Forty-Eighth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 29 - October 1, 2010

Providing Secrecy when the Eavesdropper Channel is Arbitrarily Varying: a Case for Multiple Antennas

Xiang He Aylin Yener
Wireless Communications and Networking Laboratory
Electrical Engineering Department
The Pennsylvania State University, University Park, PA 16802
xxh119@psu.edu yener@ee.psu.edu

***Abstract*—Despite recent attempts with compound channel settings, guaranteeing information theoretically secure communication rates with no knowledge of the eavesdropper channel states has been elusive to date. In this work, we address this problem, and consider the worst case scenario where the eavesdropper's channel is arbitrarily varying, and is known only by itself. We show that it is possible to provide non-zero secure communication rates with the aid of multiple antennas. Under the mild assumption that the number of antennas at the eavesdropper is limited, we provide the strong secrecy rate for the MIMO wiretap channel. Specifically, it is proved that there exists a universal coding scheme that secures the confidential message against any sequence of channel states experienced by the eavesdropper. We show that the result is tight in terms of secure degrees of freedom.**

I. Introduction

The notion of information theoretic secrecy was introduced by Shannon in [1]. Wyner, in [2], used this notion to study the wiretap channel where an eavesdropper had a noisy observation of the signals sent by the transmitter and showed that a positive rate could be supported for transmitting confidential messages without requiring the communicating parties to share a key. The wiretap model was generalized by Csiszár and Körner in [3].

The wiretap channel model in [2]–[4] has inspired considerable effort toward identifying secure communication limits of various channel models, e.g. [5]–[10]. In these works, it is assumed that the transmitter(s) has (have) perfect knowledge of the eavesdropper channel states, which may be difficult to obtain in a practical system, since the eavesdropper is by nature a passive entity. To resolve this issue, recent works attempt to relax this condition by assuming the transmitter only has *partial* knowledge about the channel states of the eavesdropper. Notably, this line of work includes the compound setting, where the eavesdropper channel can only be taken from a *finite* selection [11]–[14], and the fading channel, where the transmitter only knows the distribution of the eavesdropper channel [15]. These each call for different types of codebook design. For example, in [11]–[13], the coding scheme depends on the possible channel gains of the eavesdropper included in the finite set. For the fading setting [15], the rate of the codebook depends on the fading parameter of the eavesdropper, e.g., the variance of the Rayleigh distribution. Given the absence of a robustness analysis toward understanding how sensitive the achievable secrecy rate is to errors in the aforementioned modeling parameters in [11]–[13], [15], it is difficult to ascertain how close these can model a realistic secure system design based on information theoretic guarantees.

In this work, we study the secure communication problem where the legitimate parties have no knowledge of the channel states of the eavesdropper and show that positive secrecy rates can be guaranteed using multiple antennas under the mild assumption that the number of antennas at the eavesdropper is limited. The channel state sequence of the eavesdropper is assumed to be independent from the input and output of the wiretap channel. Conditioned on any given sequence of channel states, we assume that the eavesdropper channel is memoryless. No restriction is placed on the statistics of the channel model observed by the eavesdropper. That is, the eavesdropper's channel is totally unknown to the legitimate parties and can vary arbitrarily over time.

The main contribution of this work is to prove the existence of a *universal* coding scheme that secures the confidential message against any sequence of eavesdropper channel states for the MIMO wiretap setting described above. This means the coding scheme could withstand the presence of infinitely many eavesdroppers as long as they do not collude. The universal nature of the coding scheme is what sets this work apart from the previous work that considered a similar setting with fading [16]. Additionally, unlike [17] which considered the discrete arbitrarily varying wiretap channel, this work considers a Gaussian setting which does not lend itself to a direct extension from its discrete counterpart.

For this universal coding scheme, an achievable rate is derived for the MIMO wiretap channel with the *strong* secrecy requirements [18]. The achieved rate matches with the converse in terms of secure degrees of freedom (s.d.o.f.), which is a high SNR characteristic of the secrecy capacity.

978-1-4244-8216-0/10/$26.00 ©2010 IEEE

II. The (N_T, N_R, N_E) MIMO Wiretap Channel

The channel from the transmitter to the intended receiver, i.e., *the main channel*, is assumed to be static. Let $A(i)$ denote the value of the signal A during the ith channel use. The input and output of the main channel during the ith channel use are related as:

$$\mathbf{Y}_{N_R\times 1}(i) = \mathbf{H}_{N_R\times N_T}\mathbf{X}_{N_T\times 1}(i) + \mathbf{Z}_{N_R\times 1}(i) \quad (1)$$

where the subscripts denote the dimension of each term. $\mathbf{H}$ denotes the $N_R \times N_T$ channel matrix with complex entries[1]. It is assumed that $\mathbf{H}$ has full rank. $\mathbf{Z}$ is a $N_R \times 1$ vector representing the additive noise. $\mathbf{Z}$ is composed of independent rotationally invariant complex Gaussian random variables, each with zero mean and unit variance. $\mathbf{X}$ and $\mathbf{Y}$ are the transmitted and received signals respectively.

The channel from the transmitter to the eavesdropper, i.e., *the eavesdropper channel*, is an arbitrarily varying channel. It can be expressed as:

$$\tilde{\mathbf{Y}}_{N_E\times 1}(i) = \tilde{\mathbf{H}}_{N_E\times N_T}(i)\mathbf{X}_{N_T\times 1}(i) \quad (2)$$

where $\tilde{\mathbf{Y}}(i)$ denotes the signals received by the eavesdropper during the ith channel use. $\tilde{\mathbf{H}}_{N_E\times N_T}(i)$ is the channel state matrix for the eavesdropper channel during the ith channel use. We use $\tilde{\mathbf{H}}^n$ to denote $\tilde{\mathbf{H}}(1), ..., \tilde{\mathbf{H}}(n)$. $\tilde{\mathbf{H}}^n$ is *not* known at the legitimate parties. We also assume $\tilde{\mathbf{H}}^n$ is independent from $\mathbf{X}^n$.

Note that we assume the eavesdropper's channel is noiseless. This is obviously a worst case assumption, and if the eavesdropper's signals are corrupted by additive noise, they can always be considered as a degraded version of the signals received by the eavesdropper considered in this work.

Let W denote the confidential message transmitted to the intended receiver, over n channel uses using $\mathbf{X}^n$. In addition, we assume that there is a local random source M which is only known to the transmitter. $\mathbf{X}^n$ is computed by the transmitter from W and M using the following encoding function f_n:

$$\mathbf{X}^n = f_n(W, M) \quad (3)$$

Note that f_n does not depend on $\tilde{\mathbf{H}}$, since the transmitter does not know the channel state of the eavesdropper.

