Secure Transmission Using MIMO Precoding

Chia-Hua Lin, Shang-Ho (Lawrence) Tsai, *Senior Member, IEEE*, and Yuan-Pei Lin, *Senior Member, IEEE*

***Abstract*— In this paper, we propose an MIMO precoding and postcoding transceiver to achieve data secrecy at the physical layer. When full channel state information (CSI) of the legitimate receiver is known to the transmitter, the proposed precoder can simultaneously maximize to receive signal-to-noise ratio (SNR) and attain secrecy. When only partial CSI is available, a modified Lloyd algorithm is proposed to construct codebooks for quantizing the precoder. As will be explained in this paper, the use of the proposed codebooks does not affect the secrecy of the proposed system. In addition, we propose a low-complexity postcoder to compensate the SNR loss when full CSI is not available. Furthermore, the performance of the proposed system is analyzed from the view points of both recovery rate and secrecy. Based on the analyzed results, we show how to achieve full recovery rate and perfect secrecy for the proposed system. Finally, simulation results corroborate the theoretical results, and show that the proposed system enjoys different advantages when different recovery algorithms are used.**

***Index Terms*— Physical layer secrecy, precoding and postcoding, transceiver design, compressive sensing, MIMO wiretap channel, codebook design.**

I. Introduction

DATA security has long been considered as an important issue in many applications including multimedia and communications systems. In wireless communications, security issue becomes more pronounced because the transmitted data can be accessed by some unauthorized users. Nowadays most of the communications systems encrypt data at the network layer, where key-based encryption techniques are adopted to protect data from stealing. Recently there have been several interesting results on attaining security at the *physical layer*, see *e.g.,* [1]–[7]. A main motivation for physical-layer security is that the channels for different users are generally different. The channel discrepancy can be used to encrypt data in a natural way. That is, the channel characteristics of individual users can be treated as "unique key" to encrypt confidential information. The number of keys is theoretically infinite, because the coefficients of the baseband channel are complex numbers. On the other hand, the number of keys for data encryption at the network layer is generally finite. Moreover, since the channel is time-variant in wireless environments, the key changes with time. All these properties prevent the eavesdroppers from breaking the key by extensive computations, which is an intuitive way to break the encryption key at the network layer. Key-based encryption system are also studied in coding theory. For example, [8] introduces a cryptographic system based on the difficulty of decoding linear codes, *e.g.,* Goppa code. In addition, [9] mentions that linear codes have good ability to resist quantum Fourier sampling attacks. However these schemes are channel independent.

Wyner in [11] investigated the scenario that the transmitter sends information to the legitimate receiver but the information is intercepted by an eavesdropper through a so called wiretap channel. He analyzed the information secrecy using Shannon's theory, and the results are widely used for research on physical layer security. Extending Wyner's results, the authors in [12] characterized the secrecy capacity for the non-degraded discrete memoryless wiretap channel. Several existing precoding techniques for secrecy over MIMO wiretap channels were reviewed in [13].

When the transmitter knows full channel state information (CSI) of the legitimate receiver and all the eavesdroppers, the secrecy capacity R is maximized in [2], [3], [11], [12], and [14]. That is, let the transmission rate between the transmitter and the legitimate receiver be R_b and that between the transmitter and the eavesdroppers be R_e, the value $R = R_b - R_e$ can be maximized. In this case, the eavesdroppers could still obtain certain amount of data, because $R_e > 0$. Moreover, it is somewhat impractical to assume that the transmitter knows the full CSI of the legitimate receiver and all the eavesdroppers in some applications, *e.g.,* in military communications systems, because the eavesdroppers usually steal data secretly, and would not send back their channel information. On the other hand, if the transmitter knows full CSI of the legitimate receiver but only partial CSI of the eavesdroppers, the authors in [6] propose to add artificial noise (AN) for degrading the channel quality of the eavesdroppers. The AN is designed so that it lies on the null space of the legitimate receiver's channel. Bashar et. al. in [15] analyze and present the secrecy performance of a codebook beamforming transmission and further explain the relationship between the AN and the codebook beamforming schemes.

Literature has pointed out that both encryption and data compression can be attained via compressive sensing (CS), see *e.g.,* [7] and [16]. Furthermore, the CS scheme proposed in [17] can achieve perfect secrecy *i.e.,* the mutual information is equal to zero. It is worth pointing out that these works may not be dedicated to communications systems because they do not consider the randomness characteristic of the channels

Manuscript received September 4, 2013; revised December 23, 2013; accepted February 19, 2014. Date of publication February 28, 2014; date of current version April 10, 2014. This work was supported by the National Science Council of Taiwan through the Cooperative Agreement under Grant 102-2221-E-009-017-MY3. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Kah Chan Teh.

The authors are with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: chlin.ece98g@nctu.edu.tw; shanghot@alumni.usc.edu; ypl@mail.nctu.edu.tw).

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

Digital Object Identifier 10.1109/TIFS.2014.2309211

for increasing the security. In addition, these researches do not include precoding and postcoding designs, and thus they do not consider SNR maximization for improving the recovery performance. Nevertheless, these works inspire us that communication secrecy may be attained by utilizing the CS techniques.

From the discussion above, we are motivated to design the transceivers that achieve security at the physical layer. Our goals for the transceiver are as follows: *1)* The transmitter does not need to know any CSI information of the eavesdroppers. Even if the transmitter only knows partial CSI from the legitimate receiver and no CSI from eavesdroppers, the transceiver can still achieve security. *2)* The transceiver attains perfect secrecy defined by Shannon in [18] that the mutual information between the transmitter and the eavesdroppers is zero, *i.e.,* $R_e = 0$.

In this paper, we assume that the transmitter knows only the CSI information of the legitimate user. In addition, the key or the precoding information is known only to the transmitter and the legitimate user, but not to the eavesdroppers, as in [19]–[21]. We propose a MIMO precoding and postcoding system that can achieve the goals. The proposed system can be modeled as an underdetermined linear system. Thus the recovery algorithms designed for CS can be used to reconstruct the transmitted signals. More specially, when the transmitter knows full CSI of the legitimate receiver (CSI of the eavesdroppers is not needed), we propose an optimal precoder for maximizing the instantaneous SNR. That is, the proposed precoder can simultaneously maximize the SNR as well as attain secrecy. On the other hand, when only partial CSI of the legitimate receiver is known to the transmitter, we propose a modified Lloyd algorithm that constructs codebooks using the concept of Grassmann manifold [22]. We show that using the proposed codebooks only affects the SNR performance and does not affect the secrecy of the proposed encryption system. On the other hand, to compensate the SNR loss when only partial CSI of the legitimate receiver is available, a low-complexity postcoder at the legitimate receiver is proposed. In this case a performance comparable to that with full CSI can be achieved. Furthermore, the performance of the proposed system is analyzed. The analysis is divided into two parts. First, for the legitimate receiver, for which the recovery rate is of concern, we analyze the recovery rate as a function of several system parameters. Secondly, for the eavesdroppers, for which secrecy is of concern, we use the results in [16]-[18], [23], and [24] and show how to design the proposed system to achieve perfect secrecy. It is worth to emphasize that the proposed system can enjoy perfect secrecy in each transmission. On the other hand, the systems in [6] and [15] can guarantee that the secrecy capacity is greater than a specific rate, and the performance is measured in terms of outage probability, where the eavesdroppers still have chance to obtain certain information.

Combining the analysis for both the legitimate receiver and the eavesdroppers, we show how to design the system parameters to achieve both full recovery rate and perfect secrecy. It is worth to point out that the proposed system has different advantages when the recovery algorithms with different complexity are used. Therefore, several recovery algorithms, from the simplest Orthogonal Matching Pursuit (OMP) [25], to the most complicated Dantzig selector [26], are conducted to explain these advantages. Simulation results corroborate the theoretical results.

The paper is organized as follows. Section II introduces the system model for the legitimate receiver and the eavesdroppers separately. Section III shows how to design the precoders for maximizing the instantaneous received SNR and how this precoder achieves the optimal performance if the transmitter knows full CSI. On the other hand, if only partial CSI is available, we propose a modified Lloyd algorithm to construct the codebook. Moreover, a postcoder to compensate the performance loss due to the lack of full CSI is also proposed in this section. Section IV analyzes the performance of the proposed system considering both the recovery performance and the system secrecy, and the results provide parameter settings for practical designs. Simulation results are provided in Section V, and conclusions are made in Section VI.

Notations. All vectors are in lowercase boldface and matrices are in uppercase boldface. $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ denote the transpose, conjugate and conjugate transpose of a matrix, respectively. $\mathbf{X}^{(i,j)}$ denotes the element in the i-th row and j-th column of an matrix $\mathbf{X}$. $\mathrm{tr}(\cdot)$ is the trace of a square matrix. $\mathbb{E}\{\cdot\}$ denotes expectation. $||\cdot||_i$ with $i = \{0, 1, 2\}$ is the ℓ_i vector norm, $||\cdot||_F$ denotes the matrix Frobenius norm. $|\mathcal{S}|$ is the size of a set $\mathcal{S}$.

