Secure Broadcasting: The Secrecy Rate Region

Ghadamali Bagherikaram, Abolfazl S. Motahari, Amir K. Khandani
Coding and Signal Transmission Laboratory,
Department of Electrical and Computer Engineering,
University of Waterloo, Waterloo, Ontario, N2L 3G1
Emails: {gbagheri,abolfazl,khandani}@cst.uwaterloo.ca

[1]

Abstract

In this paper, we consider a scenario where a source node wishes to broadcast two confidential messages for two respective receivers, while a wire-tapper also receives the transmitted signal. This model is motivated by wireless communications, where individual secure messages are broadcast over open media and can be received by any illegitimate receiver. The secrecy level is measured by the equivocation rate at the eavesdropper. We first study the general (non-degraded) broadcast channel with confidential messages. We present an inner bound on the secrecy capacity region for this model. The inner bound coding scheme is based on a combination of random binning, and the Gelfand-Pinsker bining. This scheme matches Marton's inner bound on the broadcast channel without confidentiality constraint. We further study the situation in which the channels are degraded. For the degraded broadcast channel with confidential messages, we present the secrecy capacity region. Our achievable coding scheme is based on Cover's superposition scheme and random binning. We refer to this scheme as Secret Superposition Scheme. In this scheme, we show that randomization in the first layer increases the secrecy rate of the second layer. This capacity region matches the capacity region of the degraded broadcast channel without security constraint. It also matches the secrecy capacity for the conventional wire-tap channel. Our converse proof is based on a combination of the converse proof of the conventional degraded broadcast channel and Csiszar Lemma. We then assume that the channels are Additive White Gaussian Noise (AWGN) and show that secret superposition scheme with Gaussian codebook is optimal. The converse proof is based on the generalized entropy power inequality. Finally, we use a broadcast strategy for the slowly fading wire-tap channel when only the eavesdropper's channel is fixed and known at the transmitter. We derive the optimum power allocation for the layers which maximizes the total average rate.

I. Introduction

The notion of information theoretic secrecy in communication systems was first introduced by Shannon in [1]. The information theoretic secrecy requires that the received signal of the eavesdropper not provide even a single bit information about the transmitted messages. Shannon considered a pessimistic situation where both the intended receiver and the eavesdropper have direct access to the transmitted signal (which is called ciphertext). Under these circumstances, he proved a negative result showing that perfect secrecy can be achieved only when the entropy of the secret key is greater than, or equal to the entropy of the message. In modern cryptography, all practical cryptosystems are based on Shannnon's pessimistic assumption. Due to practical constraints, secret keys are much shorter than messages. Therefore, these practical cryptosystems are theoretically susceptible of breaking by attackers. The goal of designing such practical ciphers, however, is to guarantee that no efficient algorithm exists for breaking them.

Wyner in [2] showed that the above negative result is a consequence of Shannon's restrictive assumption that the adversary has access to precisely the same information as the legitimate receiver. Wyner considered a scenario in which a wire-tapper receives the transmitted signal over a degraded channel with respect to the legitimate receiver's channel. He further assumed that the wire-tapper has no computational limitations and knows the codebook used by the transmitter. He measured the level of ignorance at the eavesdropper by its equivocation and characterized the capacity-equivocation region. Interestingly, a non-negative perfect secrecy capacity is always achievable for this scenario.

The secrecy capacity for the Gaussian wire-tap channel is characterized by Leung-Yan-Cheong in [3]. Wyner's work is then extended to the general (non-degraded) broadcast channel with confidential messages by Csiszar and Korner [4]. They considered transmitting confidential information to the legitimate receiver while transmitting common information to both the legitimate receiver and the wire-tapper. They established a capacity-equivocation region of this channel. The BCC has recently been further studied in [5]–[7], where the source node transmits a common message for both receivers, along with two additional confidential messages for two respective receivers. Here,the confidentiality of each message is measured with respect to the other user, and there is no external eavesdropper.

The fading wire-tap channel is investigated in [8] where the source-to destination channel and the source-to-eavesdropper channel are corrupted by multiplicative fading gain coefficients, in addition to additive white Gaussian noise terms. In this work, channels are fast fading and the channel state information of the legitimate receiver is available at the transmitter. The

[1] Financial support provided by Nortel and the corresponding matching funds by the Natural Sciences and Engineering Research Council of Canada (NSERC), and Ontario Centres of Excellence (OCE) are gratefully acknowledged.

perfect secrecy capacity is derived for two different scenarios regarding the availability of the eavesdropper's CSI. Moreover, the optimal power control policy is obtained for the different scenarios. The effect of the slowly fading channel on the secrecy capacity of a conventional wire-tap channel was studied in [9], [10]. In these works, it is assumed that the fading is quasi-static and the transmitter does not know the fading gains. The outage probability, which is the probability that the main channel is stronger than the eavesdropper's channel, is defined in these works. In an outage strategy, the transmission rate is fixed and the information is detected when the instantaneous main channel is stronger than the instantaneous eavesdropper's channel; otherwise, either nothing is decoded at the legitimate receiver, or the information is leaked to the eavesdropper. The term outage capacity refers to the maximum achievable average rate. In [11], a broadcast strategy for the slowly fading Gaussian point to point channel is introduced. In this strategy, the transmitter uses a layered coding scheme and the receiver is viewed as a continuum of ordered users.

In [12], the wire-tap channel is extended to the parallel broadcast channels and the fading channels with multiple receivers. Here, the secrecy constraint is a perfect equivocation for each of the messages, even if all the other messages are revealed to the eavesdropper. The secrecy sum capacity for a reverse broadcast channel is derived subject to this restrictive assumption. The notion of the wire-tap channel is also extended to multiple access channels [13]–[16], relay channels [17]–[20], parallel channels [21] and MIMO channels [22]–[27]. Some other related works on the communication of confidential messages can be found in [28]–[32].

In this paper, we consider a scenario where a source node wishes to broadcast two confidential messages for two respective receivers, while a wire-tapper also receives the transmitted signal. This model is motivated by wireless communications, where individual secure messages are broadcast over shared media and can be received by any illegitimate receiver. In fact, we simplify the restrictive constraint imposed in [12] and assume that the eavesdropper does not have access to the other messages. We first study the general broadcast channel with confidential messages. We present an achievable rate region for this channel. Our achievable coding scheme is based on a combination of random binning and the Gelfand-Pinsker bining [33]. This scheme matches Marton's inner bound [34] on the broadcast channel without confidentiality constraint. We further study the situation wherein the channels are physically degraded and characterize the secrecy capacity region. Our achievable coding scheme is based on Cover's superposition coding [35] and the random binning. We refer to this scheme as the Secret Superposition Coding. This capacity region matches the capacity region of the degraded broadcast channel without anys security constraint. It also matches the secrecy capacity of the wire-tap channel. We also characterize the secrecy capacity region when the channels are additive white Gaussian noise. We show that the secret superposition of Gaussian codebooks is the optimal choice. Based on the rate characterization of the secure broadcast channel, we then use broadcast strategy for the slow fading wire-tap channel when only the eavesdropper's channel is fixed and known at the transmitter. In broadcast strategy, a source node sends secure layers of coding and the receiver is viewed as a continuum of ordered users. We derive optimum power allocation for the layers which maximizes the total average rate.

The rest of the paper is organized as follows: in section II we introduce the system model. In section III we provide an inner bound on the secrecy capacity region when the channels are not degraded. In section IV we specialize our channel to the degraded ones and establish the secrecy capacity region. In section V we derive the secrecy capacity region when the channels are AWGN. Based on the secrecy capacity region of the AWGN channel, in section VI we use a broadcast strategy for the slow fading wire-tap channel when the transmitter only knows the eavesdropper's channel. Section VII concludes the paper.

II. Preliminaries

In this paper, random variables are denoted by capital letters (e.g. X) and their realizations are denoted by corresponding lower case letters (e.g. x). The finite alphabet of a random variable is denoted by a script letter (e.g. $\mathcal{X}$) and its probability distribution is denoted by $P(x)$. The vectors will be written as $x^n = (x_1, x_2, ..., x_n)$, where subscripted letters denote the components and superscripted letters denote the vector. Bold capital letters represent matrices (e.g. $\mathbf{A}$). The notation x^{i-1} denotes the vector $(x_1, x_2, ..., x_{i-1})$ and the notation $\widetilde{x}^i$ denotes the vector $(x_i, x_{i+1}, ..., x_n)$. A similar notation will be used for random variables and random vectors.