We represent $\mathbf{X}^n$ as a $N_T \times n$ matrix, with the average power constraint[2]

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{trace}(\mathbf{X}^n(\mathbf{X}^n)^H) \le \bar{P}. \quad (4)$$

Let $\hat{W}$ denote the decoder output of the intended receiver from $\mathbf{Y}^n$. We require

$$\lim_{n\to\infty}\Pr(W \neq \hat{W}) = 0 \quad (5)$$

[1]Since we assume that the main channel is static, $\mathbf{H}$ remains fixed for all channel uses.

[2]trace denotes the sum of the diagonal elements of a square matrix.

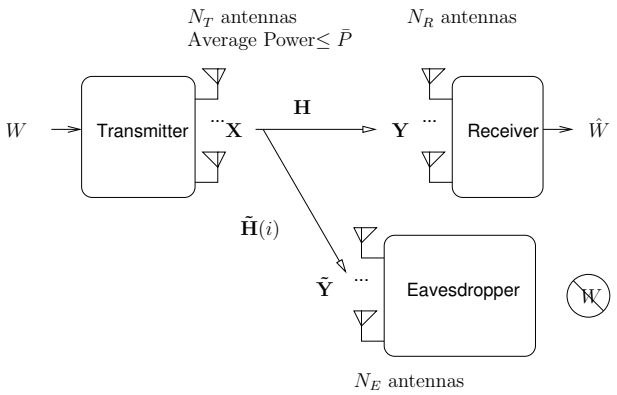

Fig. 1. The MIMO Wiretap Channel.

for reliable communication. Additionally, we require the following *strong* secrecy constraint [18] to hold for any distribution of $\tilde{\mathbf{H}}^n$.

$$\lim_{n\to\infty} I(W; \tilde{\mathbf{Y}}^n, \tilde{\mathbf{H}}^n) = 0 \quad (6)$$

When designing the encoder f_n, we need a uniform bound on $I(W; \tilde{\mathbf{Y}}^n, \tilde{\mathbf{H}}^n)$ over all possible distributions of eavesdropper channel state sequences for a given n, which will determine the minimal codeword length required to achieve a certain level of secrecy. Hence, it is important for the convergence in (6) to be *uniform* for all possible statistics of the eavesdropper channel state sequence.

Since $\tilde{\mathbf{H}}^n$ is independent from $\mathbf{X}^n$, (6) can be written as:

$$\lim_{n\to\infty} I(W; \tilde{\mathbf{Y}}^n | \tilde{\mathbf{H}}^n) = 0 \quad (7)$$

On the other hand, since we do not want the secrecy constraint to rely on the distribution of $\tilde{\mathbf{H}}^n$, we require that for any given sequence $\tilde{\mathbf{h}}^n$,

$$\lim_{n\to\infty} I(W; \tilde{\mathbf{Y}}^n | \tilde{\mathbf{H}}^n = \tilde{\mathbf{h}}^n) = 0. \quad (8)$$

In the sequel, to simplify the notation, we use a scalar γ to index the sequence $\tilde{\mathbf{h}}^n$ and use $\tilde{\mathbf{Y}}^n_\gamma$ to represent the random variable $\tilde{\mathbf{Y}}^n$ when $\tilde{\mathbf{H}}^n$ equals the sequence $\tilde{\mathbf{h}}^n$ indexed by γ. Then (8) can be written as:

$$\lim_{n\to\infty} I(W; \tilde{\mathbf{Y}}^n_\gamma) = 0, \quad \forall\gamma \quad (9)$$

Secrecy rate R_s

$$R_s = \lim_{n\to\infty}\frac{1}{n}H(W) \quad (10)$$

for the MIMO wiretap channel is said to be achievable if for each n there exists a *fixed* encoding function f_n as defined by (3), such that for any given the eavesdropper channel $\tilde{\mathbf{H}}$, (5), (9) and (10) holds for R_s. The supremum of all possible values for R_s is the *secrecy capacity* of this channel model.

High SNR behavior of the secrecy rate is characterized by:

$$\text{s.d.o.f.} = \limsup_{\bar{P}\to\infty}\frac{R_s}{\log_2(\bar{P})} \quad (11)$$

We use the term *the secure degrees of freedom* of a channel to represent the largest possible value of (11).

III. Main Results

The main results of this work are the following:

Theorem 1: Let $N_{T,R} = \min\{N_T, N_R\}$. Let s_i be the $N_{T,R}$ singular values of $\mathbf{H}$. Define P as

$$P = \max\{\bar{P} - N_{T,R}, 0\} \tag{12}$$

Define $C(x)$ as $\log_2(1+x)$. Then any positive secrecy rate R_s smaller than

$$\max\left\{\left(\sum_{i=1}^{N_{T,R}} C\left(\frac{s_i^2 P}{(s_i^2+1)N_{T,R}}\right)\right) - N_E C\left(P\right), 0\right\} \tag{13}$$

is achievable for the MIMO-wiretap channel described in Section II.

From Theorem 1, we have the following Corollary.

Corollary 1: If $\mathbf{H}$ has rank $\min\{N_T, N_R\}$, then the s.d.o.f. of the MIMO-wiretap channel in Section II is

$$\max\{\min\{N_T, N_R\} - N_E, 0\} \tag{14}$$

The proofs of Theorem 1 and Corollary 1 are provided in Section IV.

IV. Proof of the Main Results

A. Notation

We use $p_W(w)$ to denote the probability mass function (p.m.f.) of a random variable W evaluated at w. $f_{\gamma,A}(a)$ denotes the probability density function (p.d.f.) of a random variable A at value a with parameter γ. $f_{\gamma,A|B}(a|b)$ denotes the conditional p.d.f. of a random variable A conditioned on a random variable B when $A = a, B = b$ with parameter γ.

For a vector x^n, we let $\|x^n\|$ denote its L_2-norm. For a matrix $\mathbf{A}$, we let $\|\mathbf{A}\|^2$ denote the sum of the L_2-norm squared of all the row vectors of $\mathbf{A}$. $\mathrm{E}_B[A]$ denotes the expectation of A averaged over B.

