V. Finally, we summarize our results and discuss future work in Section VI.

## II. Parallel CMAC-CM

### A. Channel Model

We consider the discrete memoryless parallel CMAC-CM with $L$ independent subchannels (see Fig. 1(b)). Each subchannel is assumed to connect users 1 and 2 to the destination, and user 2 can receive the channel output from each subchannel, and hence may obtain information sent by user 1. We assume that user 2 does not use its channel outputs to adapt its transmission. The channel transition probability distribution is given by

$$p(y_{[1,L]}, y_{2,[1,L]} | x_{1,[1,L]}, x_{2,[1,L]}) = \prod_{j=1}^{L} p(y_j, y_{2,j} | x_{1,j}, x_{2,j})$$

where we use the notation $y_{[1,L]} := (y_1, \ldots, y_L)$, etc.

In this model, a common message $W_0 \in \mathcal{W}_0$ is known to both the primary user (user 2) and the secondary user (user 1), and hence both users cooperate to transmit $W_0$ to the destination. Moreover, the secondary user (user 1) also has confidential message $W_1 \in \mathcal{W}_1$ intended for the destination. User 1 views user 2 as an eavesdropper and wishes to keep its confidential information as secret as possible from user 2. The (stochastic) encoder at user 1, i.e., $f_1 : \mathcal{W}_0 \times \mathcal{W}_1 \to \mathcal{X}_1^n$, maps each message pair $(w_0, w_1) \in \mathcal{W}_0 \times \mathcal{W}_1$ to a codeword $x_1^n \in \mathcal{X}_1^n$; and the encoder at user 2, i.e., $f_2 : \mathcal{W}_0 \to \mathcal{X}_2^n$, maps each message $w_0 \in \mathcal{W}_0$ to a codeword $x_2^n \in \mathcal{X}_2^n$. The decoder at the destination, i.e., $g : \mathcal{Y}^n \to \mathcal{W}_0 \times \mathcal{W}_1$, maps a received sequence $y^n$ to a message pair $(\hat{w}_0, \hat{w}_1) \in \mathcal{W}_0 \times \mathcal{W}_1$.

In this paper, we focus on the case in which perfect secrecy is required, i.e., user 2 should not obtain any information about the message $W_1$. More formally, this condition is characterized by (e.g., see [1], [2], and [4])

$$\frac{1}{n} I(W_1; Y_2^n, X_2^n, W_0) \to 0$$

where $X_2^n := (X_{2,1}, \ldots, X_{2,n})$ and $Y_2^n := (Y_{2,1}, \ldots, Y_{2,n})$ are the input and output sequences of user 2, respectively, and the limit is taken as the block length $n \to \infty$. The goal is to characterize the *secrecy capacity region* $\mathcal{C}_s$ that contains all rate pairs $(R_0, R_1)$ achievable by some coding scheme with asymptotically small probability of decoding error and perfect secrecy.

### B. Secrecy Capacity Region of the Parallel CMAC-CM

For the parallel CMAC-CM, we obtain the following secrecy capacity region.

*Theorem 1:* For the parallel CMAC-CM, the secrecy capacity region is given by

$$\mathcal{C}_s^{[\mathrm{P}]} = \bigcup_{\substack{\prod_j p(q_j, x_{2,j}) p(u_j|q_j) p(x_{1,j}|u_j) \\ p(y_j, y_{2,j}|x_{1,j}, x_{2,j})}} \left\{ \begin{array}{l} (R_0, R_1) : \\ R_0 \geq 0, \ R_1 \geq 0; \\ R_1 \leq \sum_{j=1}^{L} [I(U_j; Y_j | X_{2,j}, Q_j) \\ \qquad -I(U_j; Y_{2,j} | X_{2,j}, Q_j)] \\ R_0 \leq \sum_{j=1}^{L} I(Q_j, X_{2,j}; Y_j) \end{array} \right\} \quad (1)$$

where $Q_j$ and $U_j$'s are auxiliary random variables, and $Q_j$ can be chosen to be a deterministic function of $U_j$ for $j = 1, \ldots, L$.

*Proof:* The achievable scheme for the parallel CMAC-CM is based on that for the CMAC-CM provided in [4]. The basic idea is that user 1 and user 2 cooperatively transmit $W_0$ by generating correlated inputs, and then user 1 generates the input for $W_1$ via superposition coding. To keep $W_1$ secret from user 1, the coding scheme for $W_1$ adopts stochastic encoding as in [1], i.e., for each message $W_1$, a codeword is randomly selected from a set of codewords (i.e., codewords in a bin) for transmission. We refer the reader to [4] for details. Based on this coding scheme, the achievable region for the parallel CMAC-CM can be obtained as detailed in Appendix A. The converse proof is also provided in Appendix A to establish optimality of the achievable rate region. ■

Theorem 1 implies that having independent inputs for each subchannel is optimal. This fact does not follow directly from the single-letter result on the secrecy capacity region of the CMAC-CM (i.e., the case when $L = 1$) given in [4]. Hence, a converse proof is needed, which is provided in Appendix A.

### C. Parallel CMAC-CM With Degraded Subchannels

We consider the parallel CMAC-CM with degraded subchannels, in which each subchannel is either degraded such that given the input of user 2, the output at user 2 is a conditionally degraded version of the output at the destination, or reversely degraded such that given the input of user 2, the output at the destination is a conditionally degraded version of the output at user 2.

Following [4], we define the conditionally degraded subchannels as follows. Let $\mathcal{A}$ denote the index set that includes all indices of subchannels such that given $x_{2,j}$, the output at user 2 is a conditionally degraded version of the output at the destination, i.e., for $j \in \mathcal{A}$

$$p(y_j, y_{2,j} | x_{1,j}, x_{2,j}) = p(y_j | x_{1,j}, x_{2,j}) p(y_{2,j} | y_j, x_{2,j}). \quad (2)$$

We further define $\bar{\mathcal{A}}$ to be the complement of the set $\mathcal{A}$, and $\bar{\mathcal{A}}$ includes all indices of subchannels such that given $x_{2,j}$, the output at the destination is a conditionally degraded version of the output at user 2, i.e., for $j \in \bar{\mathcal{A}}$

$$p(y_j, y_{2,j}|x_{1,j}, x_{2,j}) = p(y_{2,j}|x_{1,j}, x_{2,j})p(y_j|y_{2,j}, x_{2,j}). \quad (3)$$

Hence, the channel transition probability distribution is given by

$$\begin{aligned} p(y_{[1,L]}, y_{2,[1,L]}|x_{1,[1,L]}, x_{2,[1,L]}) &= \prod_{j\in\mathcal{A}} p(y_j|x_{1,j}, x_{2,j})p(y_{2,j}|y_j, x_{2,j}) \\ &\quad \cdot \prod_{j\in\bar{\mathcal{A}}} p(y_{2,j}|x_{1,j}, x_{2,j})p(y_j|y_{2,j}, x_{2,j}). \quad (4) \end{aligned}$$

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

*Theorem 2:* For the parallel CMAC-CM with degraded subchannels, the secrecy capacity region is given by

$$\mathcal{C}_s^{[\mathrm{D}]} = \bigcup_{\substack{\prod_j p(q_j,x_{2,j})p(x_{1,j}|q_j) \\ p(y_j,y_{2,j}|x_{1,j},x_{2,j})}} \left\{ \begin{array}{l} (R_0,\, R_1): \\ R_0 \geq 0,\ R_1 \geq 0; \\ \quad R_1 \leq \sum_{j\in\mathcal{A}}[I(X_{1,j};Y_j|X_{2,j},Q_j) \\ \qquad\qquad -I(X_{1,j};Y_{2,j}|X_{2,j},Q_j)] \\ \quad R_0 \leq \sum_{j\in\mathcal{A}} I(Q_j, X_{2,j}; Y_j) \\ \qquad\qquad +\sum_{j\in\bar{\mathcal{A}}} I(X_{1,j}, X_{2,j}; Y_j) \end{array} \right\} \quad (5)$$

where $Q_j$, for $j \in \mathcal{A}$, are auxiliary random variables that satisfy the Markov chain relationships

$$Q_j \rightarrow (X_{1,j}, X_{2,j}) \rightarrow (Y_j, Y_{2,j}).$$

*Proof:* See Appendix B. ■

It can be seen that the common message $W_0$ is sent over all subchannels, and the confidential message $W_1$ of user 1 is sent only over the subchannels for which the output at user 2 is a *conditionally degraded* version of the output at the destination. Furthermore, user 1 sends the common message $W_0$ and the confidential message $W_1$ by using superposition encoding.

We note that compared to Theorem 1, Theorem 2 has a simpler form by replacing the auxiliary random variable $U_j$ by $X_{1,j}$ due to the degradedness condition. Such simplification is beneficial for the analysis of the Gaussian case and the numerical computation in the later sections.

## III. Parallel Gaussian CMAC-CM

### A. Channel Model

In this section, we consider the parallel Gaussian CMAC-CM in which the channel outputs at the destination and user 2 are corrupted by additive Gaussian noise terms. The channel input-output relationship is given by

$$\begin{aligned} Y_{j,i} &= X_{1,j,i} + X_{2,j,i} + Z_{j,i} \\ \text{and} \qquad Y_{2,j,i} &= X_{1,j,i} + X_{2,j,i} + Z_{2,j,i} \end{aligned} \quad (6)$$

where $i$ is the time index, and for $j = 1, \ldots, L$, the noise processes $\{Z_{j,i}\}$ and $\{Z_{2,j,i}\}$ are independent and identically distributed (i.i.d.) with the components being zero-mean Gaussian random variables with variances $\nu_j$ and $\mu_j$, respectively. We assume $\nu_j < \mu_j$ for $j \in \mathcal{A}$ and $\nu_j \geq \mu_j$ for $j \in \bar{\mathcal{A}}$. The channel input sequences $X^n_{1,[1,L]}$ and $X^n_{2,[1,L]}$ are subject to average power constraints $P_1$ and $P_2$, respectively, i.e.,

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{L} \mathsf{E}[X^2_{1,j,i}] &\leq P_1 \\ \text{and} \qquad \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{L} \mathsf{E}[X^2_{2,j,i}] &\leq P_2. \end{aligned}$$

### B. Secrecy Capacity Region

We now apply Theorem 2 to obtain the secrecy capacity region of the parallel Gaussian MAC. It can be seen from (6) that the subchannels of the parallel Gaussian MAC are not physically degraded. We consider the following subchannels, for $j \in \mathcal{A}$:

$$Y_{j,i} = X_{1,j,i} + X_{2,j,i} + Z_{j,i}, \;\; Y_{2,j,i} = Y_{j,i} + Z'_{2,j,i}; \quad (7)$$

and, for $j \in \bar{\mathcal{A}}$

$$Y_{j,i} = Y_{2,j,i} + Z'_{j,i}, \;\; Y_{2,j,i} = X_{1,j,i} + X_{2,j,i} + Z_{2,j,i} \quad (8)$$

where $\{Z'_{j,i}\}$ and $\{Z'_{2,j,i}\}$ are i.i.d. random processes with components being zero-mean Gaussian random variables with variances $\nu_j - \mu_j$ for $j \in \bar{\mathcal{A}}$ and $\mu_j - \nu_j$ for $j \in \mathcal{A}$, respectively. Moreover, $\{Z'_{j,i}\}$ is independent of $\{Z_{2,j,i}\}$, and $\{Z'_{2,j,i}\}$ is independent of $\{Z_{j,i}\}$. We notice that the channel defined in (7)–(8) is a parallel Gaussian MAC with physically degraded subchannels. Since the channel (7)–(8) has the same marginal distributions $p(y|x_1, x_2)$ and $p(y_2|x_1, x_2)$ as the parallel Gaussian MAC defined in (6), these two channels have the same secrecy capacity region, because the reliability and secrecy conditions for the secrecy capacity region depend only on the marginals. Further details about this argument can be found in, e.g., [4, Lemma 1].