Consider a Broadcast Channel with Confidential Messages (BCCM) as depicted in Fig. 4. In this confidential setting, the transmitter wishes to send two independent messages (W_1, W_2) to the respective receivers in n uses of the channel and prevent the eavesdropper from having any information about the messages. A discrete memoryless broadcast channel with confidential messages is represented by $(\mathcal{X}, P, \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Z})$ where, $\mathcal{X}$ is the finite input alphabet set, $\mathcal{Y}_1$, $\mathcal{Y}_2$ and $\mathcal{Z}$ are three finite output alphabet sets, and P is the channel transition probability $P(y_1, y_2, z|x)$. The input of the channel is $x^n \in \mathcal{X}^n$ and the outputs are $y_1^n \in \mathcal{Y}_1^n$, $y_2^n \in \mathcal{Y}_2^n$, and $z^n \in \mathcal{Z}^n$ for Receiver 1, Receiver 2, and the eavesdropper, respectively. The channel is discrete memoryless in the sense that

$$P(y_1^n, y_2^n, z^n|x^n) = \prod_{i=1}^{n} P(y_{1,i}, y_{2,i}, z_i|x_i). \tag{1}$$

A $((2^{nR_1}, 2^{nR_2}), n)$ code for a broadcast channel with confidential messages consists of a stochastic encoder

$$f : (\{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_2}\}) \to \mathcal{X}^n, \tag{2}$$

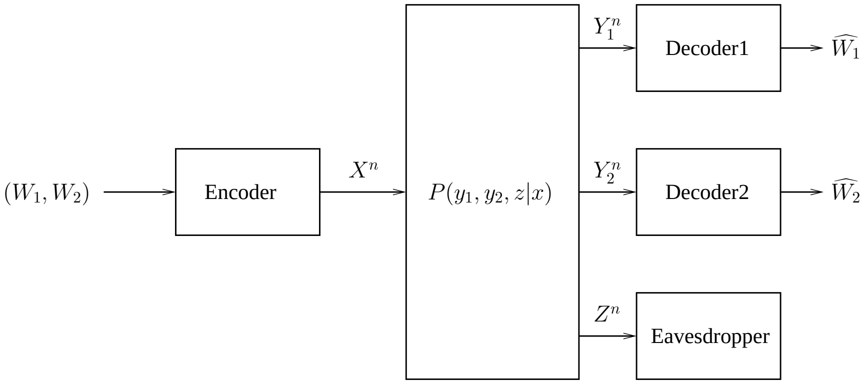

Fig. 1. Broadcast Channel with Confidential Messages

and two decoders,

$$g_1 : \mathcal{Y}_1^n \rightarrow \{1, 2, ..., 2^{nR_1}\} \tag{3}$$

and

$$g_2 : \mathcal{Y}_2^n \rightarrow \{1, 2, ..., 2^{nR_2}\}. \tag{4}$$

The average probability of error is defined as the probability that the decoded messages are not equal to the transmitted messages; that is,

$$P_e^{(n)} = P(g_1(Y_1^n) \neq W_1 \cup g_2(Y_2^n) \neq W_2). \tag{5}$$

The knowledge that the eavesdropper gets about W_1 and W_2 from its received signal Z^n is measured by

$$I(Z^n, W_1) = H(W_1) - H(W_1|Z^n), \tag{6}$$

$$I(Z^n, W_2) = H(W_2) - H(W_2|Z^n), \tag{7}$$

and

$$I(Z^n, (W_1, W_2)) = H(W_1, W_2) - H(W_1, W_2|Z^n). \tag{8}$$

Perfect secrecy revolves around the idea that the eavesdropper cannot get even a single bit information about the transmitted messages. Perfect secrecy thus requires that

$$I(Z^n, W_1) = 0 \Leftrightarrow H(W_1) = H(W_1|Z^n),$$
$$I(Z^n, W_2) = 0 \Leftrightarrow H(W_2) = H(W_2|Z^n),$$

and

$$I(Z^n, (W_1, W_2)) = 0 \Leftrightarrow H(W_1, W_2) = H(W_1, W_2|Z^n).$$

The secrecy levels of confidential messages W_1 and W_2 are measured at the eavesdropper in terms of equivocation rates which are defined as follows.

Definition 1 *The equivocation rates R_{e1}, R_{e2} and R_{e12} for the Broadcast channel with confidential messages are:*

$$R_{e1} = \frac{1}{n} H(W_1|Z^n),$$
$$R_{e2} = \frac{1}{n} H(W_2|Z^n),$$
$$R_{e12} = \frac{1}{n} H(W_1, W_2|Z^n).$$

The perfect secrecy rates R_1 and R_2 are the amount of information that can be sent to the legitimate receivers in a reliable and confidential manner.

Definition 2 *A secrecy rate pair (R_1, R_2) is said to be achievable if for any $\epsilon > 0, \epsilon_1 > 0, \epsilon_2 > 0, \epsilon_3 > 0$, there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ codes, such that for sufficiently large n, we have:*

$$P_e^{(n)} \leq \epsilon, \tag{9}$$

$$R_{e1} \geq R_1 - \epsilon_1, \tag{10}$$

$$R_{e2} \geq R_2 - \epsilon_2, \tag{11}$$

$$R_{e12} \geq R_1 + R_2 - \epsilon_3. \tag{12}$$

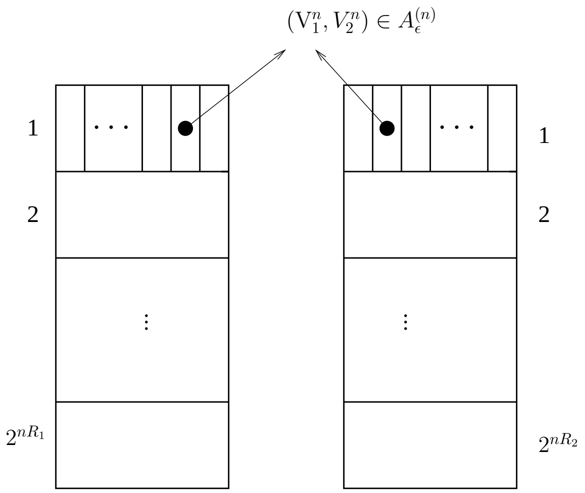

Fig. 2. The Stochastic Encoder

In the above definition, the first condition concerns the reliability, while the other conditions guarantee perfect secrecy for each individual message and both messages as well. The capacity region is defined as follows.

Definition 3 *The capacity region of the broadcast channel with confidential messages is the closure of the set of all achievable rate pairs* (R_1, R_2)*.*

III. Achievable Rates for General BCCM

In this section, we consider the general broadcast channel with confidential messages and present an achievable rate region. Our achievable coding scheme is based on a combination of the random binning and Gelfand-Pinsker bining schemes [33]. The following theorem illustrates the achievable rate region for this channel.

Theorem 1 *Let* $\mathbb{R}_I$ *denote the union of all non-negative rate pairs* (R_1, R_2) *satisfying*

$$\begin{aligned} R_1 &\leq I(V_1;Y_1) - I(V_1;Z), \\ R_2 &\leq I(V_2;Y_2) - I(V_2;Z), \\ R_1 + R_2 &\leq I(V_1;Y_1) + I(V_2;Y_2) - I(V_1,V_2;Z) - I(V_1;V_2), \end{aligned}$$

over all joint distributions $P(v_1,v_2)P(x|v_1,v_2)P(y_1,y_2,z|x)$*. Any rate pair* $(R_1,R_2) \in \mathbb{R}_I$ *is then achievable for the broadcast channel with confidential messages.*

Remark 1 *If we remove the secrecy constraints by removing the eavesdropper, then the above rate region becomes Marton's achievable region for the general broadcast channel.*

Remark 2 *If we remove one of the users, e.g. user* 2*, then we get Csiszar and Korner's secrecy capacity for the other user.*

Proof:

1) *Codebook Generation*: The structure of the encoder is depicted in Fig.2. Fix $P(v_1)$, $P(v_2)$ and $P(x|v_1,v_2)$. The stochastic encoder generates $2^{n(I(V_1;Y_1)-\epsilon)}$ independent and identically distributed sequences v_1^n according to the distribution $P(v_1^n) = \prod_{i=1}^n P(v_{1,i})$. Next, randomly distribute these sequences into 2^{nR_1} bins such that each bin contains $2^{n(I(V_1;Z)-\epsilon)}$ codewords. Similarly, it generates $2^{n(I(V_2;Y_2)-\epsilon)}$ independent and identically distributed sequences v_2^n according to the distribution $P(v_2^n) = \prod_{i=1}^n P(v_{2,i})$. Randomly distribute these sequences into 2^{nR_2} bins such that each bin contains $2^{n(I(V_2;Z)-\epsilon)}$ codewords. Index each of the above bins by $w_1 \in \{1,2,...,2^{nR_1}\}$ and $w_2 \in \{1,2,...,2^{nR_2}\}$ respectively.

2) *Encoding*: To send messages w_1 and w_2, the transmitter looks for v_1^n in bin w_1 of the first bin set and looks for v_2^n in bin w_2 of the second bin set, such that $(v_1^n, v_2^n) \in A_\epsilon^{(n)}(P_{V_1,V_2})$ where $A_\epsilon^{(n)}(P_{V_1,V_2})$ denotes the set of jointly typical sequences v_1^n and v_2^n with respect to $P(v_1,v_2)$. The rates are such that there exist more than one joint typical pair. The transmitter randomly chooses one of them and then generates x^n according to $P(x^n|v_1^n,v_2^n) = \prod_{i=1}^n P(x_i|v_{1,i},v_{2,i})$. This scheme is equivalent to the scenario in which each bin is divided into subbins and the transmitter randomly chooses one of the subbins of bin w_1 and one of the subbins of bin w_2. It then looks for a joint typical sequence (v_1^n, v_2^n) in the corresponding subbins and generates x^n.

3) *Decoding*: The received signals at the legitimate receivers, y_1^n and y_2^n, are the outputs of the channels $P(y_1^n|x^n) = \prod_{i=1}^n P(y_{1,i}|x_i)$ and $P(y_2^n|x^n) = \prod_{i=1}^n P(y_{2,i}|x_i)$, respectively. The first receiver looks for the unique sequence v_1^n such that (v_1^n, y_1^n) is jointly typical and declares the index of the bin containing v_1^n as the message received. The second receiver uses the same method to extract the message w_2.

4) *Error Probability Analysis*: Since the region of (9) is a subset of Marton's region, then the error probability analysis is the same as [34].

5) *Equivocation Calculation*: The proof of secrecy requirement for each individual message (10) and (11) is straightforward and may therefore be omitted.

