Directional Modulation via Symbol-Level Precoding: A Way to Enhance Security

Ashkan Kalantari *Member, IEEE*, Mojtaba Soltanalian *Member, IEEE*, Sina Maleki *Member, IEEE*, Symeon Chatzinotas *Senior Member, IEEE*, and Björn Ottersten, *Fellow, IEEE*

***Abstract*—Wireless communication provides a wide coverage at the cost of exposing information to unintended users. As an information-theoretic paradigm, secrecy rate derives bounds for secure transmission when the channel to the eavesdropper is known. However, such bounds are shown to be restrictive in practice and may require exploitation of specialized coding schemes. In this paper, we employ the concept of directional modulation and follow a signal processing approach to enhance the security of multi-user MIMO communication systems when a multi-antenna eavesdropper is present. Enhancing the security is accomplished by increasing the symbol error rate at the eavesdropper. Unlike the information-theoretic secrecy rate paradigm, we assume that the legitimate transmitter is not aware of its channel to the eavesdropper, which is a more realistic assumption. We examine the applicability of MIMO receiving algorithms at the eavesdropper. Using the channel knowledge and the intended symbols for the users, we design security enhancing symbol-level precoders for different transmitter and eavesdropper antenna configurations. We transform each design problem to a linearly constrained quadratic program and propose two solutions, namely the iterative algorithm and one based on non-negative least squares, at each scenario for a computationally-efficient modulation. Simulation results verify the analysis and show that the designed precoders outperform the benchmark scheme in terms of both power efficiency and security enhancement.**

***Keywords*—Array processing, directional modulation, M-PSK modulation, physical layer security, symbol-level precoding.**

I. Introduction

A. Motivation

Wireless communications allows information flow through broadcasting; however, unintended receivers may also receive these information, with eavesdroppers amongst them. To derive a bound for secure transmission, Wyner proposed the secrecy rate concept in his seminal paper [2] for discrete memoryless channels. The secrecy rate defines the bound for secure transmission and proper coding is being developed to achieve this bound [3]. However, the secrecy rate can restrict the communication system in some aspects. Primarily, the secrecy rate requires perfect or statistical knowledge of the eavesdropper's channel state information (CSI) [2], [4]–[6], however, it may not be possible to acquire the perfect or statistical CSI of a passive eavesdropper in practice. In addition, in the secrecy rate approach, the transmission rate has to be lower than the achievable rate, which may conflict with the increasing rate demands in wireless communications. Furthermore, the transmit signal usually is required to follow a Gaussian distribution which is not the case in current digital communication systems.

Recently, there has been a growing research interest on directional modulation technology and its security enhancing ability. As a pioneer, [7] implements a directional modulation transmitter using parasitic antenna. This system creates the desired amplitude and phase in a specific direction by varying the length of the reflector antennas for each symbol while scrambling the symbols in other directions. The authors of [8] suggest using a phased array at the transmitter and employ the genetic algorithm to derive the phase values of a phased array in order to create symbols in a specific direction. The directional modulation concept is later extended to directionally modulating symbols to more than one destination. In [9], the singular value decomposition (SVD) is used to directionally modulate symbols in a two user system. The authors of [10] derive the array weights to create two orthogonal far field patterns to directionally modulate two symbols to two different locations and [11] uses least-norm to derive the array weights and directionally modulate symbols towards multiple destinations in a multi-user multi-input multi-output (MIMO) system. The authors in [1] design the array weights of a directional modulation transmitter in a MIMO system to minimize the power consumption while keeping the signal-to-noise ratio (SNR) of each received signal above a specific level. The directional modulation literature focuses on practical implementation and the security enhancing characteristics of this technology. On top of the works in the directional modulation literature where antennas excitation weights change on a symbol basis, the symbol-level precoding to create constructive interference between the transmitted symbols has been developed in [12]–[16] by focusing on the digital processing of the signal before being fed to the antenna array. The main

This work was supported by the National Research Fund (FNR) of Luxembourg under AFR grant for the project "Physical Layer Security in Satellite Communications (ref. 5798109)", SeMIGod, and SATSENT. Ashkan Kalantari, Sina Maleki, Symeon Chatzinotas, and Björn Ottersten are with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), The University of Luxembourg, 4 rue Alphonse Weicker, L-2721 Luxembourg-Kirchberg, Luxembourg, (E-mails: {ashkan.kalantari,sina.maleki,symeon.chatzinotas,bjorn.ottersten}@uni.lu).
M. Soltanalian is with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL 60607, E-mail: (msol@uic.edu).
A part of this work was presented at the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2016 [1].

difference between the directional modulation and the digital symbol-level precoding for constructive interference is that the former focuses on applying array weights in the analog domain such that the received signals on the receiving antennas have the desired amplitude and phase, whereas the latter uses symbol-level precoding for digital signal design at the transmitter to create constructive interference at the receiver. Furthermore, directional modulation was originally motivated by physical layer security, whereas symbol-level precoding by energy efficiency.

B. Contributions

In this paper, we study and design the optimal precoder for a directional modulation transmitter in order to enhance the security in a quasi-static fading MIMO channel where a multi-antenna eavesdropper is present. Here, enhancing the security means increasing the symbol error rate (SER) at the eavesdropper. In directional modulation, users' MIMO channel and symbols meant for the users are used to design the precoder. The precoder is designed to induce the symbols on the receiver antennas rather than generating the symbols at the transmitter and sending them, which is the case in the conventional transmit precoding [17], [18]. In other words, in the directional modulation, the modulation happens in the radio frequency (RF) level while the arrays' emitted signals pass through the wireless channel. This way, we simultaneously communicate multiple interference-free symbols to multiple users. Also, the precoder is designed such that the receiver antennas can directly recover the symbols without CSI knowledge and equalization. Therefore, assuming the eavesdropper has a different channel compared to the users, it receives scrambled symbols. In fact, the channels between the transmitter and users act as secret keys [19] in the directional modulation. Furthermore, since the precoder depends on the symbols, the eavesdropper cannot calculate it. In contrast to the information theoretic secrecy rate paradigm, the directional modulation enhances the security by considering more practical assumptions. Particularly, directional modulation does not require the eavesdropper's CSI to enhance the security; in addition, it does not reduce the transmission rate and signals are allowed to follow a non-Gaussian distribution. In light of the above, our contributions in this paper can be summarized as follows:

1) We design the optimal symbol-level precoder for a security enhancing directional modulation transmitter in a MIMO fading channel to communicate with arbitrary number of users through symbol streams. In addition, we derive the necessary condition for the existence of the precoder, which is novel compared to the digital symbol-level precoding works in [12]–[16]. The directional modulation literature mostly includes LoS analysis with one or limited number of users, and multi-user works do not design the optimal precoder to communicate symbols with arbitrary multi-antenna users from a power efficiency point of view.
2) We analyze the applicability of various MIMO receiving algorithms at the eavesdropper. Since the imposed SER on the eavesdropper depends on the difference between the number of transmitter and the eavesdropper antennas, we consider the cases when the eavesdropper has less or more antennas than the transmitter and design a specific precoder for each case. We minimize the transmission power for the former case and maximize the SER at the eavesdropper for the latter case to prevent or suppress successful decoding at the eavesdropper. This is done while keeping the SNR of users' received signals above a predefined threshold and thus the users' rate demands are satisfied. The analysis of different MIMO receiving algorithms at the eavesdropper and designing a precoder to maximize the SER at the eavesdropper are absent in the available directional modulation literature and digital symbol-level precoding works [12]–[16].
3) We show that the SER imposed on the eavesdropper in the conventional precoding depends on the difference between the number of antennas of the eavesdropper and the receiver. In our design, the SER imposed on the eavesdropper depends on the difference between the number of eavesdropper and transmitter antennas since the precoder depends on both the channels and symbols. The transmitter, e.g., a base station, probably has more antennas than the receiver, hence, it is more likely to preserve the security in directional modulation, especially in a massive MIMO system.
4) We simplify the power and SNR minimization precoder design problems into a linearly-constrained quadratic programming problem. For faster design, we introduce new auxiliary variable to transform the constraint into equality and propose two different ways to solve the design problems. In the first way, we use the penalty method to get an unconstrained problem and solve it by proposing an iterative algorithm. Also, we prove that the algorithm converges to the optimal point. In the second one, we use the constraint to get a non-negative least squares design problem. For the latter, there are already fast techniques to solve the problem.

C. Additional Related Works to Directional Modulation

Array switching at the symbol rate is used in [20], [21] to induce the desired symbols. In connection with [7], [22] studies the far field area coverage of a parasitic antenna and shows that it is a convex region. The technique of [8] is implemented in [23] using a four element microstrip patch array where symbols are directionally modulated for Q-PSK modulation. The authors of [24] propose an iterative nonlinear optimization approach to design the array weights which minimizes the distance between the desired and the directly modulated symbols in a specific direction. The Fourier transform is used in [25], [26] to create the optimal constellation pattern for Q-PSK directional modulation. In [9], [27]–[29] directional modulation is employed along with noise injection. The authors of [27], [28] utilize an orthogonal vector approach to derive the array weights in order to directly modulate the data and inject the artificial noise in the direction of the eavesdropper. The work

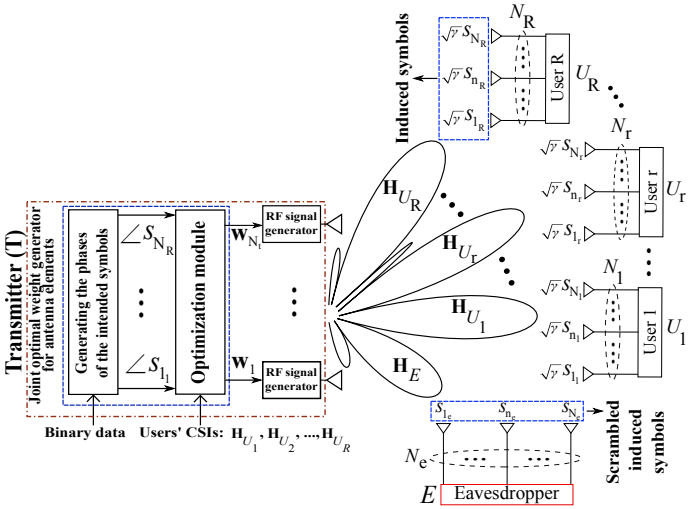

Fig. 1. Generic architecture of a directional modulation transmitter, including the optimal security enhancing antenna weight generator using the proposed algorithms.

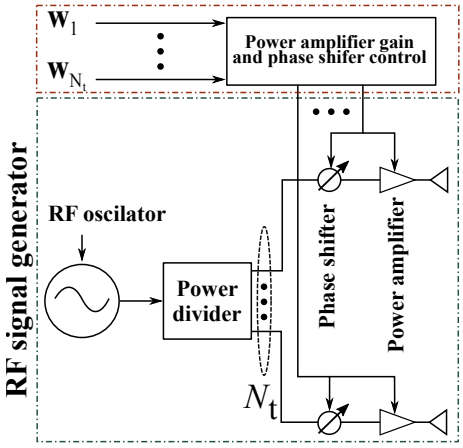

Fig. 2. RF signal generation using actively driven elements, including power amplifiers and phase shifters.

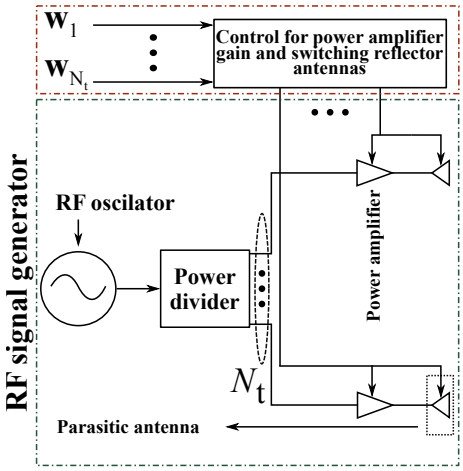

Fig. 3. RF signal generation using power amplifiers and parasitic antennas.

of [27] is extended to retroactive arrays[1] in [29] for a multi-path environment. An algorithm including exhaustive search is used in [30] to adjust two-bit phase shifters for directionally modulating information.

D. Organization

The remainder of the paper is organized as follows. In Section II, transmitter architectures, network configuration, and the signal model are introduced. The security of the directional modulation is studied in Section III. In Section IV, the optimal precoders for the directional modulation are designed and the benchmark scheme is mentioned. The complexity of our scheme and the benchmark method are studied in Section V. In Section VI, we present the simulation results. Finally, the conclusions are drawn in Section VII.