II. System Model and Problem Formulation

The proposed system model is introduced in this section. For legitimate receiver, we explain how to use precoding techniques to achieve encryption. On the other hand, for eavesdroppers, we explain that they could not obtain information due to the lack of precoding matrices and CSI. The legitimate receiver and eavesdroppers are discussed separately as follows:

A. Problem Formulation for the Legitimate Receiver

The block diagram of the proposed system is shown in Fig. 1. At the first stage, each of the elements of a $K \times 1$ symbol vector $\mathbf{x}$ is randomly allocated to K elements of an $L \times 1$ vector $\mathbf{s}$; the other $L - K$ elements are inserted zeros. Assume $L > K$, $\mathbf{s}$ is therefore a sparse vector. For example, let $K = 3$ and the elements of the symbol vector $\mathbf{x} \in \{-1, 1\}$. If $L = 35$ and $\mathbf{x} = \begin{bmatrix} 1 & -1 & -1 \end{bmatrix}^T$, by randomly allocating the elements of $\mathbf{x}$ to $\mathbf{s}$, a possible $\mathbf{s}$ can be $\mathbf{s} = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 & \cdots & 0 & -1 \end{bmatrix}^T$. A vector $\mathbf{s}$ which only has K nonzero elements is usually called K-sparsity, *i.e.,* $\|\mathbf{s}\|_0 = K$. The bit rate of $\mathbf{s}$ is defined as

$$I_r = \frac{\log_2 2^K \binom{L}{K}}{L} = \frac{K + \log_2 \binom{L}{K}}{L}. \tag{1}$$

Let N_t and N_r be the numbers of transmit and receive antennas respectively, and let $N_t > N_r$. The complex MIMO channel $\mathbf{H}_c \in \mathbb{C}^{N_r \times N_t}$ is assumed to be independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$; therefore, the magnitude of the

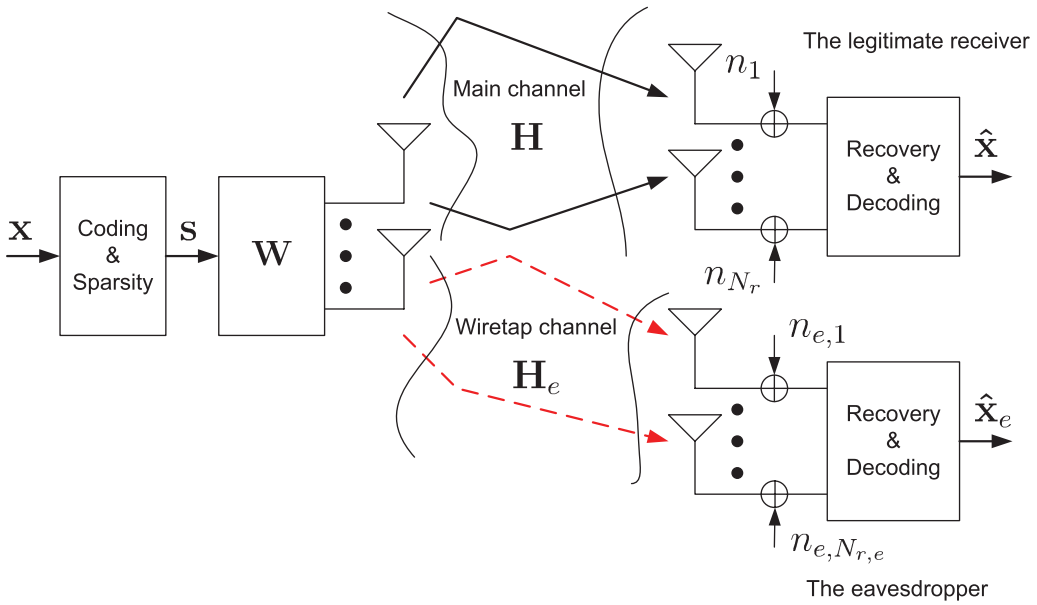

Fig. 1. The proposed system with MIMO wiretap channel.

channel coefficients is Rayleigh distributed. The encryption problem for the legitimate receiver can then be formulated as

$$\mathbf{y}_c = \mathbf{H}_c\mathbf{W}_c\mathbf{s} + \mathbf{n}_c, \tag{2}$$

where $\mathbf{n}_c \in \mathbb{C}^{N_r\times 1}$ is a complex noise vector whose entries are i.i.d. $\mathcal{CN}(0, \sigma_n^2)$, and $\mathbf{W}_c \in \mathbb{C}^{N_t\times L}$ is a complex precoder. For the proposed system, we choose $L > 2N_r$ so that (2) becomes an underdetermined linear model and $\Psi = \mathbf{H}_c\mathbf{W}_c \in \mathbb{C}^{N_r\times L}$ is called the *sensing matrix*.

Since $\mathbf{s}$ is a sparse signal and (2) is an underdetermined linear model, $\mathbf{s}$ can be recovered by using the CS recovery techniques if the sensing matrix Ψ satisfies the RIP (see [29], [30]), which is defined as follows,

Definition 1 *Given the index sets $\mathcal{I} \subset \{1, 2, \ldots, L\}$, a vector $\mathbf{a} \in \mathbb{R}^{|\mathcal{I}|}$ and a matrix $\Psi \in \mathbb{R}^{M\times L}$, the matrix Ψ is said to satisfy the Restricted Isometry Property (RIP) with parameter (K, δ) and $K \le M$, if the following inequality holds.*

$$(1-\delta)\|\mathbf{a}\|_2^2 \le \|\Psi_{\mathcal{I}}\mathbf{a}\|_2^2 \le (1+\delta)\|\mathbf{a}\|_2^2,$$

where $0 \le \delta \le 1$, and $\Psi_{\mathcal{I}}$ consists of the columns of Ψ with indices $\mathcal{I}$ and $|\mathcal{I}| \le K$.

For systems satisfying the RIP, the recovery algorithm such as linear programming (LP) can be used for solving the ℓ_1 optimization problem, and yielding an exact solution in noiseless channels. A popular family of sensing matrices is the $M \times L$ (real- or complex-valued) random matrices, which satisfy the RIP and lead to a high probability of recovery rate. This paper uses random matrices with i.i.d. entries. The distribution of the entries can be Gaussian or Bernoulli distribution with zero mean and variance $1/L$. It is mentioned in [31] that if the entries of the sensing matrices are generated in this way, the RIP holds and the underdetermined linear model can be perfectly recovered using the ℓ_1 optimization solutions whenever

$$K \le \beta\frac{M}{\ln(L/M)}, \tag{3}$$

where β is a constant, and now $M = N_r$. In current communication systems, the number N_r of receive antennas is generally not large enough to make Ψ meet the RIP in (3) for a moderate K. From (3), the number K of sparsity increases as M increases. The reason that we use the real-value system is because currently most of the mature research results in compressive sensing are developed based on the real-value system. Therefore we transformed the complex-value system into real-value system to best utilize the existing results of the compressive sensing. The vector $\mathbf{s}$ is repeatedly transmit by T times, which are described separately as follows:

We reformulate the complex-valued system into a real-valued system. More specifically, by rearranging (2), we have

$$\underbrace{\begin{bmatrix}\Re\{\mathbf{y}_c\}\\ \Im\{\mathbf{y}_c\}\end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix}\Re\{\mathbf{H}_c\} & -\Im\{\mathbf{H}_c\}\\ \Im\{\mathbf{H}_c\} & \Re\{\mathbf{H}_c\}\end{bmatrix}}_{\mathbf{H}}\underbrace{\begin{bmatrix}\Re\{\mathbf{W}_c\}\\ \Im\{\mathbf{W}_c\}\end{bmatrix}}_{\mathbf{W}}\mathbf{s} + \underbrace{\begin{bmatrix}\Re\{\mathbf{n}_c\}\\ \Im\{\mathbf{n}_c\}\end{bmatrix}}_{\mathbf{n}}, \tag{4}$$

where $\mathbf{y} \in \mathbb{R}^{2N_r\times 1}$, $\mathbf{H} \in \mathbb{R}^{2N_r\times 2N_t}$, $\mathbf{W} \in \mathbb{R}^{2N_t\times L}$ and $\mathbf{n} \in \mathbb{R}^{2N_r\times 1}$. Now the sensing matrix is $\Phi = \mathbf{HW} \in \mathbb{R}^{2N_r\times L}$, and its number of rows is doubled compared to the complex-valued system. According to the results in [10], transforming the complex-valued system into the real-valued system does not degrade the ability of carrying sparsity if the condition of the RIP property is satisfied. As a result, the number of sparsity K increases. The real-valued precoder $\mathbf{W}$ in (4) can be expressed as

$$\mathbf{W} = \begin{bmatrix}\mathbf{P}_1 & \mathbf{P}_2 & \cdots & \mathbf{P}_\alpha\end{bmatrix}, \tag{5}$$

where $\mathbf{P}_i \in \mathbb{R}^{2N_t\times R}$, $1 \le i \le \alpha$ is a sub-precoder and R is the rank of the channel matrix $\mathbf{H}$. α is a positive constant and

is defined as

$$\alpha = \left\lfloor \frac{L}{R} \right\rfloor.$$

Note that α should be designed to satisfy the RIP in (3), and we will explain later once $\mathbf{P}_1$ is determined, the other sub-precoders $\mathbf{P}_i, i \neq 1$, can be determined from $\mathbf{P}_1$ easily. Assume that $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = (K/L)\sigma_s^2\mathbf{I}_L$, where σ_s^2 is the variance of the sparse vector $\mathbf{s}$. From (4), the instantaneous signal-to-noise ratio (SNR) γ is defined as

$$\gamma = \frac{1}{\sigma_n^2}\frac{\mathbb{E}\{\|\mathbf{H}\mathbf{W}\mathbf{s}\|_2^2\}}{KN_t}. \tag{6}$$

The proposed system could be regarded as a single-stream MIMO system. Thus using SNR as a design criteria is reasonable. The extension to the multi-stream MIMO systems would need to consider the design criterion such as maximizing sum-rate (throughput) or minimizing bit-error-rate (BER), which are not well developed in the field of compressive sensing. Therefore, this problem is still open.

The proposed system repeatedly transmits the same sparse vector using different precoders to increase the number of rows of the sensing matrix. Let the repeating number be T. From (4), the proposed system with repeated transmission can be formulated as

$$\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_T \end{bmatrix} = \overline{\mathbf{y}} = \begin{bmatrix} \mathbf{H}_1 & & \\ & \ddots & \\ & & \mathbf{H}_T \end{bmatrix} \begin{bmatrix} \mathbf{W}_1 \\ \vdots \\ \mathbf{W}_T \end{bmatrix} \mathbf{s} + \begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_T \end{bmatrix}$$
$$= \overline{\mathbf{H}}\overline{\mathbf{W}}\mathbf{s} + \bar{\mathbf{n}} = \overline{\Phi}\mathbf{s} + \bar{\mathbf{n}}, \tag{7}$$

where $\overline{\mathbf{H}} \in \mathbb{R}^{2TN_r \times 2TN_t}$ is a block diagonal matrix with different $2N_r \times 2N_t$ real channel matrices on the diagonal, $\overline{\mathbf{W}} \in \mathbb{R}^{2TN_t \times L}$ is an equivalent precoder with repeating factor T, and $\overline{\Phi} \in \mathbb{R}^{2TN_r \times L}$ is an equivalent sensing matrix. In a slow fading environment, the repeated sparse vectors of $\mathbf{s}$ may experience similar MIMO channels. This does not affect the RIP of the proposed system, and may not affect the long-term time average recovery performance. However this may affect the short-term time average performance, *e.g.,* channel with serious fading and thus resulting in poor performance during this period. To gain time diversity for the short-term time average performance, the ith repeated vector of the jth data block can be transmitted via a uniformly interlaced way at time index $t(i-1)+j$, where t is the channel coherent time and $i = 1, 2, \ldots, T$. The penalty is that the decoding latency becomes long, which is limited by the channel coherent time. Now $M = 2TN_r$, by properly choosing T, the sensing matrix satisfies (3) and can recover a sparse vector with high probability. Due to the repeating transmission, the equivalent instantaneous SNR Γ can be expressed as

$$\Gamma = \frac{1}{\sigma_n^2}\frac{\mathbb{E}\{\|\overline{\mathbf{H}}\overline{\mathbf{W}}\mathbf{s}\|_2^2\}}{TKN_t}. \tag{8}$$

Similar to (1), the bit rate with repeating transmission becomes

$$I_R = \frac{I_r}{T} = \frac{\log_2 2^K \binom{L}{K}}{TL} = \frac{K + \log_2 \binom{L}{K}}{TL}. \tag{9}$$

We will introduce the design criterion and method for the precoder $\overline{\mathbf{W}}$ later in Sec. III.