To prove the requirement of (12) from $H(W_1, W_2|Z^n)$, we have

$$\begin{aligned}
nR_{e12} &= H(W_1, W_2|Z^n) \\
&= H(W_1, W_2, Z^n) - H(Z^n) \\
&= H(W_1, W_2, V_1^n, V_2^n, Z^n) - H(V_1^n, V_2^n|W_1, W_2, Z^n) - H(Z^n) \\
&= H(W_1, W_2, V_1^n, V_2^n) + H(Z^n|W_1, W_2, V_1^n, V_2^n) - H(V_1^n, V_2^n|W_1, W_2, Z^n) - H(Z^n) \\
&\overset{(a)}{\geq} H(W_1, W_2, V_1^n, V_2^n) + H(Z^n|W_1, W_2, V_1^n, V_2^n) - n\epsilon_n - H(Z^n) \\
&\overset{(b)}{=} H(W_1, W_2, V_1^n, V_2^n) + H(Z^n|V_1^n, V_2^n) - n\epsilon_n - H(Z^n) \\
&\overset{(c)}{\geq} H(V_1^n, V_2^n) + H(Z^n|V_1^n, V_2^n) - n\epsilon_n - H(Z^n) \\
&= H(V_1^n) + H(V_2^n) - I(V_1^n; V_2^n) - I(V_1^n, V_2^n; Z^n) - n\epsilon_n \\
&\overset{(d)}{\geq} I(V_1^n; Y_1^n) + I(V_2^n; Y_2^n) - I(V_1^n; V_2^n) - I(V_1^n, V_2^n; Z^n) - n\epsilon_n \\
&\geq nR_1 + nR_2 - n\epsilon_n,
\end{aligned}$$

where (a) follows from Fano's inequality, which states that for sufficiently large n, $H(V_1^n, V_2^n|W_1, W_2, Z^n) \leq h(P_{we}^{(n)})$ $+nP_{we}^n R_w \leq n\epsilon_n$. Here P_{we}^n denotes the wiretapper's error probability of decoding (v_1^n, v_2^n) in the case that the bin numbers w_1 and w_2 are known to the eavesdropper and $R_w = I(V_1; Z) + I(V_2; Z) \leq I(V_1, V_2; Z) + I(V_1; V_2)$. Since the sum rate is small enough, then $P_{we}^n \to 0$ for sufficiently large n. (b) follows from the following Markov chain: $(W_1, W_2) \to (V_1^n, V_2^n) \to Z^n$. Hence, we have $H(Z^n|W_1, W_2, V_1^n, V_2^n) = H(Z^n|V_1^n, V_2^n)$. (c) follows from the fact that $H(W_1, W_2, V_1^n, V_2^n) \geq H(V_1^n, V_2^n)$. (d) follows from that fact that $H(V_1^n) \geq I(V_1^n; Y_1^n)$ and $H(V_2^n) \geq I(V_2^n; Y_2^n)$.

■

IV. The Capacity Region of the Degraded BCCM

In this section, we consider the degraded broadcast channel with confidential messages and establish its secrecy capacity region.

Definition 4 *A broadcast channel with confidential messages is said to be physically degraded, if $X \to Y_1 \to Y_2 \to Z$ forms a Markov chain. In other words, we have*

$$P(y_1, y_2, z|x) = P(y_1|x)P(y_2|y_1)P(z|y_2).$$

Definition 5 *A broadcast channel with confidential messages is said to be stochastically degraded if its conditional marginal distributions are the same as that of a physically degraded broadcast channel, i.e., if there exist two distributions $P^{'}(y_2|y_1)$ and $P^{'}(z|y_2)$, such that*

$$\begin{aligned}
P(y_2|x) &= \sum_{y_1} P(y_1|x)P^{'}(y_2|y_1), \\
P(z|x) &= \sum_{y_2} P(y_2|x)P^{'}(z|y_2).
\end{aligned}$$

Lemma 1 *The secrecy capacity region of a broadcast channel with confidential messages depends only on the conditional marginal distributions $P(y_1|x)$, $P(y_2|x)$ and $P(z|x)$.*

Proof: It suffices to show that the error probability $P_e^{(n)}$ and the equivocations $H(W_1|Z^n)$, $H(W_2|Z^n)$ and $H(W_1, W_2|Z^n)$ are only functions of marginal distributions when we use the same codebook and encoding schemes. Note that

$$\max\{P_{e,1}^{(n)}, P_{e,2}^{(n)}\} \leq P_e^{(n)} \leq P_{e,1}^{(n)} + P_{e,2}^{(n)}.$$

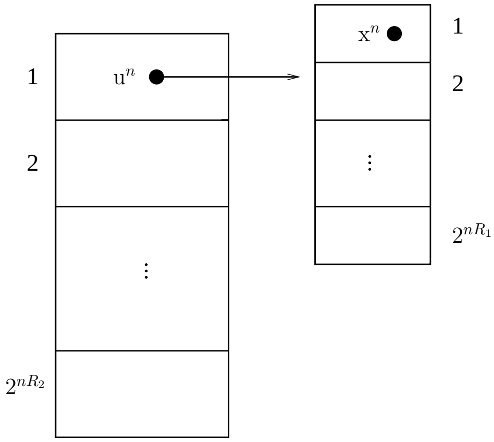

Fig. 3. Secret Superposition structure

Hence, $P_e^{(n)}$ is small if and only if both $P_{e,1}^{(n)}$ and $P_{e,2}^{(n)}$ are small. On the other hand, for a given codebook and encoding scheme, the decoding error probabilities $P_{e,1}^{(n)}$ and $P_{e,2}^{(n)}$ and the equivocation rates depend only on marginal channel probability densities $P_{Y_1|X}$, $P_{Y_2|X}$ and $P_{Z|X}$. Thus, the same code and encoding scheme gives the same $P_e^{(n)}$ and equivocation rates. ■

In the following theorem, we fully characterize the capacity region of the physically degraded broadcast channel with confidential messages.

Theorem 2 *The capacity region for transmitting independent secret information over the degraded broadcast channel is the convex hull of the closure of all (R_1, R_2) satisfying*

$$R_1 \leq I(X;Y_1|U) + I(U;Z) - I(X;Z), \tag{13}$$

$$R_2 \leq I(U;Y_2) - I(U;Z), \tag{14}$$

for some joint distribution $P(u)P(x|u)P(y_1, y_2, z|x)$.

Remark 3 *If we remove the secrecy constraints by removing the eavesdropper, then the above theorem becomes the capacity region of the degraded broadcast channel.*

The coding scheme is based on Cover's superposition coding and random bining. We refer to this scheme as the Secure Superposition Coding scheme. The available resources at the encoder are used for two purposes: to confuse the eavesdropper so that perfect secrecy can be achieved for both layers, and to transmit the messages into the main channels. To satisfy confidentiality, the randomization used in the first layer is fully exploited in the second layer. This makes an increase of $I(U;Z)$ in the bound of R_1.

Proof:

Achievablity: The formal proof of the achievablity is as follows:

1) *Codebook Generation*: The structure of the encoder is depicted in Fig.7. Let us fix $P(u)$ and $P(x|u)$. We generate $2^{n(I(U;Y_2)-\epsilon)}$ independent and identically distributed sequences u^n according to the distribution $P(u^n) = \prod_{i=1}^n P(u_i)$. Next, we randomly distribute these sequences into 2^{nR_2} bins such that each bin contains $2^{n(I(U;Z)-\epsilon)}$ codewords. We index each of the above bins by $w_2 \in \{1, 2, ..., 2^{nR_2}\}$. For each codeword of u^n, we also generate $2^{n(I(X;Y_1|U)-\epsilon)}$ independent and identically distributed sequences x^n according to the distribution $P(x^n|u^n) = \prod_{i=1}^n P(x_i|u_i)$. We randomly distribute these sequences into 2^{nR_1} bins such that each bin contains $2^{n(I(X;Z)-I(U;Z)-\epsilon)}$ codewords. We index each of the above bins by $w_1 \in \{1, 2, ..., 2^{nR_1}\}$.

2) *Encoding*: To send messages w_1 and w_2, the transmitter randomly chooses one of the codewords in bin w_2, say u^n. Then given u^n, the transmitter randomly chooses one of x^n in bin w_1 of the second layer and sends it.

3) *Decoding*: The received signal at the legitimate receivers, y_1^n and y_2^n, are the outputs of the channels $P(y_1^n|x^n) = \prod_{i=1}^n P(y_{1,i}|x_i)$ and $P(y_2^n|x^n) = \prod_{i=1}^n P(y_{2,i}|x_i)$, respectively. Receiver 2 determines the unique u^n such that (u^n, y_2^n) are jointly typical, and declares the index of the bin containing u^n as the message received. If there is none of such messages or more than of one such, an error is declared. Receiver 1 looks for the unique (u^n, x^n) such that (u^n, x^n, y_1^n) are jointly typical and declares the indices of the bins containing u^n and x^n as the messages received. If there is none of such or more than of one such, an error is declared.

4) *Error Probability Analysis*: Since each rate pair of (13) is in the capacity region of the degraded broadcast channel without confidentiality constraint, then it can be readily shown that the error probability is arbitrarily small, c.f. [35].

