Secure Linear Precoding in Overloaded Wiretap MU-MIMO Networks

Abstract—**We propose a secure linear precoding scheme for the downlink of a multiuser MIMO (MU-MIMO) network that is tapped by an external eavesdropper (Eve). No knowledge on the location of Eve is assumed at the access point (Alice). The information signals for downlink users (Bobs) are accompanied by bogus signals (aka. friendly jamming) that are generated from Alice. The network is studied for both when it is underloaded and overloaded. In an underloaded (overloaded) network, the number of antennas at Alice is larger (smaller) than the total number of Bobs' antennas. In the overloaded setting, traditional methods of creating friendly jamming (FJ), such as zero-forcing-based methods, are infeasible. Our linear precoding scheme relaxes such infeasibility in overloaded MU-MIMO networks. In the worst-case scenario where Eve has knowledge of the channels between Alice and Bobs, we show that our method imposes the most stringent condition on the number of antennas required at Eve to cancel out FJ signals. We verify our analyses with simulations, and it turns out that choosing the number of independent streams to be sent to Bobs has an important role in achieving a trade-off between security, reliability and the achievable rate of the Bobs.**

Index Terms—**multiuser MIMO, linear precoding, physical-layer security, friendly jamming**

I. Introduction

The ever-increasing growth of wireless systems has made them an essential part of our daily life. Many users rely heavily on wireless networks for important/secret transmissions, such as financial information, emergency services, and health data. The proliferation of wireless communication devices also opens new doors for many security breaches ranging from eavesdropping to jamming attacks. Such disadvantage stems from the broadcast nature of wireless transmissions, which creates an exposed environment. Out of all malicious activities that can be done in a wireless network, our main focus in this paper is on eavesdropping attacks. While cryptographic techniques have been exploited to thwart eavesdropping attacks and thus realize a secure wireless network, recent advances in computational abilities of commercial electronic devices casts doubt on solely relying on cryptography for security. A number of new studies that have challenged common cryptographic techniques were mentioned in [1]. Cryptography also demands excessive overhead to establish necessary key distribution and authentication protocols. In future when wireless networks become loaded with massive amounts of nodes, conventional key agreement schemes may not scale well, which calls for fundamentally new approaches in maintaining the secrecy of wireless networks.

Physical-layer (PHY-layer) security (in its information-theoretic sense) has been introduced as a promising candidate to guarantee secrecy regardless of computational abilities of eavesdropper(s). Moreover, PHY-layer security techniques demand simple control protocols, thus have the potential to obviate traditional security protocols. In PHY-layer security, the main focus is to boost the signal strength of a legitimate transmitter (Alice) at its corresponding receiver (Bob) in a way that a nearby eavesdropper (Eve) receives a degraded version of that signal. In his seminal paper, Wyner showed that such a scenario is in line with achieving *perfect secrecy*, i.e., even when Eve has unlimited computational capability, she cannot do better than random guessing to decode the signal [1].

Of all the PHY-layer techniques proposed over the recent decade, the method of *friendly jamming(FJ)* gained considerable attention due to its practicality and no assumption on Eves' locations. Along with numerous fundamental and theoretical studies about its performance (e.g., [2], [3]), FJ techniques were also implemented in several recent studies (e.g., [4] and reference therein). In FJ techniques, Alice creates a bogus signal along with her secret message to deliberately garble the signal received at Eve but keep the signal strength at Bob intact. This idea is shown to be easily achievable using multiple antennas at Alice because Alice can create the FJ signal such that it falls in the null space of the channel between Alice and Bob. FJ was also shown to be realizable in other types of networks, such as relay networks [3], interference networks [2], and broadcast networks [5][1].In this paper, we focus on the application of TxFJ techniques in the downlink of a broadcast network[2]. Alice and Bobs, all have multiple antennas, resulting in a *multiuser MIMO (MU-MIMO) network*. MU-MIMO networks have been the subject of numerous studies, and several standards such as 802.11ac and LTE have been pushed to support this network architecture at least for downlink communications. The use of multiple antennas in MU-MIMO networks grants the best use of spectral resources by simultaneously servicing Bobs in downlink/uplink communications. Precoding approaches proposed over the last two decades have come a long way to approach the capacity of MU-MIMO networks. The theoretical precoding method of *dirty-paper coding* guarantees to achieve the capacity of these networks [7]. However, complicated and nonlinear design procedure of this method declines the

[1] Note that the notion of *broadcast network* in this paper refers to a network of one Alice and many Bobs where each Bob has his own separate message to be sent from Alice. The type of signaling is different from *multicast signaling* where Alice sends a common message to all Bobs (see [6]).

[2] A *broadcast network* refers to a network of one Alice and many Bobs, where each Bob receives his own separate message from Alice.

feasibility of implementing it in real-world systems. Instead, linear precoding schemes, such as the ones based on *zero forcing (ZF)* and *minimum mean square error (MMSE)* [8] criteria, have been extensively used in practical realizations of MU-MIMO networks. The PHY-layer secrecy of MU-MIMO networks has also been studied in the literature, and several precoders have been designed to create TxFJ in such networks [5], [6]. We narrow down our focus to an MU-MIMO network where Bobs are not malicious nodes, i.e., Bobs are not interested in transmissions of their neighbors. Instead, an external Eve exists in the network. The lower and upper bounds on the secrecy capacity of such networks were derived in [9]. The authors in [5] introduced TxFJ techniques for MU-MIMO networks. The study of MU-MIMO networks when massive number of antennas exist in Alice side was done in [10]. Other interesting problems related to the secrecy performance of FJ, such as the case where spatial correlation exists between Alice's antennas and power allocation between FJ and information signals were considered in [11], [12], respectively.

We are primarily interested in linear precoding design approaches, as nonlinear designs are not suitable for practical implementation. In conventional ZF-based methods for MU-MIMO networks, the number of antennas at Alice must be greater than or equal to the total number of antennas at Bobs so as to generate interference-free signals on all Bobs [13]. We refer to this condition as *information rate rank constraint (IRRC)*. The case where IRRC is met is referred to as the *underloaded* scenario. If IRRC is violated, the network is *overloaded*, and hence the ZF-based and MMSE-based precoder designs are infeasible. To satisfy IRRC in overloaded networks, scheduling algorithms have been used to select a subset of Bobs, thus creating an underloaded network. When no information on Eve's location is known (hence FJ techniques are typically used), the ZF method requires the MU-MIMO network to be underloaded to allow for creation of FJ signals [5]. We refer to this condition as the *secrecy rank constraint (SRC)*.