For the channel defined in (7)–(8), we can apply Theorem 2 to obtain the following secrecy capacity region. In particular, the degradedness of the subchannels allows the use of the entropy power inequality in the proof of the converse. We can thus obtain the secrecy capacity region for the parallel Gaussian CMAC-CM.

*Theorem 3:* For the parallel Gaussian CMAC-CM, the secrecy capacity region is given by

$$\mathcal{C}_s^{[\mathrm{G}]} = \bigcup_{\underline{p}\in\mathcal{P}} \left\{ \begin{array}{l} (R_0,\, R_1): \\ R_0 \geq 0,\; R_1 \geq 0; \\ R_1 \leq \sum_{j\in\mathcal{A}} \left[\frac{1}{2}\log\left(1+\frac{b_j}{\nu_j}\right) - \frac{1}{2}\log\left(1+\frac{b_j}{\mu_j}\right)\right] \\ R_0 \leq \sum_{j\in\mathcal{A}} \frac{1}{2}\log\left(1+\frac{a_j+p_{2,j}+2\sqrt{a_j p_{2,j}}}{b_j+\nu_j}\right) \\ \qquad + \sum_{j\in\bar{\mathcal{A}}} \frac{1}{2}\log\left(1+\frac{a_j+p_{2,j}+2\sqrt{a_j p_{2,j}}}{\nu_j}\right) \end{array} \right\} \tag{9}$$

where $\underline{p}$ is the power allocation vector, which consists of $(a_j, b_j, p_{2,j})$ for $j \in \mathcal{A}$ and $(a_j, 0, p_{2,j})$ for $j \in \bar{\mathcal{A}}$ as components, and the set $\mathcal{P}$ includes all power allocation vectors $\underline{p}$ that satisfy the power constraint

$$\mathcal{P} := \left\{ \underline{p} : \sum_{j=1}^{L}(a_j+b_j) \leq P_1 \text{ and } \sum_{j=1}^{L} p_{2,j} \leq P_2 \right\}. \tag{10}$$

*Proof:* See Appendix C. ■

We note that $\underline{p}$ denotes the power allocation among all subchannels. In particular, for $j \in \mathcal{A}$, since user 1 needs to transmit both common and confidential information, the pair $(a_j, b_j)$ controls the power allocation between the common message $W_0$ and the confidential message $W_1$. For $j \in \bar{\mathcal{A}}$, user 1 transmits only the common information, and $b_j = 0$ indicates that the power is allocated to transmit the common message $W_0$ only.

We also note that although the Gaussian model in (6) is not given in the general form of the Gaussian channel (which has arbitrary scaling factors for the channel inputs), our analysis and result (i.e., Theorem 3) for this model can be extended in a natural way to the general Gaussian channel. In fact, the fading channel that we study in Section IV can be viewed as a general Gaussian channel for each fading state, and Corollary 1 provides an extension of the above theorem to the general Gaussian channel.

### C. Optimal Power Allocation

To characterize the secrecy capacity region of the parallel Gaussian CMAC-CM given in (9), we need to characterize every boundary point and the power allocation vector that achieves each boundary point. Since the secrecy capacity region $\mathcal{C}_s^{[\mathrm{G}]}$ is convex, for every boundary point $(R_0^\star, R_1^\star)$, there exists $\gamma_1 \geq 0$ such that $(R_0^\star, R_1^\star)$ is the solution to the optimization problem

$$\max_{(R_0,R_1)\in\mathcal{C}_s^{[\mathrm{G}]}} \left[R_0 + \gamma_1 R_1\right]. \tag{11}$$

Note that the optimization problem (11) serves as a complete characterization of the corresponding boundary of the secrecy capacity region, and the solution to (11) provides the power allocations that achieve the boundary of the secrecy capacity region. We characterize the optimal power allocation $\underline{p}$ that solves (11) in the following theorem.

*Theorem 4:* Let $\underline{p}^\star$ be an optimal solution to the optimization problem of (11) that achieves the boundary of the secrecy capacity region of the parallel Gaussian CMAC-CM. Then, $\underline{p}^\star$ can be written as follows.

For $j \in \mathcal{A}$, if

$$\frac{2\lambda_1^2 \ln 2}{\lambda_1+\lambda_2} < \frac{\gamma_1(\mu_j-\nu_j)-\mu_j}{\mu_j\nu_j}$$

then

$$\begin{aligned} a_j^\star &= \frac{\lambda_2^2}{(\lambda_1+\lambda_2)^2}\left(s_{1,j}-\phi_j\right)^+ \\ b_j^\star &= \left(\min\left[s_{2,j},\, \phi_j\right]\right)^+ \\ \text{and} \quad p_{2,j}^\star &= \frac{\lambda_1^2}{(\lambda_1+\lambda_2)^2}\left(s_{1,j}-\phi_j\right)^+ \end{aligned} \tag{12}$$

alternatively, if

$$\frac{2\lambda_1^2 \ln 2}{\lambda_1+\lambda_2} \geq \frac{\gamma_1(\mu_j-\nu_j)-\mu_j}{\mu_j\nu_j}$$

then

$$\begin{aligned} a_j^\star &= \frac{\lambda_2^2}{(\lambda_1+\lambda_2)^2}\left(s_{1,j}\right)^+ \\ b_j^\star &= 0 \\ \text{and} \quad p_{2,j}^\star &= \frac{\lambda_1^2}{(\lambda_1+\lambda_2)^2}\left(s_{1,j}\right)^+ \end{aligned} \tag{13}$$

for $j \in \bar{\mathcal{A}}$

$$\begin{aligned} a_j^\star &= \frac{\lambda_2^2}{(\lambda_1+\lambda_2)^2}\left(s_{1,j}\right)^+ \\ \text{and} \quad p_{2,j}^\star &= \frac{\lambda_1^2}{(\lambda_1+\lambda_2)^2}\left(s_{1,j}\right)^+ \end{aligned} \tag{14}$$

where $(x)^+ = \max(0,\, x)$, $\gamma_1 \geq 0$

$$\begin{aligned} s_{1,j} &= \frac{\lambda_1+\lambda_2}{2\lambda_1\lambda_2\ln 2} - \nu_j \\ s_{2,j} &= \frac{1}{2}\left[\sqrt{(\mu_j-\nu_j)\left(\mu_j-\nu_j+\frac{2\gamma_1}{\lambda_1\ln 2}\right)} - (\mu_j+\nu_j)\right] \\ \phi_j &= -\frac{1}{2}\left(\mu_j+\nu_j+\frac{1}{\omega}\right) + \\ &\quad \frac{1}{2}\sqrt{\left(\mu_j+\nu_j+\frac{1}{\omega}\right)^2 - 4\left[\mu_j\nu_j - \frac{\gamma_1(\mu_j-\nu_j)-\mu_j}{\omega}\right]} \\ \omega &= (2\ln 2)\frac{\lambda_1^2}{\lambda_1+\lambda_2} \end{aligned} \tag{15}$$

and the pair $(\lambda_1, \lambda_2)$ is chosen to satisfy the power constraint

$$\sum_{j=1}^{L}(a_j+b_j) \leq P_1 \text{ and } \sum_{j=1}^{L} p_{2,j} \leq P_2. \tag{16}$$

*Proof:* The optimization problem is nonconvex. Our proof applies the KKT conditions (as necessary conditions), which help to express the Lagrangian in the form of an integral. This specific structure of the problem is then exploited to obtain a closed-form solution for the optimal power allocation strategy, which follows the technique introduced in [20]. The details can be found in Appendix D. ■

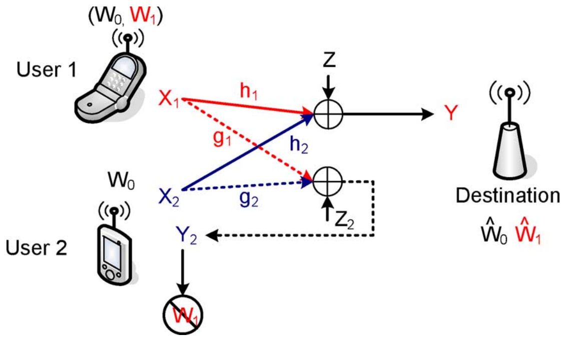


Fig. 2. Fading CMAC-CM.

## IV. Fading CMAC-CM

In this section, we study the fading CMAC-CM, where both the user-to-destination and the user-to-user channels are corrupted by multiplicative fading gain processes in addition to additive white Gaussian processes. The channel input-output relationship is given by

$$\begin{aligned} Y_i &= h_{1,i}X_{1,i} + h_{2,i}X_{2,i} + Z_i \\ \text{and} \quad Y_{2,i} &= g_{1,i}X_{1,i} + g_{2,i}X_{2,i} + Z_{2,i} \end{aligned} \tag{17}$$

where $i$ is the time index, $X_{1,i}$ and $X_{2,i}$ are channel inputs at the time instant $i$ from user 1 and user 2, respectively, $Y_i$ and $Y_{2,i}$ are channel outputs at the time instant $i$ at the destination and the receiver of user 2, respectively; $\underline{h}_i := (h_{1,i}, h_{2,i})$ and $\underline{g}_i := (g_{1,i}, g_{2,i})$ are proper complex random channel attenuation pairs imposed on the destination and the receiver of user 2; and the noise processes $\{Z_i\}$ and $\{Z_{2,i}\}$ are i.i.d. with the components being zero-mean proper complex Gaussian random variables with variances $\nu$ and $\mu$, respectively. The input sequences $\{X_{1,i}\}$ and $\{X_{2,i}\}$ are subject to the average power constraint $P_1$ and $P_2$, i.e.,

$$\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}[X_{1,i}^2] \le P_1 \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}[X_{2,i}^2] \le P_2.$$

We assume that the CSI (i.e., the realization of $(\underline{h}, \underline{g})$) is known at both the transmitters and the receivers instantaneously. Depending on the CSI, each user can dynamically change its transmission power and rate to achieve better performance. In this section, we assume that there is no delay constraint on the transmitted messages, and that the secrecy capacity region is an average over all channel states, which is referred to as the *ergodic* secrecy capacity region.