5) *Equivocation Calculation*: To prove the secrecy requirement of (10), we have

$$\begin{aligned}
nR_{e1} &= H(W_1|Z^n)\\
&\geq H(W_1|Z^n,U^n)\\
&= H(W_1,Z^n|U^n)-H(Z^n|U^n)\\
&= H(W_1,X^n,Z^n|U^n)-H(Z^n|U^n)-H(X^n|W_1,Z^n,U^n)\\
&\overset{(a)}{=} H(W_1,X^n|U^n)+H(Z^n|W_1,U^n,X^n)-H(Z^n|U^n)-n\epsilon_n\\
&\overset{(b)}{\geq} H(X^n|U^n)+H(Z^n|X^n)-H(Z^n|U^n)-n\epsilon_n\\
&\overset{(c)}{\geq} I(X^n;Y_1^n|U^n)+I(U^n;Z^n)-I(X^n;Z^n)-n\epsilon_n\\
&\overset{(d)}{\geq} nR_1-n\epsilon_n,
\end{aligned}$$

where (a) follows from Fano's inequality, which states that $H(X^n|W_1,Z^n,U^n)\leq h(P_{we}^{(n)})+nP_{we}^nR_w\leq n\epsilon_n$ for sufficiently large n. Here P_{we}^n denotes the wiretapper's error probability of decoding x^n given that the bin number and the codeword u^n are known to the eavesdropper and $R_w=I(X;Z)-I(U;Z)$. Since the rate R_w is less than $I(X;Z)$, then $P_{we}^n\to 0$ for sufficiently large n. (b) follows from the fact that $(W_1,U^n)\to X^n\to Z^n$ forms a Markov chain. Thus, we have $H(Z^n|W_1,U^n,X^n)=H(Z^n|X^n)$. (c) follows from two identities: $H(X^n|U^n)\geq I(X^n;Y_1^n|U^n)$ and $H(Z^n|X^n)-H(Z^n|U^n)=I(U^n;Z^n)-I(X^n;Z^n)$. Similarly, we can prove (11). Thus, we only need to prove (12).

$$\begin{aligned}
nR_{e12} &= H(W_1,W_2|Z^n)\\
&= H(W_1,W_2,Z^n)-H(Z^n)\\
&= H(W_1,W_2,U^n,X^n,Z^n)-H(U^n,X^n|W_1,W_2,Z^n)-H(Z^n)\\
&= H(W_1,W_2,U^n,X^n)+H(Z^n|W_1,W_2,U^n,X^n)-H(U^n|W_1,W_2,Z^n)-H(X^n|W_1,W_2,Z^n,U^n)-H(Z^n)\\
&\overset{(a)}{\geq} H(W_1,W_2,U^n,X^n)+H(Z^n|W_1,W_2,U^n,X^n)-n\epsilon_n-H(Z^n)\\
&\overset{(b)}{=} H(W_1,W_2,U^n,X^n)+H(Z^n|U^n,X^n)-n\epsilon_n-H(Z^n)\\
&\overset{(c)}{\geq} H(U^n,X^n)+H(Z^n|U^n,X^n)-n\epsilon_n-H(Z^n)\\
&= H(U^n)+H(X^n|U^n)-I(U^n,X^n;Z^n)-n\epsilon_n\\
&\overset{(d)}{=} I(U^n;Y_2^n)+I(X^n;Y_1^n|U^n)-I(X^n;Z^n)-I(U^n;Z^n|X^n)-n\epsilon_n\\
&\geq nR_1+nR_2-n\epsilon_n,
\end{aligned}$$

where (a) follows from Fano's inequality that $H(U^n|W_1,W_2,Z^n)\leq h(P_{we1}^{(n)})+nP_{we1}^nR_{w1}\leq n\epsilon_n/2$ and $H(X^n|W_1,W_2,Z^n,U^n)\leq h(P_{we2}^{(n)})+nP_{we2}^nR_{w2}\leq n\epsilon_n/2$ for sufficiently large n. Here P_{we1}^n and P_{we2}^n denotes the wiretapper's error probability of decoding u^n and x^n in the case that the bin numbers w_1 and w_2 are known to the eavesdropper, respectively. The eavesdropper first looks for the unique u^n in bin w_2 of the first layer, such that it is jointly typical with z^n. As the number of candidate codewords is small enough, the probability of error is arbitrarily small for a sufficiently large n. Next, given u^n, the eavesdropper looks for the unique x^n in the bin w_1 which is jointly typical with z^n. Similarly, since the number of available candidates is small enough, then the probability of error decoding is arbitrarily small. (b) follows from the fact that $(W_1,W_2)\to U^n\to X^n\to Z^n$ forms a Markov chain. Therefore, we have $I(W_1,W_2;Z^n|U^n,X^n)=0$, where it is implied that $H(Z^n|W_1,W_2,U^n,X^n)=H(Z^n|U^n,X^n)$. (c) follows from the fact that $H(W_1,W_2,U^n,X^n)\geq H(U^n,X^n)$. (d) follows from that fact that $H(U^n)=I(U^n;Y_2^n)$ and $H(X^n|U^n)=I(X^n;Y_1^n|U^n)$. This completes the achievablity proof of (13) and (14).

Converse: The transmitter sends two independent secret messages W_1 and W_2 to Receiver 1 and Receiver 2, respectively. Let us define $U_i=(W_2,Y_1^{i-1})$. The following Lemma bounds the secrecy rates for a general case of $(W_1,W_2)\to X^n\to Y_1^nY_2^nZ^n$:

Lemma 2 *For the broadcast channel with confidential messages, the perfect secrecy rates are bounded as follows,*

$$\begin{aligned}
nR_1&\leq\sum_{i=1}^{n}I(W_1;Y_{1i}|W_2,Z_i,Y_1^{i-1},\widetilde{Z}^{i+1})+n\delta_1+n\epsilon_3,\\
nR_2&\leq\sum_{i=1}^{n}I(W_2;Y_{2i}|Z_i,Y_2^{i-1},\widetilde{Z}^{i+1})+n\delta_1+n\epsilon_2.
\end{aligned}$$

Proof: We need to prove the second bound. The first bound can similarly be proven. nR_2 is bounded as follows:

$$\begin{aligned} nR_2 &\overset{(a)}{\leq} H(W_2|Z^n) + n\epsilon_2 \\ &\overset{(b)}{\leq} H(W_2|Z^n) - H(W_2|Y_2^n) + n\delta_1 + n\epsilon_2 \\ &= I(W_2;Y_2^n) - I(W_2;Z^n) + n\delta_1 + n\epsilon_2 \end{aligned}$$

where (a) follows from the secrecy constraint that $H(W_2|Z^n) \geq H(W_2) - n\epsilon_2$. (b) follows from Fano's inequality that $H(W_2|Y_2^n) \leq n\delta_1$. Next, we expand $I(W_2;Y_2^n)$ and $I(W_2;Z^n)$ as follows.

$$\begin{aligned} I(W_2;Y_2^n) &= \sum_{i=1}^{n} I(W_2;Y_{2i}|Y_2^{i-1}) \\ &= \sum_{i=1}^{n} I(W_2,\widetilde{Z}^{i+1};Y_{2i}|Y_2^{i-1}) - I(\widetilde{Z}^{i+1};Y_{2i}|W_2,Y_2^{i-1}) \\ &= \sum_{i=1}^{n} I(W_2;Y_{2i}|Y_2^{i-1},\widetilde{Z}^{i+1}) + I(\widetilde{Z}^{i+1};Y_{2i}|Y_2^{i-1}) - I(\widetilde{Z}^{i+1};Y_{2i}|W_2,Y_2^{i-1}) \\ &= \sum_{i=1}^{n} I(W_2;Y_{2i}|Y_2^{i-1},\widetilde{Z}^{i+1}) + \Delta_1 - \Delta_2, \end{aligned}$$

where, $\Delta_1 = \sum_{i=1}^{n} I(\widetilde{Z}^{i+1};Y_{2i}|Y_2^{i-1})$ and $\Delta_2 = \sum_{i=1}^{n} I(\widetilde{Z}^{i+1};Y_{2i}|W_2,Y_2^{i-1})$. Similarly, we have,

$$\begin{aligned} I(W_2;Z^n) &= \sum_{i=1}^{n} I(W_2;Z_i|\widetilde{Z}^{i+1}) \\ &= \sum_{i=1}^{n} I(W_2,Y_2^{i-1};Z_i|\widetilde{Z}^{i+1}) - I(Y_2^{i-1};Z_i|W_2,\widetilde{Z}^{i+1}) \\ &= \sum_{i=1}^{n} I(W_2;Z_i|Y_2^{i-1},\widetilde{Z}^{i+1}) + I(Y_2^{i-1};Z_i|\widetilde{Z}^{i+1}) - I(Y_2^{i-1};Z_i|W_2,\widetilde{Z}^{i+1}) \\ &= \sum_{i=1}^{n} I(W_2;Z_i|Y_2^{i-1},\widetilde{Z}^{i+1}) + \Delta_1^* - \Delta_2^*, \end{aligned}$$

where, $\Delta_1^* = \sum_{i=1}^{n} I(Y_2^{i-1};Z_i|\widetilde{Z}^{i+1})$ and $\Delta_2^* = \sum_{i=1}^{n} I(Y_2^{i-1};Z_i|W_2,\widetilde{Z}^{i+1})$. According to Lemma 7 of [4], $\Delta_1 = \Delta_1^*$ and $\Delta_2 = \Delta_2^*$. Thus, we have,

$$\begin{aligned} nR_2 &\leq \sum_{i=1}^{n} I(W_2;Y_{2i}|Y_2^{i-1},\widetilde{Z}^{i+1}) - I(W_2;Z_i|Y_2^{i-1},\widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_2 \\ &= \sum_{i=1}^{n} H(W_2|Z_i,Y_2^{i-1},\widetilde{Z}^{i+1}) - H(W_2|Y_{2i},Y_2^{i-1},\widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_2 \\ &\overset{(a)}{\leq} \sum_{i=1}^{n} H(W_2|Z_i,Y_2^{i-1},\widetilde{Z}^{i+1}) - H(W_2|Y_{2i},Z_i,Y_2^{i-1},\widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_2 \\ &= \sum_{i=1}^{n} I(W_2;Y_{2i}|Z_i,Y_2^{i-1},\widetilde{Z}^{i+1}) + n\delta_1 + n\epsilon_2, \end{aligned}$$

where (a) follows from the fact that conditioning always decreases the entropy. ■