Antenna selection and scheduling are two different approaches to satisfy either IRRC or SRC in MU-MIMO networks. In fact, antenna selection decreases the number of data streams that Bobs can receive by selecting a subset of their antennas, while scheduling aims to reduce the total number of serviced Bobs without removing any of their antennas/streams. In an extensive recent study done by Björnson et.al [14], it was shown that in MU-MIMO networks where several multi-antenna Bobs exist, it is more beneficial (in terms of lowering the bit-error-rate) to decrease the number of streams for each Bob and service many Bobs than to decrease the number of Bobs (by scheduling). Henceforth, we focus on schemes where the number of streams are kept low to serve more Bobs.

While antenna selection can force the network to satisfy IRRC and SRC, selecting a subset of antennas is a difficult integer programming problem [13]. Antenna selection also requires RF switchers. These components can impose delay on receivers' operations if the wireless channels are sufficiently far from being slowly fading channels [15]. RF switchers also increases the cost of production [16]. Lastly, antenna selection may reduce the combining capabilities of Bobs. Specifically, when Bobs switch on a few number of antennas (or RF chains), they cannot increase the diversity as much as when all their antennas are functioning.

Motivated by these challenges, we propose a new linear precoding scheme for the downlink of a MU-MIMO network which uses FJ for achieving secrecy but relies on using a few streams per Bob to function in overloaded settings. To do this, we relax IRRC conditions, allowing for multi-user interference (MUI) between downlink users. However, we aim to minimize MUI at each downlink user via a specific precoder design. Our scheme offers the same complexity as the combination of a ZF-based (or MMSE-based) precoding with a suboptimal antenna selection algorithm. However, the sum-rate of our algorithm is the same as that of ZF-based precoding schemes merged with the optimal antenna selection algorithm.

It turns out that allowing MUI between downlink users not only enables our scheme to operate in overloaded settings, but also imposes the most stringent condition on the number of antennas that Eve requires to cancel out the FJ signals. Overall, the contributions of this paper are as follows:

- We propose a linear precoding scheme for the downlink of MU-MIMO networks that relies on minimizing the interference leakage caused from downlink signals. Our precoders are different from ZF-based precoders, as we relax the zero interference leakage condition to improve on the feasibility conditions of traditional precoders in over/fully loaded MU-MIMO networks.
- We also create FJ signals using the linear precoders that we designed for minimizing MUI. Compared to traditional methods of FJ, our approach demands the same complexity but imposes the most stringent condition on the number of antennas that Eve requires to cancel out FJ signals. Using simulations, we show that the freedom in choosing rank of our precoding matrices enables us to establish a trade-off between secrecy, reliability and sum rate of the network.

1) Notation: Boldface uppercase/lowercase letters denote matrices/vectors. $\mathbf{A}^{(:,a:b)}$ and $\mathbf{A}^{(a:b,:)}$, respectively denote matrices comprised of columns a to b of $\mathbf{A}$ and rows a to b of $\mathbf{A}$. $\mathbf{I}$ and $\mathbf{0}$ denote the identity matrix and the zero matrix (i.e., matrix with zero entries) of appropriate sizes. E[$\bullet$], $\bullet^\dagger$, Tr($\bullet$) are respectively, the expected value, conjugate transpose, and trace operators. Lastly, $\mathbb{C}$ is the set of complex numbers.

2) General System Model: Consider a network where Alice has M antennas and communicates with Q Bobs, $Q \geq 2$. Let $\mathcal{Q} = \{1, 2, \ldots, Q\}$. Bob$_q$ has $N_q < M$ antennas, $q \in \mathcal{Q}$. Without loss of generality, assume that all Bobs have the same number of antennas, i.e., $N_q = N < M$, $\forall q \in \mathcal{Q}$. An external Eve with L antennas also exists in the range of communications[3]. The setting where $M = NQ$ is referred to

[3] A single Eve with L antennas can also represent several multi-antenna colluding Eves.

as the fully-loaded scenario. When $M < NQ$, the network is overloaded, and when $M > NQ$ the network is underloaded. Bob$_q$, $q \in \mathcal{Q}$, receives K_q independent streams from Alice, where $K_q \leq N$. Without loss of generality, assume that $K_q = K, \ \forall q \in \mathcal{Q}$. The number of streams determines how the antennas at Alice and Bobs are exploited. For example, $K = N$ indicates that the signals intended for Bobs have the maximum number of streams, thus the antennas are used to provide spatial multiplexing. In contrast, $K = 1$ signifies that the combining features of Bobs are used to increase the diversity (thus reliability) of transmissions.

II. Conventional Precoder Design

To better understand our method, we first explain the ZF method used in designing the precoding matrices. The received signal at Bob$_q$, $q \in \mathcal{Q}$, can be expressed as

$$\mathbf{y}_q = \mathbf{H}_q(\mathbf{u} + \mathbf{f}) + \mathbf{n} \tag{1}$$

where $\mathbf{y}_q \in \mathbb{C}^N$, $\mathbf{H}_q \in \mathbb{C}^{N \times M}$ is the complex channel between Alice and Bob$_q$, $\mathbf{u} \in \mathbb{C}^M$ is the signal containing information from Alice, $\mathbf{f} \in \mathbb{C}^M$ is the FJ signal, and $\mathbf{n} \in \mathbb{C}^N$ is the AWGN which has i.i.d. zero-mean-circularly-symmetric-complex Gaussian- (ZMCSCG-) distributed entries with E$[\mathbf{n}\mathbf{n}^\dagger] = N_0/N\mathbf{I}$. The signal $\mathbf{u}$ is expressed as

$$\mathbf{u} \triangleq \sum_{q=1}^{Q} \mathbf{u}_q \triangleq \sum_{q=1}^{Q} \mathbf{T}_q \mathbf{s}_q \tag{2}$$

where $\mathbf{u}_q \in \mathbb{C}^M$ is the signal intended for Bob$_q$. $\mathbf{T}_q$ is the precoder that is responsible for cancelling the MUI generated from $\mathbf{u}_q$. $\mathbf{s}_q \in \mathbb{C}^K$ is the K-dimensional information signal (K streams of data) intended for Bob$_q$.

Assume that E$[\mathbf{s}_q\mathbf{s}_q^\dagger] = \phi P_q/K\mathbf{I}$, where P_q is the power of Alice allocated to Bob$_q$'s signal and ϕ is the portion of Alice's total power allocated to all information signals. Let $P \triangleq \sum_{q=1}^{Q} P_q$, where P is the Alice's total power. Alice allocates ϕP of her total power to all information signals. The rest of the power (i.e., $(1-\phi)P$) goes to the FJ signal.

