MIMO TRANSMISSIONS WITH INFORMATION-THEORETIC SECRECY FOR SECRET-KEY AGREEMENT IN WIRELESS NETWORKS

Xiaohua(Edward) Li [*],

Department of Electrical and Computer Engineering
State University of New York at Binghamton
Binghamton, NY 13902
xli@binghamton.edu

E. Paul Ratazzi

Air Force Research Laboratory
AFRL/IFGB
Rome, NY 13441
paul.ratazzi@afrl.af.mil

ABSTRACT

Information-theoretic secrecy gives the strongest information security that is important for military wireless networks. In this paper, we show that randomized MIMO transmissions can achieve such secrecy. A method is proposed to randomize the MIMO transmission coefficients. The received signal of the adversary has some special statistical distributions, based on which the indeterminacy of the adversary's blind deconvolution is proved. Receiving error rates and secret channel capacity are analyzed and simulated. This paper points out a practical way to realize the well-known wire-tap channel for perfect secrecy, which is one of the interesting challenges in information theory. The proposed method is also useful for key management in military wireless networks.

1. INTRODUCTION

Advanced wireless techniques such as multi-input multi-output (MIMO) techniques are important to broadband and highly dynamic wireless communication networks that are essential for military operations. Such techniques are developed with efficiency instead of security as the primary criterion, or even without security consideration at all. Because wireless transmissions are lack of physical boundary (any adversary can receive the signal within the range), the lack of security in these techniques may potentially result in a weak physical-layer security, which may even weaken the end-to-end network security.

[*]This work was supported by US AFRL under grant FA8750-05-1-0233.

Innovative cross-layer security designs with both physical-layer security and upper-layer security techniques are desirable for military wireless networks. While the physical-layer may rely on upper-layer encryption techniques for security, it is interesting to study whether the physical-layer can have built-in security and whether physical-layer security techniques can assist upper-layer security designs.

The built-in security of the physical-layer is defined as that physical-layer transmissions guarantee low-probability-of-interception (LPI) based on transmission properties such as modulations, signals and channels, without resorting to source data encryption. No secret keys are required before transmission. This is in contrast to the traditional techniques such as spread spectrum that rely on secret codes shared between the transmitter and the receiver.

One of the fundamental issues for physical-layer built-in security is the capacity of the transmission channel when built-in security is guaranteed. Such capacity is named secret channel capacity. The secrecy is defined as information-theoretic secrecy, i.e., the adversary's received signal gives no more information for eavesdropping than purely guessing. Information-theoretic secrecy is in fact equivalent to perfect secrecy [1]. Practically, it means negligibly low interception probability.

One of the recent attempts on specifying secret channel capacity is [2], where the MIMO secret channel capacity is analyzed under the assumption that the adversary does not know even his own channel. Unfortunately, such an assumption does not seem practical if considering deconvolution or blind deconvo-

lution techniques. As a matter of fact, almost all existing results on secret channel capacity are based on some kinds of assumptions that appear impractical [3], [4], [5]. It has been a challenge in information theory for decades (after [3]) to find practical ways to realize information-theoretic secrecy.

Recently, we have found that antenna array transmissions may provide valid ways to realize information-theoretic secrecy when the transmission redundancy and the limit of blind deconvolution are appropriately exploited [6]. LPI can be guaranteed by using some extra antennas with some more transmission power or bandwidth. This innovative concept of secure waveform design is completely different from the traditional techniques such as spread spectrum.

This paper extends the results of [6] into multiple-input multiple-output transmissions so that the complexity of exhaustive search, the only way left for the adversary, is an order of magnitude higher. In contrast to [6] where the secrecy is discussed qualitatively, we derive a quantitative analysis of secrecy based on a formal proof of the indeterminacy of the blind deconvolution. We develop the new scheme within the framework of secret-key agreement in order to show that physical-layer security techniques can indeed assist upper-layer encryption techniques.

This paper is organized as follows. In Section 2, a framework of MIMO transmission for secret-key agreement is introduced. In Section 3, we develop the new MIMO transmission scheme, whose secrecy is proved in Section 4. Simulations are given in Section 5 and conclusions are presented in Section 6.

2. MIMO TRANSMISSION FOR SECRET-KEY AGREEMENT

2.1. Secret-key agreement

In cryptography, secret-key agreement denotes the procedure by which Alice sends messages to Bob for the latter to extract a secret key. During this procedure, the adversary Eve can not obtain sufficient information about the key.