B. Channel Model Transformation

Consider a general channel matrix $\mathbf{H}$. We can perform singular value decomposition (SVD) on $\mathbf{H}$ and canceling the right and left unitary matrices of its SVD decomposition at the transmitter and intended receiver respectively. After this cancellation, if $N_T > N_R$, the main channel is equivalent to a MIMO channel whose channel state matrix is $[\mathbf{D}_{N_R \times N_R}, \mathbf{0}_{N_R \times (N_T - N_R)}]$, where $\mathbf{D}$ is a diagonal matrix composed of singular values of $\mathbf{H}$. In this case, we simply use the first N_R antennas at the transmitter only in this equivalent channel when designing the achievable scheme. The remaining $N_T - N_R$ antennas transmit signals of value 0. If $N_T < N_R$, the main channel is then equivalent to a MIMO link whose channel state matrix is $[\mathbf{D}_{N_T \times N_T}, \mathbf{0}_{N_T \times (N_R - N_T)}]^T$, where $(\,)^T$ means transpose operation. In this case, we simply discard the signals observed at the last $N_R - N_T$ antennas at the receiver in this equivalent channel when designing the achievable scheme. In both cases, from the perspective of designing the achievable scheme, the resulting main channel is equivalent to a MIMO link where the transmitter and the receiver both have $\min\{N_T, N_R\}$ antennas and the channel matrix is diagonal. Thus, without loss of generality, we assume $N_T = N_R$ and $\mathbf{H}$ is a diagonal matrix.

We next observe that in order to prove Theorem 1, we only need to consider $N_E < \min\{N_T, N_R\}$. When $N_E \geq \min\{N_T, N_R\}$, the achievable secrecy rate in Theorem 1 is zero, by virtue of the maximum of the two terms on the right hand side of (13) being zero.

Since $N_E < \min\{N_T, N_R\}$, $\tilde{\mathbf{H}}(i)$ has the following form of SVD decomposition:

$$\tilde{\mathbf{H}}(i) = [\mathbf{I}_{N_E \times N_E}, \mathbf{0}_{N_E \times (N_T - N_E)}]\mathbf{U}(i) \tag{15}$$

where $\mathbf{U}(i)$ is a $N_T \times N_T$ unitary matrix. $\mathbf{I}$ is an identity matrix. This can be achieved by canceling the left unitary matrix of the SVD decomposition of $\tilde{\mathbf{H}}(i)$ and normalizing the singular value at the eavesdropper. Note that the transmitter does not know $\mathbf{U}(i)$.

Remark 1: Note that if $\tilde{\mathbf{H}}(i)(\mathbf{U}(i))^{-1}$ has all zero rows, we can alter appropriate entries to be 1s so that the resulting channel matrix has the form in (15). The signals received by the original eavesdropper are always degraded compared to the signals received by the eavesdropper after this modification. Hence, it is sufficient to consider the eavesdroppers with $\tilde{\mathbf{H}}(i)$ in the form given by (15). □

Following [16], we introduce artificial noise at the transmitter. Thus, we have

$$\mathbf{X}(i) = \tilde{\mathbf{X}}(i) + \mathbf{N}(i). \tag{16}$$

$\mathbf{N}$ is the $N_T \times 1$ artificial noise vector consisting of independent rotationally invariant complex Gaussian random variables with zero mean and unit variance. Coding is over $\tilde{\mathbf{X}}$.

Define $\bar{\mathbf{N}}$ and $\tilde{\mathbf{N}}(i)$ as

$$\bar{\mathbf{N}}(i) = \mathbf{H}\mathbf{N}(i) \tag{17}$$

$$\tilde{\mathbf{N}}(i) = \tilde{\mathbf{H}}(i)\mathbf{N}(i) \tag{18}$$

Viewing $\tilde{\mathbf{X}}$ as the input to the channel, the channel model can be expressed as:

$$\mathbf{Y}(i) = \mathbf{H}\tilde{\mathbf{X}}(i) + \bar{\mathbf{N}}(i) + \mathbf{Z}(i) \tag{19}$$

$$\tilde{\mathbf{Y}}(i) = \tilde{\mathbf{H}}(i)\tilde{\mathbf{X}}(i) + \tilde{\mathbf{N}}(i) \tag{20}$$

From (15) and (18), we observe that $\tilde{\mathbf{N}}$ has zero mean and is Gaussian distributed. The covariance matrix of $\tilde{\mathbf{N}}$ is

$$\mathrm{E}[\tilde{\mathbf{H}}(i)\mathbf{N}(i)(\mathbf{N}(i))^H(\tilde{\mathbf{H}}(i))^H] \tag{21}$$

$$= \tilde{\mathbf{H}}(i)\mathrm{E}[\mathbf{N}(i)(\mathbf{N}(i))^H](\tilde{\mathbf{H}}(i))^H \tag{22}$$

$$= \tilde{\mathbf{H}}(i)(\tilde{\mathbf{H}}(i))^H \tag{23}$$

$$= \mathbf{I}_{N_E \times N_E} \tag{24}$$

C. Codebook Construction

The *codebook ensemble* we use is constructed as follows:

Recall that P was defined in (12). We choose the input distribution for $\tilde{\mathbf{X}}$, $Q_{\tilde{\mathbf{X}}}(x)$, as rotationally invariant zero mean complex Gaussian with covariance matrix $(\frac{P(1-\varepsilon_P)}{N_T})\mathbf{I}_{N_T \times N_T}$.

The codebook ensemble is composed of the codebooks constructed as described in [19, Section 7.3]. In the context of this work, this means defining the n-letter distribution $Q_{\tilde{\mathbf{X}}^n}(x^n)$ as follows: Let x_i denote the ith component of x^n. $Q_{\tilde{\mathbf{X}}^n}(x^n)$ is given by:

$$Q_{\tilde{\mathbf{X}}^n}(x^n) = \mu_{n,\varepsilon_P}^{-1} \varphi(x^n) \prod_{i=1}^{n} Q_{\tilde{\mathbf{X}}}(x_i) \tag{25}$$

where

$$\varphi(x^n) = \begin{cases} 1, & \text{if } \frac{1}{n}\|x^n\|^2 \le P \\ 0, & \text{otherwise} \end{cases} \tag{26}$$

$$\mu_{n,\varepsilon_P} = \int \varphi(x^n) \prod_{i=1}^{n} Q_{\tilde{\mathbf{X}}}(x_i) dx^n \tag{27}$$

Note that $0 < \mu_{n,\varepsilon_P} < 1$, and for a given $\varepsilon_P > 0$, there exists an $\alpha(\varepsilon_P) > 0$, such that [20, (B2)]

$$1 - \mu_{n,\varepsilon_P} \le e^{-n\alpha(\varepsilon_P)} \tag{28}$$

$$\lim_{\varepsilon_P \to 0} \alpha(\varepsilon_P) = 0 \tag{29}$$

Any codebook in the ensemble is constructed by sampling 2^{nR} sequences from the distribution $Q_{\tilde{\mathbf{X}}^n}$ in independent and identically distributed (i.i.d.) fashion, where R is given by:

$$R = I(\tilde{\mathbf{X}}; \mathbf{Y}) - \delta' \tag{30}$$

The mutual information in (30) is evaluated when $\tilde{\mathbf{X}}$ has distribution $Q_{\tilde{\mathbf{X}}}$. δ' is a positive constant that can be arbitrarily small.