We assume that Alice knows all $\mathbf{H}_i$, $\forall i \in \mathcal{Q}$, and Bob$_q$ only knows $\mathbf{H}_q$. In the channel estimation phase, Alice sends pilot signals to Bobs, so that Bob$_q$ can estimate $\mathbf{H}_q$ and feed it back to Alice. Substituting (2) in (1), the effective channel that Bob$_q$ sees from Alice would be $\mathbf{H}_q\mathbf{T}_q$. Hence, Alice can apply another precoder for each Bob to optimize her transmissions. Specifically, assume that $\mathbf{T}_q \in \mathbb{C}^{M \times \tau}, \ K < \tau \leq N$. Then, Alice can assign an extra precoder $\mathbf{W}_q \in \mathbb{C}^{\tau \times K}$, so that $\mathbf{y}_q$ can be written as

$$\mathbf{y}_q = \mathbf{H}_q \Big(\sum_{q=1}^{Q} \mathbf{T}_q \mathbf{W}_q \mathbf{s}_q + \mathbf{f} \Big) + \mathbf{n}. \tag{3}$$

Bob$_q$ also applies a linear combiner to estimate the transmitted information signal. In particular, Bob$_q$ applies $\mathbf{D}_q \in \mathbb{C}^{K \times N}$ to have the following estimate of $\mathbf{s}_q$:

$$\hat{\mathbf{s}}_q \triangleq \mathbf{D}_q \mathbf{y}_q = \mathbf{D}_q \Bigg(\mathbf{H}_q \Big(\sum_{q=1}^{Q} \mathbf{T}_q \mathbf{W}_q \mathbf{s}_q + \mathbf{f} \Big) + \mathbf{n} \Bigg). \tag{4}$$

Let $\mathbf{H}_q\mathbf{T}_q = \mathbf{U}_q\mathbf{\Sigma}_q\mathbf{V}_q^\dagger$ be the singular-value decomposition (SVD) of $\mathbf{H}_q\mathbf{T}_q$, where $\mathbf{U}_q$ and $\mathbf{V}_q$ are the unitary matrices of left and right singular vectors, and $\mathbf{\Sigma}_q$ is the matrix of singular values. Therefore, if Alice sets $\mathbf{W}_q = \mathbf{V}_q^{(:,1:K)}$ and Bob$_q$ sets $\mathbf{D}_q = \mathbf{U}_q^{(:,1:K)\dagger}$, the optimal precoder/combiner duo to estimate $\mathbf{s}_q$ at Bob$_q$ can be established [5].

We now focus on the design of $\mathbf{T}_q$ and $\mathbf{f}$. The ZF method is based on nullifying both the FJ signal and MUI on unintended Bobs. Formally, the following conditions must be satisfied:

$$\mathbf{H}_r\mathbf{T}_q = \mathbf{0}, \ r \neq q, \ \forall r, q \in \mathcal{Q} \tag{5a}$$

$$\mathbf{H}_q\mathbf{f} = \mathbf{0}, \ \forall q \in \mathcal{Q} \tag{5b}$$

The precoder $\mathbf{T}_q$ can be determined as follows. Define $\bar{\mathbf{H}}_q \triangleq [\mathbf{H}_1^\dagger, \ldots, \mathbf{H}_{q-1}^\dagger, \mathbf{H}_{q+1}^\dagger, \ldots, \mathbf{H}_Q^\dagger]^\dagger \in \mathbb{C}^{N(Q-1)\times M}$, and let $\bar{\mathbf{H}}_q = \mathbf{L}_q\mathbf{J}_q\mathbf{R}_q$ be the SVD of $\bar{\mathbf{H}}_q$, where $\mathbf{L}_q$ and $\mathbf{R}_q$ denote the matrices of left and right singular vectors, and $\mathbf{J}_q$ denotes the matrix of singular values. Provided that $M > N(Q-1)$, $\bar{\mathbf{H}}_q$ has a nontrivial null-space, which can be exploited to meet condition (5a). Specifically, if $M > N(Q-1)$, Alice sets $\mathbf{T}_q = \mathbf{R}_q^{(:,B:B+\tau)} \in \mathbb{C}^{M \times \tau}$, where $B = N(Q-1)+1$, to satisfy (5a) for all $q \in \mathcal{Q}$. The condition

$$M \geq N(Q-1) + \tau \tag{6}$$

constitutes the IRRC in the downlink of the ZF method. The FJ signal mentioned in (1) has the following structure in the ZF method. Define $\tilde{\mathbf{H}} \triangleq [\mathbf{H}_1^\dagger, \ldots, \mathbf{H}_Q^\dagger]^\dagger \in \mathbb{C}^{NQ \times M}$. Let $\tilde{\mathbf{H}} = \mathbf{L}\mathbf{J}\mathbf{R}$ be the SVD of $\tilde{\mathbf{H}}$, where $\mathbf{L}$ and $\mathbf{R}$ denote the matrices of left and right singular vectors, and $\mathbf{J}$ denotes the matrix of singular values. To satisfy (5b), $\tilde{\mathbf{H}}$ must have a nontrivial null-space, which requires $M > NQ$. Hence, the inequality $M > NQ$ is the SRC for the ZF method. We choose $\tau = N$, as IRRC in (6) is dominated by SRC. The FJ signal is expressed as $\mathbf{f} = \mathbf{Z}\mathbf{v}$, where $\mathbf{Z}$ is the associated precoder for FJ, which spans the null space of $\tilde{\mathbf{H}}$, and $\mathbf{v}$ is the vector of artificial noise that has the same characteristics of AWGN except that Tr$[\mathbf{v}\mathbf{v}^\dagger] = (1-\phi)P$. If SRC is violated, the creation of FJ signal becomes infeasible.

III. Proposed Signaling scheme

In this section, we introduce our proposed signaling scheme. Although the precoding design in this section is not much different from previous section, the signaling scheme that we propose here will play an important role in the design of our precoders in the next section. We first modify the signal model at Bobs and Eve in (3) and (12). Specifically, the received signal at Bob$_q$, $q \in \mathcal{Q}$ can be expressed as

$$\mathbf{y}_q = \mathbf{H}_q\mathbf{u}' + \mathbf{n} \tag{7}$$

where $\mathbf{u}'$ is Alice's signal in our proposed signaling scheme:

$$\mathbf{u}' = \sum_{q=1}^{Q} \left(\mathbf{u}'_q + \mathbf{f}'_q \right) \tag{8}$$

where $\mathbf{u}'_q$ is the signal intended for Bob$_q$, $q \in \mathcal{Q}$, and $\mathbf{f}'_q$ is the FJ signal designed to protect Alice's transmissions that are intended for Bob$_q$. In fact, compared to (1), the main change in the signal model is the decomposition of the FJ signal (i.e., convert $\mathbf{f}$ to $\mathbf{f}'_q$, $q \in \mathcal{Q}$) in a way that each FJ signal exclusively protects the transmissions intended for one Bob.