The Shannon secrecy model [1] is traditionally used for secret-key agreement protocols, where Eve is assumed to receive messages identical to Bob. It is well known that perfect secrecy is impractical under the Shannon model. Instead, with computational secrecy, keys can be generated based on some intractable computational problems such as factorizing integers in a feasible time. Nevertheless, such intractability assumption is unproven, so does the secrecy [7].

In contrast, information-theoretic secrecy is provably realizable under some models that are more relaxed but also more practical than the Shannon model. In wireless communications, Bob and Eve usually have different received signals, which is different from the Shannon model assumptions. In particular, the noise difference between Bob and Eve has been studied for information-theoretic secrecy. The first of such work is the famous wire-tap channel [3]. If Bob and Eve have different receiving bit-error-rate (BER) ϵ and δ, respectively, then the secret channel capacity from Alice to Bob can be [4]

$$C_1 = \begin{cases} h(\delta) - h(\epsilon), & \text{if } \delta > \epsilon \\ 0, & \text{else} \end{cases} \tag{1}$$

where $h(\epsilon)$ denotes the binary entropy function $h(\epsilon) = -\epsilon \log_2 \epsilon - (1-\epsilon)\log_2(1-\epsilon)$. In addition, if Bob and Alice can exchange information over another public and insecure channel (messages on this channel may be known to Eve), the secret channel capacity becomes [5]

$$C_2 = h(\epsilon + \delta - 2\epsilon\delta) - h(\epsilon), \tag{2}$$

which is always positive unless $\epsilon = 0.5$ or $\delta = 0.5$. Unfortunately, if noise is considered as the only source of receiving error, it is possible for then Eve's error rate δ to be much less than Bob's ϵ. In this case, C_1 is zero whereas C_2 becomes too small to be useful.

2.2. MIMO transmission model

We consider the case where Alice transmits to Bob using J transmitting antennas. Bob uses K receiving antennas. During this procedure, Bob extracts a secret key whereas Eve should be deprived of signal interception capability, as illustrated in Fig. 1.

In this paper, we consider the secret channel only, where Alice transmits a secret sequence to Bob using the J transmitting antennas. Other than that, they may communicate with each other using another insecure public channel shown in Fig. 1. Though messages in the insecure channel may be known to Eve, this channel guarantees a higher secret channel capacity (2)

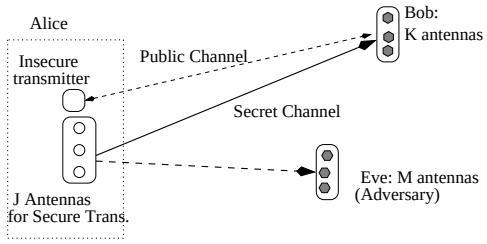

Fig. 1. System model for secret-key agreement between Alice and Bob.

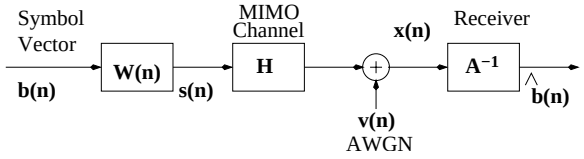

Fig. 2. Block diagram of secure MIMO transmission.

than (1). Obviously, it can also be used to transmit encrypted data for higher efficiency, but with keys generated by the secret channel. In addition, Alice can use an extra transmitting antenna for piloting purpose, i.e., for Bob to synchronize in both time and frequency.

In the secret channel, a MIMO transmission scheme shown in Fig. 2 is used by Alice and Bob. A vector symbol sequence $\{\mathbf{b}(n)\}$ is processed by Alice with a corresponding matrix sequence $\{\mathbf{W}(n)\}$, which gives the transmitted signal vector sequence $\{\mathbf{s}(n)\}$. Specifically, we have

$$\mathbf{s}(n) = \mathbf{W}(n)\mathbf{b}(n), \tag{3}$$

where $\mathbf{s}(n)$, $\mathbf{W}(n)$, and $\mathbf{b}(n)$ have dimensions $J \times 1$, $J \times K$ and $K \times 1$, respectively.

The signal vector $\mathbf{s}(n)$ is transmitted through the J transmitting antennas in the n^{th} symbol interval. Assume the propagation channel be Rayleigh flat fading. The signal received by Bob is

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{v}(n), \tag{4}$$

where $\mathbf{x}(n)$ is the $K \times 1$ received sample vector, $\mathbf{v}(n)$ denotes the $K \times 1$ AWGN vector with zero-mean. The channel matrix $\mathbf{H}$ is $K \times J$, whose coefficients are independent complex circular symmetric Gaussian distributed with zero-mean and unit variance.