Now according to the above Lemma, the secrecy rates are bound as follows:

$$
\begin{aligned}
nR_1 &\overset{(a)}{\leq} \sum_{i=1}^{n} I(W_1;Y_{1,i}|W_2,Z_i,Y_1^{i-1},\widetilde{Z}^{i+1})+n\delta_1+n\epsilon_3 \\
&= \sum_{i=1}^{n} I(W_1;Y_{1,i}|U_i,Z_i,\widetilde{Z}^{i+1})+n\delta_1+n\epsilon_3 \\
&\overset{(b)}{\leq} \sum_{i=1}^{n} I(X_i;Y_{1,i}|U_i,Z_i,\widetilde{Z}^{i+1})+n\delta_1+n\epsilon_3 \\
&\overset{(c)}{=} \sum_{i=1}^{n} I(X_i;Y_{1,i},U_i,Z_i|\widetilde{Z}^{i+1})-I(X_i;Z_i|\widetilde{Z}^{i+1})-I(X_i;U_i|Z_i,\widetilde{Z}^{i+1})+n\delta_1+n\epsilon_3 \\
&\overset{(d)}{=} \sum_{i=1}^{n} I(X_i;Y_{1,i}|U_i,\widetilde{Z}^{i+1})+I(X_i;U_i|\widetilde{Z}^{i+1})-I(X_i;Z_i|\widetilde{Z}^{i+1})-I(X_i;U_i|Z_i,\widetilde{Z}^{i+1})+n\delta_1+n\epsilon_3 \\
&\overset{(e)}{=} \sum_{i=1}^{n} I(X_i;Y_{1,i}|U_i,\widetilde{Z}^{i+1})-I(X_i;Z_i|\widetilde{Z}^{i+1})+I(Z_i;U_i|\widetilde{Z}^{i+1})-I(Z_i;U_i|X_i,\widetilde{Z}^{i+1})+n\delta_1+n\epsilon_3 \\
&\overset{(f)}{=} \sum_{i=1}^{n} I(X_i;Y_{1,i}|U_i,\widetilde{Z}^{i+1})-I(X_i;Z_i|\widetilde{Z}^{i+1})+I(Z_i;U_i|\widetilde{Z}^{i+1})+n\delta_1+n\epsilon_3,
\end{aligned}
$$

where (a) follows from the Lemma (2). (b) follows from the data processing theorem. (c) follows from the chain rule. (d) follows from the fact that $I(X_i;Y_{1,i},U_i,Z_i|\widetilde{Z}^{i+1})=I(X_i;U_i|\widetilde{Z}^{i+1})+I(X_i;Y_{1,i}|U_i,\widetilde{Z}^{i+1})+I(X_i;Z_i|Y_{1,i},U_i,\widetilde{Z}^{i+1})$ and from the fact that $\widetilde{Z}^{i+1}U_i \rightarrow X_i \rightarrow Y_{1,i} \rightarrow Y_{2,i} \rightarrow Z_i$ forms a Markov chain, which means that $I(X_i;Z_i|Y_{1,i},U_i,\widetilde{Z}^{i+1})=0$. (e) follows from the fact that $I(X_i;U_i|\widetilde{Z}^{i+1})-I(X_i;U_i|Z_i,\widetilde{Z}^{i+1})=I(Z_i;U_i|\widetilde{Z}^{i+1})-I(Z_i;U_i|X_i,\widetilde{Z}^{i+1})$. (f) follows from the fact that $\widetilde{Z}^{i+1}U_i \rightarrow X_i \rightarrow Z_i$ forms a Markov chain. Thus, $I(Z_i;U_i\widetilde{Z}^{i+1}|X_i)=0$ which implies that $I(Z_i;U_i|X_i,\widetilde{Z}^{i+1})=0$.

For the second receiver, we have

$$
\begin{aligned}
nR_2 &\overset{(a)}{\leq} \sum_{i=1}^{n} I(W_2;Y_{2,i}|Y_2^{i-1},Z_i,\widetilde{Z}^{i+1})+n\delta_2+n\epsilon_1 \\
&= \sum_{i=1}^{n} H(Y_{2,i}|Y_2^{i-1},Z_i,\widetilde{Z}^{i+1})-H(Y_{2,i}|W_2,Y_2^{i-1},Z_i,\widetilde{Z}^{i+1})+n\delta_2+n\epsilon_1 \\
&\overset{(b)}{\leq} \sum_{i=1}^{n} H(Y_{2,i}|Z_i,\widetilde{Z}^{i+1})-H(Y_{2,i}|W_2,Y_1^{i-1},Y_2^{i-1},Z_i,\widetilde{Z}^{i+1})+n\delta_2+n\epsilon_1 \\
&\overset{(c)}{=} \sum_{i=1}^{n} H(Y_{2,i}|Z_i,\widetilde{Z}^{i+1})-H(Y_{2,i}|U_i,Z_i,\widetilde{Z}^{i+1})+n\delta_2+n\epsilon_1 \\
&= \sum_{i=1}^{n} I(Y_{2,i};U_i|Z_i,\widetilde{Z}^{i+1})+n\delta_2+n\epsilon_1 \\
&= \sum_{i=1}^{n} I(Y_{2,i};U_i|\widetilde{Z}^{i+1})+I(Y_{2,i};Z_i|U_i,\widetilde{Z}^{i+1})-I(Y_{2,i};Z_i|\widetilde{Z}^{i+1})+n\delta_2+n\epsilon_1 \\
&= \sum_{i=1}^{n} I(Y_{2,i};U_i|\widetilde{Z}^{i+1})-I(Z_i;U_i|\widetilde{Z}^{i+1})+I(Z_i;U_i|Y_{2,i},\widetilde{Z}^{i+1})+n\delta_2+n\epsilon_1 \\
&\overset{(d)}{=} \sum_{i=1}^{n} I(Y_{2,i};U_i|\widetilde{Z}^{i+1})-I(Z_i;U_i|\widetilde{Z}^{i+1})+n\delta_2+n\epsilon_1,
\end{aligned}
$$

where (a) follows from the lemma (2). (b) follows from the fact that conditioning always decreases the entropy. (c) follows from the fact that $Y_2^{i-1} \rightarrow W_2\widetilde{Z}^{i+1}Y_1^{i-1} \rightarrow Y_{2i} \rightarrow Z_i$ forms a Markov chain. (d) follows from the fact that $\widetilde{Z}^{i+1}U_i \rightarrow Y_{2,i} \rightarrow Z_i$ forms a Markov chain. Thus $I(Z_i;U_i\widetilde{Z}^{i+1}|Y_{2i})=0$ which implies that $I(Z_i;U_i|Y_{2i},\widetilde{Z}^{i+1})=0$. Now, following [35], let us define the time sharing random variable Q which is uniformly distributed over $\{1,2,...,n\}$ and independent of $(W_1,W_2,X^n,Y_1^n,Y_2^n)$. Let us define $U=U_Q,\ V=(\widetilde{Z}^{Q+1},Q),\ X=X_Q,\ Y_1=Y_{1,Q},\ Y_2=Y_{2,Q},\ Z=Z_Q$, then R_1 and R_2 can be written as

$$
\begin{aligned}
R_1 &\leq I(X;Y_1|U,V)+I(U;Z|V)-I(X;Z|V), && (15)\\
R_2 &\leq I(U;Y_2|V)-I(U;Z|V). && (16)
\end{aligned}
$$

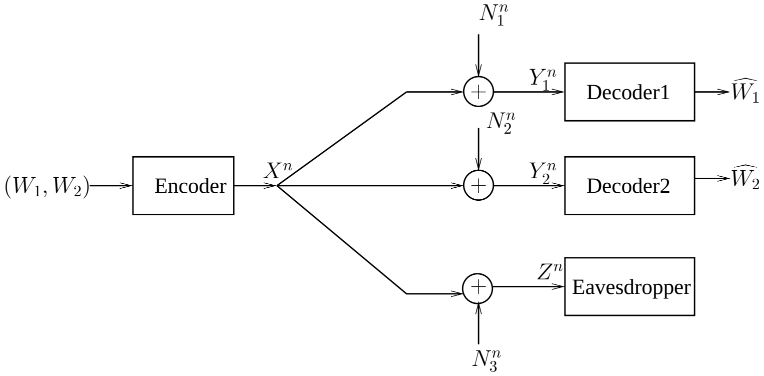

Fig. 4. Gaussian Broadcast Channel with Confidential Messages (G-BCCM)

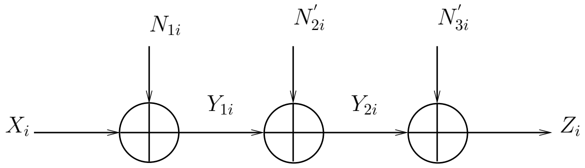

Fig. 5. Equivalent Channels for the G-BCCM

Since conditional mutual information are average of unconditional ones, the largest region is achieved when V is a constant. This proves the converse part. ■

Remark 4 *As Lemma 2 bounds the secrecy rates for the general broadcast channel with confidential messages then, Theorem 2 is true when only the legitimate receivers are degraded.*

V. Capacity Region of Gaussian BCCM

In this section, we consider the Gaussian Broadcast Channel with Confidential Messages (G-BCCM). Note that optimizing (13) and (14) for AWGN channels involves solving a nonconvex functional. Usually nontrivial techniques and strong inequalities are used to solve the optimization problems of this type. In [3], Leung-Yan-Cheong successfully evaluated the capacity expression of the wire-tap channel by using the entropy power inequality. Alternatively, it can also be evaluated using a classical result from the Estimation Theory and the relationship between mutual information and minimum mean-squared error estimation. On the other hand, the entropy power inequality is sufficient to establish the converse proof of a Gaussian broadcast channel without secrecy constraint. Unfortunately, the traditional entropy power inequality does not extend to the secure multi-user case. Here, by using the generalized version of the entropy power inequality, we show that secret superposition coding with Gaussian codebook is optimal.