Each time we sample a codeword, we label it with (i, j). Define N_i and N_j as the range of i and j as follows:

$$N_i = 2^{n(R - I(\tilde{\mathbf{X}};\tilde{\mathbf{Y}}) - \delta_n)} \tag{31}$$

$$N_j = 2^{n(I(\tilde{\mathbf{X}};\tilde{\mathbf{Y}}) + \delta_n)} \tag{32}$$

$\{\delta_n\}$ is a positive sequence whose details will be specified later. Again the mutual information in (32) is evaluated when $\tilde{\mathbf{X}}$ has distribution $Q_{\tilde{\mathbf{X}}}$.Note that we drop the subscript γ in this expression since *the value of the mutual information does not depend on γ when $\tilde{\mathbf{X}}$ has the distribution $Q_{\tilde{\mathbf{X}}}$.*

The label i takes values from $1, ..., N_i$. j takes values from $1, ..., N_j$. The initial value for i and j are both 1. After we label a codeword, if $j < N_j$, then we increase j by one. Otherwise, we increase i by one and reset j to 1.

Let $\mathcal{C}$ denote a codebook in the codebook ensemble $\{\mathcal{C}\}$. Let $x_{i,j}^n$ denote the codeword in the codebook $\mathcal{C}$ that is labeled with (i, j).

D. Encoder

The encoder f_n uses K codebooks $\mathcal{C}_1, ..., \mathcal{C}_K$ sampled from the codebook ensemble. The choice of K and the existence of K good codebooks shall be determined later.

For each codebook $\mathcal{C}$, we define its associated encoder $f_{n,\mathcal{C}}$ as follows: Let the confidential message W be uniformly distributed over the set of $\{i\}$. Given $W = i$, $f_{n,\mathcal{C}}$ selects a codeword from all the codewords with label i in codebook $\mathcal{C}$ according to a uniform distribution. With this encoder, we observe that (i, j) has a uniform distribution.

The encoder f_n is constructed from $f_{n,\mathcal{C}_1}, ..., f_{n,\mathcal{C}_K}$ as in the two stage transmission scheme from [21]:

1) In the first stage, the transmitter chooses the value for an integer K' from $\{1, ..., K\}$ according to a uniform distribution. Given $W = i$, f_n outputs the codeword with label (i, j) computed by $f_{n,\mathcal{C}_{K'}}$.
2) In the second stage, K' is transmitted to the intended receiver using a good channel codebook for the main channel.

E. Decoder

The decoder of the intended receiver first decode K', then decode the confidential message using $\psi_{n,\mathcal{C}_{K'}}$.

As in [19], for a given codebook $\mathcal{C}$, the intended receiver uses a maximum likelihood decoder: Upon receiving $\mathbf{Y}^n = y^n$, the decoder $\psi_{\mathcal{C}}(y^n)$ is given by

$$\psi_{\mathcal{C}}(y^n) = \arg \max_{i,j: x_{i,j}^n \in \mathcal{C}} \|y^n - \mathbf{H}^n x^n\| \tag{33}$$

The probability of decoding error for each codeword, and the average probability of decoding error for each codebook $\mathcal{C}$ and the codebook ensembles are defined as:

$$\lambda_{\mathcal{C},i,j} = \Pr\left(\psi_{\mathcal{C}}(\tilde{\mathbf{Y}}^n) \ne (i, j) | \tilde{\mathbf{X}}^n = x_{i,j}^n\right) \tag{34}$$

$$\lambda_{\mathcal{C}} = \frac{1}{N_i N_j} \sum_{i,j} \lambda_{\mathcal{C},i,j} \tag{35}$$

$$\lambda = \mathrm{E}_{\mathcal{C}}[\lambda_{\mathcal{C}}] \tag{36}$$

Since the codebook ensemble is constructed as in [19], we know that for some n_0 there exists a positive error exponent $E(R(\delta'))$ that [19]

$$\lambda \le 5 \exp(-nE(R(\delta'))), \forall n > n_0 \tag{37}$$

F. Existence of $\mathcal{C}_1, ..., \mathcal{C}_K$

In this section, we show that there exists K good codebooks which can be used by the encoder to achieve a strong secrecy rate. We begin by introducing some notations. Let $\tilde{\mathbf{X}}_G^n$ denote $\tilde{\mathbf{X}}^n$ when it is sampled in an i.i.d. fashion from the input distribution $Q_{\tilde{\mathbf{X}}}(x)$ instead of the codebook. Let $\tilde{\mathbf{X}}_T^n$ denote $\tilde{\mathbf{X}}^n$ when it is sampled in an i.i.d. fashion from the n-letter truncated Gaussian input distribution $Q_{\tilde{\mathbf{X}}^n}$ instead of the codebook.

Let $\tilde{\mathbf{Y}}_G^n, \tilde{\mathbf{Y}}_T^n, \tilde{\mathbf{Y}}_{\mathcal{C}}^n$ denote $\tilde{\mathbf{Y}}^n$ when $\tilde{\mathbf{X}}^n$ is $\tilde{\mathbf{X}}_G^n$, $\tilde{\mathbf{X}}_T^n$ or uniformly distributed over the codebook $\mathcal{C}$ respectively.

Let γ index a sequence of the eavesdropper channel states over n channel uses: $\tilde{\mathbf{H}}_\gamma(1), ..., \tilde{\mathbf{H}}_\gamma(n)$. We use $d_{\gamma,\mathcal{C}}$ to denote the variational distance between two distribution $p_W f_{\gamma,\tilde{\mathbf{Y}}_{\mathcal{C}}^n}$ and $p_W f_{\gamma,\tilde{\mathbf{Y}}_{\mathcal{C}}^n|W}$, which is defined as:

$$d_{\gamma,\mathcal{C}} = \sum_w p_W(w) \int |f_{\gamma,\tilde{\mathbf{Y}}_{\mathcal{C}}^n}(z^n) - f_{\gamma,\tilde{\mathbf{Y}}_{\mathcal{C}}^n|W}(z^n|w)| dz^n \tag{38}$$

As we shall see in Lemma 3, a sufficiently small variational distance can imply strong secrecy. Hence the goal of the proof is to find a good bound on the variational distance.