We assume that $\mathbf{H}$ is block fading [2], i.e., it is constant or slowly time-varying during the transmission of a block of symbol vectors, but may change randomly between blocks. The symbols (i.e., elements in $\mathbf{b}(n)$) are independent uniformly distributed with zero-mean and unit variance. The transmission power is thus determined by the processing weights $\mathbf{W}(n)$.

Eve may use multiple receiving antennas for better interception, and the interception becomes much easier when the propagation is flat-fading. Therefore, we consider the worst case (to Alice and Bob) where Eve receives signals from M receiving antennas,

$$\mathbf{x}_u(n) = \mathbf{H}_u\mathbf{s}(n) + \mathbf{v}_u(n), \tag{5}$$

where $\mathbf{x}_u(n)$ and $\mathbf{H}_u$ are with dimension $M \times 1$ and $M \times J$, respectively. The vector $\mathbf{v}_u(n)$ is AWGN with zero-mean and covariance matrix $\sigma_v^2\mathbf{I}_M$. The notations are similar to (4) except that $(\cdot)_u$ is used to denote the unauthorized user Eve. Each element of $\mathbf{H}_u$ has the same distribution as, but is independent from, those of $\mathbf{H}$. From the extensive studies on MIMO channels, we know that as long as the distance between Bob and Eve is larger than several carrier wavelengths, then their channels can be considered as independent.

Eve does not know the channels $\mathbf{H}$ and $\mathbf{H}_u$, but she may try blind or non-blind methods to estimate $\mathbf{H}_u$ from $\mathbf{x}_u(n)$. Alice and Bob do not know both channels either, and in particular, they can not estimate $\mathbf{H}_u$.

Note that although we study one direction secrecy only, the same scheme can be set up on the other direction. Or, as one direction is secured, the other direction can easily be secured too. In addition, although we consider a single-point to single-point transmission in this paper, the scheme also fits single-point to multi-point transmission (multi-user communication or broadcasting) scenarios. For each specific user, all other ones can be treated as adversaries because different users have different channels.

3. MIMO TRANSMISSION PROCEDURE

Since we can not depend on noise as the only source of receiving error for Eve, we create some intentional difference between Bob and Eve's signals by exploiting the redundancy of antenna array transmissions. Traditionally, such redundancy is used completely to optimize transmission efficiency. When some of the redundancy is used for secrecy, the optimal transmission efficiency may not be available. This indicates

the fundamental trade-off between the transmission efficiency and secrecy. Our objective is not to make our scheme competing against the traditional encryption-based schemes in terms of transmission efficiency. But rather, we are interested in information-theoretic secrecy for a security level higher than the latter. Nevertheless, the efficiency loss can be small.

3.1. Transmission and receiving from Alice to Bob

As shown in [6], a valid way to guarantee a high δ is to prevent Eve from channel estimation. In terms of channel estimation, Bob has no advantage over Eve. Therefore, our objective is to design a transmission scheme so that Bob can detect signals without channel knowledge, which can be realized by shifting the channel estimation task from the receiver Bob to the transmitter Alice. Once Alice has the channel knowledge, she can adjust the MIMO transmission so that Bob does not need to estimate channel in order for symbol estimation.

There are various ways for Alice to estimate the channel $\mathbf{H}$ [6]. We use the reciprocity of the forward and backward channels in this paper. Bob first transmits a pilot signal to Alice using the same carrier frequency as the secret channel, during which Alice can estimate the backward channel, and use it for array transmission. Note that this procedure is required only once as an initialization, and gives no useful information to Eve.

From (3)(4), Alice design the matrix $\mathbf{W}(n)$ so that

$$\mathbf{HW}(n) = \mathbf{A}, \tag{6}$$

where $\mathbf{A}$ is a $K \times K$ diagonal matrix determined by Alice but unknown to Bob and Eve. The key point is to make $\mathbf{A}$ to have positive diagonal elements so that Bob can estimate them easily. Without loss of generality, we assume that $\mathbf{H}$ is full row rank. Otherwise the system simply reduces to the one with a smaller K.

From the received signal $\mathbf{x}(n) = \mathbf{A}\mathbf{b}(n) + \mathbf{v}(n)$, Bob can easily detect signals as

$$\hat{\mathbf{b}}(n) = \mathbf{A}^{-1}\mathbf{x}(n). \tag{7}$$

where $\mathbf{A}$ can be estimated from the received signal power because $\mathbf{A}$ is diagonal with positive elements. Since no channel estimation is required, it is not necessary for Alice to transmit training sequences. As a result, Eve has no training available either, and has to rely on blind deconvolution.