Fig.4 shows the channel model. At time i the received signals are $Y_{1i} = X_i + N_{1i}$, $Y_{2i} = X_i + N_{2i}$ and $Z_i = X_i + N_{3i}$, where N_{ji} is Gaussian random variable with zero mean and $Var(N_{ji}) = \sigma_j^2$ for $j = 1, 2, 3$. Here $\sigma_1^2 \leq \sigma_2^2 \leq \sigma_3^2$. Assume that the transmitted power is limited to $E[X^2] \leq P$. Since the channels are degraded, the received signals can alternatively be written as $Y_{1i} = X_i + N_{1i}$, $Y_{2i} = Y_{1i} + N_{2i}^{'}$ and $Z_i = Y_{2i} + N_{3i}^{'}$, where N_{1i}'s are i.i.d $\mathcal{N}(0, \sigma_1^2)$, $N_{2i}^{'}$'s are i.i.d $\mathcal{N}(0, \sigma_2^2 - \sigma_1^2)$, and $N_{3i}^{'}$'s are i.i.d $\mathcal{N}(0, \sigma_3^2 - \sigma_2^2)$. Fig. 5 shows the equivalent channels for the G-BCCM. The following theorem illustrates the secrecy capacity region of G-BCCM.

Theorem 3 *The secrecy capacity region of the G-BCCM is given by the set of rates pairs (R_1, R_2) satisfying*

$$R_1 \leq C\left(\frac{\alpha P}{\sigma_1^2}\right) - C\left(\frac{\alpha P}{\sigma_3^2}\right), \tag{17}$$

$$R_2 \leq C\left(\frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\right) - C\left(\frac{(1-\alpha)P}{\alpha P + \sigma_3^2}\right). \tag{18}$$

for some $\alpha \in [0, 1]$.

Proof:

Achievability: Let $U \sim \mathcal{N}(0, (1-\alpha)P)$ and $X^{'} \sim \mathcal{N}(0, \alpha P)$ be independent and $X = U + X^{'} \sim \mathcal{N}(0, P)$. Now consider the following secure superposition coding scheme:

1) *Codebook Generation*: Generate $2^{nI(U;Y_2)}$ i.i.d Gaussian codewords u^n with average power $(1-\alpha)P$ and randomly distribute these codewords into 2^{nR_2} bins. Then index each bin by $w_2 \in \{1, 2, ..., 2^{nR_2}\}$. Generate an independent set of $2^{nI(X^{'};Y_1)}$ i.i.d Gaussian codewords $x^{'n}$ with average power αP. Then, randomly distribute them into 2^{nR_1} bins. Index each bin by $w_1 \in \{1, 2, ..., 2^{nR_1}\}$.

2) *Encoding*: To send messages w_1 and w_2, the transmitter randomly chooses one of the codewords in bin w_2, (say u^n) and one of the codewords in bin w_1 (say $x^{'n}$). The transmitter then simply transmits $x^n = u^n + x^{'n}$.

3) *Decoding*: The received signal at the legitimate receivers are y_1^n and y_2^n respectively. Receiver 2 determines the unique u^n such that (u^n, y_2^n) are jointly typical and declares the index of the bin containing u^n as the message received. If there is none of such or more than of one such, an error is declared. Receiver 1 uses the successive cancelation method; it first decodes u^n and subtracts it from y_1^n and then looks for the unique $x^{'n}$ such that $(x^{'n}, y_1^n - u^n)$ are jointly typical and declares the index of the bin containing $x^{'n}$ as the message received.

It can be shown that if R_1 and R_2 satisfy (17) and (18), the error probability analysis and equivocation calculation is straightforward and may therefore be omitted.

Converse: According to the previous section, R_2 is bound as follows:

$$nR_2 \leq I(Y_2^n; U^n|Z^n) = h(Y_2^n|Z^n) - h(Y_2^n|U^n, Z^n), \tag{19}$$

where h is the differential entropy. The classical entropy power inequality states that:

$$2^{\frac{2}{n}h(Y_2^n + N_3^{'n})} \geq 2^{\frac{2}{n}h(Y_2^n)} + 2^{\frac{2}{n}h(N_3^{'n})}$$

Therefore, $h(Y_2^n|Z^n)$ may be written as follows:

$$\begin{aligned} h(Y_2^n|Z^n) &= h(Z^n|Y_2^n) + h(Y_2^n) - h(Z^n) \\ &= \frac{n}{2}\log 2\pi e(\sigma_3^2 - \sigma_2^2) + h(Y_2^n) - h(Y_2^n + N_3^{'n}) \\ &\leq \frac{n}{2}\log 2\pi e(\sigma_3^2 - \sigma_2^2) + h(Y_2^n) - \frac{n}{2}\log(2^{\frac{2}{n}h(Y_2^n)} + 2\pi e(\sigma_3^2 - \sigma_2^2)). \end{aligned}$$

On the other hand, for any fixed $a \in \mathcal{R}$, the function

$$f(t, a) = t - \frac{n}{2}\log(2^{\frac{2}{n}t} + a)$$

is concave in t and has a global maximum at the maximum value of t. Thus, $h(Y_2^n|Z^n)$ is maximized when Y_2^n (or equivalently X^n) has Gaussian distribution. Hence,

$$\begin{aligned} h(Y_2^n|Z^n) &\leq \frac{n}{2}\log 2\pi e(\sigma_3^2 - \sigma_2^2) + \frac{n}{2}\log 2\pi e(P + \sigma_2^2) - \frac{n}{2}\log 2\pi e(P + \sigma_3^2) \\ &= \frac{n}{2}\log\left(\frac{2\pi e(\sigma_3^2 - \sigma_2^2)(P + \sigma_2^2)}{P + \sigma_3^2}\right). \end{aligned} \tag{20}$$

Now consider the term $h(Y_2^n|U^n, Z^n)$. This term is lower bounded with $h(Y_2^n|U^n, X^n, Z^n) = \frac{n}{2}\log 2\pi e(\sigma_2^2)$ which is greater than $\frac{n}{2}\log 2\pi e(\frac{\sigma_2^2(\sigma_3^2 - \sigma_2^2)}{\sigma_3^2})$. Hence,

$$\frac{n}{2}\log 2\pi e(\frac{\sigma_2^2(\sigma_3^2 - \sigma_2^2)}{\sigma_3^2}) \leq h(Y_2^n|U^n, Z^n) \leq h(Y_2^n|Z^n). \tag{21}$$

Inequalities (20) and (21) imply that there exists an $\alpha \in [0, 1]$ such that

$$h(Y_2^n|U^n, Z^n) = \frac{n}{2}\log\left(\frac{2\pi e(\sigma_3^2 - \sigma_2^2)(\alpha P + \sigma_2^2)}{\alpha P + \sigma_3^2}\right). \tag{22}$$

Substituting (22) and (20) into (19) yields the desired bound

$$\begin{aligned} nR_2 &\leq h(Y_2^n|Z^n) - h(Y_2^n|U^n, Z^n) \\ &\leq \frac{n}{2}\log\left(\frac{(P + \sigma_2^2)(\alpha P + \sigma_3^2)}{(P + \sigma_3^2)(\alpha P + \sigma_2^2)}\right) \\ &= nC\left(\frac{(1-\alpha)P}{\alpha P + \sigma_2^2}\right) - nC\left(\frac{(1-\alpha)P}{\alpha P + \sigma_3^2}\right). \end{aligned} \tag{23}$$

Note that the left hand side of (22), can be written as $h(Y_2^n, Z^n|U^n) - h(Z^n|U^n)$ which implies that

$$h(Y_2^n|U^n) - h(Z^n|U^n) = \frac{n}{2}\log\left(\frac{\alpha P + \sigma_2^2}{\alpha P + \sigma_3^2}\right). \tag{24}$$

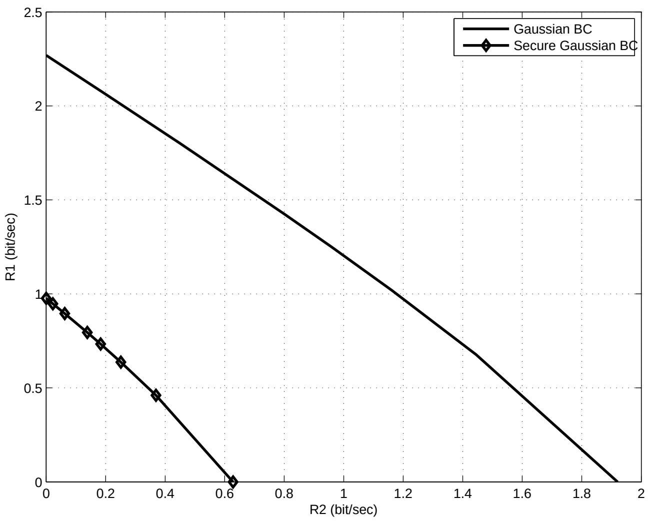

Fig. 6. Secret/Non-Secret Capacity Region of a Degraded Broadcast Channel

Since $\sigma_1^2 \leq \sigma_2^2 \leq \sigma_3^2$, there exists a $0 \leq \beta \leq 1$ such that $\sigma_2^2 = \beta\sigma_1^2 + (1-\beta)\sigma_3^2$ or equivalently $Y_2^n = \beta Y_1^n + (1-\beta)Z^n$. According to the entropy power inequality and the fact that $h(aX^n) = h(X^n) + \log(a^n)$, we have

$$\begin{aligned} &\frac{n}{2}\log\left(\beta 2^{\frac{2}{n}h(Y_1^n|U^n)} + (1-\beta)2^{\frac{2}{n}h(Z^n|Uv)}\right) - h(Z^n|U^n) \\ &\leq \frac{n}{2}\log\left(\frac{\alpha P + \sigma_2^2}{\alpha P + \sigma_3^2}\right). \end{aligned} \tag{25}$$