Notation: Upper-case and lower-case bold-faced letters are used to denote matrices and column vectors, respectively. The superscripts $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^\dagger$ represent transpose, conjugate, Hermitian, and Moore-Penrose pseudo inverse operators, respectively. $\mathbf{I}_{N\times N}$ denotes an N by N identity matrix, $\mathrm{diag}(\mathbf{a})$ denotes a diagonal matrix where the elements of the vector $\mathbf{a}$ are its diagonal entries, $\mathbf{a}\circ\mathbf{b}$ is the element-wise Hadamard product, $\mathbf{a}_+$ denotes a vector where negative elements of the vector $\mathbf{a}$ are replaced by zero, $\mathbf{0}$ is the all zero vector, $\|\cdot\|$ is the Frobenius norm, and $|\cdot|$ represents the absolute value of a scalar. $\mathrm{Re}\,(\cdot)$, $\mathrm{Im}\,(\cdot)$, and $\arg(\cdot)$ represent the real valued part, imaginary valued part, and angle of a complex number, respectively.

II. Signal and System Model

We consider a communication network with a multi-antenna transmitter denoted by T, R multi-antenna users denoted by U_r for $r = 1, ..., R$ where the r-th user has N_r antennas, and a multi-antenna eavesdropper[2] denoted by E with N_e antennas, as shown in Fig. 1. In addition, all the communication channels are considered to be quasi-static block fading. Two possible architectures for the RF signal generator block of Fig. 1 are presented in Figures. 2 and 3. In Fig. 2, power amplifiers and phase shifters are used in each RF chain to adjust the gain and the phase of the transmitted signal from each antenna. In Fig. 3, we adapt the technique of [7] to adjust the phase using parasitic antennas in each RF chain. A parasitic antenna is comprised of a dipole antenna and multiple reflector antennas. Near field interactions between the dipole and reflector antennas creates the desired amplitude and phase in the far filed, which can be adjusted by switching the proper MOSFETs. When using parasitic antennas, the channel from each parasitic antenna to the far field needs to be LoS, and we need to acquire the CSI of the fading channel from the far field of each parasitic antenna

[1] A retroactive antenna can retransmit a reference signal back along the path which it was incident despite the presence of spatial and/or temporal variations in the propagation path.

[2] The same system model and solution holds for multiple colluding single-antenna eavesdroppers.

to the receiving antennas. For simplicity, we only consider the amplitude and phase of the received signals and drop $e^{j2\pi ft}$, which is the carrier frequency part.

After applying the optimal coefficients to array elements, the received signals by U_r and E are

$$\mathbf{y}_{U_r} = \mathbf{H}_{U_r}\mathbf{w} + \mathbf{n}_{U_r},\ r = 1, ..., R \tag{1}$$

$$\mathbf{y}_E = \mathbf{H}_E\mathbf{w} + \mathbf{n}_E, \tag{2}$$

where the signal $\mathbf{y}_{U_r}$ is an $N_r \times 1$ vector denoting the received signals by U_r, $\mathbf{y}_E$ is an $N_e \times 1$ vector denoting the received signals by E, $\mathbf{H}_{U_r} = [\mathbf{h}_{1_r}, ..., \mathbf{h}_{n_r}, ..., \mathbf{h}_{N_r}]^T$ is an $N_r \times N_t$ matrix denoting the channel from T to U_r, $\mathbf{h}_{n_r}$ is an $N_t \times 1$ vector containing the channel coefficients from the transmitter antennas to the n-th antenna of the r-th user, the channel for all users is an $N_U \times N_t$ matrix defined as $\mathbf{H}_U = [\mathbf{H}_{U_1}, ..., \mathbf{H}_{U_r}, ..., \mathbf{H}_{U_R}]^T$, $\mathbf{H}_E$ is an $N_e \times N_t$ matrix denoting the channel from T to E, and $\mathbf{w}$ denotes the transmit precoding vector. In directional modulation, the elements of $\mathbf{H}_{U_r}\mathbf{w} = \left[\sqrt{\gamma}s_{1_r}, ..., \sqrt{\gamma}s_{n_r}, ..., \sqrt{\gamma}s_{N_r}\right]^T$ are the induced M-PSK symbols on the antennas of the r-th user, s_{n_r} is the induced M-PSK symbol on the n-th antenna of the r-th user with instantaneous unit energy, i.e., $|s_{n_r}|^2 = 1$, γ is the SNR of the induced symbol, and M is the M-PSK modulation order. To detect the received symbols, U_r can apply conventional detectors on each antenna. The random variables $\mathbf{n}_{U_r}$ and $\mathbf{n}_E$ denote the additive white Gaussian noise at U_r and E, respectively. The Gaussian random variables $\mathbf{n}_{U_r}$ and $\mathbf{n}_E$ are independent and identically distributed (i.i.d.) with $\mathbf{n}_{U_r} \sim \mathcal{CN}(\mathbf{0}, \sigma^2_{n_{U_r}}\mathbf{I}_{N_r \times N_r})$, and $\mathbf{n}_E \sim \mathcal{CN}(\mathbf{0}, \sigma^2_{n_E}\mathbf{I}_{N_e \times N_e})$, respectively, where $\mathcal{CN}$ denotes a complex and circularly symmetric random variable.

Throughout the paper, we assume that T knows only $\mathbf{H}_U$ while E knows both $\mathbf{H}_U$ and $\mathbf{H}_E$. In the following, we analyze the conditions under which we can enhance the system security.

III. Security analysis of directional modulation

In this section, we discuss different MIMO receiving algorithms and investigate whether E can use them to estimate the received signals by the users or not. We assume that E's channel is independent from those of the users, and to consider the worst case, we assume that $\mathbf{H}_E$ is full rank. Hence, the element numbers of $\mathbf{H}_E\mathbf{w}$, i.e., received signals on E's antennas, are different from those of $\mathbf{H}_{U_r}\mathbf{w}$, i.e., received signals on receiver antennas, for $r = 1, ..., R$. Since $\mathbf{w}$ depends on the symbols, E cannot directly calculate it. In the following, we analyze the capability of E in using MIMO receiving algorithms to estimate $\mathbf{w}$.