We note that for a given fading state, i.e., a realization of $(\underline{h}, \underline{g})$, the fading CMAC-CM is a Gaussian CMAC-CM. Hence, the fading CMAC-CM can be viewed as a parallel Gaussian CMAC-CM with each fading state corresponding to one subchannel. Thus, the following secrecy capacity region of the fading CMAC-CM follows from Theorem 3.

In the following, for each channel state $(\underline{h}, \underline{g})$, we use $p_1(\underline{h}, \underline{g})$ and $p_2(\underline{h}, \underline{g})$ to denote the powers allocated at users 1 and 2, respectively. We further define

$$p(\underline{h}, \underline{g}) := \big(a(\underline{h}, \underline{g}), b(\underline{h}, \underline{g}), p_2(\underline{h}, \underline{g})\big).$$

Let $\mathcal{P}$ denote the set that includes all power allocations satisfying the power constraint

$$\begin{aligned} \mathcal{P} := \big\{p(\underline{h}, \underline{g}) : &\mathsf{E}[a(\underline{h}, \underline{g}) + b(\underline{h}, \underline{g})] \le P_1 \\ &\text{and} \quad \mathsf{E}[p_2(\underline{h}, \underline{g})] \le P_2\big\} \end{aligned} \tag{18}$$

and $\mathcal{A}$ denote a set of channel states as follows:

$$\mathcal{A} := \left\{(\underline{h}, \underline{g}) : \frac{|h_1|^2}{\nu} > \frac{|g_1|^2}{\mu}\right\}.$$

The set $\mathcal{A}$ includes all channel states at which the destination receives better transmission from user 1 than user 2.

*Corollary 1:* The secrecy capacity region of the fading CMAC-CM is given by

$$\mathcal{C}_s^{[\mathrm{F}]} = \bigcup_{p(\underline{h},\underline{g})\in\mathcal{P}} \left\{ \begin{array}{l} (R_0,\, R_1): \\ R_0 \ge 0,\; R_1 \ge 0; \\ R_1 \le \mathsf{E}_{(\underline{h},\underline{g})\in\mathcal{A}} \left[\log\left(1 + \frac{b(\underline{h},\underline{g})|h_1|^2}{\nu}\right) \right. \\ \qquad\qquad \left. - \log\left(1 + \frac{b(\underline{h},\underline{g})|g_1|^2}{\mu}\right)\right] \\ R_0 \le \mathsf{E}_{(\underline{h},\underline{g})\in\mathcal{A}} \log\left(1 + \frac{\chi(\underline{h},\underline{g})}{b(\underline{h},\underline{g})|h_1|^2+\nu}\right) \\ \qquad + \mathsf{E}_{(\underline{h},\underline{g})\in\bar{\mathcal{A}}} \log\left(1 + \frac{\chi(\underline{h},\underline{g})}{\nu}\right) \end{array} \right\} \tag{19}$$

where

$$\chi(\underline{h}, \underline{g}) = \left[\sqrt{a(\underline{h}, \underline{g})}|h_1| + \sqrt{p_2(\underline{h}, \underline{g})}|h_2|\right]^2 \tag{20}$$

and the random vector pair $(\underline{h}, \underline{g})$ has the same distribution as the marginal distribution of the process $\{(\underline{h}_i, \underline{g}_i)\}$ at a single time instant.

*Remark 1:* For each given fading state, the model given in (17) is a more general version of the Gaussian channel than the one given in (6). Hence, Corollary 1 generalizes the result on the secrecy capacity region given in Theorem 3. The proof for this more general result is omitted, because it follows the same steps and arguments for the proof of Theorem 3.

The secrecy capacity region given in Corollary 1 is established for fading processes $(\underline{h}, \underline{g})$ for which only ergodic and stationary conditions are assumed. The fading process $(\underline{h}, \underline{g})$ can be correlated across time, and is not necessarily Gaussian. This generality is due to the assumption that the CSI is known at the transmitters. In the noncoherent case, channel correlation across time may be important for designing transmission schemes, and may thus affect the secrecy capacity region, as being demonstrated in [13].

Since users are assumed to know the CSI, they can allocate their powers according to the instantaneous channel realization. The optimal power allocation that achieves the boundary of the secrecy capacity region for the fading CMAC-CM can be derived from Theorem 4 and is given in the following corollary.

*Corollary 2:* Let $p^\star(\underline{h}, \underline{g})$ be an optimal power allocation that achieves the boundary of the secrecy capacity region of the fading CMAC-CM. Then, $p^\star(\underline{h}, \underline{g})$ is given as follows:

- for $(\underline{h}, \underline{g}) \in \mathcal{A}$, if

$$\frac{\lambda_1^2|h_2|^2 \ln 2}{\lambda_1|h_2|^2+\lambda_2|h_1|^2} < \frac{\gamma_1\left(\mu|h_1|^2-\nu|g_1|^2\right)-\mu|h_1|^2}{\mu\nu}$$

then

$$\begin{aligned} a^\star(\underline{h}, \underline{g}) &= \frac{\lambda_2^2|h_1|^2\left[s_1(\underline{h}, \underline{g})-\phi(\underline{h}, \underline{g})\right]^+}{\left(\lambda_1|h_2|^2+\lambda_2|h_1|^2\right)^2} \\ b^\star(\underline{h}, \underline{g}) &= \left(\min\left[s_2(\underline{h}, \underline{g}),\, \phi(\underline{h}, \underline{g})\right]\right)^+ \\ \text{and} \quad p_2^\star(\underline{h}, \underline{g}) &= \frac{\lambda_1^2|h_2|^2\left[s_1(\underline{h}, \underline{g})-\phi(\underline{h}, \underline{g})\right]^+}{\left(\lambda_1|h_2|^2+\lambda_2|h_1|^2\right)^2} \end{aligned} \quad (21)$$

alternatively, if

$$\frac{\lambda_1^2|h_2|^2 \ln 2}{\lambda_1|h_2|^2+\lambda_2|h_1|^2} \geq \frac{\gamma_1\left(\mu|h_1|^2-\nu|g_1|^2\right)-\mu|h_1|^2}{\mu\nu}$$

then

$$\begin{aligned} a^\star(\underline{h}, \underline{g}) &= \frac{\lambda_2^2|h_1|^2\left[s_1(\underline{h}, \underline{g})\right]^+}{\left(\lambda_1|h_2|^2+\lambda_2|h_1|^2\right)^2} \\ b^\star(\underline{h}, \underline{g}) &= 0 \\ \text{and} \quad p_2^\star(\underline{h}, \underline{g}) &= \frac{\lambda_1^2|h_2|^2\left[s_1(\underline{h}, \underline{g})\right]^+}{\left(\lambda_1|h_2|^2+\lambda_2|h_1|^2\right)^2} \end{aligned} \quad (22)$$

- for $(\underline{h}, \underline{g}) \in \bar{\mathcal{A}}$,

$$\begin{aligned} a^\star(\underline{h}, \underline{g}) &= \frac{\lambda_2^2|h_1|^2\left[s_1(\underline{h}, \underline{g})\right]^+}{\left(\lambda_1|h_2|^2+\lambda_2|h_1|^2\right)^2} \\ \text{and} \quad p_2^\star(\underline{h}, \underline{g}) &= \frac{\lambda_1^2|h_2|^2\left[s_1(\underline{h}, \underline{g})\right]^+}{\left(\lambda_1|h_2|^2+\lambda_2|h_1|^2\right)^2} \end{aligned} \quad (23)$$

where $\gamma_1 \geq 0$, $s_1(\underline{h}, \underline{g})$, $s_2(\underline{h}, \underline{g})$, $\phi(\underline{h}, \underline{g})$ and $\omega(\underline{h}, \underline{g})$ are defined in (24) at the bottom of this page, and the pair $(\lambda_1, \lambda_2)$ is chosen to satisfy the power constraint

$$\mathsf{E}[a(\underline{h}, \underline{g})+b(\underline{h}, \underline{g})] \leq P_1 \quad \text{and} \quad \mathsf{E}[p_2(\underline{h}, \underline{g})] \leq P_2. \quad (25)$$

## V. Numerical Examples

In this section, we study two numerical examples to illustrate the secrecy capacity regions of the parallel Gaussian CMAC-CM and the fading CMAC-CM, respectively.