After some manipulation on (25), we have

$$\begin{aligned} &h(Y_1^n|U^n) - h(Z^n|U^n) \\ &\leq \frac{n}{2}\log\left(\frac{\alpha P + \sigma_2^2 + (\beta-1)(\alpha P + \sigma_3^2)}{\beta(\alpha P + \sigma_3^2)}\right) \\ &= \frac{n}{2}\log\left(\frac{\alpha P + \sigma_1^2}{\alpha P + \sigma_3^2}\right). \end{aligned} \tag{26}$$

The rate R_1 is bounded as follows

$$\begin{aligned} nR_1 &\leq I(X^n;Y_1^n|U^n) - I(X^n;Z^n) + I(U^n;Z^n) \qquad (27) \\ &= h(Y_1^n|U^n) - h(Y_1^n|X^n,U^n) + h(Z^n|X^n) - h(Z^n|U^n) \\ &= h(Y_1^n|U^n) - h(Z^n|U^n) + \frac{n}{2}\log(\frac{\sigma_3^2}{\sigma_1^2}) \\ &\overset{(a)}{\leq} \frac{n}{2}\log\left(\frac{\alpha P + \sigma_1^2}{\alpha P + \sigma_3^2}\frac{\sigma_3^2}{\sigma_1^2}\right) \\ &= nC\left(\frac{\alpha P}{\sigma_1^2}\right) - nC\left(\frac{\alpha P}{\sigma_3^2}\right), \end{aligned}$$

where (a) follows from (26). ■

Fig. 6 shows the capacity region of a degraded Gaussian broadcast channel with and without secrecy constraint. In this figure $P = 20$, $N_1 = 0.9$, $N_2 = 1.5$ and $N_3 = 4$.

VI. A Multilevel Coding Approach to the Slowly Fading Wire-Tap Channel

In this section, we use the secure degraded broadcast channel from the previous section to develop a new broadcast strategy for a slow fading wire-tap channel. This strategy aims to maximize the average achievable rate where main channel state information is not available at the transmitter. By assuming that there are infinite number of ordered receivers which are related to different channel realizations, we propose a secret multilevel coding that maximizes the objection. First, some preliminaries and definitions are given, and then the multilevel coding approach is described. Here, we follow the steps of the broadcast strategy for the slowly fading point-to-point channel of [11].

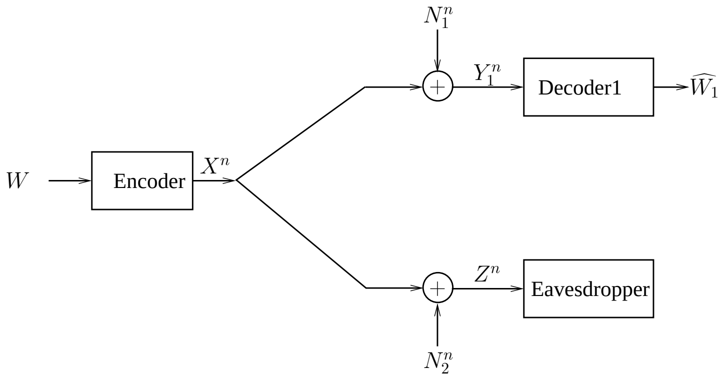

Fig. 7. Gaussian Wire-tap Channel

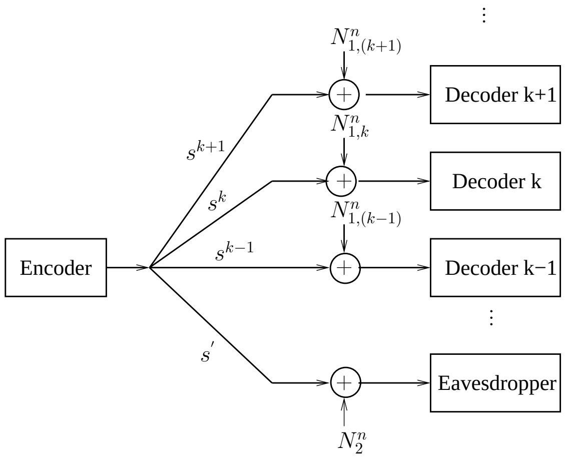

Fig. 8. Equivalent Broadcast Channel Model.

A. Channel Model

Consider a wire-tap channel as depicted in Fig.7. The transmitter wishes to communicate with the destination in the presence of an eavesdropper. At time i, the signal received by the destination and the eavesdropper are given as follows

$$\begin{aligned} Y_i &= h_M X_i + N_{1i} \\ Z_i &= h_E X_i + N_{2i} \end{aligned} \tag{28}$$

where X_i is the transmitted symbol and h_M, h_E are the fading coefficients from the source to the legitimate receiver and the eavesdropper, respectively. The fading power gains of the main and eavesdropper channels are denoted by $s = |h_M|^2$ and $s^{'} = |h_E|^2$ respectively. N_{1i}, N_{2i} are the additive noise samples, which are Gaussian i.i.d with zero mean and unit variance. We assume that the channels are slowly fading, and also assume that the transmitter knows only channel state information of the eavesdropper channel. For each realization of h_M there is an achievable rate. Since the transmitter has no information about the main channel and the channels are slowly fading, then the system is non-ergodic. Here, we are interested in the average rate for various independent transmission blocks. The average shall be calculated over the distribution of h_M.

B. The Secret Multilevel Coding Approach

An equivalent broadcast channel for our channel is depicted in Fig. 8. where, the transmitter sends an infinite number of secure layers of coded information. The receiver is equivalent to a continuum of ordered users. For each channel realization h_M^k with the fading power gain s^k, the information rate is $R(s^k)$. We drop the superscript k, and the realization of the fading power random variable S is denoted by s. Therefore, the transmitter views the main channel as a secure degraded Gaussian broadcast channel with infinite an number of receivers. The result of the previous section for the two receivers can easily be extended to an arbitrary number of users. According to theorem 3, the incremental differential secure rate is then given by

$$dR(s) \quad = \quad \left[\frac{1}{2}\log\left(1+\frac{s\rho(s)ds}{1+sI(s)}\right) - \frac{1}{2}\log\left(1+\frac{s^{'}\rho(s)ds}{1+s^{'}I(s)}\right)\right]^+, \tag{29}$$

where $\rho(s)ds$ is the transmit power of a layer parameterized by s, intended for receiver s. The $\log$ function may be discarded according to the justifications of [37]. The function $I(s)$ represents the interference noise of the receivers indexed by $u > s$ which cannot be canceled at receiver s. The interference at receiver s is therefore given by

$$I(s) = \int_{s}^{\infty} \rho(u)d(u). \tag{30}$$

The total transmitted power is the summation of the power assigned to the layers

$$P = I(0) = \int_{0}^{\infty} \rho(u)d(u). \tag{31}$$

The total achievable rate for a fading realization s is an integration of the incremental rates over all receivers, which can successfully decode the respective layer

$$R(s) = \frac{1}{2}\int_{0}^{s} \left[\frac{u\rho(u)du}{1+uI(u)} - \frac{s^{'}\rho(u)du}{1+s^{'}I(u)}\right]^{+}. \tag{32}$$

Our goal is to maximize the total average rate over all fading realizations with respect to the power distribution $\rho(s)$ (or equivalently, with respect to $I(u)$, $u \geq 0$) under the power constraint of 31. The optimization problem may be written as

$$\begin{aligned} R_{\max} &= \max_{I(u)} \int_{0}^{\infty} R(u)f(u)du, \\ s.t& \\ P &= I(0) = \int_{0}^{\infty} \rho(u)d(u), \end{aligned} \tag{33}$$

where $f(u)$ is the probability distribution function (pdf) of the power gain S. Nothing that the cumulative distribution function (cdf) is $F(u) = \int_{0}^{u} f(a)da$, the optimization problem may be written as

$$\begin{aligned} R_{\max} &= \frac{1}{2}\max_{I(u)} \int_{0}^{\infty} (1-F(u))G(u)du, \\ s.t& \\ P &= I(0) = \int_{0}^{\infty} \rho(u)d(u), \end{aligned} \tag{34}$$

where $G(u) = \left[\frac{u}{1+uI(u)} - \frac{s^{'}}{1+s^{'}I(u)}\right]^{+}\rho(u)$. Note that $\rho(u) = -I^{'}(u)$. The functional of (34), therefore, may be written as

$$\begin{aligned} & J(x, I(x), I^{'}(x)) = \\ & -(1-F(x))\left[\frac{x}{1+xI(x)} - \frac{s^{'}}{1+s^{'}I(x)}\right]^{+} I^{'}(x). \end{aligned} \tag{35}$$

The necessary condition for maximization of an integral of J over x is

$$J_I - \frac{d}{dx}J_{I'} = 0, \tag{36}$$

where J_I means derivation of function J with respect to I, and similarly $J_{I'}$ is the derivation of J with respect to $I^{'}$. After some manipulations, the optimum $I(x)$ is given by

$$I(x) = \begin{cases} \frac{1-F(x)-(x-s^{'})f(x)}{s^{'}(1-F(x))+x(x-s^{'})f(x)}, & \max\{s^{'}, x_0\} \leq x \leq x_1; \\ 0, & \text{otherwise,} \end{cases}$$

where x_0 is determined by $I(x_0) = P$, and x_1 by $I(x_1) = 0$.