3.2. Transmission weights design

Although (6) looks similar to transmit beamforming, the major difference is that $\mathbf{W}(n)$ changes randomly in each symbol interval n so that $\mathbf{H}_u\mathbf{W}(n)$ becomes random, which prevents Eve from channel/symbol estimation. This can be realized by selecting randomly the elements of $\mathbf{W}(n)$ while satisfying the constraint (6). Obviously, we need $J > K$, i.e., more transmitting antennas than receiving antennas.

The criteria for designing $\mathbf{W}(n)$ include satisfying (6), preventing Eve from detecting $\mathbf{b}(n)$ based on (5), limiting the total transmission power, balancing the transmission power among the transmitting antennas, and finally, being computationally efficient.

Considering those criteria, we use a procedure similar to [6] to determine $\mathbf{W}(n)$. We first select randomly K columns from $\mathbf{H}$ to form a $K \times K$ submatrix $\mathbf{H}_0$ (with sufficiently large $\|\mathbf{H}_0\|$). Without loss of generality, we assume that the first K columns of $\mathbf{H}$ are chosen, i.e.,

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_1 \end{bmatrix}. \tag{8}$$

We can subdivide the matrix $\mathbf{W}(n)$ accordingly into $\mathbf{W}_0(n)$ and $\mathbf{W}_1(n)$ so that (6) becomes

$$\mathbf{H}_0\mathbf{W}_0(n) + \mathbf{H}_1\mathbf{W}_1(n) = \mathbf{A}. \tag{9}$$

Then the MIMO processing matrix $\mathbf{W}(n)$ is

$$\mathbf{W}(n) = \begin{bmatrix} \mathbf{H}_0^{-1}[\mathbf{A} - \mathbf{H}_1\mathbf{W}_1(n)] \\ \mathbf{W}_1(n) \end{bmatrix}. \tag{10}$$

Therefore, the transmission weights design procedure is to first select $\mathbf{H}_0$, then generate randomly $\mathbf{W}_1(n)$, and finally calculate $\mathbf{W}(n)$.

The computational complexity is $O(J^3)$ per transmission. Nevertheless, the matrix inversion is required only once for each channel realization. Transmission power can be adjusted by choosing $\mathbf{W}_1(n)$, $\mathbf{A}$ and selecting $\mathbf{H}_0$ appropriately.

As a comparison, the optimal $\mathbf{W}(n)$ in the traditional MIMO is the so-called eigen-beamforming. In this case,

$$\mathbf{W}_{\text{opt}}(n) = \mathbf{V}_{\text{opt}} \begin{bmatrix} \mathbf{D}_{\text{opt}}^{-1}\mathbf{U}_{\text{opt}}^{H}\mathbf{A}_{\text{opt}} \\ \mathbf{B} \end{bmatrix}, \tag{11}$$

with the singular-value decomposition (SVD) $\mathbf{H} = \mathbf{U}_{\rm opt}[\mathbf{D}_{\rm opt},\ \mathbf{0}]\mathbf{V}_{\rm opt}^H$, $\mathbf{A}_{\rm opt} = \mathbf{I}_K/\sqrt{\mathrm{tr}(\mathbf{D}^{-2})}$, and $\mathbf{B} = \mathbf{0}$. Unfortunately, even though $\mathbf{B}$ can be random, this design is susceptible to the blind deconvolution of Eve, similarly as shown in [6].

4. TRANSMISSION SECRECY

4.1. A randomization scheme

Since the matrix $\mathbf{W}_1(n)$ is with dimension $(J-K)\times K$, there are $K(J-K)$ degrees of freedom in designing $\mathbf{W}(n)$, which can be exploited to randomize $\mathbf{W}(n)$ so that the transmitted signal vector $\mathbf{s}(n)$ has distribution unknown to Eve (and Bob). It is well known that blind deconvolution requires some *a priori* knowledge about the statistics of the transmitted sequence $\{\mathbf{s}(n)\}$, such as independence, non-Gaussian distribution, and/or distinct power spectral [6]. From (10), we can choose $\mathbf{W}_1(n)$ appropriately to violate all such conditions in order to prevent Eve from blind deconvolution.

Nevertheless, Eve knows that Alice and Bob depends on (6) for secret transmission although she does not know $\mathbf{H}$ and the diagonal elements of $\mathbf{A}$. This makes it non-trivial to design and analyze the secret transmission.