A more detailed representation of $\mathbf{u}'$ can be given as

$$\mathbf{u}' = \sum_{q=1}^{Q} \mathbf{T}'_q(\mathbf{W}'_q\mathbf{s}_q + \mathbf{Z}'_q\mathbf{v}'_q) \tag{9}$$

with $\mathbf{u}'_q = \mathbf{T}'_q\mathbf{W}'_q\mathbf{s}_q$ and $\mathbf{f}'_q = \mathbf{T}'_q\mathbf{Z}'_q\mathbf{v}'_q$. The precoder $\mathbf{T}'_q$ is responsible for cancelling MUI and FJ on unintended Bobs, $\mathbf{W}'_q$ is the precoder to boost signal strength on Bob$_q$ (same as $\mathbf{W}_q$ in previous section), $\mathbf{Z}'_q$ is the precoder for the FJ signal that protects Bob$_q$, and $\mathbf{v}'_q$ is the vector of artificial noise. As before, $\mathbf{s}_q$ is the K-stream information signal intended for Bob$_q$. Because precoder $\mathbf{T}'_q$ is applied to both information and FJ signals (compare (9) and (2)), we are ensured that FJ will have no effect on unintended Bobs. As in (4), a linear receiver $\mathbf{D}'_q$ is applied at Bob$_q$ to recover $\mathbf{s}_q$. Using (7) and (9), Bob$_q$ has the following estimate of $\mathbf{s}_q$

$$\hat{\mathbf{s}}_q \triangleq \mathbf{D}'_q\mathbf{y}_q = \mathbf{D}'_q\left(\mathbf{H}_q\Big(\sum_{q=1}^{Q}\mathbf{T}'_q(\mathbf{W}'_q\mathbf{s}_q + \mathbf{Z}'_q\mathbf{v}'_q)\Big) + \mathbf{n}\right). \tag{10}$$

The conditions for completely nullifying the MUI and FJ signals for the signal model in this section are as follows:

$$\mathbf{H}_r\mathbf{T}'_q = \mathbf{0},\ r \neq q,\ \forall r,q \in \mathcal{Q} \tag{11a}$$

$$\mathbf{D}'_q\mathbf{H}_q\mathbf{T}'_q\mathbf{Z}'_q\mathbf{v}'_q = \mathbf{0},\ \forall q \in \mathcal{Q} \tag{11b}$$

The design of $\mathbf{T}'_q$, $\mathbf{W}'_q$, and $\mathbf{D}'_q$ would be the same as those of $\mathbf{T}_q$, $\mathbf{W}_q$ and $\mathbf{D}_q$ in the previous section. Therefore, the IRRC of our method is the same as that of conventional ZF. All FJ signals are removed by a combination of (11a) and (11b). Notice that (11b) is different from (5b) in that $\mathbf{Z}'_q$ in (11b) is designed so that only $\mathbf{v}'_q$ is nullified at Bob$_q$ with the help of $\mathbf{D}'_q$. The rest of FJ signals (i.e., $\mathbf{v}'_r,\ r \neq q$) are removed by $\mathbf{T}'_q$ that satisfies (11a). Therefore, the SRC of our method is determined by the condition that is the most dominant in (9). Due to keeping the same design of the conventional ZF method for $\mathbf{T}'_q$, the SRC is the same as IRRC in our method, i.e., $M \geq NQ$ given that $\tau = N$ (see (6)).

Because we use a different procedure to nullify the FJ signal, the design of $\mathbf{Z}'_q$ is different from $\mathbf{Z}$ of the previous section in that $\mathbf{Z}'_q$ is designed for each Bob$_q$. Let $\mathbf{H}_q\mathbf{T}'_q = \mathbf{U}'_q\boldsymbol{\Sigma}'_q\mathbf{V}_q'^{\dagger}$ be the SVD of $\mathbf{H}_q\mathbf{T}'_q$, where $\mathbf{U}'_q$ and $\mathbf{V}'_q$ are the unitary matrices of left and right singular vectors, and $\boldsymbol{\Sigma}'_q$ is the matrix of singular values. Therefore, if Alice sets $\mathbf{W}'_q = \mathbf{V}_q'^{(:,1:K)}$, $\mathbf{D}'_q = \mathbf{U}_q'^{(:,1:K)\dagger}$ (same as previous section), and $\mathbf{Z}'_q = \mathbf{V}_q^{(:,K+1:\tau)}$, then (11b) is also satisfied (compare with the design of $\mathbf{Z}$).

A. Security Analysis of the Proposed Method

The received signal at Eve can be expressed as

$$\mathbf{z} = \mathbf{G}\mathbf{u}' + \mathbf{e} = \mathbf{G}\Big(\sum_{q=1}^{Q}\big(\mathbf{u}'_q + \mathbf{f}'_q\big)\Big) + \mathbf{e} \tag{12}$$

where $\mathbf{G} \in \mathbb{C}^{L\times M}$ is the channel between Alice and Eve, and $\mathbf{e}$ has the same characteristics as $\mathbf{n}$ in (1). Eve has to first combat the MUI to be able to wiretap ongoing communications. Eve does so by applying a linear combiner. For example, to eavesdrop on signals intended for Bob$_q$, Eve first applies $\mathbf{A}'_q$ on the signal she receives. Define $\mathbf{z}_q \triangleq \mathbf{A}'_q\mathbf{z}$. Upon cancelling MUI with $\mathbf{A}'_q$, Eve applies $\mathbf{B}'_q$ on $\mathbf{z}_q$ to estimate $\mathbf{s}_q$. In other words, Eve's estimation from $\mathbf{s}_q$ is $\tilde{\mathbf{s}}_q = \mathbf{B}'_q\mathbf{z}_q$. We assume the worst-case scenario where Eve knows $\mathbf{G}$. For instance, Eve can use the pilot signals sent from Alice in the channel estimation phase to estimate $\mathbf{G}$. Moreover, because Bobs have to explicitly feed back the channel estimates to Alice, Eve can snoop on the channel estimation feedback from Bobs to gain knowledge of all $\mathbf{H}_q$, $\forall q \in \mathcal{Q}$. Note, however, that neither Alice nor Bobs have any knowledge of $\mathbf{G}$, i.e., Eve is a passive eavesdropper.