We first consider a parallel Gaussian CMAC-CM with $L = 10$. We assume that the source power constraints of users 1 and 2 are

$$P_1 = 12\,\text{dB} \quad \text{and} \quad P_2 = 10\,\text{dB}$$

and the noise variances at the receivers of the destination and of user 2 are given by

$$\begin{aligned} \underline{\nu} &= [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] \\ \text{and} \quad \underline{\mu} &= [5, 3, 4, 9, 1, 10, 8, 7, 2, 6]. \end{aligned}$$

The solid curve in Fig. 3 illustrates the boundary of the secrecy capacity region for this channel. We note that achieving this secrecy capacity region requires that users 1 and 2 perform coherent combining (i.e., use correlated inputs) for transmitting the common message. However, coherent combining may not always be possible in practice due to the complexity in code design. For comparison, we also consider the asynchronous case, in which users 1 and 2 send the common message $W_0$ without coherent combining, i.e., inputs $X_1$ and $X_2$ are independently chosen. In this case, the secrecy rate region is given by

$$\mathcal{R}_s^{[\mathrm{G}]} = \bigcup_{\underline{p} \in \mathcal{P}} \left\{ \begin{array}{l} (R_0,\, R_1): \\ R_0 \geq 0,\ R_1 \geq 0; \\ R_1 \leq \sum_{j \in \mathcal{A}} \left[\frac{1}{2}\log\left(1+\frac{b_j}{\nu_j}\right)\right. \\ \qquad \left. -\ \frac{1}{2}\log\left(1+\frac{b_j}{\mu_j}\right)\right] \\ R_0 \leq \sum_{j \in \mathcal{A}} \frac{1}{2}\log\left(1+\frac{a_j+p_{2,j}}{b_j+\nu_j}\right) \\ \qquad +\sum_{j \in \bar{\mathcal{A}}} \frac{1}{2}\log\left(1+\frac{a_j+p_{2,j}}{\nu_j}\right) \end{array} \right\} \quad (26)$$

where $\underline{p}$ is the power allocation vector, which consists of $(a_j, b_j, p_{2,j})$ for $j \in \mathcal{A}$ and $(a_j, 0, p_{2,j})$ for $j \in \bar{\mathcal{A}}$ as components, and the set $\mathcal{P}$ includes all power allocation vectors $\underline{p}$ that satisfy the power constraint (16). The boundary of the region $\mathcal{R}_s^{[\mathrm{G}]}$ is also plotted in Fig. 3 as a dashed curve. The boundary points of this region are obtained by finding the optimal power allocations via techniques similar to that for the synchronous case. We observe that the synchronous transmission mode

$$\begin{aligned} s_1(\underline{h}, \underline{g}) &= \frac{\lambda_1|h_2|^2+\lambda_2|h_1|^2}{\lambda_1\lambda_2 \ln 2} - \nu \\ s_2(\underline{h}, \underline{g}) &= \frac{1}{2}\left[\sqrt{\left(\frac{\mu}{|g_1|^2}-\frac{\nu}{|h_1|^2}\right)\left(\frac{\mu}{|g_1|^2}-\frac{\nu}{|h_1|^2}+\frac{2\gamma_1}{\lambda_1 \ln 2}\right)} - \left(\frac{\mu}{|g_1|^2}+\frac{\nu}{|h_1|^2}\right)\right] \\ \phi(\underline{h}, \underline{g}) &= -\frac{1}{2}\left(\frac{\mu}{|g_1|^2}+\frac{\nu}{|h_1|^2}+\frac{1}{\omega(\underline{h}, \underline{g})}\right) \\ &\quad + \frac{1}{2}\sqrt{\left(\frac{\mu}{|g_1|^2}+\frac{\nu}{|h_1|^2}+\frac{1}{\omega(\underline{h}, \underline{g})}\right)^2 - 4\left[\frac{\mu}{|g_1|^2}\frac{\nu}{|h_1|^2} - \frac{\gamma_1\left(\frac{\mu}{|g_1|^2}-\frac{\nu}{|h_1|^2}\right)-\frac{\mu}{|g_1|^2}}{\omega(\underline{h}, \underline{g})}\right]} \\ \text{and} \quad \omega(\underline{h}, \underline{g}) &= (\ln 2)\frac{\lambda_1^2|h_2|^2}{\lambda_1|h_2|^2+\lambda_2|h_1|^2} \end{aligned} \quad (24)$$

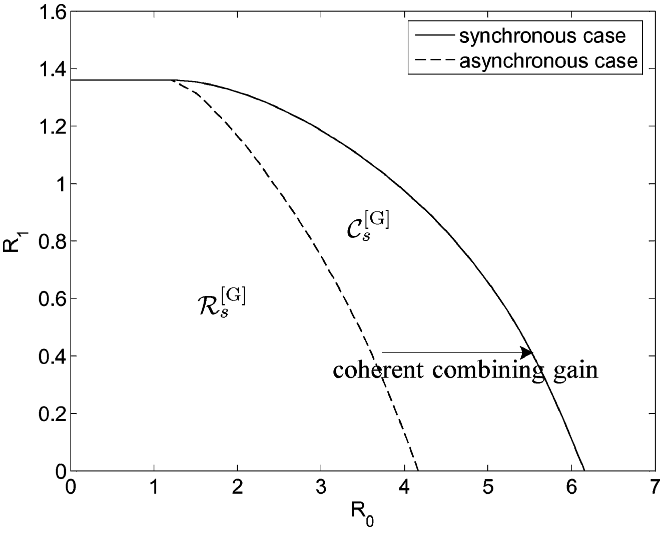


Fig. 3. Secrecy capacity region versus asynchronous secrecy rate region for the example $L = 10$ parallel Gaussian CMAC-CM.

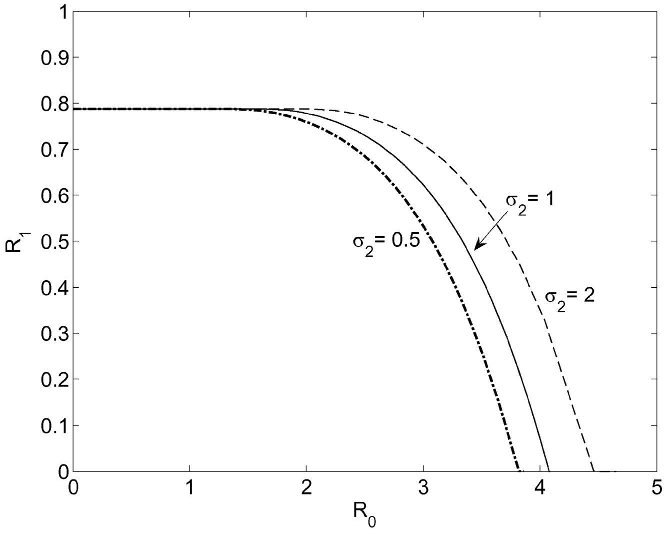


Fig. 5. Secrecy capacity regions for the example fading CMAC-CMs ($P_1 = P_2 = 10$ dB, $\nu = \mu = 2$, and $\sigma_1 = \sigma_3 = 1$).

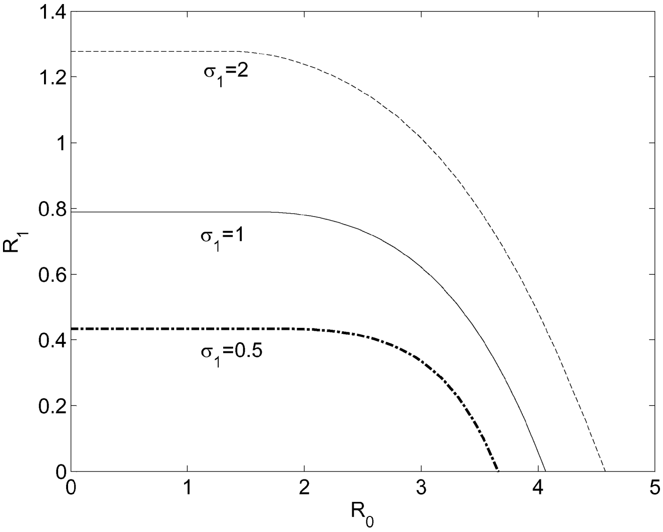


Fig. 4. Secrecy capacity regions for the example fading CMAC-CMs ($P_1 = P_2 = 10$ dB, $\nu = \mu = 2$, and $\sigma_2 = \sigma_3 = 1$).

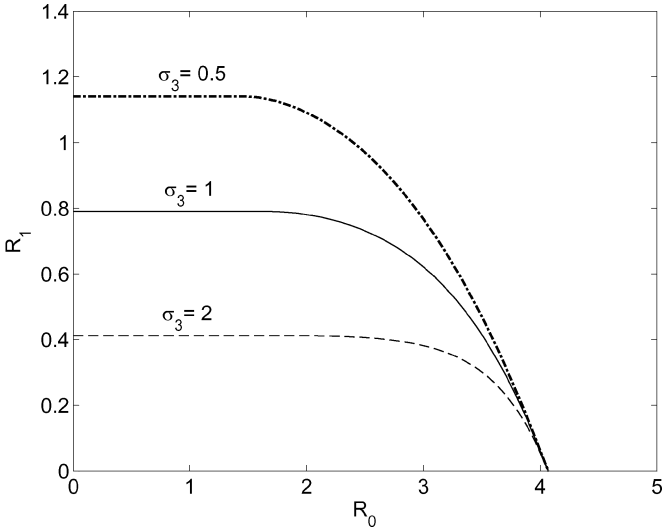


Fig. 6. Secrecy capacity regions for the example fading CMAC-CMs ($P_1 = P_2 = 10$ dB, $\nu = \mu = 2$, and $\sigma_1 = \sigma_2 = 1$).

significantly increases the rate $R_0$ of the common message since coherent combining detection can be employed at the destination.

Next, we consider the Rayleigh-fading CMAC-CM, where $h_1$, $h_2$ and $g_1$ are zero-mean proper complex Gaussian random variables. Hence, $|h_1|^2$, $|h_2|^2$ and $|g_1|^2$ are exponentially distributed with means $\sigma_1$, $\sigma_2$ and $\sigma_3$. We assume that the power constraints of users 1 and 2 are $P_1 = P_2 = 10$ dB, and the noise variances at the receivers of the destination and of user 2 are $\nu = \mu = 2$. In Fig. 4, we plot the boundaries of the secrecy capacity regions corresponding to $\sigma_1 = 0.5,\ 1,\ 2$ and fixed $\sigma_2 = \sigma_3 = 1$. It can been seen that as $\sigma_1$ increases, both the secrecy rate $R_1$ of the confidential message $W_1$ and the rate $R_0$ of the common message $W_0$ improve. This is because larger $\sigma_1$ implies a better channel from user 1 to the destination. In Fig. 5, we plot the boundaries of the secrecy capacity regions corresponding to $\sigma_2 = 0.5,\ 1,\ 2$ and fixed $\sigma_1 = \sigma_3 = 1$. It can been seen that as $\sigma_2$ increases, only the rate $R_0$ of the common message $W_0$ improves. In Fig. 6, we plot the boundaries of the secrecy capacity regions corresponding to $\sigma_3 = 0.5,\ 1,\ 2$ and fixed $\sigma_1 = \sigma_2 = 1$. It can been seen that as $\sigma_3$ decreases, only the rate $R_1$ of the confidential message $W_1$ improves.

## VI. Conclusion

We have established the secrecy capacity region of the parallel CMAC-CM, in which it is seen that having independent inputs to each subchannel is optimal. From this result, we have derived the secrecy capacity region for the parallel Gaussian CMAC-CM and the ergodic secrecy capacity region for the fading CMAC-CM. We have illustrated that, when both users know the CSI, they can dynamically adapt their transmission powers with the channel realization to achieve the optimal performance.

This work suggests several interesting directions for further research. For example, it is of interest to study the block-ergodic fading case similar to [13] and to analyze achievable rates that do not require the CSI about the eavesdropper's channel at the transmitter (user 1). Another interesting extension is to study the MIMO case similar to [21]. It may also be possible to strengthen our results for strong secrecy, for which the techniques in [22] may be relevant. We note that the studies in [23] and [24] have established the benefits of fading for secrecy in the context of practical secrecy code design. The results derived here are mostly of theoretical interest at this point, and an interesting future topic is to address secure code design for the MAC.