We propose that Alice simply designs $\mathbf{W}_1(n)$ such that $\mathbf{s}_1(n) = \mathbf{W}_1(n)\mathbf{b}(n)$ is $(J-K)$-variate Gaussian distributed [8] with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, i.e., $\mathbf{s}_1(n) \sim \mathcal{N}_{J-K}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$E[\mathbf{s}_1(n)] = \boldsymbol{\mu}, E[(\mathbf{s}_1(n)-\boldsymbol{\mu})(\mathbf{s}_1(n)-\boldsymbol{\mu})^H] = \boldsymbol{\Sigma}. \tag{12}$$

In particular, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ can be made unknown to both Eve and Bob.

From (10) and (3), we have

$$\mathbf{s}(n) = \left[\begin{array}{c} \mathbf{s}_0(n) \\ \mathbf{s}_1(n) \end{array}\right] = \left[\begin{array}{c} \mathbf{H}_0^{-1}[\mathbf{A}\mathbf{b}(n) - \mathbf{H}_1\mathbf{s}_1(n)] \\ \mathbf{s}_1(n) \end{array}\right]. \tag{13}$$

The total transmission power is

$$\begin{aligned}\mathrm{tr}\{E[\mathbf{s}(n)\mathbf{s}^H(n)]\} &= \mathrm{tr}\{\boldsymbol{\mu}\boldsymbol{\mu}^H + \boldsymbol{\Sigma}\}+ \\ &\mathrm{tr}\{\mathbf{H}_0^{-1}[\mathbf{A}\mathbf{A}^H + \mathbf{H}_1(\boldsymbol{\mu}\boldsymbol{\mu}^H + \boldsymbol{\Sigma})\mathbf{H}_1^H]\mathbf{H}_0^{-H}\},\end{aligned} \tag{14}$$

whereas the diagonal entry of $E[\mathbf{s}(n)\mathbf{s}^H(n)]$ gives the transmission power of each antenna. We need both to reduce the total power and to balance the power among the transmitting antennas. This can be conducted by choosing properly $\mathbf{A}$, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Details will be reported elsewhere due to space limit. Instead, we focus on the transmission secrecy analysis in this paper.

4.2. Indeterminacy of Eve's blind deconvolution

Since $\mathbf{s}_1(n)$ is multi-variate Gaussian $\mathcal{N}_{J-K}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, from (13), for a given symbol vector $\mathbf{b}(n)$, the signal vector

$$\begin{aligned}\mathbf{s}_0(n) &= -\mathbf{H}_0^{-1}\mathbf{H}_1\mathbf{s}_1(n) + \mathbf{H}_0^{-1}\mathbf{A}\mathbf{b}(n) \\ &\triangleq \mathbf{G}\mathbf{s}_1(n) + \mathbf{g}\end{aligned} \tag{15}$$

is K-variate Gaussian $\mathcal{N}_K(\mathbf{G}\boldsymbol{\mu} + \mathbf{g}, \mathbf{G}\boldsymbol{\Sigma}\mathbf{G}^H)$.

From (5) and (13), Eve's received signal becomes

$$\mathbf{x}_u(n) = \left[\begin{array}{cc} \mathbf{H}_u\mathbf{F} & \mathbf{I}_M \end{array}\right] \left[\begin{array}{c} \mathbf{s}_1(n) \\ \mathbf{v}_u(n) \end{array}\right] + \mathbf{H}_u\mathbf{f}, \tag{16}$$

where

$$\mathbf{F} = \left[\begin{array}{c} \mathbf{G} \\ \mathbf{I}_{J-K} \end{array}\right], \quad \mathbf{f} = \left[\begin{array}{c} \mathbf{g} \\ \mathbf{0} \end{array}\right]. \tag{17}$$

Obviously, for a given symbol vector $\mathbf{b}(n)$,

$$\left[\begin{array}{c} \mathbf{s}_1(n) \\ \mathbf{v}_u(n) \end{array}\right] \sim \mathcal{N}_{M+J-K}\left(\left[\begin{array}{c} \boldsymbol{\mu} \\ \mathbf{0} \end{array}\right], \left[\begin{array}{cc} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \sigma_v^2\mathbf{I}_M \end{array}\right]\right). \tag{18}$$

Then Eve's signal is M-variate Gaussian distributed,

$$\begin{aligned}&\mathbf{x}_u(n) \sim \\ &\mathcal{N}_M\left(\mathbf{H}_u\mathbf{F}\boldsymbol{\mu} + \mathbf{H}_u\mathbf{f}, \mathbf{H}_u\mathbf{F}\boldsymbol{\Sigma}\mathbf{F}^H\mathbf{H}_u^H + \sigma_v^2\mathbf{I}_M\right)\end{aligned} \tag{19}$$

Since the distribution (19) depends on $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, the unknown $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ contribute to the indeterminacy of $\mathbf{H}_u$.