We now describe how Eve chooses her combiners to decode Alice's transmissions. We also show how many antennas Eve requires to decode all messages. Using (12), $\mathbf{z}_q = \mathbf{A}'_q\mathbf{z}$, and the linear estimate $\tilde{\mathbf{s}}_q = \mathbf{B}'_q\mathbf{z}_q$, we have the following

$$\tilde{\mathbf{s}}_q = \mathbf{B}'_q\mathbf{A}'_q\left(\mathbf{G}\Big(\sum_{q=1}^{Q}\big(\mathbf{u}'_q + \mathbf{f}'_q\big)\Big) + \mathbf{e}\right). \tag{13}$$

Eve cancels MUI by designing a combiner $\mathbf{A}'_q$ such that

$$\mathbf{A}'_q\mathbf{G}\left(\mathbf{u}'_r + \mathbf{f}'_r\right) = 0,\ r \neq q,\ \forall r,q \in \mathcal{Q} \tag{14a}$$

$$\mathbf{A}'_q\mathbf{G}\,\mathbf{f}'_q = 0,\ \forall q \in \mathcal{Q} \tag{14b}$$

Using (8), (9) and (13), Eve first constructs the following blocked matrix

$$\mathbf{G}'_q = [\boldsymbol{\Omega}'_1, \ldots, \boldsymbol{\Omega}'_{q-1}, \boldsymbol{\Omega}'_{q+1}, \ldots, \boldsymbol{\Omega}'_Q, \boldsymbol{\Gamma}'_q] \tag{15}$$

where $\boldsymbol{\Omega}'_q = \mathbf{G}\mathbf{T}'_q \in \mathbb{C}^{L\times\tau}$ and $\boldsymbol{\Gamma}'_q = \mathbf{G}\mathbf{T}'_q\mathbf{Z}'_q \in \mathcal{C}^{L\times\tau-K}$. Eve sets $\mathbf{A}'_q$ to be the last K columns of the matrix of left singular values of $\mathbf{G}'_q$. For such a choice of $\mathbf{A}'_q$ that allows Eve to cancel MUI and FJ, the minimum value of L is derived by counting the column of $\mathbf{G}'_q$, i.e.,

$$\Psi' = \tau(Q-1) + (\tau - K) + K = \tau Q \tag{16}$$

Setting $\tau = N$, we have $\Psi' = NQ$. The first term in the right hand side (RHS) of (16) is the number of antennas that $\Omega_r,\ r \neq q,\ r \in \mathcal{Q}$ occupies in establishing $\mathbf{G}'_q$ in (15). The second term in (16) is the number of antennas that $\boldsymbol{\Gamma}'_q$ occupies in (15). Finally, the third term is the number of antennas that are required to recover $\mathbf{s}_q$ after nullifying MUI and FJ. The same security analysis can be done for the ZF method, and it can be shown that if Alice uses the conventional ZF method, Eve requires at least $\Psi = M - (N-K)Q$ antennas.

B. Security Comparison Between Conventional ZF Method and the Proposed Method

We now compare required the number of Eve's antennas for both the ZF and the proposed method to decode all messages in an underloaded scenario, i.e., we compare Ψ and Ψ' when $M > NQ$. Consider the conditions when $\Psi > \Psi'$, i.e., $M - (N-K)Q > NQ$. In other words, we examine when the ZF method is better than our approach. Clearly such a comparison depends on K, which is analyzed as follows:

- For $K = N$, we end up with $M > NQ$, which is always true in the underloaded scenario, so in the case of using all streams (i.e., spatial multiplexing), the ZF method imposes a more stringent condition than our method.
- For $K < N$, the simplified inequality is $2N - K < \frac{M}{Q}$. By lowering the number of streams (K), it can be deduced that the ZF method imposes more antennas on Eve than our method only when the network is *sufficiently* underloaded. To clarify, take the extreme example of $K = 1$; In this case, $M-(N-K)Q > NQ$ is reduced to $M > (2N-1)Q$ which is more demanding than an underloaded network (i.e., $M > NQ$) with $N > 1$.

Overall, when a few streams are selected for each Bob, the ZF method does not impose more antennas on Eve than our proposed method unless the network is sufficiently underloaded. Normally, a sufficiently underloaded is not preferred, as the MU-MIMO network would not be fully utilized.

C. Antenna Selection for Zero Forcing Precoding

To compensate for the absence of FJ in over/fully loaded scenarios, antenna selection algorithms can be used to decrease the number of functioning receive antennas at Bobs from N to N', so that SRC can be satisfied, i.e., $M > N'Q$. We mainly focus on *capacity-based* antenna selection algorithms, but our analysis can be simply extended to other types of antenna selection algorithms. We introduce antenna selection for when the the network is over/fully loaded, i.e., the number of Bobs is large enough that $M \leq NQ$.

The capacity of the channel between Alice and Bob$_q$, $q \in \mathcal{Q}$ can be expressed as[4]

$$C_q = \log\det(\mathbf{I} + \phi P_q \mathbf{H}_q \mathbf{H}_q^\dagger) \tag{17}$$

Using antenna selection, we are interested in switching on only $K \leq N' < N$ antennas of Bob$_q$ such that $M > N'Q$. Denote $\bar{\mathbf{H}}_q$ as a matrix comprised of N' columns of $\mathbf{H}_q$. Denote $\mathcal{S}(\mathbf{H}_q)$ as the the set of matrices that are formed using N' rows of $\mathbf{H}_q$. Therefore, the problem of antenna selection can be formulated as

$$\bar{\mathbf{H}}_q^* = \underset{\mathcal{S}(\mathbf{H}_q)}{\arg\max}\Bigg(\log\det(\mathbf{I} + \phi P_q \bar{\mathbf{H}}_q \bar{\mathbf{H}}_q^\dagger)\Bigg) \tag{18}$$

where $\bar{\mathbf{H}}_q^* \in \mathbb{C}^{N' \times M}$. The optimal antenna selection is a difficult integer programming problem, thus suboptimal algorithms such as [17] can be used which are based on maximizing the upper bounds of the capacity. After performing the antenna selection at each Bob, the SRC is expected to be met (i.e., $M > N'Q$), which allows for creation of FJ. Hence, by replacing N with N', we can deduce that the number of antennas required at Eve to cancel out FJ and MUI in the ZF method with antenna selection would be $M - (N' - K)$[5].

IV. Proposed Precoding Method

The current precoder design for $\mathbf{T}'_q$ in our proposed signaling scheme has two issues. First, the IRRC condition is still the same as that of the conventional ZF method, which prohibits our signaling scheme from operating in overloaded scenarios. Second, after implementation of these precoders, although for $K < N$ our signaling scheme can impose more antennas on Eve to decode the ongoing messages –by adding more columns to matrix $\mathbf{G}'_q$ in (13), see Section III-B–, it turns out that the rank of $\mathbf{G}'_q$ does not increase with the added columns. Therefore, Eve can still decode the signals with fewer antennas than what our proposed signaling scheme claims. In this section, we modify the design of $\mathbf{T}'_q$ to resolve these issues.