B. Problem Formulation for Eavesdropper

Communication security over wiretap channel is a well-established area. A classical problem in this field is the Wyner wiretap channel [11]. Recently communication security is investigated in wireless MIMO environments (see [32]). The wiretap channel may be applied to the proposed system and the overall system is shown in Fig. 1, where the communications is eavesdropped. To steal data, the eavesdroppers need to know the repetition number T, and the numbers of transmit and receive antennas, *i.e.* N_t and N_r respectively. This increases the decoding effort for eavesdroppers to obtain these parameters before reconstructing the received signals. If the eavesdropper knows T, N_t and N_r, the received signal $\mathbf{y}_e \in \mathbb{R}^{2N_r \times 1}$ of the eavesdroppers can be represented as

$$\mathbf{y}_e = \mathbf{H}_e\mathbf{W}\mathbf{s} + \mathbf{n}_e = \Phi_e\mathbf{s} + \mathbf{n}_e, \tag{10}$$

where $\Phi_e = \mathbf{H}_e\mathbf{W}$ is the equivalent sensing matrix to the eavesdroppers. The subscription 'e' is added to reflect the fact that the experienced channels of the eavesdroppers are not the same as the legitimate receiver. When T is known to the eavesdroppers, the problem after repetition is formulated as

$$\bar{\mathbf{y}}_e = \overline{\mathbf{H}}_e\overline{\mathbf{W}}\mathbf{s} + \bar{\mathbf{n}}_e = \overline{\Phi}_e\mathbf{s} + \bar{\mathbf{n}}_e, \tag{11}$$

where $\overline{\Phi}_e \in \mathbb{R}^{2TN_r \times L}$ is the overall equivalent sensing matrix of the eavesdropper. The secrecy of the proposed system is attained as explained as follows: The CSI of individual users is generally different unless their positions are very close in distance, *i.e.,* within one half distance of the wavelength. For instance letting the carrier frequency be 2.3 GHz, the distance is $\lambda/2 = c/(2f) = 3 \times 10^{10}/(2 \times 2.3 \times 10^9) \approx 6.5$ cm. However if the distance is only 6.5 cm, the legitimate receiver is alert to the eavesdroppers easily. Therefore, by using the TDD scheme and the reciprocity assumption, the eavesdroppers cannot obtain the CSI between the transmitter and the legitimate receiver. As a result, it is reasonable to assume $\overline{\mathbf{H}}_e \neq \overline{\mathbf{H}}$.

Since the precoding matrix $\overline{\mathbf{W}}$ is highly related to the $\overline{\mathbf{H}}$, and is used as the sensing matrix for decoding, it is very unlikely that the eavesdroppers can reconstruct the signals without knowing the CSI of the legitimate receiver. In addition, later we will discuss that the optimal design of the precoder is not unique, and thus a channel-independent random matrix can be included into the precoder without affecting the optimality. This random matrix can be generated by a unique key (or seed) known only to the transmitter and the legitimate receiver. As a result, this property further enhances the encryption, and the eavesdroppers do not have chance to steal data without knowing the generation key (seed) or the CSI of the legitimate receiver $\overline{\mathbf{H}}$.

III. Proposed Precoder and Postcoder

In this section, we describe how to design the precoders to encrypt the transmitted information. The design criterion for the precoder is maximizing the received SNR. Moreover, the precoder designs with full CSI and partial CSI are both considered. Furthermore, for systems with partial CSI, we propose a postcoder design for improving the decoding ability of the proposed system.

A. Precoder Design for Maximizing Received SNR

The goal of designing the precoder is to simultaneously achieve high receive SNR and attain encryption in wiretap channel. Now we show how to design the sub-precoders $\mathbf{P}_i$ to maximize the instantaneous SNR γ, and the equivalent instantaneous SNR Γ defined in (6) and (8), respectively.

Letting $\varepsilon_s^2 = \mathbb{E}\{\mathbf{s}\mathbf{s}^H\}$, the instantaneous SNR γ in (6) can be shown to be

$$\gamma = \frac{1}{L}\frac{\varepsilon_s^2}{\sigma_n^2}\frac{\|\mathbf{HW}\|_F^2}{N_t}.$$

From (5), we rewrite γ as

$$\gamma = \frac{1}{L}\frac{\varepsilon_s^2}{\sigma_n^2}\frac{\sum_{i=1}^{\alpha}\mathrm{tr}(\mathbf{P}_i{}^H\mathbf{H}^H\mathbf{H}\mathbf{P}_i)}{N_t} = \frac{1}{L}\frac{\varepsilon_s^2}{\sigma_n^2}\frac{\sum_{i=1}^{\alpha}\|\mathbf{H}\mathbf{P}_i\|_F^2}{N_t}. \tag{12}$$

From (12), maximizing γ is equivalent to maximizing the following objective function:

$$\max\gamma = \max\|\mathbf{HW}\|_F^2 = \max\sum_{i=1}^{\alpha}\|\mathbf{H}\mathbf{P}_i\|_F^2.$$

If $\mathbf{P}_i$, $1 \le i \le \alpha$, are designed independently (we will explain this is true later), the terms $\|\mathbf{H}\mathbf{P}_i\|_F^2$ and $\|\mathbf{H}\mathbf{P}_j\|_F^2$ can be maximized independently; that is, the maximization problem does not need to be solved jointly. Also, since $\|\cdot\|_F^2$ is positive, it yields $\max\sum_{i=1}^{\alpha}\|\mathbf{H}\mathbf{P}_i\|_F^2 = \sum_{i=1}^{\alpha}\max\|\mathbf{H}\mathbf{P}_i\|_F^2$. From the above results, the instantaneous SNR in (6) can be maximized by designing the sub-precoders $\mathbf{P}_i$ using the following relationships:

$$\max\gamma = \max\|\mathbf{H}\mathbf{P}_i\|_F^2. \tag{13}$$

Next we show that the equivalent instantaneous SNR Γ in (8) is maximized if the instantaneous SNR γ for every transmission is maximized. For notational convenience, let $\gamma^{(j)}$ be the instantaneous SNR at the jth transmission. From (8), Γ is expressed as

$$\Gamma = \frac{K}{L}\frac{\varepsilon_s^2}{\sigma_n^2}\frac{\|\overline{\mathbf{H}}\,\overline{\mathbf{W}}\|_F^2}{TKN_t} = \frac{1}{L}\frac{\varepsilon_s^2}{\sigma_n^2}\frac{\sum_{j=1}^{T}\|\mathbf{H}_j\mathbf{W}_j\|_F^2}{TN_t}. \tag{14}$$

From (13), $\max\sum_{j=1}^{T}\|\mathbf{H}_j\mathbf{W}_j\|_F^2 \equiv \max\sum_{j=1}^{T}\gamma^{(j)}$ for $\{\gamma^{(j)}, 1 \le j \le T\}$. Moreover, individual transmissions are assumed to be independent because channels are independent. Since $\|\cdot\|_F^2$ are positive value, $\max\sum_{j=1}^{T}\gamma^{(j)} \equiv \sum_{j=1}^{T}\max\gamma^{(j)}$. Thus we obtain the following relationships:

$$\max\Gamma \equiv \sum_{j=1}^{T}\max\gamma^{(j)}. \tag{15}$$

From (13) and (15), we have the following proposition.

Proposition 1 *The SNR Γ in (8) is maximized if the sub-precoder $\{\mathbf{P}_i, 1 \le i \le \alpha\}$ in (5) is designed to maximize γ for every transmission. That is, Γ is maximized if $\mathbf{P}_i$ is designed to maximize the following value*

$$\sum_{i=1}^{\alpha}\max\|\mathbf{H}\mathbf{P}_i\|_F^2, \tag{16}$$

for all T transmissions. ■

The result in Proposition 1 shows that the proposed precoder is to maximize the Frobenius norm of Φ. This result is similar to that in [33], where the authors shows that a lower bound on the mean-squared error is achieved when $\|\Phi\|_F^2$ is maximized. Next let us explain how to design the precoders when the transmitter side knows full CSI or only partial CSI separately.