We shall also use the notion of normalized variational distance: For a given codebook $\mathcal{C}$, and a given eavesdropper channel state sequence $\{\tilde{\mathbf{H}}_\gamma(1),...,\tilde{\mathbf{H}}_\gamma(n)\}$, the normalized variational distance $d'_{\gamma,\mathcal{C}}$ is defined as:

$$d'_{\gamma,\mathcal{C}} = \frac{1}{2}\sum_w p_W(w)\int_{z^n} |f_{\gamma,\tilde{\mathbf{Y}}_G^n}(z^n) - f_{\gamma,\tilde{\mathbf{Y}}_{\mathcal{C}}^n|W}(z^n|w)|dz^n \tag{39}$$

which satisfies:

$$0 \le d'_{\gamma,\mathcal{C}} \le 1 \tag{40}$$

It can be shown that $d'_{\gamma,\mathcal{C}}$ is related to $d_{\gamma,\mathcal{C}}$ via [22]:

$$d_{\gamma,\mathcal{C}} \le 4d'_{\gamma,\mathcal{C}} + 8e^{-n\alpha(\varepsilon_P)} \tag{41}$$

With these definitions and notations, proving the existence of K codebooks that yield small variational distance universally is done by the following steps:

1) We begin by proving for any given sequence of the eavesdropper channel states, the variational distance averaged over an ensemble of wiretap codebooks decreases uniformly and exponentially fast with respect to the code length n.
2) Then, we create a finite a subset of channel state sequences by quantizing them. We use the correlation elimination argument from [21] to show that there exists a small number of codebooks in the codebook ensemble such that the variational distance averaged over these codebooks is small when the eavesdropper channel state sequence is in this finite set.
3) Then we show that when the eavesdropper channel state sequence is outside the finite set, the variational distance averaged over this small set of codebooks can be approximated by the variational distance when the eavesdropper channel state sequence is in the finite set and hence is also small.

The first step of the proof outline is shown by the following lemma whose proof can be found in [22].

Lemma 1: [22] For a given $\varepsilon > 0$, if δ_n is chosen as:

$$\delta_n \ge \max\{2\varepsilon, \varepsilon + \frac{\alpha(\varepsilon_P)}{2}\log_2 e\} \tag{42}$$

then there exists a constant c' such that for sufficiently large n, we have:

$$\mathrm{E}_{\mathcal{C}}\left[d'_{\gamma,\mathcal{C}}\right] \le \exp(-c'n) \tag{43}$$

The value of c' depends only on ε and ε_P. The minimum n for (43) to hold depends only on ε_P.
Applying (43) to (41), we complete the first step in the proof outline.

We then define the finite set S_M of γ in the second step of proof outline as follows: If the real and imaginary parts of each element in $M\tilde{\mathbf{H}}_\gamma(i)$ are integers for all $i = 1,...,n$, then $\gamma \in S_M$. Note that from (15), $\tilde{\mathbf{H}}(i)(\tilde{\mathbf{H}}(i))^H = \mathbf{I}_{N_E\times N_E}$, hence the absolute value of the real and imaginary parts of each element in $\tilde{\mathbf{H}}(i)$ can not exceed 1. Therefore S_M is a finite set with at most $(2M+1)^{2N_TN_En}$ components.

We next use the correlation elimination argument from [21] and consider K codebooks, $\mathcal{C}_1,...,\mathcal{C}_k$, each generated as described in Section IV-C. $\mathcal{C}_k$ is a random variable, and so is $\frac{1}{K}\sum_{k=1}^K d'_{\gamma,\mathcal{C}_k}$. Since for different k, $\mathcal{C}_k$ are i.i.d., $d'_{\gamma,\mathcal{C}_k}$ are also i.i.d.. These facts, along with (40), mean that the derivation in [21, (4.1)-(4.5)] can be applied here. In particular, the j, T_j, ε and R in [21] corresponds to k, $d'_{\gamma,\mathcal{C}_k}$, c' and K here respectively. Consider a positive sequence $\{\epsilon_n\}$. Reference [21, (4.1)-(4.5)] shows that if (43) and (40) holds, for $\alpha' > 0$ and for n that satisfies:

$$1 + e^{\alpha'}e^{-c'n} \le e^{\epsilon_n} \tag{44}$$

we have:

$$\Pr\left(\frac{1}{K}\sum_{k=1}^K d'_{\gamma,\mathcal{C}_k} \ge \epsilon_n\right) \le e^{-(\alpha'-1)K\epsilon_n}. \tag{45}$$

Let $\alpha' = 2$. Then we have

$$\Pr\left(\frac{1}{K}\sum_{k=1}^K d'_{\gamma,\mathcal{C}_k} \ge \epsilon_n\right) \le e^{-\epsilon_n K}. \tag{46}$$

Let $|S_M|$ denote the size of the set S_M. Then we have:

$$\Pr\left(\frac{1}{K}\sum_{k=1}^K d'_{\gamma,\mathcal{C}_k} < \epsilon_n, \forall\gamma \in S_M\right) \tag{47}$$

$$\ge 1 - |S_M|\Pr\left(\frac{1}{K}\sum_{k=1}^K d'_{\gamma,\mathcal{C}_k} > \epsilon_n\right) \tag{48}$$

$$= 1 - |S_M|e^{-\epsilon_n K} \tag{49}$$

We next consider the third step in the proof outline. When $\gamma \notin S_M$, we have the following lemma, whose proof is provided in [22].

Lemma 2: [22] For $\varepsilon > 0$, define r' and r as:

$$(r')^2 = \frac{2N_TN_EP}{M^2} \qquad r = r' + \sqrt{N_E(1+\varepsilon)} \tag{50}$$

Define $g(r,r')$ as

$$g(r,r') = r'(2r + r') \tag{51}$$

if we can choose M with respect to n such that

$$ng(r,r') < 1 \tag{52}$$

then there must exist $\gamma' \in S_M$ such that

$$d'_{\gamma,\mathcal{C}} \le d'_{\gamma',\mathcal{C}} + e^{-n\alpha(\varepsilon)} + ng(r,r') \tag{53}$$

where $\alpha(\varepsilon)$ is a positive constant that only depends on ε.
From Lemma 2, it follows that we have:

$$\frac{1}{K}\sum_{k=1}^K d'_{\gamma,\mathcal{C}_k} \le \frac{1}{K}\sum_{k=1}^K d'_{\gamma',\mathcal{C}_k} + e^{-n\alpha(\varepsilon)} + ng(r,r'). \tag{54}$$

To conclude the second and third steps in the proof outline, we still need to choose ϵ_n, the number of codebooks K and the variable M, which controls the size of the set $|S_M|$ carefully such that for sufficiently large n,

1) (44) is satisfied.

2) $\frac{1}{K}\sum_{k=1}^{K} d'_{\gamma,\mathcal{C}_k}$ in (47) vanishes with high probability for $\gamma \in S_M$. This can be established if ϵ_n in (47) converges to 0 and (49) converges to 1 as n goes to ∞.
3) $\frac{1}{K}\sum_{k=1}^{K} d'_{\gamma,\mathcal{C}_k}$ on the left hand side of (54) vanishes for $\gamma \notin S_M$. In order to use the bound (54) to establish this, (52) must be satisfied and the right hand side of (54) must vanish as well, which relies on 2).
4) The average probability of decoding error for each codebook, $\lambda_{\mathcal{C}_k}, k = 1, ..., K$, vanishes with high probability.