To do so, we relax condition (11a) in a way that MUI created from $\mathbf{s}_q$ inflicts the least amount of damage on the reception of other Bobs. Formally, we design the precoder $\mathbf{T}'_q$, $q \in \mathcal{Q}$ using an optimization problem that is detailed later on. Before presenting this optimization problem, we formulate the ZF method as a variant of a family of optimization problems. Consider the following optimization problem

$$\begin{aligned} \underset{\mathbf{T}'_q}{\text{maximize}} \quad & \frac{||\mathbf{H}_q \mathbf{T}'_q||_F}{\sum_{\substack{r=1 \\ r \neq q}}^{Q} ||\mathbf{H}_r \mathbf{T}'_q||_F + \frac{N_0}{\phi P_q}} \\ \text{s.t.} \quad & {\mathbf{T}'_q}^\dagger \mathbf{T}'_q = \mathbf{I} \end{aligned} \tag{19}$$

where $||\bullet||_F$ is the Ferobnius norm. In problem (19) the precoder for Bob$_q$ must be designed in a way that the interference generated from $\mathbf{s}_q$ (i.e., denominator of the objective in (19)) is minimized while the strength of $\mathbf{s}_q$ at Bob$_q$ (i.e., the numerator of the objective) is maximized. The constraint on $\mathbf{T}'_q$ causes the product $\mathbf{H}_q \mathbf{T}'_q$ to have the same statistical properties of $\mathbf{H}_q$. Problem (19) is identified as a Rayleigh quotient problem [18]. It is easy to see that when $N_0 << \phi P_q$ (i.e., high SNR scenario), the solution to (19) reduces to the ZF method from the previous section because the maximum objective value would be achieved if the denominator goes to zero, which is in line with condition (5a) or (11a). In moderate SNRs, the solution to (19) reduces to MMSE-based precoding methods [19]. Also notice that problem (19) does not impose any rank constraint on its solution. We now examine (11a) again. This condition imposes the result of $\mathbf{H}_q \mathbf{T}'_q$ to have entries with the minimum possible value. We decompose (11a) as follows:

$$\mathbf{H}_r {\mathbf{T}'_q}^{(:,n)} = \mathbf{0}, \;\; r \neq q, \;\; \forall r, q \in \mathcal{Q} \;\&\; \forall n \tag{20}$$

[4] Note that such a capacity can be achieved with dirty-paper coding scheme, which is a nonlinear precoding method [7].

[5] Clearly, antenna selection can also be performed in situations where IRRC is also violated. However, for the sake of brevity, we only apply antenna selection to satisfy SRC.

where ${\mathbf{T}'_q}^{(:,n)}$ is the nth column of $\mathbf{T}'_q$. In fact (20) suggests the same condition in (11a) but is represented on a column-by-column basis. Also notice that since we have not explicitly designed $\mathbf{T}'_q$ yet, we do not impose any constraints on its rank, thus no information is yet available on the values that n in (20) can take. For now, assume that $n \in \{1, \dots, \tau\}$ where $K \leq \tau \leq N$. Instead of (19), we propose our precoding method by formulating the following optimization problem

$$\underset{\mathbf{T}'_q}{\text{maximize}} \quad \frac{||\mathbf{H}_q {\mathbf{T}'_q}^{(:,n)}||_F}{\sum_{\substack{r=1\\ r\neq q}}^{Q} ||\mathbf{H}_r {\mathbf{T}'_q}^{(:,n)}||_F + \frac{N_0}{\phi P_q N}}$$
$$\text{s.t.} \quad {\mathbf{T}'_q}^{(:,n)} {{\mathbf{T}'_q}^{(:,n)}}^{\dagger} = \frac{1}{\tau}, \; n \in \{1, \dots, \tau\}. \qquad (21)$$

Problem (21) is still a Rayleigh quotient problem, but the difference with (19) is that in (21) we find the solution on a column-by-column basis. The constraint in (21) ensures that the resulting precoder does not violate the power constraint. In fact, because we assumed that $E[\mathbf{s}_q \mathbf{s}_q^\dagger] = \phi P_q / K\mathbf{I}$, we must also ensure that ideally, $E[\mathbf{T}'_q \mathbf{s}_q \mathbf{s}_q^\dagger {\mathbf{T}'_q}^{\dagger}] = \phi P_q / K\mathbf{I}$ (see (2) and description of $\mathbf{s}_q$ below it). The solution to (21) is given by [20]

$$ {\mathbf{T}'_q}^{*(:,n)} = \frac{1}{\sqrt{\tau}} \frac{\mathbf{\Delta}^{(:,n)}}{||\mathbf{\Delta}^{(:,n)}||_F} \qquad (22)$$

where $\mathbf{\Delta}$ is the matrix of generalized eigenvectors corresponding to τ non-zero generalized eigenvalues of numerator and denominator of the objective in (21), i.e.,

$$\mathbf{\Delta} \triangleq \text{eig}_{max,\tau}\Big(\mathbf{H}_q^\dagger \mathbf{H}_q, \sum_{\substack{r=1\\ r\neq q}}^{Q} \mathbf{H}_r^\dagger \mathbf{H}_r + \frac{N_0}{\phi P_q N}\Big) \qquad (23)$$

where $\text{eig}_{max,\tau}$ is the operator for extracting τ generalized eigenvector that correspond to τ non-zero generalized eigenvalues. From the properties of generalized eigenvalue problems, it can be deduced that there are N eigenvectors that correspond to non-zero generalized eigenvalues in (23) [20]. Hence, $\mathbf{\Delta} \in \mathbb{C}^{M\times\tau}$.

Solving problem (21) allows us to relax the condition in (11a). Interestingly, there is no guarantee on the solution of (21) to satisfy the constraint of (19), which makes (19) and (21) to be essentially not equivalent to each other. Even in high SNR scenario, there is no guarantee on the equivalence of the solutions of (19) and (21). In fact, that the resulting precoders of (21) are do not necessarily have diagonal covariance matrices to satisfy the constraint in (19). However, the constraint in (21) ensures that ${\mathbf{T}'_q}^{*}$ does not violate the power constraint at Alice. Specifically, in ${{\mathbf{T}'_q}^{*}}^{\dagger} {\mathbf{T}'_q}^{*}$, we have the following

$${{\mathbf{T}'_q}^{*(:,r)}}^{\dagger} {\mathbf{T}'_q}^{*(:,n)} = \frac{1}{\tau} \frac{{\mathbf{\Delta}^{(:,r)}}^{\dagger} \mathbf{\Delta}^{(:,n)}}{||\mathbf{\Delta}^{(:,r)}||_F ||\mathbf{\Delta}^{(:,n)}||_F} \leq \frac{1}{\tau}. \qquad (24)$$

Therefore, $||{{\mathbf{T}'_q}^{*(:,n)}}^{\dagger} {\mathbf{T}'_q}^{*}||_F \leq 1$ is guaranteed, ensuring that our proposed precoding in (21) does not violate the power constraint, i.e., $E[{\mathbf{T}'_q}^{*} \mathbf{s}_q \mathbf{s}_q^\dagger {\mathbf{T}'_q}^{*\dagger}] \leq \phi P_q / K\mathbf{I}$[6]. In summary, our proposed method in (21) relaxes the general shape of the ZF-based and MMSE-based precoders that are known from problem (19), such that the MUI is still minimized as much as possible.