## Appendix

### *A) Proof of Theorem 1:*

*Achievability:* The achievability follows from [4, Corollary 3] by setting

$$\begin{aligned} Q &:= (Q_1, \ldots, Q_L) \\ U &:= (U_1, \ldots, U_L) \\ X_1 &:= (X_{1,1}, \ldots, X_{1,L}) \\ X_2 &:= (X_{2,1}, \ldots, X_{2,L}) \\ Y &:= (Y_1, \ldots, Y_L) \\ \text{and} \quad Y_2 &:= (Y_{2,1}, \ldots, Y_{2,L}) \end{aligned} \tag{27}$$

with $Q$, $U$, $X_1$, and $X_2$ having independent components. Furthermore, we choose the components of these random vectors to satisfy the condition, for $j = 1, \ldots, L$,

$$\begin{aligned} &p(q_j, u_j, x_{1,j}, x_{2,j}, y_j, y_{2,j}) \\ &\quad = p(q_j, x_{2,j}) p(u_j|q_j) p(x_{1,j}|u_j) p(y_j, y_{2,j}|x_{1,j}, x_{2,j}). \end{aligned}$$

Using the above definition, we have the following achievable region:

$$\mathcal{R}_s^{[\mathrm{P}]} := \bigcup_{\substack{\prod_j p(q_j,x_{2,j})p(u_j|q_j)p(x_{1,j}|u_j) \\ p(y_j,y_{2,j}|x_{1,j},x_{2,j})}} \left\{ \begin{array}{l} (R_0, R_1): \\ R_0 \geq 0,\ R_1 \geq 0; \\ R_1 \leq \sum_{j=1}^L [I(U_j; Y_j | X_{2,j}, Q_j) - \\ \quad I(U_j; Y_{2,j} | X_{2,j}, Q_j)] \\ R_0 + R_1 \leq \sum_{j=1}^L [I(U_j, X_{2,j}, Q_j; Y_j) - \\ \quad I(U_j; Y_{2,j} | X_{2,j}, Q_j)] \end{array} \right\}. \tag{28}$$

Note that

$$\begin{aligned} &[I(U_j; Y_j | X_{2,j}, Q_j) - I(U_j; Y_{2,j} | X_{2,j}, Q_j)] \\ &+ I(X_{2,j}, Q_j; Y_j) \\ &\quad = I(U_j, X_{2,j}, Q_j; Y_j) - I(U_j; Y_{2,j} | X_{2,j}, Q_j) \end{aligned}$$

and hence, any rate pair $(R_0, R_1) \in \mathcal{C}_s^{[\mathrm{P}]}$ must also satisfy $(R_0, R_1) \in \mathcal{R}_s^{[\mathrm{P}]}$. This implies that the secrecy rate region $\mathcal{C}_s^{[\mathrm{P}]}$ is achievable.

*Converse:* By Fano's inequality [18, Chapter 2.11], we have

$$H\left(W_0, W_1 | Y_{[1,L]}^n\right) \leq n(R_0, +R_1)\epsilon + 1 := n\delta \tag{29}$$

where $\delta \to 0$ if $\epsilon \to 0$. On the other hand, the information theoretic secrecy implies that

$$H(W_1) \leq H\left(W_1 | Y_{2,[1,L]}^n, X_{2,[1,L]}^n, W_0\right) + n\epsilon. \tag{30}$$

Now, we consider the upper bound on the secrecy rate $R_1$ as

$$\begin{aligned} nR_1 &= H(W_1) \\ &\leq H\left(W_1 | Y_{2,[1,L]}^n, X_{2,[1,L]}^n, W_0\right) + n\epsilon \qquad (31) \\ &\leq H\left(W_1 | Y_{2,[1,L]}^n, X_{2,[1,L]}^n, W_0\right) \\ &\quad - H\left(W_1 | Y_{[1,L]}^n, X_{2,[1,L]}^n, W_0\right) + n(\epsilon + \delta) \qquad (32) \\ &= I\left(W_1; Y_{[1,L]}^n | X_{2,[1,L]}^n, W_0\right) \\ &\quad - I\left(W_1; Y_{2,[1,L]}^n | X_{2,[1,L]}^n, W_0\right) + n(\epsilon + \delta) \\ &= \sum_{j=1}^L \left[ I\left(W_1; Y_j^n | Y_{[1,j-1]}^n, X_{2,[1,L]}^n, W_0\right) \right. \\ &\quad \left. - I\left(W_1; Y_{2,j}^n | Y_{2,[j+1,L]}^n, X_{2,[1,L]}^n, W_0\right) \right] + n(\epsilon + \delta) \\ &\qquad (33) \end{aligned}$$

where (31) follows from the secrecy constraint (30), (32) follows from Fano's inequality (29), and (33) follows from the chain rule of mutual information [18, Chapter 2.5]. We define

$$M_j := \left(Y_{[1,j-1]}^n, Y_{2,[j+1,L]}^n, X_{2,[1,L]}^n, W_0\right). \tag{34}$$

By applying [2, Lemma 7], we can rewrite (33) as follows:

$$\begin{aligned} nR_1 &\leq \sum_{j=1}^L \left[ I\left(W_1; Y_j^n | M_j\right) - I\left(W_1; Y_{2,j}^n | M_j\right) \right] + n(\epsilon + \delta) \\ &= \sum_{j=1}^L \sum_{i=1}^n \left[ I\left(W_1; Y_{j,i} | Y_j^{i-1}, X_{2,j,i}, M_j\right) \right. \\ &\quad \left. - I\left(W_1; Y_{2,j,i} | Y_{2,j,i+1}^n, X_{2,j,i}, M_j\right) \right] + n(\epsilon + \delta) \quad (35) \end{aligned}$$

where (35) follows from the chain rule of mutual information [18, Chapter 2.5] and $X_{2,j,i}$ appears in the conditioning because it is a component in $M_j$. Let

$$Q_{j,i} := \left(Y_j^{i-1}, Y_{2,j,i+1}^n, M_j\right). \tag{36}$$

We notice that the definitions of (34) and (36) imply the Markov chain relationship

$$X_{2,j,i} \to Q_{j,i} \to (W_1, Q_{j,i}) \to X_{1,j,i}. \tag{37}$$

Applying [2, Lemma 7] again, we obtain

$$\begin{aligned} nR_1 &\leq \sum_{j=1}^L \sum_{i=1}^n \left[ I\left(W_1; Y_{j,i} | X_{2,j,i}, Q_{j,i}\right) \right. \\ &\quad \left. - I\left(W_1; Y_{2,j,i} | X_{2,j,i}, Q_{j,i}\right) \right] + n(\epsilon + \delta). \end{aligned} \tag{38}$$

We also can write

$$\begin{aligned}
nR_0 &= H(W_0)\\
&\leq I(W_0; Y^n_{[1,L]}) + n\delta && (39)\\
&= \sum_{j=1}^{L}\sum_{i=1}^{n} I(W_0; Y_{j,i}|Y_j^{i-1}, Y^n_{[1,j-1]}) + n\delta && (40)\\
&\leq \sum_{j=1}^{L}\sum_{i=1}^{n}\\
&\quad I(W_0, Y_j^{i-1}, Y^n_{[1,j-1]}, Y^n_{2,j,i+1}, Y^n_{2,[j+1,L]}, X^n_{2,[1,L]}; Y_{j,i})\\
&\quad + n\delta\\
&= \sum_{j=1}^{L}\sum_{i=1}^{n} I(Q_{j,i}, X_{2,j,i}; Y_{j,i}) + n\delta && (41)
\end{aligned}$$

where (39) follows from Fano's inequality (29), (40) follows from the chain rule, and (41) follows from the definition of $Q_{j,i}$ in (36).

We introduce a time-sharing random variable $T$[18, Chapter 14.3] that is independent of all other random variables in the model, and uniformly distributed over $\{1, \ldots, n\}$. Define $Q_j = (T, Q_{j,T})$, $U_j = (Q_j, W_1)$, $X_{1,j} = X_{1,j,T}$, $X_{2,j} = X_{2,j,T}$, $Y_{2,j} = Y_{2,j,T}$, and $Y_j = Y_{j,T}$ for $j = 1, \ldots, L$. Note that $(Q_j, X_{1,j}, X_{2,j}, Y_j, Y_{2,j})$ satisfies the following Markov chain relationship, for $j = 1, \ldots, L$,

$$Q_j \to U_j \to (X_{1,j}, X_{2,j}) \to (Y_j, Y_{2,j}).$$

Using the above definition, (38) and (41) become

$$\begin{aligned}
R_1 &\leq \sum_{j=1}^{L} [I\left(U_j; Y_j|X_{2,j}, Q_j\right)\\
&\quad - I\left(U_j; Y_{2,j}|X_{2,j}, Q_j\right)] + (\epsilon + \delta).\\
\text{and} \qquad R_0 &\leq \sum_{j=1}^{L} I\left(X_{2,j}, Q_j; Y_j\right) + \delta.
\end{aligned}$$

*B) Proof of Theorem 2:* The achievability follows from Theorem 1 by setting

$$\begin{aligned}
U_j &= X_{1,j}, && \text{for} \quad j \in \mathcal{A}\\
\text{and} \qquad Q_j &= U_j = X_{1,j}, && \text{for} \quad j \in \bar{\mathcal{A}}.
\end{aligned}$$

To show the converse, we first consider the upper bound on $R_0$. By using (1) in Theorem 1, we have

$$\begin{aligned}
R_0 &\leq \sum_{j=1}^{L} I(Q_j, X_{2,j}; Y_j)\\
&= \sum_{j\in\mathcal{A}} I(Q_j, X_{2,j}; Y_j) + \sum_{j\in\bar{\mathcal{A}}} I(Q_j, X_{2,j}; Y_j)\\
&\leq \sum_{j\in\mathcal{A}} I(Q_j, X_{2,j}; Y_j) + \sum_{j\in\bar{\mathcal{A}}} I(X_{1,j}, X_{2,j}; Y_j) && (42)
\end{aligned}$$

where (42) follows from the Markov chain relationship:

$$Q_j \to (X_{1,j}, X_{2,j}) \to Y_j.$$

Now, we consider the upper bound on $R_1$. By applying Theorem 1, we obtain

$$\begin{aligned}
R_1 &\leq \sum_{j=1}^{L} [I(U_j; Y_j|X_{2,j}, Q_j) - I(U_j; Y_{2,j}|X_{2,j}, Q_j)]\\
&= \sum_{j\in\mathcal{A}} [I(U_j; Y_j|X_{2,j}, Q_j) - I(U_j; Y_{2,j}|X_{2,j}, Q_j)]\\
&\quad + \sum_{j\in\bar{\mathcal{A}}} [I(U_j; Y_j|X_{2,j}, Q_j) - I(U_j; Y_{2,j}|X_{2,j}, Q_j)]. && (43)
\end{aligned}$$