A proper choice for ϵ_n, K and M is as follows: For a positive constant ε' such that

$$\varepsilon' < c' \tag{55}$$
$$\varepsilon' < \alpha(\varepsilon) \tag{56}$$
$$\varepsilon' < \alpha(\varepsilon_P) \tag{57}$$
$$2\varepsilon' < E(R(\delta')) \tag{58}$$

ϵ_n, K and M are chosen to be:

$$\epsilon_n = e^{-\varepsilon' n} \tag{59}$$
$$K = e^{2\varepsilon' n} \tag{60}$$
$$M = e^{2\varepsilon' n} \tag{61}$$

c' in (55) was given in (43).

We first check if these choices satisfy (44). We observe that, since $\epsilon_n > 0$, the right hand side of (44) is lower bounded as:

$$e^{\epsilon_n} \geq 1 + \epsilon_n \tag{62}$$

which, due to (59), equals:

$$1 + e^{-\varepsilon' n} \tag{63}$$

Due to (55), we find (63) is greater than the left hand side of (44) for sufficiently large n such that

$$1 + e^2 e^{-c' n} < 1 + e^{-\varepsilon' n} \tag{64}$$

Hence, (44) is satisfied.

We next check whether requirement 2) is satisfied. We observe from (59) and (60) that

$$e^{-\epsilon_n K} = e^{-e^{-n\varepsilon'} e^{2n\varepsilon'}} = e^{-e^{n\varepsilon'}} \tag{65}$$

We also observe that, due to (61), for sufficiently large n,

$$2M + 1 \leq e^{4\varepsilon' n} \tag{66}$$

Hence,

$$|S_M| = (2M+1)^{2N_T N_E n} < e^{8N_T N_E \varepsilon' n^2} \tag{67}$$

Therefore, from (65) and (67), we have:

$$\lim_{n\to\infty} |S_M| e^{-\epsilon_n K} = 0 \tag{68}$$

This means (49) will converge to 1 when n goes to ∞. Since ϵ_n is shown by (59) to converge to 0 when n goes to ∞, we observe, from (47)-(49), that $\frac{1}{K}\sum_{k=1}^{K} d'_{\gamma,\mathcal{C}_k}$ in (47) vanishes with high probability.

We next check whether requirement 3) is satisfied by examining (54). We observe from (50) and (51) that $g(r, r')$ decreases at the rate of $1/M$ which, according to (61), equals $e^{-2\varepsilon' n}$. Hence for sufficiently large n, we have

$$ng(r, r') < e^{-1.5\varepsilon' n} < e^{-\varepsilon' n} = \epsilon_n \tag{69}$$

holds and (52) is satisfied.

Also, due to (56), we have

$$e^{-n\alpha(\varepsilon)} < e^{-n\varepsilon'} = \epsilon_n \tag{70}$$

If, for the γ' in (54),

$$\frac{1}{K}\sum_{k=1}^{K} d'_{\gamma',\mathcal{C}_k} < \epsilon_n \tag{71}$$

then, from (69), (70) and (54), we have

$$\frac{1}{K}\sum_{k=1}^{K} d'_{\gamma,\mathcal{C}_k} < 3\epsilon_n \tag{72}$$

which means requirement 3) is fulfilled.

Finally, we consider $\lambda_{\mathcal{C}_k}$. From Markov inequality and (36)-(37), we have:

$$\Pr\left(\lambda_{\mathcal{C}_k} > 5nKe^{-Np(R(\delta'))}\right) \leq \frac{1}{nK} \tag{73}$$

Therefore:

$$\Pr\left(\exists k : \lambda_{\mathcal{C}_k} > 5nKe^{-nE(R(\delta'))}\right) \tag{74}$$
$$\leq \sum_{k=1}^{K} \Pr\left(\lambda_{\mathcal{C}_k} > 5nKe^{-nE(R(\delta'))}\right) \tag{75}$$
$$\leq \frac{1}{n} \tag{76}$$

Or equivalently

$$\Pr\left(\lambda_{\mathcal{C}_k} \leq 5nKe^{-nE(R(\delta'))}, k = 1, ..., K\right) \geq 1 - \frac{1}{n}. \tag{77}$$

Hence, from (47)-(49), (68) and (71)-(72) and (77), we observe there must exist K codebooks, where K is given by (60), such that for any k and γ,

$$\lambda_{\mathcal{C}_k} \leq 5nKe^{-nE(R(\delta'))} \tag{78}$$
$$\frac{1}{K}\sum_{k=1}^{K} d'_{\gamma,\mathcal{C}_k} < 3\epsilon_n \tag{79}$$

ϵ_n is given by (59).

Equation (79) can be used to derive a bound on $\frac{1}{K}\sum_{k=1}^{K} d_{\gamma,\mathcal{C}_k}$, which can be used to derive strong secrecy later. Due to (57), we have, from (41), for sufficiently large n:

$$12\epsilon_n = 12e^{-n\varepsilon'} > 8e^{-n\alpha(\varepsilon_P)} \tag{80}$$

Hence from (41), (79) implies

$$\frac{1}{K}\sum_{k=1}^{K} d_{\gamma,\mathcal{C}_k} < 24\epsilon_n, \quad \forall\gamma \tag{81}$$

It remains to check whether the right hand side of (78) vanishes. By applying (60) to the right hand side of (78), we find it equals:

$$5ne^{-n(E(R(\delta'))-2\varepsilon')} \tag{82}$$

Due to (58), we find that the right hand side of (78) converges to 0 when n goes to ∞. This means:

$$\lim_{n\to\infty} \lambda_{\mathcal{C}_k} = 0, \quad \forall k \tag{83}$$

Hence we conclude there exists K codebooks such that for each codebook both the average probability of decoding errors and the variational distance vanish uniformly for all possible eavesdropper channel state sequences when the code length n increases.

G. Probability of Decoding Error

Let $\hat{K}'$ be the result decoded by the intended receiver for K'. Then

$$\Pr\left(W \neq \hat{W}\right) \tag{84}$$

$$\leq \Pr\left(K' \neq \hat{K}'\right) + \Pr\left(W \neq \hat{W} | K' = \hat{K}'\right) \tag{85}$$

$$= \Pr\left(K' \neq \hat{K}'\right) + \frac{1}{K}\sum_{k=1}^{K} \lambda_{\mathcal{C}_k} \tag{86}$$

Since $\lim_{n\to\infty} \Pr\left(K' \neq \hat{K}'\right) = 0$ and (83) holds, we have $\lim_{n\to\infty} \Pr\left(W \neq \hat{W}\right) = 0$.