In case of an underloaded network (i.e., $M > NQ$), we set $\tau = N$ (i.e., same as Section II and III). In case of over/fully loaded networks (i.e., $M \leq NQ$), we set $\tau = \lceil \frac{M}{Q} \rceil$, where $\lceil \bullet \rceil$ is the ceiling function to handle the case of non-integer values of τ. Notice that in an overloaded scenario, we do not decrease Q via scheduling. Instead, we have the freedom in choosing τ and still keeping all users in the network. Using the fact that $K < \tau \leq N$, we can also determine the value of K. After designing $\mathbf{T}'_q$ and determining K, the remaining matrices in our proposed method (i.e., $\mathbf{W}'_q$, $\mathbf{D}'_q$ and $\mathbf{Z}'_q$) can be designed as in Section III. Hence, all terms in (9) and (10) are defined, and our proposed precoding method is complete.

The security analysis of our method in underloaded scenarios was already done in Section III-B, where we showed Eve requires $\Psi' = \tau Q$ antennas to decode all messages. In the case of overloaded network as mentioned before, we choose $\tau = \lceil \frac{M}{Q} \rceil$. Hence, $\Psi' = \max\{\tau Q, M\}$ which is the most stringent condition on Eve's number of antennas. The conventional ZF method is not able to generate the FJ signal in an overloaded network because condition (5b) cannot be satisfied. Hence, it can be shown that Eve only requires $\Psi = KQ$ antennas to decode all messages in ZF method. As $KQ < \max\{\tau Q, M\}$, then our method always performs better than the conventional ZF scheme in overloaded networks.

Notice that our proposed precoder design for $\mathbf{T}'_q$ in this section can also be used in the conventional ZF method to design $\mathbf{T}_q$ for overloaded scenarios and relax condition (5a). However, there will be no increase in the number of Eve's antennas required to decode Alice's messages because the design of FJ in the conventional ZF method is decoupled from the design of $\mathbf{T}_q$.

Overall, the combination of our signaling scheme in Section III and the precoder design in Section IV not only handles the overloaded scenarios (without scheduling), but also increases the rank of $\mathbf{G}'_q$ in (13), which leads to increase in the number of antennas that Eve requires to decode all messages.

Although our method and the optimal antenna selection perform equally, we already mentioned that antenna selection methods are prone to many issues which are mainly to do with requiring RF switchers. However, our approach does not require these considerations. In terms of computational complexity of our method and antenna selection, our method is dominated by the computation of generalized eigenvalues and several SVD calculations. The complexity of antenna selection methods are also dominated by the calculation of SVD and solving the optimization in (18). Our derivations –which are skipped here for the sake of brevity– show that both methods demand the same amount of computational complexity.

[6] In the ZF method, it can be easily seen that the resulting ZF precoder satisfies $||{\mathbf{T}_q^{(:,n)}}^{\dagger} \mathbf{T}_q||_F = 1$. Thus, $E[\mathbf{T}_q \mathbf{s}_q \mathbf{s}_q^\dagger {\mathbf{T}_q}^{\dagger}] = \phi P_q / K\mathbf{I}$.

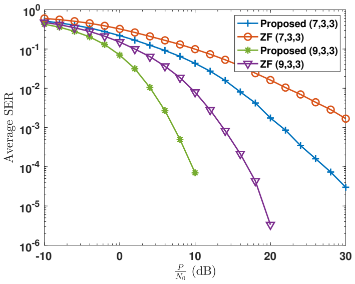

Fig. 1: Comparison of SER (Underloaded)

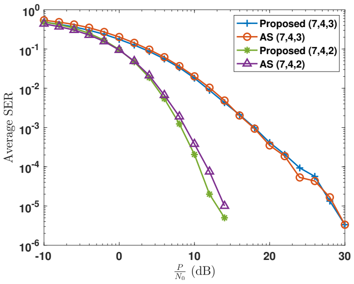

Fig. 2: Comparison of SER (Overloaded)

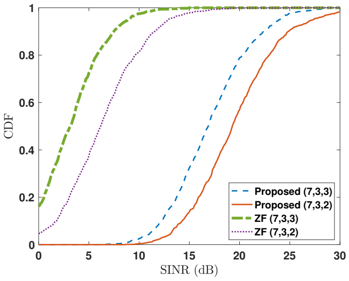

Fig. 3: Comparison of achieved SINR (Underloaded, $P = 30$ dB)

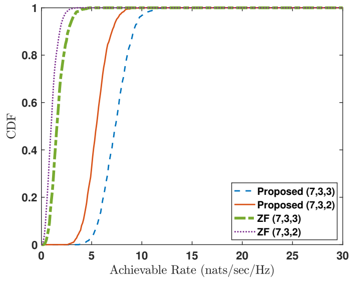

Fig. 4: Comparison of achievable rate (Underloaded)

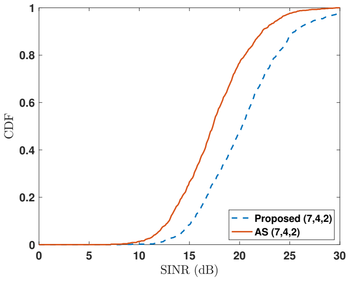

Fig. 5: Comparison of achieved SINR (Overloaded, $P = 30$ dB) S

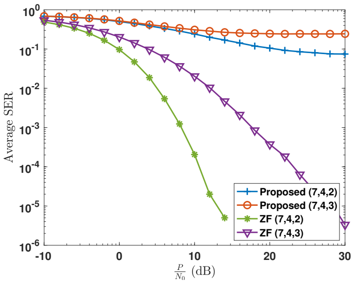

Fig. 6: Comparison of Eve's SER (Overloaded)

V. Numerical Results

We verify our theoretical analyses via simulations. All simulations are done for a network of $Q = 2$ Bobs. Similar conclusions can be drawn for networks with more Bobs and more antennas at Alice. Our *proposed method* in these simulations is the combination of the methods in Section III and Section IV, while the simulated ZF method is the scheme that we discussed in Section II. In our proposed method, the power allocated to Bob's message is divided equally between its associated information and FJ signals. The total power allocated to each Bob is also calculated using the method in [21]. Same is done for the ZF method. We use uncoded QPSK modulation for all simulations. For simulation that show SINR and achievable rate, we use Gaussian codebooks. The triplet (M, N, K) in all simulations denote number of Alice/Bob antennas and number of data streams.

Fig. 1 shows the symbol error rate (SER) of the Alice-Bob channels, averaged across all Bobs for an underloaded scenario. It can be seen that our proposed method outperforms the ZF method for both settings because our precoders are more flexible. In fact, although the precoders designed by the ZF method completely suppress MUI, they also do not contribute to the strength of the signal to the intended user.