B. Design Strategy for Full CSI

If the transmitter side has full CSI, we discuss how to design the optimal precoder $\mathbf{P}_i$ to maximize the SNR Γ. From (16), $\max\|\mathbf{H}\mathbf{P}_i\|_F^2 = \max\mathrm{tr}(\mathbf{P}_i^H\mathbf{H}^H\mathbf{H}\mathbf{P}_i)$, where $\mathbf{P}_i^H\mathbf{H}^H\mathbf{H}\mathbf{P}_i$ is a Hermitian matrix. Since $\mathbf{H}$ is not a square matrix, we need to do a little trick to obtain the optimal solution. Letting $\mathbf{Q}$ be an $N_t \times N_t$ unitary matrix and $\mathbf{Q} = [\mathbf{P}_i\ \mathbf{Q}_0]$, we can formulate $\mathrm{tr}\left(\mathbf{Q}^H\mathbf{H}^H\mathbf{H}\mathbf{Q}\right)$ as

$$\begin{aligned}\mathrm{tr}\left(\mathbf{Q}^H\mathbf{H}^H\mathbf{H}\mathbf{Q}\right) &= \mathrm{tr}\left(\begin{bmatrix}\mathbf{P}_i^H\\ \mathbf{Q}_0^H\end{bmatrix}\mathbf{H}^H\mathbf{H}\begin{bmatrix}\mathbf{P}_i & \mathbf{Q}_0\end{bmatrix}\right)\\ &= \mathrm{tr}\left(\mathbf{P}_i^H\mathbf{H}^H\mathbf{H}\mathbf{P}_i\right)+\mathrm{tr}\left(\mathbf{Q}_0^H\mathbf{H}^H\mathbf{H}\mathbf{Q}_0\right).\end{aligned} \tag{17}$$

From (17), we have the inequality

$$\mathrm{tr}\left(\mathbf{P}_i^H\mathbf{H}^H\mathbf{H}\mathbf{P}_i\right) \le \mathrm{tr}\left(\mathbf{Q}^H\mathbf{H}^H\mathbf{H}\mathbf{Q}\right) = \left(\mathbf{H}^H\mathbf{H}\right).$$

Performing the SVD for $\mathbf{H}$, *i.e.,* $\mathbf{H} = \mathbf{U}\Sigma\hat{\mathbf{V}}^H$, we have

$$\mathrm{tr}\left(\mathbf{P}_i^H\mathbf{H}^H\mathbf{H}\mathbf{P}_i\right) \le \mathrm{tr}(\Sigma^2). \tag{18}$$

Equality holds if and only if $\mathbf{P}_i = \hat{\mathbf{V}}\mathbf{U}_i$, *i.e.,* a column orthonormal matrix. That is, $\mathbf{P}_i$ is of the form $\mathbf{P}_i = \hat{\mathbf{V}}\mathbf{U}_i$ for $i = 1, \ldots, \alpha$, where $\hat{\mathbf{V}} \in \mathbb{R}^{2N_t \times R}$ is the right singular vectors corresponding to the R largest singular values of $\mathbf{H}$, and $\mathbf{U}_i$ is an $R \times R$ unitary matrix. Thus, the optimal sub-precoder $\mathbf{P}_i$ is the right singular vectors corresponding to the largest $R = \mathrm{rank}(\mathbf{H})$ singular values of the channel matrix $\mathbf{H}$. The following position summarizes the result.

Proposition 2 *The optimal sub-precoder $\mathbf{P}_i$ for maximizing the instantaneous SNR of the proposed system is of the form $\mathbf{P}_i = \hat{\mathbf{V}}\mathbf{U}_i$, and $\hat{\mathbf{V}}$ is the right singular vectors corresponding to the $R = rank(\mathbf{H})$ largest singular values of $\mathbf{H}$. Moreover, the resulting SNR is $tr\left(\Sigma^2\right)$ where Σ is the singular value matrix of $\mathbf{H}$.*

Hence, the solutions can be obtained by letting $\mathbf{P}_i = \hat{\mathbf{V}}\mathbf{U}_i$ for $i = 1, \ldots, \alpha$ and the optimal precoder $\mathbf{W}$ in (5) can be designed by

$$\begin{aligned}\mathbf{W} &= \begin{bmatrix}\hat{\mathbf{V}}\mathbf{U}_1 & \hat{\mathbf{V}}\mathbf{U}_2 & \cdots & \hat{\mathbf{V}}\mathbf{U}_\alpha\end{bmatrix}\\ &= \hat{\mathbf{V}}\begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2 & \cdots & \mathbf{U}_\alpha\end{bmatrix}.\end{aligned} \tag{19}$$

Remark 1 *The unitary matrices $\mathbf{U}_i$ can be obtained by performing the QR decomposition for several random square Gaussian matrices. In this case, $\|\mathbf{H}\hat{\mathbf{V}}\mathbf{U}_i\|_F^2 = tr(\mathbf{U}_i^H\hat{\mathbf{V}}^H\mathbf{H}^H\mathbf{H}\hat{\mathbf{V}}\mathbf{U}_i) = tr(\hat{\mathbf{V}}^H\mathbf{H}^H\mathbf{H}\hat{\mathbf{V}})$, which does not destroy the optimality in Proposition 2.*

Now consider the precoders used for repeatedly transmitting the sparse vector by T times so as to satisfy the RIP in (3).

The equivalent received signal in (7) can be rewritten as

$$\begin{aligned}\bar{\mathbf{y}} &= \overline{\mathbf{H}}\begin{bmatrix}\hat{\boldsymbol{V}}_1 & & \\ & \ddots & \\ & & \hat{\boldsymbol{V}}_T\end{bmatrix}\begin{bmatrix}\mathbf{U}_{11} & \cdots & \mathbf{U}_{1\alpha}\\ \vdots & \ddots & \\ \mathbf{U}_{T1} & & \mathbf{U}_{T\alpha}\end{bmatrix}\mathbf{s}+\bar{\mathbf{n}}\\ &= \overline{\mathbf{H}}\overline{\mathbf{V}}\overline{\mathbf{U}}\mathbf{s}+\bar{\mathbf{n}} = \overline{\Phi}\mathbf{s}+\bar{\mathbf{n}},\end{aligned} \tag{20}$$

where $\overline{\mathbf{V}}$ is the overall precoder, which is a $2TN_t \times TR$ block diagonal matrix, $\hat{\boldsymbol{V}}_i$ is the singular vectors corresponding to the R largest singular values of $\mathbf{H}_i$ for $i = 1, \ldots, T$, and $\overline{\mathbf{U}}$ is the encryption matrix that is a $TR \times \alpha R$ block matrix with every sub-block being an $R \times R$ unitary matrix obtained by using Remark 1.

Remark 2 *According to (20), encryption is attained via two aspects. First, according to Remark 1, the encryption matrix* $\overline{\mathbf{U}}$ *is independent of channels, and can be generated from random matrices with a unique key (or seed) known only to the transmitter and the legitimate receiver. Without knowing this key (or seed), it is very unlikely that the eavesdroppers can recover the received signals. Secondly with full CSI, the proposed system can achieve the maximum SNR by using the optimal precoder. Meanwhile the system is automatically encrypted because* $\hat{\boldsymbol{V}}_i$ *is from the unique CSI between the transmitter and the legitimate receiver.*

It is worth noting that in Shannon's fundamental paper [18], he assumed that the eavesdroppers may know the family of encryption function and the probability of the choosing key, but not the exact encryption function and key. Thus it may be reasonable to assume that the eavesdroppers do not know the exact precoding matrices $\overline{\mathbf{V}}$ and $\overline{\mathbf{U}}$ in encryption systems.

C. Precoder Design for Partial CSI

In FDD (frequency division multiplexing) systems, the transmitter knows only partial CSI. In this case, the overall precoder $\overline{\mathbf{V}}$ in (20) should be represented by a finite set of B-bit codebook $\mathcal{W} = \{\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_{2^B}\} \subset \mathcal{U}(2N_t, R)$, where $\mathcal{U}(2N_t, R)$ is the set of $2N_t \times R$ real matrices with orthonormal columns, and the legitimate receiver only sends the index back to the transmitter. We briefly review two popular codebook design methods, and then introduce the proposed modified codebook design:

1) Codebook Design by Grassmannian Subspace Packing: The authors in [22] designed the codebooks by maximizing the minimum subspace distance, which is known as the Grassmannian subspace packing and the following two distances are introduced for designing codebooks.

The projection two-norm distance:

$$d_{proj}(\mathbf{C}_1, \mathbf{C}_2) = \|\mathbf{C}_1\mathbf{C}_1^H - \mathbf{C}_2\mathbf{C}_2^H\|_2 = \sqrt{1-\lambda^2_{\min}\{\mathbf{C}_1^H\mathbf{C}_2\}}, \tag{21}$$

where $\lambda^2_{\min}\{\cdot\}$ is the square of the minimum singular value of the matrix in the brace.

The Fubini-Study distance:

$$d_{FS}(\mathbf{C}_1, \mathbf{C}_2) = \arccos|\det(\mathbf{C}_1^H\mathbf{C}_2)|. \tag{22}$$

The two distance functions may form a criterion for constructing the codebooks given by

$$\mathcal{W}^\star = \arg\max_{\mathcal{C}\in\mathcal{U}(M_t,M)} \min_{1\le i<j\le 2^B} d(\mathbf{C}_i, \mathbf{C}_j). \tag{23}$$

where $d(\cdot,\cdot)$ is the distance function either from (21) or (22).

2) Codebook Design by Modified Llyod Algorithm: An alternative to generate the codebook is by using the *vector quantization* (VQ) technique [34], which applies different geometry viewpoint from Grassmannian subspace packing. The Lloyd algorithm is a popular approach in VQ, see *e.g.,* [27] and [28]. Generally codebook design using Lloyd algorithm focuses on capacity loss, and the codewords are vectors. However, the proposed system aims to maximize the instantaneous SNR, and the codewords are matrices with orthonormal columns instead of vectors. Hence, we utilize the distance concept of Grassmannian packing in (21) or (22), and propose a modified Lloyd algorithm to construct the codebooks. The main idea is to iteratively compute the local optimal solutions by maximizing the instantaneous SNR. Multiple iteration threads with different initial settings lead to a near global-optimal solution. The proposed algorithm is summarized as follows:

Proposed modified Lloyd algorithm:

- Condition 1: Nearest neighborhood condition (NNC)
  Letting the initial codewords be $\{\mathbf{C}_\alpha \mid \alpha = 1, 2, \ldots, 2^B\}$ and using (23), the optimal partition region (Voronoi cell) $\mathcal{A}_i$ of the ith codeword is

$$\begin{aligned}\mathcal{A}_\alpha = \{&\hat{\mathbf{V}}\ \text{ofthetrainingchannel}\ \widetilde{\mathbf{H}}\ \text{by SVD}:\\ &d(\hat{\mathbf{V}}, \mathbf{C}_\alpha) \le d(\hat{\mathbf{V}}, \mathbf{C}_\beta), \forall \alpha \ne \beta\}\end{aligned}$$

- Condition 2: Centroid condition (CC)
  Given the partition regions $\{\mathcal{A}_\alpha \mid \alpha = 1, 2, \ldots, 2^B\}$, the optimal codeword $\mathbf{C}_\alpha$ is determined according to the following criterion:

$$\begin{aligned}\mathbf{C}_\alpha &= \arg\max_{\|\mathbf{C}\|_F^2=1} \mathbb{E}\{\|\widetilde{\mathbf{H}}\mathbf{C}\|_F^2 | \widetilde{\mathbf{H}} \in \mathcal{A}_\alpha\}\\ &= \arg\max_{\|\mathbf{C}\|_F^2=1} \mathbf{C}^H\mathbf{R}_\alpha\mathbf{C}\\ &= \text{eigenvectors corresponding to}\\ &\qquad \text{the } R \text{ largest eignvalues of } \mathbf{R}_\alpha,\end{aligned}$$

where $\mathbf{R}_\alpha$ is the autocorrelation matrix of local training channel of the αth region $\mathcal{A}_\alpha$. The above two conditions are iterated until $\mathbb{E}\{\|\widetilde{\mathbf{H}}\mathbf{C}\|_F^2 | \widetilde{\mathbf{H}} \in \mathcal{A}_\alpha\}$ converges.