Proposition 1. For an unknown symbol vector $\mathbf{b}(n)$ and unknown $\mathbf{H}$, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, from the distribution of $\mathbf{x}_u(n)$, the channel matrix $\mathbf{H}_u$ is indistinguishable from $\mathbf{H}_u\mathbf{P}$ with a $J\times J$ matrix

$$\mathbf{P} = \left[\begin{array}{cc} \mathbf{U} & \mathbf{G}\mathbf{V} - \mathbf{U}\mathbf{G} \\ \mathbf{0} & \mathbf{V} \end{array}\right], \tag{20}$$

where $\mathbf{U}$ and $\mathbf{V}$ are arbitrary nonsingular $K\times K$ and $(J-K)\times(J-K)$ matrices, respectively.

Proof. From (15) and (17), we redefine the distribution (19) as a function $f(\mathbf{H}_u, \mathbf{H}_0, \mathbf{H}_1, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. First,

for any matrices $\mathbf{U}$ and $\mathbf{H}_0$, there exists nonsingular matrix $\mathbf{Z}$ such that $\mathbf{U}^{-1}\mathbf{H}_0^{-1} = \mathbf{H}_0^{-1}\mathbf{Z}$, which can be verified by simply letting $\mathbf{Z} = \mathbf{H}_0\mathbf{U}^{-1}\mathbf{H}_0^{-1}$.

Define

$$\begin{aligned}\tilde{\mathbf{F}} &= \begin{bmatrix} -\mathbf{U}^{-1}\mathbf{H}_0^{-1}\mathbf{Z}^{-1}\mathbf{H}_1 \\ \mathbf{I}_{J-K} \end{bmatrix}, \\ \tilde{\mathbf{f}} &= \begin{bmatrix} \mathbf{U}^{-1}\mathbf{H}_0^{-1}\mathbf{A}\mathbf{b}(n) \\ \mathbf{0} \end{bmatrix} \\ \tilde{\boldsymbol{\mu}} &= \mathbf{V}^{-1}\boldsymbol{\mu}, \quad \tilde{\boldsymbol{\Sigma}} = \mathbf{V}^{-1}\boldsymbol{\Sigma}\mathbf{V}^{-H}.\end{aligned}$$

Then we can verify that

$$\begin{aligned}&\mathbf{H}_u\mathbf{P}\tilde{\mathbf{F}}\tilde{\boldsymbol{\mu}} \\ &= \mathbf{H}_u\mathbf{P}\begin{bmatrix} -\mathbf{U}^{-1}\mathbf{H}_0^{-1}\mathbf{Z}^{-1}\mathbf{H}_1 \\ \mathbf{I}_{J-K} \end{bmatrix}\mathbf{V}^{-1}\boldsymbol{\mu} \\ &= \mathbf{H}_u\begin{bmatrix} -\mathbf{H}_0^{-1}\mathbf{H}_1\mathbf{V} \\ \mathbf{V} \end{bmatrix}\mathbf{V}^{-1}\boldsymbol{\mu} = \mathbf{H}_u\mathbf{F}\boldsymbol{\mu}.\end{aligned}$$

Similarly,

$$\mathbf{H}_u\mathbf{P}\tilde{\mathbf{f}} = \mathbf{H}_u\mathbf{P}\begin{bmatrix} \mathbf{U}^{-1}\mathbf{H}_0^{-1}\mathbf{A}\mathbf{b}(n) \\ \mathbf{0} \end{bmatrix} = \mathbf{H}_u\mathbf{f}.$$

In addition, it is easy to show that

$$\begin{aligned}&\mathbf{H}_u\mathbf{P}\tilde{\mathbf{F}}\tilde{\boldsymbol{\Sigma}}\tilde{\mathbf{F}}^H\mathbf{P}^H\mathbf{H}_u^H \\ &= \mathbf{H}_u\mathbf{F}\mathbf{V}\tilde{\boldsymbol{\Sigma}}\mathbf{V}^H\mathbf{F}^H\mathbf{H}_u^H = \mathbf{H}_u\mathbf{F}\boldsymbol{\Sigma}\mathbf{F}^H\mathbf{H}_u^H.\end{aligned}$$