For $j \in \bar{\mathcal{A}}$, the subchannel satisfies

$$p(y_j, y_{2,j}|x_{1,j}, x_{2,j}) = p(y_{2,j}|x_{1,j}, x_{2,j}) p(y_j|y_{2,j}, x_{2,j}).$$

This implies that

$$\begin{aligned}
I(U_j; Y_j|X_{2,j}, Q_j) &- I(U_j; Y_{2,j}|X_{2,j}, Q_j)\\
&\leq I(U_j; Y_j|X_{2,j}, Q_j, Y_{2,j})\\
&\leq I(Q_j, U_j; Y_j|X_{2,j}, Y_{2,j})\\
&= 0 \quad \text{for } j \in \bar{\mathcal{A}} && (44)
\end{aligned}$$

where the last equality of (44) follows from the Markov chain relationship, for $j \in \bar{\mathcal{A}}$

$$(Q_j, U_j) \to (X_{1,j}, X_{2,j}) \to (Y_{2,j}, X_{2,j}) \to Y_j. \qquad (45)$$

On the other hand, for $j \in \mathcal{A}$, the subchannel satisfies

$$p(y_j, y_{2,j}|x_{1,j}, x_{2,j}) = p(y_j|x_{1,j}, x_{2,j}) p(y_{2,j}|y_j, x_{2,j}). \qquad (46)$$

Hence, we obtain, for $j \in \mathcal{A}$

$$\begin{aligned}
&I(U_j; Y_j|X_{2,j}, Q_j) - I(U_j; Y_{2,j}|X_{2,j}, Q_j)\\
&\quad = I(U_j; Y_j, Y_{2;j}|X_{2,j}, Q_j) - I(U_j; Y_{2,j}|X_{2,j}, Y_j, Q_j)\\
&\qquad - I(U_j; Y_{2,j}|X_{2,j}, Q_j)\\
&\quad = I(U_j; Y_{2;j}|X_{2,j}, Q_j) + I(U_j; Y_j|Y_{2,j}, X_{2,j}, Q_j)\\
&\qquad - I(U_j; Y_{2,j}|X_{2,j}, Q_j) && (47)\\
&\quad = I(U_j; Y_j|X_{2,j}, Q_j, Y_{2,j})\\
&\quad \leq I(U_j, X_{1,j}; Y_j|X_{2,j}, Q_j, Y_{2,j})\\
&\quad = I(X_{1,j}; Y_j|X_{2,j}, Q_j, Y_{2,j}) && (48)\\
&\quad = I(X_{1,j}; Y_j, Y_{2,j}|X_{2,j}, Q_j) - I(X_{1,j}; Y_{2,j}|X_{2,j}, Q_j) && (49)\\
&\quad = I(X_{1,j}; Y_j|X_{2,j}, Q_j) - I(X_{1,j}; Y_{2,j}|X_{2,j}, Q_j) && (50)
\end{aligned}$$

where (47) follows because $I(U_j; Y_{2,j}|X_{2,j}, Y_j, Q_j) = 0$ due to the degradedness condition, (48) follows from the Markov chain relationship

$$(Q_j, U_j, Y_{2,j}) \to (X_{1,j}, X_{2,j}) \to Y_j.$$

Equation (49) follows from the chain rule of mutual information, and (50) follows from the conditional degradedness (46). Now, substituting (44) and (50) into (43), we obtain the bound on $R_1$ given in (5). This concludes the proof of the converse.

*C) Proof of Theorem 3:* As we comment in Section III-B, we need to prove Theorem 3 only for the channel defined by (7)–(8).

*Achievability:* The achievability follows by applying Theorem 2 and choosing the input distribution as follows:

$$\begin{aligned} Q_j &= \text{constant} \\ X_{2,j} &\sim \mathcal{N}(0, p_{2,j}) \\ X'_{1,j} &\sim \mathcal{N}(0, (1-\alpha_j)p_{1,j}) \\ X'_{1,j} &\text{ is independent of } X_{2,j} \\ \text{and} \quad X_{1,j} &= \sqrt{\frac{\alpha_j p_{1,j}}{p_{2,j}}} X_{2,j} + X'_{1,j}. \end{aligned} \tag{51}$$

We note that although Theorem 2 is derived for the channel without input constraints, such a result can be easily extended to the case with input cost constraints. Moreover, by the fact $\alpha_j = 1$ for $j \in \bar{\mathcal{A}}$, we conclude that the secrecy rate region $\mathcal{C}_s^{[\mathrm{G}]}$ is achievable.

*Converse:* Here, we derive a tight upper bound on the achievable weighted sum rate

$$R_0 + \gamma_1 R_1$$

using Theorem 2 as the starting point. Since a capacity region is always convex (via a time-sharing argument), an exact characterization of all the achievable weighted sum rates for all nonnegative $\gamma_1$ provides an exact characterization of the entire secrecy capacity region. By Theorem 2, any achievable rate pair $(R_0, R_1)$ must satisfy

$$\begin{aligned} R_0 + \gamma_1 R_1 &\le \sum_{j \in \mathcal{A}} [I(Q_j, X_{2,j}; Y_j) + \gamma_1 I(X_{1,j}; Y_j | X_{2,j}, Q_j) \\ &\quad - \gamma_1 I(X_{1,j}; Y_{2,j} | X_{2,j}, Q_j)] + \sum_{j \in \bar{\mathcal{A}}} I(X_{1,j}, X_{2,j}; Y_j). \end{aligned}$$

For the subchannel $j \in \bar{\mathcal{A}}$, we are concerned only with the term

$$I(X_{1,j}, X_{2,j}; Y_j). \tag{52}$$

The maximum-entropy theorem [18] implies that (52) is maximized when $X_{1,j}$ and $X_{2,j}$ are jointly Gaussian with variances $p_{1,j}$ and $p_{2,j}$, respectively, and are aligned, i.e.,

$$X_{1,j} = \sqrt{p_{1,j}/p_{2,j}} X_{2,j}.$$

Hence, we have, for $j \in \bar{\mathcal{A}}$

$$I(X_{1,j}, X_{2,j}; Y_j) \le \frac{1}{2} \log \left( 1 + \frac{p_{1,j} + p_{2,j} + 2\sqrt{p_{1,j} p_{2,j}}}{\nu_j} \right). \tag{53}$$

For the subchannel $j \in \mathcal{A}$, we focus on the terms

$$\begin{aligned} &I(Q_j, X_{2,j}; Y_j) + \gamma_1 I(X_{1,j}; Y_j | X_{2,j}, Q_j) \\ &\quad - \gamma_1 I(X_{1,j}; Y_{2,j} | X_{2,j}, Q_j). \end{aligned}$$

Based on the channel model defined in (7)–(8), we have

$$\begin{aligned} &I(Q_j, X_{2,j}; Y_j) + \gamma_1 I(X_{1,j}; Y_j | X_{2,j}, Q_j) \\ &\quad - \gamma_1 I(X_{1,j}; Y_{2,j} | X_{2,j}, Q_j) \\ &= h(Y_j) + (\gamma_1 - 1) h(Y_j | X_{2,j}, Q_j) \\ &\quad - \gamma_1 h(Y_{2,j} | X_{2,j}, Q_j) + \frac{\gamma_1}{2} \log \frac{\mu_j}{\nu_j}. \end{aligned} \tag{54}$$

Now, we consider the following two cases.

*Case 1:* $\gamma_1 \le 1$. In this case, note that

$$\begin{aligned} h(Y_j | X_{2,j}, Q_j) &\ge h(Y_j | X_{1,j}, X_{2,j}, Q_j) \\ &= \frac{1}{2} \log 2\pi e \nu_j \\ h(Y_j | X_{2,j}, Q_j) &\ge h(Y_{2,j} | X_{1,j}, X_{2,j}, Q_j) \\ &= \frac{1}{2} \log 2\pi e \mu_j \\ \text{and} \quad h(Y_j) &\le \frac{1}{2} \log 2\pi e \left( p_{1,j} + p_{2,j} \right. \\ &\quad \left. + 2\sqrt{p_{1,j} p_{2,j}} + \nu_j \right). \end{aligned}$$

Hence, we have, for $j \in \mathcal{A}$ and $\gamma_1 \le 1$

$$\begin{aligned} &I(Q_j, X_{2,j}; Y_j) + \gamma_1 I(X_{1,j}; Y_j | X_{2,j}, Q_j) \\ &\quad - \gamma_1 I(X_{1,j}; Y_{2,j} | X_{2,j}, Q_j) \\ &\le \frac{1}{2} \log \left( 1 + \frac{p_{1,j} + p_{2,j} + 2\sqrt{p_{1,j} p_{2,j}}}{\nu_j} \right). \end{aligned} \tag{55}$$

This result implies that when the weight of the confidential-message rate is less than the weight of the common-message rate, the optimum solution is to allocate all possible power to transmit the common message.