As for how to select a codeword from the codebook, the legitimate receiver should first obtain $\hat{\mathbf{V}}$ from the singular decomposition of $\widetilde{\mathbf{H}}$, and then finds a suitable index via

$$\mathbf{C}^\star = \arg\min_{1\le\alpha\le 2^B} d(\mathbf{C}_\alpha, \hat{\mathbf{V}}) \quad \forall\, \mathbf{C}_\alpha \in \mathcal{W}.$$

D. Postcoder Design

Since the encrypted signals are sparse, from the view point of CS, we would like the number K of sparsity be as large as possible to carry more information. Properly designing the

postcoder can change the property of the equivalent sensing matrix. As a result, it increases the number of allowable sparsity under a fixed recovery rate. We introduce the design of the postcoders. Let the postcoder be $\mathbf{E}$, which is a $2TN_r \times 2TN_r$ real square matrix. From (7), the received signals after the postcoder is expressed as

$$\mathbf{E}\bar{\mathbf{y}} = \mathbf{E}\overline{\Phi}\mathbf{s} + \mathbf{E}\bar{\mathbf{n}} = \mathbf{D}\mathbf{s} + \mathbf{E}\bar{\mathbf{n}}, \tag{24}$$

where $\mathbf{D} = \mathbf{E}\overline{\Phi}$ is a $2TN_r \times L$ matrix. Mutual coherence measures the maximal correlation and plays an important role in the success of recovery algorithms. A measure of mutual coherence is defined as

$$\mu(\mathbf{D}) = \max_{1\le i,j\le L, i\neq j} |\langle \mathbf{d}_i, \mathbf{d}_j\rangle|. \tag{25}$$

Let the Gram matrix of $\mathbf{D}$ be $\mathbf{G} = \mathbf{D}^H\mathbf{D}$. Without loss of generality, let $\mathbf{G}$ have normalized columns. The off-diagonal entries of $\mathbf{G}$ are the inner products, and the mutual coherence is the off-diagonal entry with the largest magnitude.

It has been demonstrated that the mutual coherence should be as small as possible to achieve a good recovery rate in CS (see [35]). The resulting relationship for the parameters is described in the following inequality

$$2TN_r \ge \eta K \mu^2(\mathbf{D}) \ln L, \tag{26}$$

where η is a positive constant. Observe that the smaller the coherence is, the smaller the value $2TN_r$ is to satisfy (26). As a result, T and N_r can be small if $\mu(\mathbf{D})$ is small.

There is another result about mutual coherence in [25] that shows the recovery algorithms are guaranteed to find the sparsest estimated vector $\hat{\boldsymbol{s}}$ if the following relationship holds:

$$\|\hat{\boldsymbol{s}}\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(\mathbf{D})}\right), \tag{27}$$

where $\hat{\boldsymbol{s}}$ is the solution for sensing matrix $\mathbf{D}$ and $\|\hat{\boldsymbol{s}}\|_0 = K$.

From the above discussion, it is desirable to design the postcoder $\mathbf{E}$ to minimize the mutual coherence $\mu(\mathbf{D})$, so as to attain better recovery performance.

There are several approaches for designing the postcoder $\mathbf{E}$. For example, the authors in [37] designed the postcoder by an approach that iteratively minimizes the t-averaged mutual coherence. This approach consumes quite some time to converge and gain performance improvements. In [38], the authors used unit-norm tight frames to improve mean square error performance, and designed the sensing matrix to formulate an optimization convex problem that can achieve better performance. However this approach also needs huge computational complexity to solve the postcoder $\mathbf{E}$. An algorithm that iteratively optimizes the sensing matrix by including the postcoder for arbitrary size was proposed in [39]. The authors in [39] attempted to make all subset of columns of $\mathbf{D}$ as orthonormal as possible. In other words, they try to make the Gram matrix as close as to an identity matrix. Since the proposed system uses a $2TN_r \times 2TN_r$ square postcoder, it does not need an arbitrary size postcoder. Thus, instead of iteratively optimizing the sensing matrix in [39], we exploit the design concept in [39], and derive a closed-form solution for the postcoder that can significantly reduces the computational complexity of the algorithm in [39]. Let us introduce the proposed postcoder as follows:

Consider the following Gram matrix of the equivalent sensing matrix,

$$\mathbf{G} = \overline{\Phi}^H\mathbf{E}^H\mathbf{E}\overline{\Phi}.$$

The object is to design $\mathbf{E}$ that makes the Gram matrix $\mathbf{G}$ as close as to an identity matrix. That is, the objective is to have

$$\mathbf{G} = \overline{\Phi}^H\mathbf{E}^H\mathbf{E}\overline{\Phi} \approx \mathbf{I}, \tag{28}$$

Multiplying both sides of (28) with $\overline{\Phi}$ on the left, and $\overline{\Phi}^H$ on the right, it yields

$$\overline{\Phi}\overline{\Phi}^H\mathbf{E}^H\mathbf{E}\overline{\Phi}\overline{\Phi}^H \approx \overline{\Phi}\overline{\Phi}^H. \tag{29}$$

Performing the eigenvalue decomposition (EVD) for $\overline{\Phi}\overline{\Phi}^H$, we have

$$\overline{\Phi}\overline{\Phi}^H = \mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}}\mathbf{V}_{\overline{\Phi}}^H.$$

From (29) and the eigenvalue decomposition of $\overline{\Phi}\overline{\Phi}^H$, the above equation becomes

$$\mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}}\mathbf{V}_{\overline{\Phi}}^H\mathbf{E}^H\mathbf{E}\mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}}\mathbf{V}_{\overline{\Phi}}^H \approx \mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}}\mathbf{V}_{\overline{\Phi}}^H,$$

and it can be manipulated as

$$\Lambda_{\overline{\Phi}}\mathbf{V}_{\overline{\Phi}}^H\mathbf{E}^H\mathbf{E}\mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}} \approx \Lambda_{\overline{\Phi}}.$$

Finally, this problem is formulated as the following optimization problem with respect to $\mathbf{E}$:

$$\min \|\Lambda_{\overline{\Phi}}\mathbf{V}_{\overline{\Phi}}^H\mathbf{E}^H\mathbf{E}\mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}} - \Lambda_{\overline{\Phi}}\|_F^2. \tag{30}$$

At the first glance, solving this objective function (30) may require iterative optimization method as described in [39]. However, as $\mathbf{E}$ is square, we can have a closed-form solution as stated in the following proposition.

Proposition 3 *The optimal postcoder* $\mathbf{E}$ *which optimizes the objective function in (30) has the closed-form solution:*

$$\mathbf{E}^\star = \mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}}^{-\frac{1}{2}}\mathbf{V}_{\overline{\Phi}}^H. \tag{31}$$

Proof: The derivative of the objective function is

$$\begin{aligned}\frac{\partial}{\partial\mathbf{E}}&\|\Lambda_{\overline{\Phi}}\mathbf{V}_{\overline{\Phi}}^H\mathbf{E}^H\mathbf{E}\mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}} - \Lambda_{\overline{\Phi}}\|_F^2 \\ &= 4\mathbf{E}\mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}}^2\mathbf{V}_{\overline{\Phi}}^H\mathbf{E}\mathbf{E}^H\mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}}^2\mathbf{V}_{\overline{\Phi}}^H - 4\mathbf{E}\mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}}^3\mathbf{V}_{\overline{\Phi}}^H,\end{aligned} \tag{32}$$

According to (29), we would like to have (32) as closely to zero as possible, then

$$\mathbf{E}\mathbf{E}^H = \mathbf{V}_{\overline{\Phi}}\Lambda_{\overline{\Phi}}^{-1}\mathbf{V}_{\overline{\Phi}}^H. \tag{33}$$

From (33) we obtain the optimal postcoder given in (31). ■

IV. Performance Analysis

This section analyzes the performance of the proposed system. First, we show to determine a suitable value for T. This is important because when T is large, the transmitted information decreases. On the other hand when T is small, the system may violate the RIP, and thus it results in a poor recovery performance. Moreover, after determining a suitable value of T, we show how the proposed system can achieve the perfect secrecy property.

A. The Minimum Value of T

Let us discuss how to determine the repeating factor T. Similar to the analysis in [40], we conduct a numerical simulation to obtain the minimum required number of rows M for the sensing matrix so as to satisfy the RIP in (3). In [40], the authors mentioned that when the sensing matrix is generated using Gaussian random matrices, the relationship of the system parameters can be described by

$$M = CK \ln (L/K + 1) + 1, \tag{34}$$

where C is a constant. Our goal is to determine the constant C and the minimum number M for the proposed system. This can be done by fixing an L and simulating for several different K. Under such setting, we can obtain a minimum number M such that the recovery rate is nearly 100%. Note that it may be inappropriate to claim the recovery rate achieves 100% because this is a curve fitting method instead of a theoretical result. Thus, we may use a reasonable recovery rate, say greater than 99% in the simulation. The residual non-recovery rate less than 1% may be corrected by error correction codes. However, it is worth to point out that depending on the system requirements, different recovery rates can also be used in the simulation, and it would affect the value of C [41].

More detailed parameter setting are described as follows: According to [40], the following stochastic model is used to describe the sparse vectors of length L:

- The set of sparsity index is a uniformly random set with K elements from $\{1, 2, \ldots, L\}$.
- The magnitude of sparsity is binary which can be either -1 or 1.