H. From Variational Distance to Strong Secrecy

The variational distance for this coding scheme, d_γ, is given by

$$d_\gamma = d\left(p_W p_{K'} f_{\gamma,\tilde{Y}^n_{\mathcal{C}_{K'}}}, p_W p_{K'} f_{\gamma,\tilde{Y}^n_{\mathcal{C}_{K'}}|W}\right) \tag{87}$$

$$= \frac{1}{K}\sum_{k=1}^{K} d_{\gamma,\mathcal{C}_k} \tag{88}$$

From (81) and (59), we observe that (88) decreases at the speed of $e^{-\varepsilon' n}$. Then we use the following lemma to relate variational distance with strong secrecy:

Lemma 3: [23, Lemma 1] Let $|\mathcal{W}|$ be the cardinality of the message set $\mathcal{W}$. Then we have:

$$I\left(W; \tilde{\mathbf{Y}}^n_\gamma\right) \leq d_{\gamma,\mathcal{C}} \log_2 \frac{|\mathcal{W}|}{d_{\gamma,\mathcal{C}}} \tag{89}$$

From Lemma 3, we have:

$$\lim_{n\to\infty} I\left(W; K', \tilde{\mathbf{Y}}^n_\gamma\right) = 0, \quad \forall\gamma \tag{90}$$

The convergence is uniform over all possible eavesdropper channel matrix sequences and the limit decreases *exponentially fast* with respect to n. Let n' denote the total number of channel uses. Then the second stage takes n_2 channel uses with n_2 given by:

$$n_2 = \frac{1}{R_0}\log_2 K = \frac{2\varepsilon' \log_2 e}{R_0} n \tag{91}$$

where $R_0 > 0$ is the rate of the conventional channel codebook $\mathcal{C}_0$. The first stage takes n channel uses. Therefore

$$n' = n + n_2 = \left(\frac{2\varepsilon' \log_2 e}{R_0} + 1\right) n \tag{92}$$

Let $\tilde{\mathbf{Y}}^{n_2}_\gamma$ denote the signals received by the eavesdropper during the second stage. Then

$$\lim_{n'\to\infty} I\left(W; \tilde{\mathbf{Y}}^{n'}_\gamma\right) \tag{93}$$

$$= \lim_{n'\to\infty} I\left(W; \tilde{\mathbf{Y}}^{n}_\gamma, \tilde{\mathbf{Y}}^{n_2}_\gamma\right) \tag{94}$$

$$\leq \lim_{n'\to\infty} I\left(W; K', \tilde{\mathbf{Y}}^{n}_\gamma\right) \tag{95}$$

$$= \lim_{n\to\infty} I\left(W; K', \tilde{\mathbf{Y}}^{n}_\gamma\right) \tag{96}$$

$$= 0, \quad \forall\gamma \tag{97}$$

I. Secrecy Rate

Let $c_4 = \delta' + \max\{2\varepsilon, \varepsilon + \frac{\alpha(\varepsilon_P)}{2}\log_2 e\}$. Define $c(\varepsilon')$ as

$$c(\varepsilon') = \left(\frac{2\varepsilon' \log_2 e}{R_0} + 1\right)^{-1} \tag{98}$$

which represents the rate loss due to the second stage. $c(\varepsilon')$ can be made arbitrarily close to 1 by making ε' small. The secrecy rate is then given by:

$$\lim_{n'\to\infty} \frac{1}{n'} H(W) \geq \left(I\left(\tilde{\mathbf{X}}; \mathbf{Y}\right) - N_E C\left(P(1-\varepsilon_P)\right) - c_4\right) c(\varepsilon') \tag{99}$$

From (29), we notice (99) can be made arbitrarily close to

$$I\left(\tilde{\mathbf{X}}; \mathbf{Y}\right) - N_E C(P) \tag{100}$$

Therefore, the same secrecy rate as given in (13) is achievable even when the eavesdropper channel is arbitrarily varying.

J. Converse for Corollary 1

In this section, we establish the result in Corollary 1, by providing the converse for the high SNR characterization of the secrecy rate found in (100).

Since $\tilde{\mathbf{H}}$ can be arbitrary, when $N_E \geq N_T$, we can choose $\tilde{\mathbf{H}}$ as $[\mathbf{I}_{N_T\times N_T}, \mathbf{0}_{N_T\times(N_E-N_T)}]^T$. The eavesdropper in this case has perfect knowledge of the transmitted signal. Clearly, the secrecy capacity is 0.

We next consider the case when $N_E < N_T$. We use X_i^j to denote the ith to the jth component in a vector $\mathbf{X}$. The secrecy rate is upper bounded by [4]:

$$I\left(\mathbf{X}; \mathbf{Y} | \tilde{\mathbf{Y}}\right) \tag{101}$$

When $N_T \geq N_R$, we assume $\mathbf{H} = [\mathbf{D}_{N_R\times N_R}, \mathbf{0}_{N_R\times(N_T-N_R)}]$ for a diagonal matrix $\mathbf{D}_{N_R\times N_R}$[3]. Since $\tilde{\mathbf{H}}$ is arbitrary, we choose $\tilde{\mathbf{H}}$ as $[\mathbf{I}_{N_E\times N_E}, \mathbf{0}_{N_E\times(N_T-N_E)}]$. Then (101) equals:

$$I\left(\mathbf{X}; \mathbf{D}_{N_R\times N_R} X_1^{N_R} + \mathbf{Z} | X_1^{N_E}\right) \tag{102}$$

$$= I\left(X_1^{N_R}, X_{N_R+1}^{N_T}; \mathbf{D}_{N_R\times N_R} X_1^{N_R} + \mathbf{Z} | X_1^{N_E}\right) \tag{103}$$

[3]Else, we can perform SVD on $\mathbf{H}$ and transform it into this form.

$$= I\left(X_1^{N_R}; \mathbf{D}_{N_R \times N_R} X_1^{N_R} + \mathbf{Z} | X_1^{N_E}\right) \quad (104)$$

When $N_T < N_R$, we assume $\mathbf{H} = [\mathbf{D}_{N_T \times N_T}, \mathbf{0}_{N_T \times (N_R - N_T)}]^T$ for a diagonal matrix $\mathbf{D}_{N_T \times N_T}$. We use the same $\tilde{\mathbf{H}}$ as we did in the previous case. Then (101) equals:

$$I\left(\mathbf{X}; \mathbf{D}_{N_T \times N_T} X_1^{N_T} + Z_1^{N_T}, Z_{N_T+1}^{N_R} | X_1^{N_E}\right) \quad (105)$$

$$= I\left(\mathbf{X}; \mathbf{D}_{N_T \times N_T} X_1^{N_T} + Z_1^{N_T} | X_1^{N_E}\right) \quad (106)$$