A. Zero-Forcing Estimation

As an approach to estimate $\mathbf{w}$, E can remove $\mathbf{H}_E$ through zero-forcing (ZF) estimation, and then multiply the estimated $\mathbf{w}$ by $\mathbf{H}_U$ to estimate the symbols. For $N_e < N_t$, E cannot estimate $\mathbf{H}_U\mathbf{w}$ since $\mathbf{H}_E^\dagger\mathbf{H}_E \neq \mathbf{I}$. However, when $N_e \geq N_t$, E can estimate $\mathbf{w}$ as follows

$$\widehat{\mathbf{w}} = \mathbf{G}_1\mathbf{y}_E = \mathbf{w} + \mathbf{G}_1\mathbf{n}_E, \tag{3}$$

where

$$\mathbf{G}_1 = \left(\mathbf{H}_E^H\mathbf{H}_E\right)^{-1}\mathbf{H}_E^H, \tag{4}$$

and $\widehat{\mathbf{w}}$ is the estimated $\mathbf{w}$ at E. Next, E can multiply $\widehat{\mathbf{w}}$ by $\mathbf{H}_U$ to estimate the signals at receiver antennas, $\mathbf{H}_U\widehat{\mathbf{w}}$, as

$$\mathbf{H}_U\widehat{\mathbf{w}} = \mathbf{H}_U\mathbf{w} + \mathbf{H}_U\left(\mathbf{H}_E^H\mathbf{H}_E\right)^{-1}\mathbf{H}_E^H\mathbf{n}_E. \tag{5}$$

Through (3) to (5), E virtually puts itself in the location of the users to estimate the received signal by them. The eavesdropper is capable of doing this since we assume that it knows the users' channels, $\mathbf{H}_U$. This way, E gets access to the secret key, which allows for observing the signals from users' point of view; however, the required process increases the noise at E.

B. Minimum Mean-Square Error Estimation

To avoid enhanced noise, E can estimate $\mathbf{w}$ via the minimum mean-square error (MMSE) technique. The estimated symbols at E through MMSE can be written as [31]

$$\widehat{\mathbf{w}} = \mathbf{G}_2\mathbf{y}_E, \tag{6}$$

$$\mathbf{H}_U\widehat{\mathbf{w}} = \mathbf{H}_U\mathbf{G}_2\mathbf{H}_E\mathbf{w} + \mathbf{H}_U\mathbf{G}_2\mathbf{n}_E, \tag{7}$$

with

$$\mathbf{G}_2 = \left(\mathbf{H}_E^H\mathbf{C}_\mathbf{w}^{-1}\mathbf{H}_E + \mathbf{C}_{\mathbf{N}_E}^{-1}\right)^{-1}\mathbf{H}_E^H\mathbf{C}_\mathbf{w}^{-1}, \tag{8}$$

where $\mathbf{C}_\mathbf{W}$ is the covariance matrix of the precoding vector, $\mathbf{w}$, and $\mathbf{C}_{\mathbf{N}_E}$ is the covariance matrix of the eavesdropper noise, $\mathbf{n}_E$. As we see in (8), the MMSE estimation of $\mathbf{w}$ at the eavesdropper requires the knowledge of $\mathbf{C}_\mathbf{W}$. As an approach to derive $\mathbf{C}_\mathbf{W}$, the eavesdropper can design $\mathbf{w}$ for different random sequences of $\mathbf{s}$ and channel realizations to derive multiple instantaneous covariance matrices as $(\mathbf{w} - \overline{\mathbf{w}})(\mathbf{w} - \overline{\mathbf{w}})^H$, where $\overline{\mathbf{w}}$ is the average of $\mathbf{w}$. Then, E can average over these instantaneous covariance matrices to calculate $\mathbf{C}_\mathbf{W}$. The eavesdropper can apply the MMSE estimation approach as long as the matrix $\mathbf{H}_E^H\mathbf{C}_\mathbf{w}^{-1}\mathbf{H}_E + \mathbf{C}_{\mathbf{N}_E}^{-1}$ is non-singular.

C. Successive Interference Cancellation and Sphere Decoding

The observed signal by the eavesdropper in a conventional MIMO system is

$$\mathbf{y}_E = \mathbf{H}_E\mathbf{W}\mathbf{s} + \mathbf{n}_\mathbf{E}, \tag{9}$$

where the precoding vector $\mathbf{W}$ depends only on the channel. The eavesdropper needs to estimate the symbol vector, $\mathbf{s}$, in (9) where its elements are drawn from a finite-alphabet set. When the successive interference cancellation (SIC) receiver is applied to a conventional MIMO receiver, each element of $\mathbf{s}$ is detected and reduced from the aggregated signal. This is possible since $\mathbf{s}$ is drawn from a finite-alphabet set [32]. However, in our case, the eavesdropper needs to estimate the precoding vector $\mathbf{w}$ whose elements take continuous values.

Hence, the successive interference cancellation techniques, e.g., ZF-SIC and MMSE-SIC, cannot be applied at the eavesdropper. Furthermore, the similar argument can be followed for the sphere decoding technique [33], which is based on creating a sphere around the received symbol and finding the closet member of the finite-alphabet set to it.

Note that E needs to estimate $\mathbf{w}$ whether it wants to estimate the symbols of a specific user or all the users.

We will see in Section VI that as the difference between N_t ad N_e goes higher, the imposed SER at E for both ZF and MMSE estimators increases.

Remark 1: Using a large-scale array transmitter, it is more probable to have a higher difference between N_t and N_e. Hence, the directional modulation technique seems to be a good candidate to enhance the security when the transmitter is equipped with a large-scale array. ■

D. Brute-force and maximum likelihood Approach

Apart from the previous estimation approaches, the eavesdropper can follow the brute-force approach and consider all the possible symbol combinations. For a specific modulation order and total number of users' antenna, the symbol vector, $\mathbf{s}$, has M^{N_U} different possibilities. This means that the eavesdropper needs to solve the design problems (13), (35), or (39) M^{N_U} times to make a look up table. Furthermore, note that the eavesdropper needs to recalculate the entire look up table if any element in $\mathbf{H}_U$ or $\mathbf{H}_E$ changes. Depending on the coherence time of the channel, this increases the computational complexity at the eavesdropper. If we assume the ideal case without noise, the eavesdropper needs to search in its look up table for $\mathbf{y}_E$ to find the corresponding vector $\mathbf{w}$.