Fig. 2 shows the SER for an overloaded scenario. It can be seen that our method's performance is close to that of antenna selection (AS) schemes. However, as mentioned earlier, our method does not have the problems of AS schemes (see Section I for our thorough explanation about AS schemes).

Fig. 3 shows the CDF of the achieved SINR in an underloaded scenario. Our method achieves higher SINR compared to the ZF method. This in fact decreases the SER of our scheme as shown in Fig. 1.

Fig. 4 shows the CDF of achievable information rate. As can be seen, our method also achieves a higher rate. Therefore, our method achieves a better tradeoff between diversity (i.e., SINR in Fig. 2) and multiplexing (i.e., achievable rate in Fig. 3). Moreover, in both Figs 2 and 3, it can be seen that using a higher number of streams results in a lower SINR but higher achievable rate, and vice versa, signifying that a lower number of streams exploits the diversity of multiple antennas.

Fig. 5 shows the SINR of our method in an overloaded scenario. It can be seen that our method performs better than AS schemes because in AS, by switching off $\lceil \frac{M}{Q} \rceil$ antennas, the combining capabilities of Bobs decreases, but our method does not require to turn off RF chains at Bobs. However, this achieved SINR does not result in a better BER as seen in Fig. 2. Similar results can be established for the achievable rate of our method and AS in overloaded networks.

Fig. 6 shows the SER of Eve in an overloaded scenario when $L = 6$. Both $(7, 4, 3)$ and $(7, 4, 2)$ settings represent overloaded scenarios. In both settings, we set $\tau = 4$. Clearly, no FJ can be created in these settings using the ZF method. It can be seen that our method performs significantly better than the ZF scheme in both overloaded settings because our method forces Eve to have at least $\Psi' = \max\{\tau Q, M\}$ antennas to decode all messages. However, the ZF method only imposes $\Psi = KQ$ antennas in overloaded scenarios. In both of these settings, $L = 6$ antennas would be enough to decode all messages in the ZF design. It can be seen that the setting $(7, 4, 3)$ experiences more SER because more data streams are used per user, which decrease the diversity gain.

VI. Conclusions

In this paper, we proposed a novel precoding scheme that not only manages the interference in MU-MIMO networks better than the zero-forcing method, but also enables the nodes to operate in overloaded settings. Compared to the ZF method, our scheme is able to impose more stringent conditions on Eve's number of antennas in overloaded scenarios. Our method also did not require the hardware modifications that some other methods, such as antenna selection schemes, demand in overloaded networks. Analysis of this scheme in massive MIMO networks, or with limited feedback from downlink users, or with in-band full-duplex capability in nodes are the subject of future research.

References

[1] P. A. Regalia, A. Khisti, Y. Liang, and S. Tomasin, "Secure communications via physical-layer and information-theoretic techniques," *Proc. IEEE*, vol. 103, no. 10, pp. 1698–1701, Oct. 2015.

[2] O. O. Koyluoglu, H. E. Gamal, L. Lai, and H. V. Poor, "On the secure degrees of freedom in the K-user Gaussian interference channel," in *Proc. IEEE ISIT Conf.*, Jul. 2008, pp. 384–388.

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

[4] N. Anand, S.-J. Lee, and E. W. Knightly, "STROBE: Actively securing wireless communications using zero-forcing beamforming," in *Proc. IEEE INFOCOM Conf.*, Mar. 2012, pp. 720–728.

[5] A. Mukherjee and A. L. Swindlehurst, "Utility of beamforming strategies for secrecy in multiuser MIMO wiretap channels," in *Proc. Allerton Conf. Commun., Control, Computing*, Sep. 2009, pp. 1134–1141.

[6] A. Mukherjee, S. Fakoorian, J. Huang, and A. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surv. Tut.*, vol. 16, no. 3, pp. 1550–1573, Feb. 2014.

[7] A. D. Dabbagh and D. J. Love, "Precoding for multiple antenna gaussian broadcast channels with successive zero-forcing," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3837–3850, Jul. 2007.

[8] H. Sung, S. R. Lee, and I. Lee, "Generalized channel inversion methods for multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3489–3499, Nov. 2009.

[9] E. Ekrem and S. Ulukus, "Secure broadcasting using multiple antennas," *J. Commun. Networks*, vol. 12, no. 5, pp. 411–432, Oct. 2010.

[10] G. Geraci, R. Couillet, J. Yuan, M. Debbah, and I. B. Collings, "Large system analysis of linear precoding in MISO broadcast channels with confidential messages," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1660–1671, Sep. 2013.

[11] N. Yang, G. Geraci, J. Yuan, and R. Malaney, "Confidential broadcasting via linear precoding in non-homogeneous MIMO multiuser networks," *IEEE Trans. Commun.*, vol. 62, no. 7, pp. 2515–2530, Jul. 2014.

[12] N. Li, X. Tao, H. Wu, J. Xu, and Q. Cui, "Large-system analysis of artificial-noise-assisted communication in the multiuser downlink: Ergodic secrecy sum rate and optimal power allocation," *IEEE Trans. on Veh. Technol.*, vol. 65, no. 9, pp. 7036–7050, Sep. 2016.

[13] T. M. Duman and A. Ghrayeb, *Coding for MIMO Communication Systems*. New York, NY, USA: John Wiley and Sons, Ltd, 2007.

[14] E. Bjornson, M. Kountouris, M. Bengtsson, and B. Ottersten, "Receive combining vs. multi-stream multiplexing in downlink systems with multi-antenna users," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3431–3446, Jul. 2013.

[15] H. A. A. Saleh, A. F. Molisch, T. Zemen, S. D. Blostein, and N. B. Mehta, "Receive antenna selection for time-varying channels using discrete prolate spheroidal sequences," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2616–2627, Jul. 2012.

[16] Y. Gao, H. Vinck, and T. Kaiser, "Massive MIMO antenna selection: Switching architectures, capacity bounds, and optimal antenna selection algorithms," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1346–1360, Mar. 2018.

[17] A. F. Molisch, M. Z. Win, Y.-S. Choi, and J. H. Winters, "Capacity of MIMO systems with antenna selection," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1759–1772, Jul. 2005.

[18] S. Yu, L.-C. Tranchevent, B. De Moor, and Y. Moreau, *Rayleigh Quotient-Type Problems in Machine Learning*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 27–37.

[19] P. Patcharamaneepakorn, S. Armour, and A. Doufexi, "On the equivalence between SLNR and MMSE precoding schemes with single-antenna receivers," *IEEE Commun. Lett.*, vol. 16, pp. 1034–1037, Jul. 2012.

[20] X.-D. Zhang, *Matrix Analysis and Applications*. Cambridge, UK: Cambridge University Press, 2017.

[21] B. Akgun, O. O. Koyluoglu, and M. Krunz, "Exploiting full-duplex receivers for achieving secret communications in multiuser MISO networks," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 956–968, Feb. 2017.