Define $N_m = \min\{N_T, N_R\}$. Then, in both cases, (101) can be written as:

$$I\left(X_1^{N_m}; \mathbf{D}_{N_m \times N_m} X_1^{N_m} + Z_1^{N_m} | X_1^{N_E}\right) \quad (107)$$

$$= I\left(X_1^{N_m}; Y_1^{N_m} | X_1^{N_E}\right) \quad (108)$$

which equals:

$$I(X_{N_E+1}^{N_m}; Y_1^{N_m} | X_1^{N_E}) \quad (109)$$

It can be verified that (109) is upper bounded by [22]:

$$I(X_{N_E+1}^{N_m}; Y_{N_E+1}^{N_m}) \quad (110)$$

Since we assume $\mathbf{H}$ of the original MIMO wiretap channel has a full rank, $\mathbf{D}_{N_m \times N_m}$ also has full rank. Hence the elements on the diagonal line of $\mathbf{D}$ are all positive. This means equation (110) increases at a rate of $O((\min\{N_T, N_R\} - N_E)C(\bar{P}))$. Hence we have proved the converse of Corollary 1.

V. Conclusion

In this work, we have considered secure communication in the presence of eavesdroppers whose channels are unknown to the legitimate parties and can be arbitrarily varying. We have shown that multiple antennas used in conjunction with the universal coding scheme presented in this paper can guarantee positive secrecy rates irrespective of the channels of the (possibly infinite numbers of) eavesdroppers. As an example, we derived achievable secrecy rates for the MIMO wiretap channel. It matches with the converse in terms of secure degrees of freedom.

The methods developed in this work can be extended to multi-user scenarios including the MIMO-MAC wiretap channel and MIMO-broadcast wiretap channels. Results on these two models where each legitimate node has the same number of antennas can be found in [22].

References

[1] C. E. Shannon. Communication Theory of Secrecy Systems. *Bell System Technical Journal*, 28(4):656–715, September 1949.

[2] A. D. Wyner. The Wire-tap Channel. *Bell System Technical Journal*, 54(8):1355–1387, 1975.

[3] I. Csiszár and J. Körner. Broadcast Channels with Confidential Messages. *IEEE Transactions on Information Theory*, 24(3):339–348, May 1978.

[4] S. Leung-Yan-Cheong and M. Hellman. The Gaussian Wire-tap Channel. *IEEE Transactions on Information Theory*, 24(4):451–456, July 1978.

[5] E. Tekin and A. Yener. The General Gaussian Multiple Access and Two-Way Wire-Tap Channels: Achievable Rates and Cooperative Jamming. *IEEE Transactions on Information Theory*, 54(6):2735–2751, June 2008.

[6] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates. Discrete Memoryless Interference and Broadcast Channels with Confidential Messages: Secrecy Rate Regions. *IEEE Transactions on Information Theory*, 54(6):2493–2507, June 2008.

[7] E. Ekrem and S. Ulukus. The Secrecy Capacity Region of the Gaussian MIMO Multi-receiver Wiretap Channel. Submitted to IEEE Transactions on Information Theory, March 2009, available online at http://arxiv.org/abs/0903.3096.

[8] R. Liu, T. Liu, H. V. Poor, and S. Shamai. Multiple-Input Multiple-Output Gaussian Broadcast Channels with Confidential Messages. Submitted to IEEE Transactions on Information Theory, March 2009. available online at http://arxiv.org/abs/0903.3786.

[9] A. Khisti and G. Wornell. Secure Transmission with Multiple Antennas-I: The MISOME Wiretap Channel. to appear in IEEE Transactions on Information Theory, submitted in August, 2007. available online at http://arxiv.org/abs/0708.4219.

[10] S. Shafiee, N. Liu, and S. Ulukus. Towards the Secrecy Capacity of the Gaussian MIMO Wire-tap Channel: The 2-2-1 Channel. *IEEE Transactions on Information Theory*, 55(9):4033–4039, September 2009.

[11] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai. Compound Wiretap Channels. *Eurasip Journal on Wireless Communication and Networking, Special issue in Wireless Physical Layer Security*, vol. 2009, Article ID 142374, 12 pages, 2009. doi:10.1155/2009/142374.

[12] E. Ekrem and S. Ulukus. Secrecy Capacity Region of the Degraded Compound Multi-Receiver Wiretap Channel. In *47th Allerton Conference on Communication, Control, and Computing*, September 2009.

[13] A. Khisti. Interference Alignment for the Multi-Antenna Compound Wiretap Channel. available online at http://arxiv.org/abs/1002.4548, February, 2010.

[14] M. Kobayashi, Y. Liang, S. Shamai, and M. Debbah. On the Compound MIMO Broadcast Channels with Confidential Messages. In *IEEE International Symposium on Information Theory*, June 2009.

[15] P. K. Gopala, L. Lai, and H. El-Gamal. On the Secrecy Capacity of Fading Channels. *IEEE Transactions on Information Theory*, 54(9):4687–4698, October 2008.

[16] S. Goel and R. Negi. Guaranteeing Secrecy using Artificial Noise. *IEEE Transactions on Wireless Communications*, 7(6):2180–2189, June 2008.

[17] E. MolavianJazi. Secure Communication Over Arbitrarily Varying Wiretap Channels. *Master Thesis*, December 2009. available online at http://etd.nd.edu/ETD-db/theses/available/etd-12112009-112419/unrestricted/MolavianJaziE122009.pdf.

[18] U. Maurer and S. Wolf. Information-theoretic Key Agreement: From Weak to Strong Secrecy for Free. *Lecture Notes in Computer Science*, pages 351–368, 2000.

[19] R. G. Gallager. *Information theory and reliable communication*. John Wiley & Sons, Inc. New York, NY, USA, 1968.

[20] G. Poltyrev. On coding without restrictions for the AWGN channel. *IEEE Transactions on Information Theory*, 40(2):409–417, 1994.

[21] R. Ahlswede. Elimination of correlation in random codes for arbitrarily varying channels. *Probability Theory and Related Fields*, 44(2):159–175, 1978.

[22] X. He and A. Yener. MIMO Wiretap Channels with Arbitrarily Varying Eavesdropper Channel States. Submitted to the IEEE Transactions on Information Theory, July, 2010, available online at http://arxiv.org/abs/1007.4801.

[23] I. Csiszár. Almost Independence and Secrecy Capacity. *Problems of Information Transmission*, 32(1):48–57, 1996.