*Case 2:* $\gamma_1 > 1$. Without loss of generality, we assume that the conditional covariance of $X_{1,j}$ given $(X_{2,j}, Q_j)$ is given by

$$\mathsf{cov}(X_{1,j} | X_{2,j}, Q_j) = (1 - \alpha_j) p_{1,j} \tag{56}$$

where $0 \le \alpha_j \le 1$. By applying the extremal inequality [19, Theorem 8], we have

$$\begin{aligned} &(\gamma_1 - 1) h(Y_j | X_{2,j}, Q_j) - \gamma_1 h(Y_{2,j} | X_{2,j}, Q_j) \\ &\le \frac{\gamma_1 - 1}{2} \log 2\pi e \left[ (1 - \alpha_j) p_{1,j} + \nu_j \right] \\ &\quad - \frac{\gamma_1}{2} \log 2\pi e \left[ (1 - \alpha_j) p_{1,j} + \mu_j \right]. \end{aligned} \tag{57}$$

Moreover, for a given $\alpha_j$

$$h(Y_j) \le \frac{1}{2} \log 2\pi e \left( p_{1,j} + p_{2,j} + 2\sqrt{\alpha_j p_{1,j} p_{2,j}} + \nu_j \right). \tag{58}$$

Substituting (57) and (58) into (54), we obtain, for $j \in \mathcal{A}$ and $\gamma_1 > 1$

$$\begin{aligned} & I(Q_j, X_{2,j}; Y_j) + \gamma_1 I(X_{1,j}; Y_j | X_{2,j}, Q_j) \\ & \quad - \gamma_1 I(X_{1,j}; Y_{2,j} | X_{2,j}, Q_j) \\ & \quad \le \max_{0 \le \alpha_j \le 1} \left[ \frac{\gamma_1 - 1}{2} \log 2\pi e \left(1 + \frac{(1-\alpha_j)p_{1,j}}{\nu_j}\right) \right. \\ & \quad - \frac{\gamma_1}{2} \log 2\pi e \left(1 + \frac{(1-\alpha_j)p_{1,j}}{\mu_j}\right) \\ & \quad \left. + \frac{1}{2} \log 2\pi e \left(1 + \frac{p_{1,j} + p_{2,j} + 2\sqrt{\alpha_j p_{1,j} p_{2,j}}}{\nu_j}\right) \right] \\ & \quad = \max_{0 \le \alpha_j \le 1} \left[ \frac{\gamma_1}{2} \log \left(1 + \frac{(1-\alpha_j)p_{1,j}}{\nu_j}\right) \right. \\ & \quad - \frac{\gamma_1}{2} \log \left(1 + \frac{(1-\alpha_j)p_{1,j}}{\mu_j}\right) \\ & \quad \left. + \frac{1}{2} \log \left(1 + \frac{\alpha_j p_{1,j} + p_{2,j} + 2\sqrt{\alpha_j p_{1,j} p_{2,j}}}{(1-\alpha_j)p_{1,j} + \nu_j}\right) \right]. \quad (59) \end{aligned}$$

Finally, combining (53), (55) and (59), we complete the converse proof.

*D) Proof of Theorem 4:* We need find the $\underline{p}^\star \in \mathcal{P}$ that maximizes

$$R_0 + \gamma_1 R_1$$

where $\gamma_1 \ge 0$. The Lagrangian is given by

$$\begin{aligned} \mathcal{L} = & \sum_{j \in \mathcal{A}} \left[ \frac{1}{2} \log \left(1 + \frac{a_j + p_{2,j} + 2\sqrt{a_j p_{2,j}}}{b_j + \nu_j}\right) \right. \\ & \left. + \frac{\gamma_1}{2} \log \left(1 + \frac{b_j}{\nu_j}\right) - \frac{\gamma_1}{2} \log \left(1 + \frac{b_j}{\mu_j}\right) \right] \\ & + \sum_{j \in \bar{\mathcal{A}}} \frac{1}{2} \log \left(1 + \frac{a_j + p_{2,j} + 2\sqrt{a_j p_{2,j}}}{\nu_j}\right) \\ & - \lambda_1 \left[ \sum_{j \in \mathcal{A}} (a_j + b_j) + \sum_{j \in \bar{\mathcal{A}}} a_j \right] - \lambda_2 \sum_{j=1}^{L} p_{2,j} \quad (60) \end{aligned}$$

where $\lambda_1$ and $\lambda_2$ are Lagrange multiplier.

For $j \in \bar{\mathcal{A}}$, $(a_j^\star, p_{2,j}^\star)$ needs to maximize the following $\mathcal{L}_j$:

$$\mathcal{L}_j = \frac{1}{2} \log \left(1 + \frac{a_j + p_{2,j} + 2\sqrt{a_j p_{2,j}}}{\nu_j}\right) - \lambda_1 a_j - \lambda_2 p_{2,j}. \quad (61)$$

Taking derivatives of the Lagrangian in (61) over $a_j$ and $p_{2,j}$, the KKT conditions can be written as follows:

$$\begin{aligned} & \frac{1}{2\ln 2} \frac{\theta_{1,j}(a_j, p_{2,j})}{\sqrt{a_j}} = \lambda_1 \\ \text{and} \quad & \frac{1}{2\ln 2} \frac{\theta_{1,j}(a_j, p_{2,j})}{\sqrt{p_{2,j}}} = \lambda_2 \quad (62) \end{aligned}$$

where

$$\theta_{1,j}(a_j, p_{2,j}) = \frac{\sqrt{a_j} + \sqrt{p_{2,j}}}{\nu_j + a_j + p_{2,j} + 2\sqrt{a_j p_{2,j}}}.$$

This implies that the pair $(a_j^\star, p_{2,j}^\star)$ that optimizes $\mathcal{L}_j$ must satisfy

$$p_{2,j}^\star = \left(\frac{\lambda_1}{\lambda_2}\right)^2 a_j^\star. \quad (63)$$

Let us define

$$\beta = \lambda_1 / \lambda_2.$$

On substituting (63) into (61), we obtain that

$$\begin{aligned} \mathcal{L}_j & = \frac{1}{2} \log \left[1 + \frac{a_j(1+\beta)^2}{\nu_j}\right] - \lambda_1 a_j (1+\beta) \\ & = \int_0^{a_j(1+\beta)^2} t_{1,j}(s)\, ds \quad (64) \end{aligned}$$

where

$$t_{1,j}(s) = \frac{1}{(2\ln 2)} \frac{1}{(\nu_j + s)} - \frac{\lambda_1}{1+\beta}. \quad (65)$$

We define $s_{1,j}$ to be the root of the equation $t_{1,j}(s) = 0$, i.e.,

$$\begin{aligned} s_{1,j} & = \frac{1+\beta}{2\lambda_1 \ln 2} - \nu_j \\ & = \frac{\lambda_1 + \lambda_2}{2\lambda_1 \lambda_2 \ln 2} - \nu_j. \quad (66) \end{aligned}$$

Hence, we obtain, for $j \in \bar{\mathcal{A}}$

$$\begin{aligned} a_j^\star & = \frac{1}{(1+\beta)^2} (s_{1,j})^+ \\ & = \frac{\lambda_2^2}{(\lambda_1 + \lambda_2)^2} \left(\frac{\lambda_1 + \lambda_2}{2\lambda_1 \lambda_2 \ln 2} - \nu_j\right)^+ \end{aligned}$$

and

$$\begin{aligned} p_{2,j}^\star & = \beta^2 a_j^\star \\ & = \frac{\lambda_1^2}{(\lambda_1 + \lambda_2)^2} \left(\frac{\lambda_1 + \lambda_2}{2\lambda_1 \lambda_2 \ln 2} - \nu_j\right)^+. \end{aligned}$$

For $j \in \mathcal{A}$, $(a_j^\star, b_j^\star, p_{2,j}^\star)$ needs to maximize the following $\mathcal{L}_j$:

$$\begin{aligned} \mathcal{L}_j = & \frac{1}{2} \log \left(1 + \frac{a_j + p_{2,j} + 2\sqrt{a_j p_{2,j}}}{b_j + \nu_j}\right) \\ & + \frac{\gamma_1}{2} \log \left(1 + \frac{b_j}{\nu_j}\right) - \frac{\gamma_1}{2} \log \left(1 + \frac{b_j}{\mu_j}\right) \\ & - \lambda_1 (a_j + b_j) - \lambda_2 p_{2,j}. \quad (67) \end{aligned}$$

Taking derivatives of the Lagrangian in (67) over $a_j$ and $p_{2,j}$, the KKT conditions can be written as follows:

$$\begin{aligned} & \frac{1}{2\ln 2} \frac{\theta_{2,j}(a_j, b_j, p_{2,j})}{\sqrt{a_j}} = \lambda_1 \\ \text{and} \quad & \frac{1}{2\ln 2} \frac{\theta_{2,j}(a_j, b_j, p_{2,j})}{\sqrt{p_{2,j}}} = \lambda_2 \quad (68) \end{aligned}$$

where

$$\theta_{2,j}(a_j,b_j,p_{2,j}) = \frac{\sqrt{a_j}+\sqrt{p_{2,j}}}{\nu_j+a_j+b_j+p_{2,j}+2\sqrt{a_j p_{2,j}}}.$$

This implies that the pair $(a_j^\star, p_{2,j}^\star)$ that optimizes $\mathcal{L}_j$ must satisfy

$$p_{2,j}^\star = \left(\frac{\lambda_1}{\lambda_2}\right)^2 a_j^\star = \beta^2 a_j^\star. \tag{69}$$

On substituting (69) into (67), we obtain that

$$\begin{aligned}\mathcal{L}_j =& \frac{1}{2}\log\left[1+\frac{a_j(1+\beta)^2}{b_j+\nu_j}\right] + \frac{\gamma_1}{2}\log\left(1+\frac{b_j}{\nu_j}\right)\\ &- \frac{\gamma_1}{2}\log\left(1+\frac{b_j}{\mu_j}\right) - \lambda_1[a_j(1+\beta)+b_j]\\ =& \int_{b_j}^{b_j+a_j(1+\beta)^2} t_{1,j}(s)\,ds + \int_0^{b_j} t_{2,j}(s)\,ds\\ \le& \int_0^\infty \left(\max\{t_{1,j}(s), t_{2,j}(s)\}\right)^+ \,ds \end{aligned} \tag{70}$$

where $t_{1,j}(s)$ is defined in (65) and

$$t_{2,j}(s) = \frac{\gamma_1}{2\ln 2}\left(\frac{1}{\nu_j+s} - \frac{1}{\mu_j+s}\right) - \lambda_1. \tag{71}$$

Next, we will derive $(a_j^\star, b_j^\star, p_{2,j}^\star)$ that achieves the upper bound on $\mathcal{L}_j$ in (70). We consider the point of intersection between $t_{1,j}(s)$ and $t_{2,j}(s)$. By using the definitions of $t_{1,j}(s)$ in (65) and $t_{2,j}(s)$ in (71), the point of intersection must satisfy

$$\frac{1}{2\ln 2}\frac{1}{\nu_j+s} - \frac{\lambda_1}{1+\beta} = \frac{\gamma_1}{2\ln 2}\left(\frac{1}{\nu_j+s} - \frac{1}{\mu_j+s}\right) - \lambda_1$$

i.e.,

$$s^2 + \left(\mu_j+\nu_j+\frac{1}{\omega}\right)s + \left[\mu_j\nu_j - \frac{\gamma_1(\mu_j-\nu_j)-\mu_j}{\omega}\right] = 0 \tag{72}$$

where

$$\begin{aligned}\omega &= (2\lambda_1\ln 2)\frac{\beta}{1+\beta}\\ &= (2\ln 2)\frac{\lambda_1^2}{\lambda_1+\lambda_2}.\end{aligned}$$

In the following, we consider two cases based on the relationship between $\omega$ and $(\gamma_1(\mu_j-\nu_j)-\mu_j)/(\mu_j\nu_j)$.