We performed 1000 trials. In each trial, the following procedure is executed.

- **Sparse vector**: a sparse vector $\mathbf{s}$ of length L is generated using the above stochastic model.
- **Sensing matrix**: the precoder is designed by assuming that full CSI is known. The elements of the MIMO channel $\mathbf{H}$ are i.i.d. Gaussian variables. Performing the SVD for $\mathbf{H}$ and multiplying it by the precoding matrix $\mathbf{W}$ *i.e.,*

$$\begin{aligned}\Phi = \mathbf{HW} &= \hat{\mathbf{U}}\Sigma\hat{\mathbf{V}}^H\hat{\mathbf{V}}\begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2 & \cdots & \mathbf{U}_\alpha\end{bmatrix} \\ &= \hat{\mathbf{U}}\Sigma\begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2 & \cdots & \mathbf{U}_\alpha\end{bmatrix},\end{aligned}$$

  where $\hat{\mathbf{U}}$, $\hat{\mathbf{V}}$ and $\mathbf{U}_i$ for $i = 1, \ldots, \alpha$ are Haar matrices [42]. We conjecture that the sensing matrix Φ is a Gaussian random matrix. Since this is difficult to prove analytically, we simulated 10 million data to obtain the histogram of the sensing matrix for $N_r = 4$, which is shown in Fig. 2. Observe that the elements of Φ are well approximated by Gaussian distribution $\mathcal{CN}(0, N_t/2N_r)$. We also verify that the approximation is reasonable for $N_r > 4$. Therefore, Φ is assumed to be an approximated Gaussian random matrix herein.
- **Recovery**: $\hat{s}$ is recovered by using the Dantzig selector [26], which is regarded as the most powerful recovering algorithm to date.

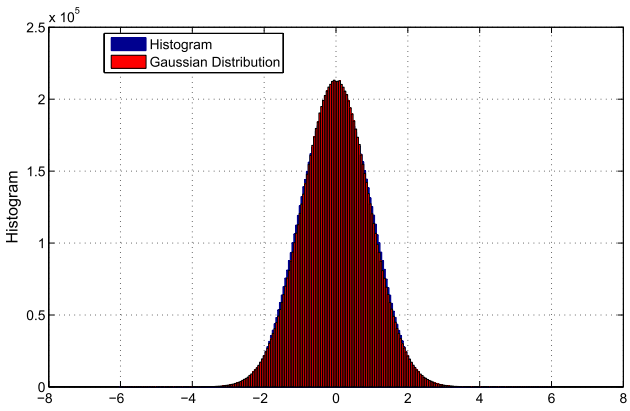

Fig. 2. The histogram of the elements in sensing matrices.

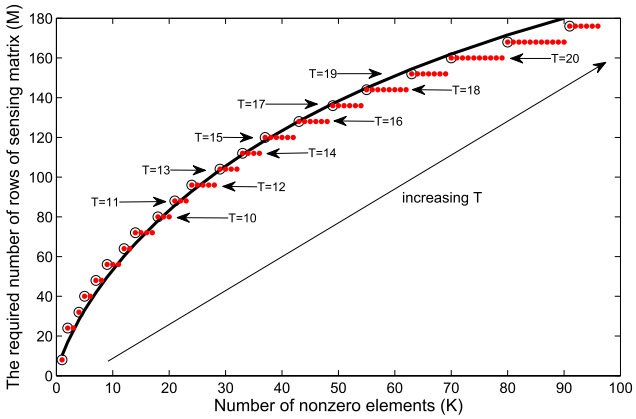

Fig. 3. The required number of rows of the sensing matrix as a function of number of nonzero elements using $L = 192$ and $N_r = 4$ is fixed for the proposed system.

A recovery is said to be successful when $\mathbf{s} = \hat{s}$. The smallest value for M is determined when the empirical successful recovery rate is greater than 99%.

Curve fitting results. Let $L = 192$ and $N_r = 4$. Fig. 3 shows the simulation results (dot and circled curves) and the curve fitting results (bold curve). From the simulation results, since $M = 2TN_r$, increasing T by 1 implies increasing M by 8. Thus, the value of M increases by a step of 8. As a result, some values of M have several possible sparsity number K. The curve is fitted (bold curve) by considering the worst case, *i.e.,* the minimum number of sparsity for a fixed M, which is shown in the circled curve. Moreover, we fit the curves for different values of T and find the results also match (34) well with C ranging between 1.71 to 1.74, which are shown in Fig. 4.

As we mentioned above, the recovery rate can be chosen arbitrarily according to the system requirement. Though not showing here, by letting the recovery rate be 99.5% instead of 99% and using the above curve fitting method, we can obtain $C = 1.76$ for $L = 192$. Hence, for the proposed system, the parameters may be determined in the following proposition:

Proposition 4 *The relationship among the number K of sparsity, the number M of rows of the sensing matrix, and*

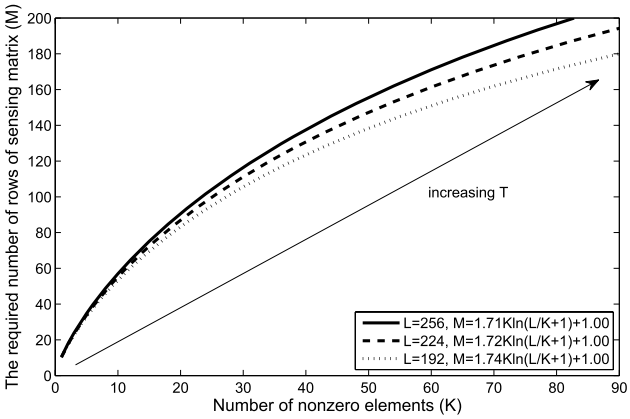

Fig. 4. The required number of rows of the sensing matrix as function of number of nonzero elements with different T and $N_r = 4$ for the proposed system.

the vector size L can be described by

$$M = CK \ln (L/K + 1) + 1.00, \tag{35}$$

where $C \in [1.71, 1.74]$.

From this proposition, M in (35) can be lower bounded by

$$M \geq 1.74K \ln (L/K + 1). \tag{36}$$

This bound becomes tight as L/K increases, and will be used in the discussion later. Since $M = 2TN_r$, using (36), the minimum repeating number T is given by

$$T^{\star} = \left\lceil \frac{M}{2N_r} \right\rceil. \tag{37}$$

It is worth to point out that the bound in (36) obeys the rule found in [40] and [41] well; that is, in [41], the relationship is shown to be $M = \mathcal{O}(K \ln (L/K))$ if the sensing matrix is a Gaussian sensing matrix. Moreover, since the sensing matrix Φ of the proposed system is assumed to be approximated random Gaussian, the bound for M in (36) matches the empirical results. Therefore, the choice of T in (37) should be reliable to achieve a high recovery rate.

B. Conditions for Perfect Secrecy

The above discussion for T is from the point of view that the proposed system achieves a high recovery rate. However, we have not considered how the eavesdroppers would affect the system parameters. That is, to achieve perfect secrecy under a fixed T, what is the maximum sparsity K? By knowing K, we can calculate the bit rate that simultaneously achieves nearly full recovery rate and perfect secrecy. This is discussed below

Definition 2 *In [16]–[18], a encryption system achieves perfect secrecy if the probability of a message conditioned on the cryptogram is equal to the a priori-probability of the message,* i.e.,

$$P(\mathbf{a} = a|\mathbf{b} = b) = P(\mathbf{a} = a),$$

where $\mathbf{b} = \Phi\mathbf{a}$.

This definition is equivalent to saying that a system achieves perfect secrecy if the mutual information $I(\mathbf{a}; \mathbf{b})$ between the transmitter and the eavesdroppers is zero, because

$$\begin{aligned} I(\mathbf{a}; \mathbf{b}) &\equiv \sum_{a,b} P(a, b) \log (P(a|b)/P(a)) \\ &= \sum_{a,b} P(a, b) \log (P(a)/P(a)) = 0, \end{aligned}$$

In other words, perfect secrecy implies that the eavesdroppers do not have a chance to steal any data information. The following lemma shows how to use the proposed system to achieve perfect secrecy.

Lemma 1 *Using the results in [17], let the sparse vector* $\mathbf{s}$ *be of K-sparsity and $M \geq 2K$, the nonzero elements of* $\mathbf{s}$ *have uniform distribution, and the sensing matrix $\Phi \in \mathbb{R}^{M \times L}$ satisfy the RIP. Assume that the eavesdroppers do not know either* $\overline{\mathbf{U}}$ *(key for generating random matrices) or* $\overline{\mathbf{V}}$ *(precoding matrix related to CSI) as mentioned in Remark 2. Then, the property of prefect secrecy can be achieved via CS for the proposed system.*

Using Lemma 1, we are able to obtain the mutual information is equal to zero *i.e.,* $I(\mathbf{a}; \mathbf{b}) = 0$. Furthermore, by Lemma 1, we have the following proposition about the bit rate.

Proposition 5 *Let $M = 2K$. If K and L are chosen such that $2 < L/K \leq e^{\frac{1}{0.87}} - 1 \approx 2.16$, the achievable bit rate, which simultaneously achieves perfect secrecy and a recovery rate more than 99%, can be expressed by*

$$I_o = \frac{1}{T^{\star}} \frac{K + \log_2 \binom{L}{K}}{L}, \tag{38}$$

where $T^{\star}$ is determined by (37).

Proof: To satisfy Proposition 4 and Lemma 1, we have the following constraint:

$$1.74K \ln(L/K+1) \leq M = 2K \Rightarrow \frac{L}{K} \leq e^{\frac{1}{0.87}} - 1 \approx 2.16. \tag{39}$$

Since the proposed system is underdetermined linear model, this yields

$$M = 2K < L \Rightarrow 2 < \frac{L}{K}. \tag{40}$$

(39) and (40) lead to the given condition in this Proposition. Under such condition, the system achieves perfect secrecy and a recovery rate more than 99%. Also, the corresponding achievable bit rate in (38) can be obtained by letting $M = 2K$, $T^{\star} = \lceil K/N_r \rceil$, and then substituting (37) into (9). Since $T^{\star}$ is the minimum required repeating factor, the resulting bit rate is the maximum achievable bit rate for the proposed system. ■

Note that the proposed codebooks in Sec. III are generated via random matrices and thus they satisfy the RIP. Once the matrices sizes are determined according to Lemma 1 and Proposition 5, it is true to state that using the proposed codebooks does not affect the secrecy of the proposed encryption system from the discussion in Lemma 1 and Proposition 5.