As a special case, consider the Rayleigh flat fading channel. The random variable S is exponentially distributed with

$$f(s) = e^{-s}, \qquad F(s) = 1 - e^{-s}, \qquad s \geq 0. \tag{37}$$

Substituting of $f(s)$ and $F(s)$ into the optimum $I(s)$ and taking the derivative with respect to the fading power s yields to the following optimum transmitter power policy

$$\rho(s) = -\frac{d}{ds}I(s) = \begin{cases} \frac{-s^2+2(s^{'}+1)s-s^{'2}}{(s^2-s^{'}s+s^{'})^2}, & \max\{s^{'}, s_0\} \leq s \leq s_1; \\ 0, & \text{otherwise,} \end{cases}$$

where s_0 is the solution of the equation $I(s_0) = P$, which is

$$s_0 = \frac{-1 + Ps^{'} + \sqrt{P^2 s'^2 + 2P(1-2P)s' + 4P + 1}}{2P},$$

and s_1 is determined by $I(s_1) = 0$, which is

$$s_1 = 1 + s^{'}.$$

VII. Conclusion

A generalization of the wire-tap channel in the case of two receivers and one eavesdropper was considered. We established an inner bound for the general (non-degraded) case. This bound matches Marton's bound on broadcast channels without security constraint. Furthermore, we considered the scenario in which the channels are degraded. We established the perfect secrecy capacity region for this case. The achievability coding scheme is a secret superposition scheme where randomization in the first layer helps the secrecy of the second layer. The converse proof combines the converse proof for the degraded broadcast channel without security constraint, and the perfect secrecy constraint. We proved that the secret superposition scheme with the Gaussian codebook is optimal in AWGN-BCCs. The converse proof is based on the generalized entropy power inequality and Csiszar lemma. Based on the rate characterization of the AWGN-BCCs, the broadcast strategy for the slowly fading wire-tap channel were used. In this strategy the transmitter only knows the eavesdropper's channel and the source node sends secure layered coding. The receiver viewed as a continuum of ordered users. We derived optimum power allocation for the layers, which maximizes the total average rate.

References

[1] C. E. Shannon, "Communication Theory of Secrecy Systems", *Bell System Technical Journal*, vol. 28, pp. 656-715, Oct. 1949.

[2] A. Wyner, "The Wire-tap Channel", *Bell System Technical Journal*, vol. 54, pp. 1355-1387, 1975.

[3] S. K. Leung-Yan-Cheong and M. E. Hellman, "Gaussian Wiretap Channel", *IEEE Trans. Inform. Theory*, vol. 24, no. 4, pp. 451-456, July 1978.

[4] I. Csiszar and J. Korner, "Broadcast Channels with Confidential Messages", *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 339-348, May 1978.

[5] R. Liu, I. Maric, P. Spasojevic and R. D. Yates, "Discrete Memoryless Interference and Broadcast Channels with Confidential Messages", *IEEE Trans. Inform. Theory*, Vol. 54, Issue: 6, pp. 2493-2507, Jun 2008.

[6] J. Xu and B. Chen, "Broadcast Confidential and Public Messages", *in Proc. 42nd Conf. Information Sciences and Systems (CISS)*, Princeton, NJ, pp. 630-635 Mar. 2008.

[7] J. Xu, Y. Cao, and B. Chen ,"Capacity Bounds for Broadcast Channels with Confidential Messages", available at http://arxiv.org/PS_cache/arxiv/pdf/0805/0805.4374v1.pdf.

[8] P. K. Gopala, L. Lai and H. El-Gamal, " On the Secrecy Capacity of Fading Channels", *in IEEE Trans. on Info. Theory*, Volume 54, Issue 10, pp. 4687-4698, Oct. 2008.

[9] P. Parada and R. Blahut, "Secrecy Capacity of SIMO and Slow Fading Channels," *in Proc. of ISIT 2005*, pp. 2152-2155, Sep. 2005.

[10] J. Barros and M. R. D. Rodrigues, "Secrecy Capacity of Wireless Channels", *in Proc. of ISIT 2006*, pp. 356-360, July 2006.

[11] S. Shamai, A. Steiner, "A Broadcast Approach for a Single-User Slowly Fading MIMO Channel", *in IEEE Trans. on Info. Theory*, Volume 49, Issue 10, pp. 2617-2635, Oct. 2003.

[12] A. Khisti, A. Tchamkerten and G. W. Wornell, "Secure Broadcasting", available at http://arxiv.org/PS_cache/cs/pdf/0702/0702093v1.pdf.

[13] E. Tekin, S. Serbetli, and A. Yener, "On secure Signaling for the Gaussian Multiple Access Wire-tap Channel", *in Proc. 2005 Asilomar Conf. On Signals, Systems, and Computers*, Asilomar, CA, pp. 1747-1751, November 2005.

[14] E. Tekin and A. Yener, "The Gaussian Multiple Access Wire-Tap Channel", *in IEEE Trans. on Info. Theory*, Volume 54, Issue 12, pp. 5747-5755, Dec. 2008.

[15] Y. Liang and V. Poor, "Generalized Multiple Access Channels with Confidential Messages", *in Proc. Of IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 952-956, July 2006.

[16] E. Tekin and A. Yener, "The Gaussian Multiple Access Wire-Tap Channel: Wireless Secrecy and Cooperative Jamming", *Information Theory and Applications Workshop*, pp. 404-413. Feb. 2007.

[17] Y. Oohama, "Coding for Relay Channels with Confidential messages", *in Proc. Of IEEE Information Theory Workshop*, pp. 87-89, Sep. 2001.

[18] Y. Oohama, "Capacity Theorems for Relay Channels with Confidential Messages ", *in Proc. of ISIT 2007*, pp. 926-930, Jun. 2007.

[19] L. Lai and H. El Gamal, "The Relay-Eavesdropper Channel: Cooperation for Secrecy", *IEEE Trans. Inf. Theory*, Volume 54, Issue 9, pp. 4005-4019, Sept. 2008.

[20] M. Yuksel and E. Erkip., "The Relay Channel with a Wiretapper", *in Proc. Forty-First Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, Mar. 2007.

[21] 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, pp. 841-848, Sep. 2006.

[22] F. Oggier, B. Hassibi, " The MIMO Wiretap Channel", *Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on.*, pp. 213-218, Mar. 2008.

[23] S. Shafiee, L. Nan and S. Ulukus, "Secrecy Capacity of the 2-2-1 Gaussian MIMO Wire-tap Channel", *Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on.*, pp. 207-212, Mar. 2008.

[24] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar, "On the Gaussian MIMO Wiretap Channel", *in Proc. IEEE Int. Symp. Information Theory (ISIT)*, Nice, France, Jun. 2007.

[25] A. Khisti and G. Wornell, "Secure Transmission with Multiple Antennas: The MISOME Wiretap Channel", available at http://arxiv.org/PS_cache/arxiv/pdf/0708/0708.4219v1.pdf.

[26] T. Liu and S. Shamai (Shitz), "A Note on the Secrecy Capacity of the Multi-antenna Wiretap Channel", available at http://arxiv.org/PS_cache/arxiv/pdf/0710/0710.4105v1.pdf.

[27] R. Liu and H. V. Poor, "Multi-Antenna Gaussian Broadcast Channels with Confidential Messages ", *in Proc. of ISIT 2008*, pp. 2202-2206, Jul. 2008.

[28] X. Tang, R. Liu, P. spasojevic and V. Poor, "The Gaussian Wiretap Channel with a Helping Interferer", *in Proc. of ISIT 2008*, pp. 389-393, Jul. 2008.

[29] C. Chan, "Success Exponent of Wiretapper: A Tradeoff between Secrecy and Reliability", available at http://arxiv.org/PS_cache/arxiv/pdf/0805/0805.3605v4.pdf.

[30] X. Tang, R. Liu, P. Spasojevic and V.Poor, “Interference-Assisted Secret Communication”, available at http://arxiv.org/PS_cache/arxiv/pdf/0804/0804.1382v1.pdf.

[31] L.Lai, H. El-Gamal, V. Poor, “The Wiretap Channel With Feedback: Encryption Over the Channel ”, *IEEE Trans. Inf. Theory*, Volume 54, Issue 11, pp. 5059-5067, Nov. 2008.

[32] O.Ozan Koyluoglu, H.El-Gamal, “On the Secure Degrees of Freedom in the K-User Gaussian Interference Channel”, *in Proc. of ISIT 2008*, pp. 384-388, Jul. 2008.

[33] S. I. Gelfand and M. S. Pinsker, “Coding for Channel with Random Parameters”, *Problemy Peredachi Informatsii*, vol. 9, no. 1, pp. 19-31, 1980.

[34] K. Marton, “A Coding Theorem for the Discrete Memoryless Broadcast Channel”, *IEEE Trans. on Inf. Theory*, vol. 25, no. 1, pp. 306-311, May 1979.

[35] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley Sons, Inc., 1991.

[36] T. Liu, P. Viswanath, “An Extremal Inequality Motivated by Multiterminal Information Theoretic Problems”, *IEEE Trans. on Inf. Theory*, vol. 53, no. 5, pp. 1839-1851, May 2007.

[37] A. J. Viterbi, “Very Low Rate Conventional Codes for Maximum Theoretical Performance of Spread-Spectrum Multiple-Access Channels”, *in IEEE J. Select. Areas Commun.*, vol. 8, pp. 641-649, May 1990.