*1) $\omega \ge \frac{\gamma_1(\mu_j-\nu_j)-\mu_j}{\mu_j\nu_j}$:* In this case, (72) implies that the point of intersection between $t_{1,j}(s)$ and $t_{2,j}(s)$ is either zero or negative. Moreover, it is easy to see, for $s \ge 0$

$$\begin{aligned}&t_{1,j}(s) - t_{2,j}(s)\\ &\quad= \frac{(\nu_j+s)(\mu_j+s)\omega - [\gamma_1(\mu_j-\nu_j)-(\mu_j+s)]}{(2\ln 2)(\nu_j+s)(\mu_j+s)} \ge 0.\end{aligned}$$

Hence, the upper bound on $\mathcal{L}_j$ in (70) is achieved by $b_j^\star = 0$

$$\begin{aligned}a_j^\star &= \frac{1}{(1+\beta)^2}(s_{1,j})^+\\ &= \frac{\lambda_2^2}{(\lambda_1+\lambda_2)^2}\left(\frac{\lambda_1+\lambda_2}{2\lambda_1\lambda_2\ln 2} - \nu_j\right)^+\end{aligned}$$

and

$$\begin{aligned}p_{2,j}^\star &= \beta^2 a_j^\star\\ &= \frac{\lambda_1^2}{(\lambda_1+\lambda_2)^2}\left(\frac{\lambda_1+\lambda_2}{2\lambda_1\lambda_2\ln 2} - \nu_j\right)^+\end{aligned}$$

where $s_{1,j}$ is defined in (66).

*2) $\omega < \frac{\gamma_1(\mu_j-\nu_j)-\mu_j}{\mu_j\nu_j}$:* In this case, (72) implies that, for $s > 0$, $t_{1,j}(s)$ and $t_{2,j}(s)$ intersect only once at

$$\begin{aligned}\phi_j =& -\frac{1}{2}\left(\mu_j+\nu_j+\frac{1}{\omega}\right)\\ &+ \frac{1}{2}\sqrt{\left(\mu_j+\nu_j+\frac{1}{\omega}\right)^2 - 4\left[\mu_j\nu_j - \frac{\gamma_1(\mu_j-\nu_j)-\mu_j}{\omega}\right]}.\end{aligned}$$

Moreover, it is easy to see that $t_{1,j}(0) < t_{2,j}(0)$. Hence, we have

$$\begin{aligned} & t_{1,j}(s) < t_{2,j}(s), && \text{for } 0 \le s < \phi_j\\ \text{and} \quad & t_{1,j}(s) \ge t_{2,j}(s), && \text{for } s \ge \phi_j.\end{aligned}$$

Let $s_{2,j}$ denote the largest root of $t_{2,j}(s) = 0$, i.e.,

$$s_{2,j} = \frac{1}{2}\left[\sqrt{(\mu_j-\nu_j)\left(\mu_j-\nu_j+\frac{2\gamma_1}{\lambda_1\ln 2}\right)} - (\mu_j+\nu_j)\right].$$

The optimal $(a_j^\star, b_j^\star, p_{2,j}^\star)$ depends on the values $t_{2,j}(0)$, $s_{1,j}$ and $\phi_j$, and falls into the following three possibilities.

(2.a) If $t_{2,j}(0) < 0$, then both $t_{1,j}(s)$ and $t_{2,j}(s)$ are negative for $s \ge 0$ (since both $t_{1,j}(s)$ and $t_{2,j}(s)$ are decreasing functions for $s \ge 0$). Then, the upper bound on $\mathcal{L}_j$ in (70) is achieved by $b_j^\star = 0$, $a_j^\star = 0$ and $p_{2,j}^\star = 0$.

(2.b) If $t_{2,j}(0) \ge 0$ and $s_{1,j} < \phi_j$, then the upper bound on $\mathcal{L}_j$ in (70) is achieved by $b_j^\star = s_{2,j}$, $a_j^\star = 0$ and $p_{2,j}^\star = 0$.

(2.c) If $t_{2,j}(0) \ge 0$ and $s_{1,j} \ge \phi_j$, then the upper bound on $\mathcal{L}_j$ in (70) is achieved by $b_j^\star = \phi_j$

$$\begin{aligned} a_j^\star &= \frac{1}{(1+\beta)^2}\left(s_{1,j}-\phi_j\right)\\ \text{and} \quad p_{2,j}^\star &= \frac{\beta^2}{(1+\beta)^2}\left(s_{1,j}-\phi_j\right).\end{aligned}$$

Combing the cases (2.a), (2.b), and (2.c), we obtain

$$\begin{aligned} a_j^\star &= \frac{\lambda_2^2}{(\lambda_1+\lambda_2)^2}\left(s_{1,j}-\phi_j\right)^+\\ b_j^\star &= \left(\min\left[\phi_j,\, s_{2,j}\right]\right)^+\\ \text{and} \quad p_{2,j}^\star &= \frac{\lambda_1^2}{(\lambda_1+\lambda_2)^2}\left(s_{1,j}-\phi_j\right)^+.\end{aligned}$$

Finally, the Lagrange parameters $\lambda_1 \ge 0$ and $\lambda_2 \ge 0$ are chosen to satisfy the power constraint (16).

## References

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

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

[3] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Foundations and Trends in Communications and Information Theory,* vol. 5, no. 4-5, pp. 355–580, 2008.

[4] 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.

[5] 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*, Seattle, WA, Jul. 2006, pp. 957–961.

[6] E. Tekin and A. Yener, "The Gaussian multiple access wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5747–5755, Dec. 2008.

[7] E. Tekin and A. Yener, "The general Gaussian multiple access and two-way wire-tap channels: achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008, Special Issue on Information Theoretic Security.

[8] E. Ekrem and S. Ulukus, "Effects of cooperation on the secrecy of multiple access channels with generalized feedback," in *Proc. Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 2008, pp. 791–796.

[9] E. Ekrem and S. Ulukus, "On the secrecy of multiple access wiretap channel," in *Proc. 46th Annu. Allerton Conf. Communications, Control and Computing*, Monticello, IL, Sep. 2008, pp. 1014–1021.

[10] R. Bassily and S. Ulukus, "A new achievable ergodic secrecy rate region for the fading multiple access wiretap channel," in *Proc. 47th Annu. Allerton Conf. Communications, Control and Computing*, Monticello, IL, Sep. 2009, pp. 819–826.

[11] O. Simeone and A. Yener, "The cognitive multiple access wire-tap channel," in *Proc. Conf. Information Sciences and Systems*, Baltimore, MD, Mar. 2009, pp. 158–163.

[12] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008, Special Issue on Information Theoretic Security.

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

[14] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "On the throughput of secure hybrid-ARQ protocols for Gaussian block-fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1575–1591, Apr. 2009.

[15] E. Ekrem and S. Ulukus, "Secrecy in cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 137–155, Jan. 2011.

[16] X. He and A. Yener, "Two-hop secure communication using an untrusted relay," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, Article ID 305146, 13 pages, 2009, Special Issue on Wireless Physical Layer Security.

[17] X. He and A. Yener, "Cooperation with an untrusted relay: A secrecy perspective," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3807–3827, Aug. 2010.

[18] T. M. Cover and J. A. Thomas*, Elements of Information Theory*. Hoboken, NJ: Wiley, 1991.

[19] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1839–1851, May 2007.

[20] D. N. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Proc. IEEE Int. Symp. Information Theory*, Ulm, Germany, Jun. 1997, p. 27.

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

[22] U. M. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *Advances in Cryptology-Eurocrypt 2000, Lecture Notes in Computer Science.* Bruges, Belgium, 2000, pp. 351–368.

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

[24] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. B. Mandayam, "Information-theoretically secret key generation for fading wireless channels," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 240–254, Jun. 2010.

**Ruoheng Liu** (S'02–M'07) received the B.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 1998 and the M.S. degree in signal and information engineering from Chinese Academy of Sciences, Beijing, in 2001. He received Ph.D. degrees in electrical engineering from Rutgers University, New Brunswick, NJ, in 2007. During his Ph.D. studies, he was with the Wireless Information Network Laboratory (WINLAB) at Rutgers University. From 2007 to 2009, he worked as a Postdoctoral Research Associate at the Department of Electrical Engineering, Princeton University, NJ. Since February 2010, he has been with Alcatel-Lucent, Murray Hill, NJ, where he is a wireless system engineer. His research interests are in the areas of wireless communications, coding theory, and information theory.

**Yingbin Liang** (S'01–M'05) received the Ph.D. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign in 2005. In 2005–2007, she was working as a postdoctoral research associate at Princeton University. In 2008–2009, she was an assistant professor at the Department of Electrical Engineering at the University of Hawaii. Since December 2009, she has been an assistant professor at the Department of Electrical Engineering and Computer Science at the Syracuse University. Dr. Liang's research interests include communications, wireless networks, information theory, and machine learning.

Dr. Liang was a Vodafone Fellow at the University of Illinois at Urbana-Champaign during 2003–2005, and received the Vodafone-U.S. Foundation Fellows Initiative Research Merit Award in 2005. She also received the M. E. Van Valkenburg Graduate Research Award from the ECE department, University of Illinois at Urbana-Champaign, in 2005. In 2009, she received the National Science Foundation CAREER Award, and the State of Hawaii Governor Innovation Award.

**H. Vincent Poor** (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is the Dean of Engineering and Applied Science, and the Michael Henry Strater University Professor of Electrical Engineering. Dr. Poor's research interests are in the areas of stochastic analysis, statistical signal processing and information theory, and their applications in wireless networks and related fields. Among his publications in these areas are *Quickest Detection* (Cambridge University Press, 2009), co-authored with Olympia Hadjiliadis, and *Information Theoretic Security* (Now Publishers, 2009), co-authored with Yingbin Liang and Shlomo Shamai.

Dr. Poor is a member of the National Academy of Engineering and of the National Academy of Sciences, a Fellow of the American Academy of Arts and Sciences, and an International Fellow of the Royal Academy of Engineering (U. K.). He is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, in 2004–07 as the Editor-in-Chief of these *Transactions*, and in 2009 as General Co-chair of the IEEE International Symposium on Information Theory, held in Seoul, South Korea. He received a Guggenheim Fellowship in 2002 and the IEEE Education Medal in 2005. Recent recognition of his work includes the 2008 Aaron D. Wyner Distinguished Service Award of the IEEE Information Theory Society, the 2009 Edwin Howard Armstrong Achievement Award of the IEEE Communications Society, the 2010 IET Ambrose Fleming Medal for Achievement in Communications, and the 2011 IEEE Eric E. Sumner Award.