Nevertheless, we have noise in practice. This requires E to compare its received signal with all the computed M^{N_U} possible cases of $\mathbf{y}_E$ to find the corresponding precoding vector $\mathbf{w}$. As we see, the possibilities increase exponentially with M and N_U. If we show the calculated possible cases of $\mathbf{w}$ as the set $w = \{\mathbf{w}_1, ..., \mathbf{w}_{M^{N_U}}\}$ where the cardinality of w is M^{N_U}, the eavesdropper can follow the maximum likelihood approach to find $\mathbf{w}$ as

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}_i \in w} \ \|\mathbf{y}_E - \mathbf{H}_E \mathbf{w}_i\|_2 \,, \tag{10}$$

where $\hat{\mathbf{w}}$ is the brute-force solution. The complexity of calculating the norm of the difference of two vectors with the length N_e is

$$\begin{aligned} c_{norm} &= 2\mathbf{N}_e O\left(n\right) + \mathbf{N}_e \left(2O\left(n^{1.465}\right) + O\left(n\right)\right) + \mathbf{N}_e O\left(n\right) \\ &= 4\mathbf{N}_e O(n) + 2\mathbf{N}_e O\left(n^{1.465}\right). \end{aligned} \tag{11}$$

Considering that the eavesdropper needs to try all the elements of the set w, the total complexity of the brute-force approach is given by $c_{brute-froce} = M^{N_U}(c_{norm} + c_{design})$, where c_{design} is the complexity of solving (25), (37), or (47), which is quantitatively mentioned in (53) and (54). The brute-force complexity increases exponentially both in modulation order and total number of receiving antennas. To further understand the amount of computational complexity of the brute-force method, we compare it with the advanced encryption security (AES) method in the following example. For $M = 32$ and $N_U = 52$, the computational complexity of the brute-force method is $2^{260}(c_{norm} + c_{design})$. The complexity of the improved biclique attack to break the largest key of the AES, which has 256 bit size, is $2^{254.27}$ [34], which is significantly lower than the complexity of the brute-force method at the eavesdropper for the mentioned example. Computation time of the brute-force method with respect to system dimension is presented in Section VI.

According to this section, we see that the optimal strategy at E is the brute-force and maximum likelihood approaches. However, we see that this comes with an extremely large computational cost.

Remark 2: Assuming that the legitimate channel is reciprocal, the users can transmit pilots to T so it can estimate $\mathbf{H}_U$. This way, we avoid the additional downlink channel estimation and the users do not have to send feedback bits to T, hence, E cannot estimate $\mathbf{H}_U$. Assuming that E knows the channel from T to itself, i.e., $\mathbf{H}_E$, it can estimate $\mathbf{w}$ as in (3) or (6), but it cannot perform (5) or (7) to estimate the received signals on the receiver antennas. ■

In the next section, optimal symbol-level precoders for the directional modulation are designed to enhance the security.

IV. Optimal Precoder Design for Directional Modulation

In this section, we define the underlaying problems to design the security enhancing symbol-level precoder for the directional modulation. Since the SER at E depends on the difference between N_t and N_e, we consider the cases $N_e < N_t$ and $N_e \geq N_t$ and design a specific precoder for each of them. The case $N_e < N_t$ focuses on energy efficiency, hence, we also perform relaxed phase analysis for this case.

A. The Case of Strong Transmitter ($N_e < N_t$, Fixed Phase)

In wireless transmission, adaptive coding and modulation (ACM) is used to enhance the link performance and the channel capacity. In ACM, the transmission power, coding rate, and the modulation order is set according to the channel signal to noise ratio (SNR) [35]. Based on this, we preserve the SNR of the induced symbol on the receiver antenna above or equal to a specific level to successfully decode it. Here, we only focus on the SNR of an uncoded signal since considering SNR of a coded transmission based on ACM is beyond the scope of this paper.

To avoid a non-convex design problem, we use the required signal properties at the receiver to formulate a convex design problem. In our design, a specific fixed phase is required for the received signal at each receiver antenna. Since the phase of the received signal at each receiving antenna, $\mathbf{h}_{n_r}^T \mathbf{w}$, is the same as the phase of the intended symbol, s_{n_r}, if the required SNR, γ, of the received signal increases, the in-phase, $\mathrm{Re}\left(\mathbf{h}_{n_r}^T \mathbf{w}\right)$, and quadrature-phase, $\mathrm{Im}\left(\mathbf{h}_{n_r}^T \mathbf{w}\right)$, parts will increase in the same proportion to satisfy the required SNR. Since the received signal by each antenna is complex valued, we separately consider amplitudes of the in-phase and

quadrature-phase parts of the received signal on the receiver antenna instead of its power. If we show the real and imaginary valued parts of s_{n_r} as $\mathrm{Re}\left(s_{n_r}\right)$ and $\mathrm{Im}\left(s_{n_r}\right)$, the required in-phase and quadrature-phase thresholds of the received signal are defined as

$$\sqrt{\gamma}\mathrm{Re}\left(s_{n_r}\right),\ \sqrt{\gamma}\mathrm{Im}\left(s_{n_r}\right). \tag{12}$$

Since $\left|s_{n_r}\right|^2 = 1$, we can see that $\gamma = \gamma\mathrm{Re}^2\left(s_{n_r}\right) + \gamma\mathrm{Im}^2\left(s_{n_r}\right)$, which satisfies the SNR constraint.

We design the directional modulation precoder to minimize the total transmit power such that 1) the signals received by the n-th antenna of the r-th user result in a phase equal to that of s_{n_r}, and 2) the signals received by the n-th antenna of the r-th user create in-phase and quadrature-phase signal levels satisfying the thresholds defined in (12). Accordingly, the precoder design problem is defined as

$$\begin{aligned}\min_{\mathbf{w}}\ & \left\|\mathbf{w}\right\|^2 \\ \text{s.t.}\ & \arg\left(\mathbf{h}_{n_r}^T\mathbf{w}\right) = \arg\left(s_{n_r}\right), && \text{(13a)}\\ & \mathrm{Re}\left(\mathbf{h}_{n_r}^T\mathbf{w}\right) \geq \sqrt{\gamma}\mathrm{Re}\left(s_{n_r}\right), && \text{(13b)}\end{aligned}$$

for $r = 1,...,R$ and $n = 1,...,N$. Since the phase of the induced symbol is fixed, we just need to put the signal level constraint over the real or imaginary part of the received signal on each receiving antenna. Hence, we have included the constraint over the value of the real part in (13b). Generally, some constraints of (13) are satisfied with inequality and the rest are satisfied with equality [36]. This depends on the difference between N_t and N_U. We will also show this through simulations in Section VI. In the case that each user is associate with a precoder, i.e., the transmitter designs $\mathbf{w}_1$,..., $\mathbf{w}_K$ for K users, the constraints are satisfied with equality at the optimal point [37]. If both sides of (13b) are negative, the signal level constraints may not be satisfied. Since (13a) holds at the optimal point, $\mathrm{Re}\left(\mathbf{h}_{n_r}^T\mathbf{w}\right)$ has the same sign as $\mathrm{Re}\left(s_{n_r}\right)$ at the optimal point. Therefore, we can multiply both sides of (13b) by $\mathrm{Re}(s_{n_r})$ to get

$$\begin{aligned}\min_{\mathbf{w}}\ & \left\|\mathbf{w}\right\|^2 \\ \text{s.t.}\ & \arg\left(\mathbf{h}_{n_r}^T\mathbf{w}\right) = \arg\left(s_{n_r}\right), && \text{(14a)}\\ & \mathrm{Re}\left(s_{n_r}\right)\mathrm{Re}\left(\mathbf{h}_{n_r}^T\mathbf{w}\right) \geq \sqrt{\gamma}\mathrm{Re}^2\left(s_{n_r}\right). && \text{(14b)}\end{aligned}$$