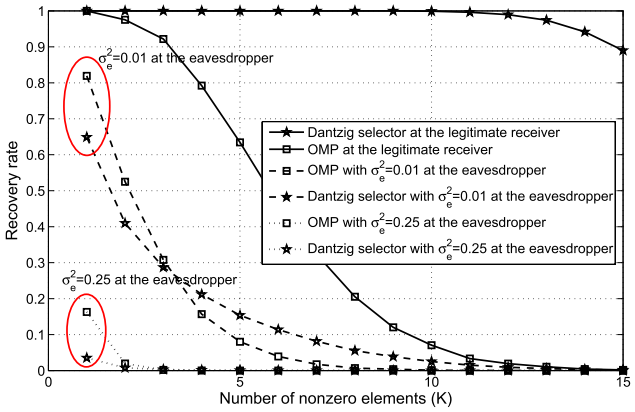

Fig. 5. The eavesdropper estimates the different levels of full CSI at the transmitter.

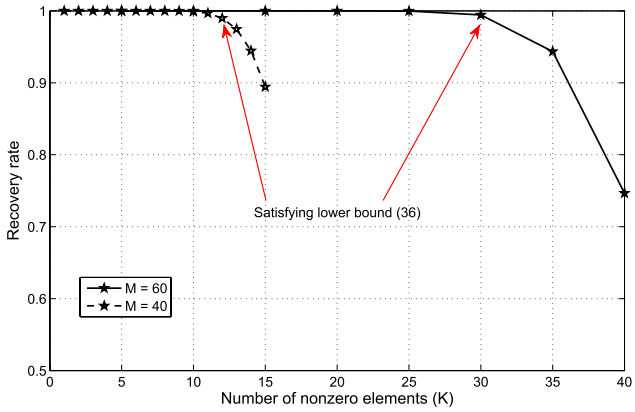

Fig. 6. Recovery rate with different M.

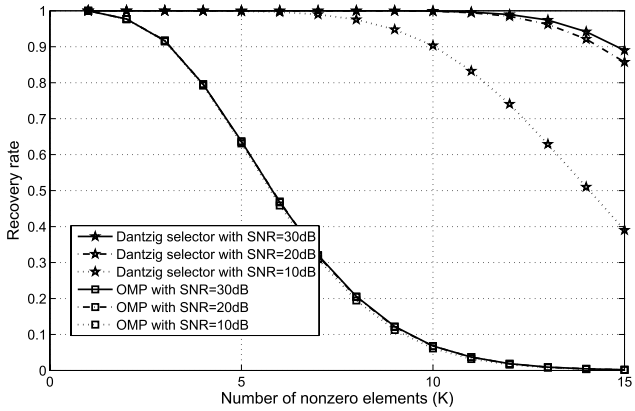

Fig. 7. Recovery rate with different SNR.

V. Simulation Results

In all experiments, the data information is encoded into sparse vectors with length $L = 64$, and the nonzero elements are ± 1. The channel coefficients of $\mathbf{H}_c$ are i.i.d. complex Gaussian distributed with $\mathcal{CN}(0, 1)$. More than 100000 iterations were conducted to compute the recovery rate. The MIMO antennas are $N_r = 4$ and $N_t = 7$. The 16 unitary matrices, *i.e.,* $\mathbf{U}_i$ were generated by conducting the QR decomposition for 16 4×4 square Gaussian random matrices. Three recovery algorithms were used to reconstruct the signals including the Orthogonal Matching Pursuit (OMP) [25], the Subspace Pursuit (SP) [43], and the Dantzig selector [26]. Note that the Dantzig selector is currently recognized as the most powerful recovery algorithm in CS and thus its performance can be regarded as a benchmark.

Experiment 1. The eavesdroppers knew $\overline{\mathbf{U}}$ and different levels of CSI. Consider the worst case that the eavesdroppers know $\overline{\mathbf{U}}$ and also different levels of noisy CSI. The channel known to the eavesdropper is $\mathbf{H}_e = \mathbf{H}_c + \sigma_e \delta$, where δ is assumed to be i.i.d. $\mathcal{CN}(0, 1)$ and σ_e^2 is the mean squared estimation error. Since the eavesdropper uses $\mathbf{H}_e$ to obtain the precoder $\overline{\mathbf{V}}_e$, which is very different from the true precoder $\overline{\mathbf{V}}$. Consequently, it results in very different equivalent sensing matrix and the recovery performance of eavesdroppers degrades significantly. Let the SNR be 30 dB. Fig. 5 shows the simulation results. Observe that the recovery rate is low for the eavesdropper. With $\sigma_e^2 = 0.01$, the recovery rate is only 65% for one sparsity using the benchmark Dantzig selector. This shows that the Dantzig selector is very sensitive to the accuracy of the CSI. For $\sigma_e^2 = 0.25$, the eavesdroppers can hardly reconstruct any data with either OMP or the Dantzig selector. Therefore, recovery performance is poor for the eavesdroppers if they do not know the exact CSI.

Experiment 2. Recovery rate for different M. This example is to verify (36), which shows the maximum number of sparsity to attain a recovery rate more than 99%. The recovery algorithm is the Dantzig selector. Fig. 6 shows the recovery rate as a function of the number K of sparsity for $M = 40$ and 60. Observe that the simulation results corroborate the theoretical result in (36). That is, from (36), the maximum numbers of sparsity to attain a 99% recovery rate are $K = 12$ for $M = 40$, and $K = 30$ for $M = 60$. The figure indeed shows that the recovery rates for $M = 40$ with $K = 12$, and $M = 60$ with $K = 30$ are both greater than 99%.

Experiment 3. Recovery rate for different SNR. Let $T = 5$, Fig. 7 shows the recovery performance for SNR = 10, 20 and 30 dB. Observe from the figure, the Dantzig selector can perfectly recover the sparse vector with $K \leq 10$ when SNR is greater than 20 dB. On the other hand, although the OMP algorithm is simple but its recovery performance is far worse than that of the Dantzig selector.

Experiment 4. Recovery rate for different repeating numbers T. Let the SNR be 25 dB, Fig. 8 shows the recovery performance for $T = 5$ and 7. Observe that for the Dantzig selector, decreasing T from 7 to 5 does not degrade the performance seriously. On the other hand, for the OMP and SP algorithms, increasing T improves the performance significantly. In addition, the Dantzig selector with $T = 5$ outperforms both the OMP and SP algorithms with $T = 7$. Similarly the SP algorithm with $T = 5$ outperforms the OMP algorithms with $T = 7$. Since increasing T would decrease the bit rate, this example shows that using powerful recovery algorithm can help increase bit rate for the proposed system.

Experiment 5. Recovery rate with full CSI and partial CSI. The performance with full CSI and partial CSI is shown in

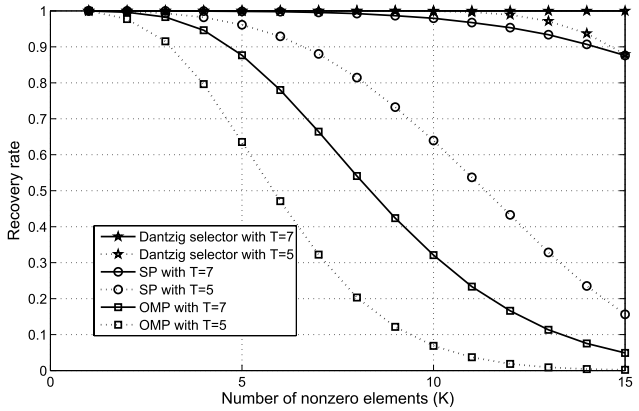

Fig. 8. Recovery rate for different T.

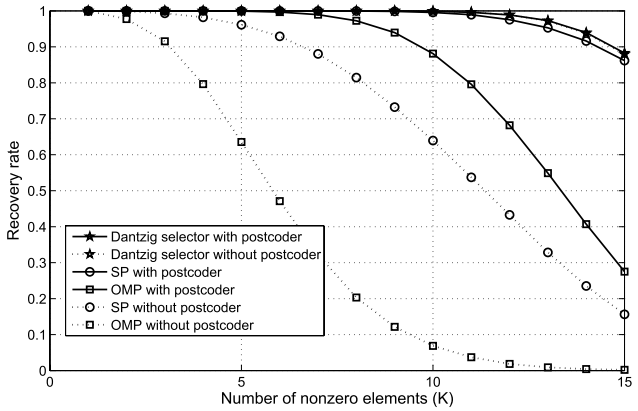

Fig. 10. Recovery rate with the proposed postcoder.

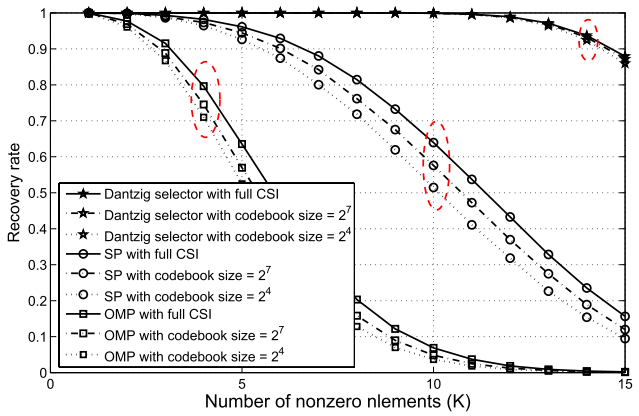

Fig. 9. Recovery rate with full CSI and partial CSI.

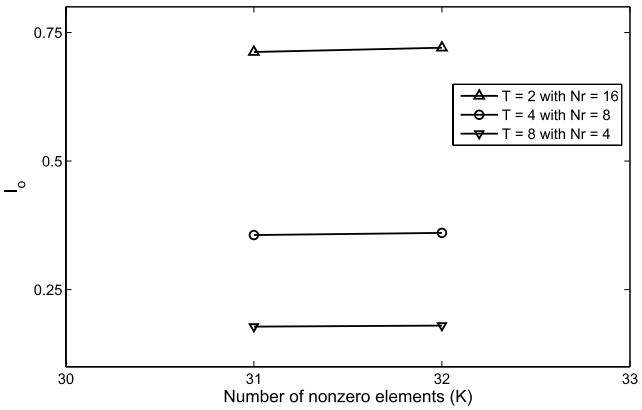

Fig. 11. The bit rate for prefect secrecy and nearly recovery rate.