Then we find that

$$\begin{aligned}&f(\mathbf{H}_u\mathbf{P}, \mathbf{H}_0\mathbf{U}, \mathbf{Z}^{-1}\mathbf{H}_1, \mathbf{V}^{-1}\boldsymbol{\mu}, \mathbf{V}^{-1}\boldsymbol{\Sigma}\mathbf{V}^{-H}) \\ &= f(\mathbf{H}_u, \mathbf{H}_0, \mathbf{H}_1, \boldsymbol{\mu}, \boldsymbol{\Sigma}).\end{aligned}$$

Therefore, from the distribution of $\mathbf{x}_u(n)$, Eve can not discriminate between $\mathbf{H}_u$ and $\mathbf{H}_u\mathbf{P}$. □

As a result, there is ambiguity of a $K \times K$ matrix $\mathbf{U}$ and a $(J-K)\times(J-K)$ matrix $\mathbf{V}$ for Eve's blind channel estimation. This ambiguity is also reflected in symbol vector estimation.

Proposition 2. Assume $\mathbf{x}_u(n)$ is generated by transmitting $\mathbf{b}(n)$. Then $\mathbf{x}_u(n)$ has identical distribution as those generated by transmitting any other symbol vector $\mathbf{d}(n)$.

Proof. The distribution of $\mathbf{x}_u(n)$ is described by (19), where the only information available about $\mathbf{b}(n)$ is the mean vector in terms of $\mathbf{H}_u\mathbf{f}$. Since $\mathbf{H}_u$ is indistinguishable from $\mathbf{H}_u\mathbf{P}$, $\mathbf{H}_u\mathbf{f}$ is indistinguishable from

$$\mathbf{H}_u\mathbf{P}\begin{bmatrix} \mathbf{g} \\ \mathbf{0} \end{bmatrix} = \mathbf{H}_u\begin{bmatrix} \mathbf{U}\mathbf{H}_0^{-1}\mathbf{A}\mathbf{b}(n) \\ \mathbf{0} \end{bmatrix}.$$

Since $\mathbf{U}$ is arbitrary (nonsingular), there exists some matrix $\tilde{\mathbf{U}}$ such that $\tilde{\mathbf{U}}\mathbf{H}_0^{-1}\mathbf{A}\mathbf{d}(n) = \mathbf{U}\mathbf{H}_0^{-1}\mathbf{A}\mathbf{b}(n)$. This can be easily verified from the fact that for any two vectors $\mathbf{x} = \mathbf{H}_0^{-1}\mathbf{A}\mathbf{b}(n)$ and $\mathbf{y} = \mathbf{H}_0^{-1}\mathbf{A}\mathbf{d}(n)$, we can of course find two matrices $\mathbf{U}$ and $\tilde{\mathbf{U}}$ such that $\mathbf{U}\mathbf{x} = \tilde{\mathbf{U}}\mathbf{y}$. Therefore, the distribution of $\mathbf{x}_u(n)$ is identical whether $\mathbf{b}(n)$ or $\mathbf{d}(n)$ is transmitted. □

Note that the distribution defined above is due to the random $\mathbf{W}_1(n)$, conditioned on the symbol vector, i.e., $P[\mathbf{x}_u(n)|\mathbf{b}(n)] = P[\mathbf{x}_u(n)|\mathbf{d}(n)]$. When considering random sequences $\{\mathbf{b}(n)\}$ and $\{\mathbf{d}(n)\}$, we have $P[\{\mathbf{x}_u(n)\}|\{\mathbf{b}(n)\}] = P[\{\mathbf{x}_u(n)\}|\{\mathbf{d}(n)\}]$ because the symbol vectors are i.i.d.

With the optimal MAP detector, Eve can estimates $\{\mathbf{b}(n)\}$ from $\arg\max P[\{\mathbf{b}(n)\}|\{\mathbf{x}_u(n)\}]$. Because

$$\begin{aligned}&P[\{\mathbf{b}(n)\}|\{\mathbf{x}_u(n)\}] \\ &= P[\{\mathbf{x}_u(n)\}|\{\mathbf{b}(n)\}]P[\{\mathbf{b}(n)\}]/P[\{\mathbf{x}_u(n)\}] \\ &= P[\{\mathbf{x}_u(n)\}|\{\mathbf{d}(n)\}]P[\{\mathbf{d}(n)\}]/P[\{\mathbf{x}_u(n)\}] \\ &= P[\{\mathbf{d}(n)\}|\{\mathbf{x}_u(n)\}], \qquad (21)\end{aligned}$$

Eve can not estimate symbols. Therefore, her receiving error rate is $\epsilon = 0.5$.