To simplify (14), we can rewrite the phase constraint in (14a) as

$$\mathrm{Re}\left(\mathbf{h}_{n_r}^T\mathbf{w}\right)\alpha_{n_r} - \mathrm{Im}\left(\mathbf{h}_{n_r}^T\mathbf{w}\right) = 0,\quad \forall n,\ \forall r, \tag{15}$$

where $\alpha_{n_r} = \tan\left(s_{n_r}\right)$. Since $\tan(\cdot)$ repeats after a π radian period[3], symbols with different phases can have the same $\tan$ value, e.g., $\tan\left(\frac{\pi}{4}\right) = \tan\left(\frac{3\pi}{4}\right)$. Therefore, replacing (14a) with (15) creates ambiguity. To avoid this, we can add the constraint

$$\mathrm{Re}\left(s_{n_r}\right)\mathrm{Re}\left(\mathbf{h}_{n_r}^T\mathbf{w}\right) \geq 0, \tag{16}$$

[3]If the phase of the M-PSK constellation falls on the points where $\tan$ function is undefined, e.g., $\frac{\pi}{2}$, we can add phase offset to the modulation.

to the design problem (14) to avoid ambiguity. Interestingly, constraint (16) is already present in (14b). Note that (15) and (16) together are equivalent to (13a), so the required conditions to go from (13) to (14) still hold. Putting together the constraints (15) and (14b) for all the users, (14) is written into the following compact form

$$\begin{aligned}\min_{\mathbf{w}}\ & \left\|\mathbf{w}\right\|^2 \\ \text{s.t.}\ & \mathbf{A}\mathrm{Re}\left(\mathbf{H}_U\mathbf{w}\right) - \mathrm{Im}\left(\mathbf{H}_U\mathbf{w}\right) = \mathbf{0}, && \text{(17a)}\\ & \mathrm{Re}\left(\mathbf{S}\right)\mathrm{Re}\left(\mathbf{H}_U\mathbf{w}\right) \geq \sqrt{\gamma}\,\mathbf{s}_r, && \text{(17b)}\end{aligned}$$

where $\mathbf{S} = \mathrm{diag}\left(\mathbf{s}\right)$, $\mathbf{s}$ is an $N_U \times 1$ vector containing all the intended M-PSK symbols for the users with $N_U = \sum_{r=1}^{R} N_r$, $\mathbf{s}_r = \mathrm{Re}\left(\mathbf{s}\right) \circ \mathrm{Re}\left(\mathbf{s}\right)$, $\mathbf{A} = \mathrm{diag}\left(\boldsymbol{\alpha}\right)$, $\boldsymbol{\alpha} = \left[\alpha_{1_1},...,\alpha_{n_r},...,\alpha_{N_R}\right]^T$.

To remove the real and imaginary valued parts from (17), we can use $\mathbf{H}_U = \mathrm{Re}\left(\mathbf{H}_U\right) + i\mathrm{Im}\left(\mathbf{H}_U\right)$ and $\mathbf{w} = \mathrm{Re}\left(\mathbf{w}\right) + i\mathrm{Im}\left(\mathbf{w}\right)$ presentations to separate the real and imaginary valued components of $\mathbf{H}_U\mathbf{w}$ as

$$\begin{aligned}\mathbf{H}_U\mathbf{w} =& \mathrm{Re}\left(\mathbf{H}_U\right)\mathrm{Re}\left(\mathbf{w}\right) - \mathrm{Im}\left(\mathbf{H}_U\right)\mathrm{Im}\left(\mathbf{w}\right) \\ & + i\left[\mathrm{Re}\left(\mathbf{H}_U\right)\mathrm{Im}\left(\mathbf{w}\right) + \mathrm{Im}\left(\mathbf{H}_U\right)\mathrm{Re}\left(\mathbf{w}\right)\right],\end{aligned} \tag{18}$$

which leads into the following expressions

$$\mathrm{Re}\left(\mathbf{H}_U\mathbf{w}\right) = \mathbf{H}_{U_1}\widetilde{\mathbf{w}},\ \mathrm{Im}\left(\mathbf{H}_U\mathbf{w}\right) = \mathbf{H}_{U_2}\widetilde{\mathbf{w}}, \tag{19}$$

where $\widetilde{\mathbf{w}} = \left[\mathrm{Re}\left(\mathbf{w}^T\right), \mathrm{Im}\left(\mathbf{w}^T\right)\right]^T$, $\mathbf{H}_{U_1} = \left[\mathrm{Re}\left(\mathbf{H}_U\right), -\mathrm{Im}\left(\mathbf{H}_U\right)\right]$, and $\mathbf{H}_{U_2} = \left[\mathrm{Im}\left(\mathbf{H}_U\right), \mathrm{Re}\left(\mathbf{H}_U\right)\right]$. Also, it is easy to see that $\left\|\widetilde{\mathbf{w}}\right\|^2 = \left\|\mathbf{w}\right\|^2$.

Using the equivalents of $\mathrm{Re}\left(\mathbf{H}_U\mathbf{w}\right)$ and $\mathrm{Im}\left(\mathbf{H}_U\mathbf{w}\right)$ derived in (19), (17) transforms into

$$\begin{aligned}\min_{\widetilde{\mathbf{w}}}\ & \left\|\widetilde{\mathbf{w}}\right\|^2 \\ \text{s.t.}\ & \left(\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}\right)\widetilde{\mathbf{w}} = \mathbf{0}, && \text{(20a)}\\ & \mathrm{Re}\left(\mathbf{S}\right)\mathbf{H}_{U_1}\widetilde{\mathbf{w}} \geq \sqrt{\gamma}\,\mathbf{s}_r. && \text{(20b)}\end{aligned}$$

Proposition 1: A necessary condition for the existence of the optimal precoder for the directional modulation is $N_t > \frac{r'}{2}$ where r' is the rank of $\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}$. If $\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}$ is full rank, the necessary condition becomes $N_t > \frac{N_U}{2}$, which means that the number of transmit antennas needs to be more than half of the total number of receiver antennas.