Fig. 9. For partial CSI, the codebooks are generated using the proposed modified Lloyd algorithm. From the figure, although the Dantzig selector needs high computational complexity, it is not that sensitive to CSI compared to the OMP and SP algorithms. Thus, if the Dantzig selector is used, the codebook size can be kept small so that only small amount of feedback information is needed. From Remark 2, the codebook may be regarded as the second encryption key. Therefore, if the eavesdroppers do not know the codebook, they are unlikely to recovery the transmitted information, even through the codebook size is small.

Experiment 6. Recovery rate with postcoder. Fig. 10 shows the performance with and without the proposed postcoder. Observe from the figure, the performance of the OMP and SP algorithms can be dramatically improved if the proposed postcoder is applied. For the SP algorithm with the postcoder, it can even achieve comparable performance with the Dantzig selector. Thus if complexity is the main concern and the Dantzig selector is not available, the proposed postcoder provides an alternative that well leverages between complexity and performance.

Experiment 7. The achievable bit rate. Let $L = 65$. Fig. 11 shows the actual bit rate which guarantees nearly full recovery rate and prefect secrecy in Proposition IV-B. The bit rate I_o can be up to 0.7 if the proposed system uses $N_r = 16$. Also, observe that increasing T more seriously affects the bit rate than decreasing K. Therefore, to increase bit rate, it is suggested to use large N_r, or decreasing K instead of increasing T. This example may provide a useful design reference to determine the bit rate for practical realizations.

VI. Conclusion

In this paper, we proposed a precoding and postcoding system to achieve data secrecy at the physical layer. The precoding procedure makes the proposed system an underdetermined linear system. Thus recovery algorithms for compressive sensing can be used to reconstruct the transmitted signals. When full CSI is available, the proposed precoder can maximize the receive SNR. At the same time, the proposed precoder can be regarded as a key to encrypt the signals. When only limited feedback is available, we proposed a modified Lloyd algorithm to construct the codebooks to represent the proposed precoder. The codebook itself can be regarded as a key. If the eavesdroppers do not know the key, it is almost unlikely for them to reconstruct the data. Moreover, when full CSI is not available, we proposed a postcoding design so that the system can achieve comparable performance with systems having full CSI. Furthermore, we analyzed the proposed system from the view points of recovery rate and secrecy, and showed how to attain full recovery rate and perfect secrecy. Finally from the simulation results, we had several interesting observations: First, if the eavesdroppers do not have full

precoding information such as the numbers of transmit and receive antennas, repeating number, and the codebook, the recovery rate is approximately zero, which shows the great security of the proposed system. Secondly, we found powerful recovery algorithms, *i.e.,* the l_1 optimization methods, can help increase the bit rate and reduce the feedback information of the proposed system. Thirdly, the proposed postcoder can compensate the SNR loss when full CSI is not available; this is more pronounced when low-complexity recovery algorithms such as OMP and SP are used.

ACKNOWLEDGMENT

The authors would like to thank all the anonymous reviewers for their constructive suggestions, which have significantly improved the quality of this work.

REFERENCES

[1] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar, "On the Gaussian MIMO wiretap channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 2471–2475.

[2] A. Khisti and G. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

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

[4] S. A. A. Fakoorian and A. L. Swindlehurst, "MIMO interference channel with confidential messages: Achievable secrecy rates and precoder design," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 640–649, Sep. 2011.

[5] P. L. Yu and B. M. Sadler, "MIMO authentication via deliberate fingerprinting at the physical layer," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 606–615, Sep. 2011.

[6] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[7] A. Orsdemir, H. O. Altun, G. Sharma, and M. F. Bocko, "On the security and robustness of encryption via compressed sensing," in *Proc. Military Commun. Conf.*, Nov. 2008, pp. 1–7.

[8] R. J. McEliece, "A public-key cryptosystem based on algebraic coding theory," DSN Progress, Brazil, Rep. 42-44, 1978, pp. 114–116.

[9] H. Dinh, C. Moore, and A. Russell. (2010, Aug.). *The McEliece Cryptosystem Resists Quantum Fourier Sampling Attacks* [Online]. Available: http://arxiv.org/bs/1008.2390

[10] Z. Yang, C. Zhang, and L. Xie, "On Phase transition of compressed sensing in the complex domain," *IEEE Signal Process. Lett.*, vol. 19, no. 1, pp. 47–50, Jan. 2012.

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

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

[13] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst. (2010, Nov.). *Principles of Physical-Layer Security in Multiuser Wireless Networks: Survey* [Online]. Available: http://arxiv.org/abs/1011.3754

[14] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4033–4040, Sep. 2009.

[15] S. Bashar, Z. Ding, and G. Y. Li, "On secrecy of codebook-based transmission beamforming under receiver limited feedback," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1212–1223, Apr. 2011.

[16] Y. Rachlin and D. Baron, "The secrecy of compressed sensing measurements," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput.*, 2009, pp. 813–817.

[17] M. R. Mayiami, B. Seyfe, and H. G. Bafghi. (2010, Nov.). *Perfect Secrecy Using Compressed Sensing* [Online]. Available: http://arxiv.org/abs/1011.3985

[18] C. E. Shannon, "Communication theory of secrecy systems," *Bell. Syst. Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

[19] A. Khisti, S. N. Diggavi, and G. W. Wornell, "Secret-key agreement with channel state information at the transmitter," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 672–681, Sep. 2011.

[20] A. Khisti, "Interactive secret key generation over reciprocal fading channels," in *Proc. 50th Annu. Allerton Conf.*, 2012, pp. 1374–1381.

[21] V. M. Prabhakaran, K. Eswaran, and K. Ramchandran, "Secrecy via sources and channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6747–6765, Nov. 2012.

[22] D. J. Love and R. W. Heath, "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.

[23] S. Agrawal and S. Vishwanath, "Secrecy using compressive sensing," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2011, pp. 563–567.

[24] G. Reeves, N. Goela, N. Milosavljevic, and M. Gastpar, "A compressed sensing wire-tap channel," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2011, pp. 548–552.

[25] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[26] E. J. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *Ann. Statist.*, vol. 35, no. 6, pp. 2313–2351, Dec. 2007.

[27] J. C. Roh and B. D. Rao, "Transmit beamforming in multiple-antenna systems with finite rate feedback: A VQ-based approach," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1101–1112, Mar. 2006.

[28] P. Xia and G. B. Giannakis, "Design and analysis of transmit-beamforming based on limited-rate feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1853–1863, May 2006.

[29] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[30] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[31] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[32] A. Hero, "Secure space-time communication," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3235–3249, Dec. 2003.

[33] E. J. Candès and M. A. Davenport, "How well can we estimate a sparse vector," *Appl. Comput. Harmon. Anal.*, vol. 1, pp. 1–3, Aug. 2012.

[34] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA, USA: Kluwer, 1992.

[35] E. J. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Probab.*, vol. 23, no. 3, pp. 969–985, Jul. 2007.

[36] T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, Jul. 2011.

[37] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007.

[38] W. Chen, M. R. D. Rodrigues, and I. J. Wassell, "On the use of unit-norm tight frames to improve the average MSE performance in compressive sensing applications," *IEEE Signal Process. Lett.*, vol. 19, no. 1, pp. 8–11, Jan. 2012.

[39] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1395–1408, Jul. 2009.

[40] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 520–544, Jan. 2010.

[41] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Construction Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[42] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[43] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

Chia-Hua Lin was born in Hsinchu, Taiwan, in 1983. He received the M.S. degree in electrical and control engineering (ECE) from National Chiao Tung University (NCTU), Hsinchu, in 2007, where he has been pursuing the Ph.D. degree in ECE since 2009. From 2007 to 2009, he was with MediaTek, Inc., Taiwan, where he participated in the digital signal processing design for global positioning system. His research interests include signal processing for communications, in particular, GPS, multiple-input-multiple-output wireless communications, precoder design, and compressive sensing.

Shang-Ho (Lawrence) Tsai (SM'12) was born in Kaohsiung, Taiwan, 1973. He received the Ph.D. degree in electrical engineering from the University of Southern California, USA, in 2005. From 1999 to 2002, he was with the Silicon Integrated Systems Corporation, where he participated the VLSI design for DMT-ADSL systems. From 2005 to 2007, he was with MediaTek Inc. and participated in the VLSI design for MIMO-OFDM systems. In 2013, he was a Visiting Fellow with the Department of Electrical Engineering, Princeton University. Since 2007, he has been with the Department of Electrical and Control Engineering (now Department of Electrical Engineering), National Chiao Tung University, where he is currently an Associate Professor. His research interests include signal processing for communications, statistical signal processing, and signal processing for VLSI designs. He was awarded a government scholarship for overseas study from the Ministry of Education, Taiwan, from 2002 to 2005.

Yuan-Pei Lin (S'93–M'97–SM'03) was born in Taipei, Taiwan, in 1970. She received the B.S. degree in control engineering from the National Chiao Tung University, Taiwan, in 1992, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, in 1993, and 1997, respectively. She was with the Department of Electrical and Control Engineering, National Chiao Tung University, in 1997. Her research interests include digital signal processing, multirate filter banks, and signal processing for digital communications. She was a recipient of the Ta-You Wu Memorial Award in 2004. She served as an Associate Editor for the IEEE TRANSACTION ON SIGNAL PROCESSING, the IEEE TRANSACTION ON CIRCUITS AND SYSTEMS II, the IEEE SIGNAL PROCESSING LETTERS, the IEEE TRANSACTION ON CIRCUITS AND SYSTEMS I, the *EURASIP Journal on Applied Signal Processing*, and the *Multidimensional Systems and Signal Processing* (Academic Press). She was a Distinguished Lecturer of the IEEE Circuits and Systems Society from 2006 to 2007. She has also coauthored two books, *Signal Processing and Optimization for Transceiver Systems* (Cambridge University Press, 2010) and *Filter Bank Transceivers for OFDM and DMT Systems* (Cambridge University Press, 2010).