On the other hand, Bob's receiving error rate δ is identical to the optimal MIMO eigen-beamforming if $\mathbf{A} = \mathbf{A}_{\text{opt}}$ is used. The cost is higher transmission power. Based on such ϵ and δ, we may quantify the secret channel capacity using either (1) or (2). Due to space limits, such results will be reported elsewhere.

The only way left for Eve is an exhaustive search of all possible channels $\mathbf{H}_u$. In fact, she just needs to guess a $K \times K$ complex detection matrix, which gives a complexity of q^{2K^2} with quantization level q. For example, the complexity of exhaustive search can be at least 2^{128} for $K = 4$ and $q = 16$ in order to obtain meaningfully low BER [6].

5. SIMULATIONS

In this section, we show the performance of the proposed transmission scheme in terms of bit-error-rate (BER) and secret channel capacity. For comparison purpose, we also evaluate the performance of the optimal eigen-beamforming (11). Eve is assumed to estimate symbols blindly, which is almost no different from guessing.

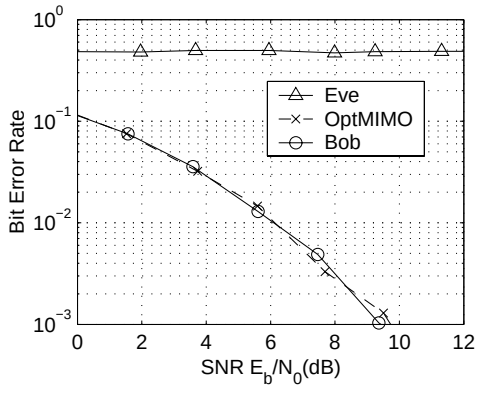

Fig. 3. Receiving performance. $J = 6, K = 4$. $\circ$: Bob in the proposed scheme. $\times$:optimal MIMO eigenbeamforming. $\triangle$: Eve in the proposed scheme.

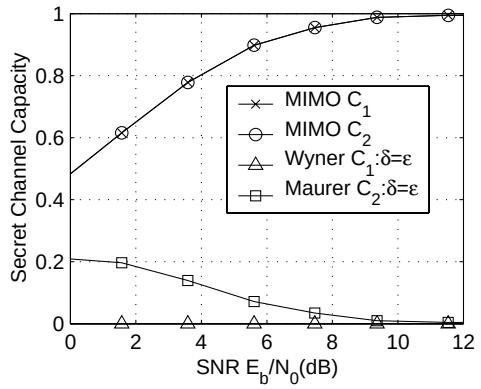

Fig. 4. Secret channel capacity. $\times$: proposed scheme with (1). $\circ$: proposed scheme with (2). $\square$:traditional transmission with (2). $\triangle$:traditional transmission with (1).

Alice transmits QPSK symbol packets. Each packet contains 400 QPSK symbols. $J = 6$ and $K = 4$ are used. We use 5000 runs to obtain each BER value. The BER results are shown in Fig. 3, from which we see that Bob can reliably receive signals while Eve can not.

Based on the BER of Bob and Eve, we can derive the secret channel capacity using (1) and (2). For comparison purpose, we also plot the secret capacity when noise is considered as the only source of error. In this case, Eve is assumed to have the same BER as Bob. As can be seen from Fig. 4, the proposal scheme gives much higher secret capacity.

6. CONCLUSIONS

In this paper, we propose to use MIMO transmissions to realize informational-theoretic secrecy. The redundancy of MIMO transmissions is exploited to create a difficult signal interception situation for the adversary. The indeterminacy of the adversary's blind deconvolution is proved, and the high error rate is shown. Sufficiently high secret channel capacity is shown by simulations. The proposed scheme indicates that physical-layer technique can assist upper-layer security designs by providing secret-key agreement with information-theoretic secrecy. It can be useful in military wireless communications where security is of primary importance.

7. REFERENCES

[1] C. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 29, pp. 656-715, 1949.

[2] A. O. Hero, III, "Secure space-time communication," *IEEE Trans. Inform. Theory*, vol. 49, no. 12, pp. 3235-3249, Dec. 2003.

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

[4] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 339-348, Mar. 1978.

[5] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 733-742, Mar. 1993.

[6] X. Li, M. Chen and P. Ratazzi, "A randomized space-time transmission scheme for secret-key agreement," *CISS'2005*, Mar. 2005.

[7] U. M. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels – Part I: definitions and a completeness result", *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 822-831, April 2003.

[8] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, 1982.