Proof: Constraint (20a) shows that $\widetilde{\mathbf{w}}$ should lie in the null space of the matrix $\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}$. If the SVD of $\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}$ is shown by $\mathbf{U}\Sigma\mathbf{V}^H$, the orthonormal basis for the null space of $\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}$ are the last $2N_t - r'$ columns of the matrix $\mathbf{V}$ with r' being the rank of $\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}$ [38]. If $\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}$ is full rank, we have $r' = N_U$. For (20) to be feasible, the mentioned null space should exist, meaning that $2N_t - r' > 0$. ■

Provided that the necessary condition of Proposition 1 is met, a sufficient condition can be proposed from a geometrical point of view; namely that the feasible set of (20) is not empty. This holds if and only if the intersection of the linear spaces in the constraint set constitutes a non-empty set.

According to Proposition 1, the null space of $\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}$ spans $\widetilde{\mathbf{w}}$ as $\widetilde{\mathbf{w}} = \mathbf{E}\boldsymbol{\lambda}$ where

$$\mathbf{E} = \left[\mathbf{v}_{r'+1}, ..., \mathbf{v}_{2N_t}\right], \ \boldsymbol{\lambda} = \left[\lambda_1, ..., \lambda_{2N_t - r'}\right]. \tag{21}$$

By replacing $\widetilde{\mathbf{w}}$ with $\mathbf{E}\boldsymbol{\lambda}$, (20) boils down into

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \quad & \|\boldsymbol{\lambda}\|^2 \\ \text{s.t.} \quad & \mathrm{Re}\left(\mathbf{S}\right)\mathbf{H}_{U_1}\mathbf{E}\boldsymbol{\lambda} \geq \sqrt{\gamma}\,\mathbf{s}_r, \end{aligned} \tag{22}$$

Problem[4] (22) is a convex linearly constrained quadratic programming problem and can be solved efficiently using standard convex optimization techniques. The design problem (22) needs to be solved once for each set of the symbols, $\mathbf{s}_T$. Using optimization packages such as CVX to solve (22) can be time consuming, hence, we propose two other approaches to solve (22).

1) Iterative solution: In this part, we propose an iterative approach to solve (22). To do so, first, we define a real valued auxiliary vector denoted by $\mathbf{u}$ to change the inequality constraint of (22) into equality as

$$\begin{aligned} \min_{\boldsymbol{\lambda},\mathbf{u}} \quad & \|\boldsymbol{\lambda}\|^2 \\ \text{s.t.} \quad & \mathbf{B}\boldsymbol{\lambda} = \sqrt{\gamma}\mathbf{s}_r + \mathbf{u}, \ \mathbf{u} \geq \mathbf{0}. \end{aligned} \tag{23}$$

where $\mathbf{B} = \mathrm{Re}\left(\mathbf{S}\right)\mathbf{H}_{U_1}\mathbf{E}$. Using the penalty method [39], we can write (23) as an unconstrained optimization problem

$$\min_{\boldsymbol{\lambda},\mathbf{u}\geq 0} \ \|\boldsymbol{\lambda}\|^2 + \eta\|\mathbf{B}\boldsymbol{\lambda} - (\sqrt{\gamma}\mathbf{s}_r + \mathbf{u})\|^2, \tag{24}$$

which is equivalent to (23) when $\eta \to \infty$. We can solve (24) using an iterative approach by first optimizing $\mathbf{u}$ and considering $\boldsymbol{\lambda}$ to be fixed, and then optimizing $\mathbf{u}$ and considering $\boldsymbol{\lambda}$ to be fixed. In the following, we mention these two optimization problems and their closed-form solutions.

When optimizing over $\mathbf{u}$ and keeping $\boldsymbol{\lambda}$ fixed, the optimization problem to be solved can be written as

$$\min_{\mathbf{u}\geq 0} \ \|\mathbf{u} - (\mathbf{B}\boldsymbol{\lambda} - \sqrt{\gamma}\mathbf{s}_r)\|^2. \tag{25}$$

The closed-form solution of (25) is given in Lemma 1.

Lemma 1: The closed-form solution of (25) is $\mathbf{u}^\star = \left(\mathbf{B}\boldsymbol{\lambda} - \sqrt{\gamma}\mathbf{s}_r\right)_+$.

Proof: To solve (25), we need to minimize the distance between the vectors $\mathbf{u}$ and $\left(\mathbf{B}\boldsymbol{\lambda} - \sqrt{\gamma}\mathbf{s}_r\right)$. Since $\boldsymbol{\lambda}$ is fixed, the elements of $\left(\mathbf{B}\boldsymbol{\lambda} - \sqrt{\gamma}\mathbf{s}_r\right)$ are known. If an element of $\mathbf{B}\boldsymbol{\lambda} - \sqrt{\gamma}\mathbf{s}_r$ is nonnegative, we pick up the same value for the corresponding element of $\mathbf{u}$. If an element of $\mathbf{B}\boldsymbol{\lambda} - \sqrt{\gamma}\mathbf{s}_r$ is negative, we pick up zero for the corresponding element of $\mathbf{u}$ since $\mathbf{u} \geq \mathbf{0}$. This is equivalent to picking up $\mathbf{u}$ as

$$\mathbf{u}^\star = \left(\mathbf{B}\boldsymbol{\lambda} - \sqrt{\gamma}\mathbf{s}_r\right)_+. \tag{26}$$

■

[4]The design problem (22) can be extended to M-QAM modulation [14] by changing the constraint into equality. A detailed derivation falls beyond the scope of this paper.

Algorithm 1 Iterative approach to solve (24)

1: Pick up $\boldsymbol{\lambda}_n \in \mathbb{R}^{2N_t}$ and $\eta \in (0, \infty]$;
2: Substitute $\boldsymbol{\lambda}_n$ in (26) to get $\mathbf{u}_n$;
3: Substitute $\mathbf{u}_n$ in (29) to get $\boldsymbol{\lambda}_{n+1}$;
4: **if** $\|\boldsymbol{\lambda}_n - \boldsymbol{\lambda}_{n+1}\| \geq \epsilon$ **then**
5: $n = n + 1$;
6: **Go to** 2;
7: **end if**

When optimizing over $\boldsymbol{\lambda}$ and keeping $\mathbf{u}$ fixed, the optimization problem is

$$\min_{\boldsymbol{\lambda}} \ \|\boldsymbol{\lambda}\|^2 + \eta\|\mathbf{B}\boldsymbol{\lambda} - (\sqrt{\gamma}\mathbf{s}_r + \mathbf{u})\|^2. \tag{27}$$

The closed-form solution of (27) is given in Lemma 2.

Lemma 2: The closed-form solution of (27) is $\boldsymbol{\lambda}^\star = \left(\frac{\mathbf{I}}{\eta} + \mathbf{B}^T\mathbf{B}\right)^{-1}\mathbf{B}^T\left(\mathbf{a} + \mathbf{u}\right)$.

Proof: First, we expand (27) as

$$\begin{aligned} f\left(\boldsymbol{\lambda}\right) =& \|\boldsymbol{\lambda}\|^2 + \eta\|\mathbf{B}\boldsymbol{\lambda} - (\sqrt{\gamma}\mathbf{s}_r + \mathbf{u})\|^2 \\ =& \boldsymbol{\lambda}^T\left(\mathbf{I} + \eta\mathbf{B}^T\mathbf{B}\right)\boldsymbol{\lambda} - 2\eta\boldsymbol{\lambda}^T\left(\mathbf{B}^T\gamma\mathbf{s}_r + \mathbf{B}^T\mathbf{u}\right) \\ & + \eta(\sqrt{\gamma}\mathbf{s}_r + \mathbf{u})^T\left(\sqrt{\gamma}\mathbf{s}_r + \mathbf{u}\right). \end{aligned} \tag{28}$$

Taking the derivative of $f\left(\boldsymbol{\lambda}\right)$ with respect to $\boldsymbol{\lambda}$ yields

$$\boldsymbol{\lambda}^\star = \left(\frac{\mathbf{I}}{\eta} + \mathbf{B}^T\mathbf{B}\right)^{-1}\mathbf{B}^T\left(\mathbf{a} + \mathbf{u}\right). \tag{29}$$

Since $\mathbf{B}^T\mathbf{B}$ is positive semidefinite, addition of $\frac{\mathbf{I}}{\eta}$ to $\mathbf{B}^T\mathbf{B}$ for $\eta \neq \infty$ leads into diagonal loading of $\mathbf{B}^T\mathbf{B}$, which makes $\frac{\mathbf{I}}{\eta} + \mathbf{B}^T\mathbf{B}$ invertible. ■

Using the closed-form solutions mentioned in Lemmas 1 and 2, we propose Algorithm 1 to solve (24), where the matrix inversion in (29) needs to be calculated once per symbol transmission.

Lemma 3: Algorithm 1 monotonically converges to the optimal point.

Proof: Let's denote the objective function in (24) by $f\left(\boldsymbol{\lambda}, \mathbf{u}\right)$. Assume $\boldsymbol{\lambda}_0$ and $\mathbf{u}_0$ are initial values of $f\left(\boldsymbol{\lambda}, \mathbf{u}\right)$. Using $\boldsymbol{\lambda}_0$ in Algorithm 1 gives us $\mathbf{u}^\star$ and $\boldsymbol{\lambda}^\star$ from (26) and (29), respectively, which results in

$$f\left(\boldsymbol{\lambda}^\star, \mathbf{u}^\star\right) \leq f\left(\boldsymbol{\lambda}_0, \mathbf{u}^\star\right) \leq f\left(\boldsymbol{\lambda}_0, \mathbf{u}_0\right). \tag{30}$$

Since fixing $\boldsymbol{\lambda}$, (25), or $\mathbf{u}$, (27), leads into a convex function, each iteration in Algorithm 1 monotonically gets closer to the optimal point. This along with the fact that $f\left(\boldsymbol{\lambda}, \mathbf{u}\right)$ is lower bounded at zero, guarantees the convergence of Algorithm 1 to the optimal point. ■

2) Non-negative least squares: We can derive $\boldsymbol{\lambda}$ using the constraint of (23) as

$$\boldsymbol{\lambda} = \mathbf{B}^\dagger\left(\sqrt{\gamma}\mathbf{s}_r + \mathbf{u}\right). \tag{31}$$

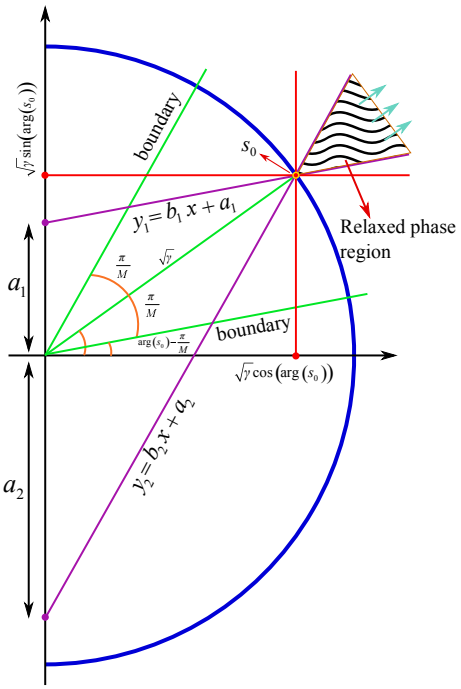

Fig. 4. Relaxed phase characterization of directional modulation design for symbol s_0 from M-PSK modulation.

Replacing the $\boldsymbol{\lambda}$ derived in (31) back into the objective of (23) yields

$$\begin{aligned} \min_{\mathbf{u}} \ & \left\| \mathbf{B}^\dagger \mathbf{u} + \sqrt{\gamma} \mathbf{B}^\dagger \mathbf{s}_r \right\|^2 \\ \text{s.t.} \ & \mathbf{u} \geq 0, \end{aligned} \tag{32}$$

which is a non-negative least squares optimization problem. Since $\mathbf{B}^\dagger$ and $\sqrt{\gamma}\mathbf{B}^\dagger \mathbf{s}_r$ are real valued, we can use the method of [40] or its fast version [41] to solve (32). We analyze the computational complexity of the non-negative least squares in Section V and mention its computational time in Section VI. Similar to Section IV-A1, $\mathbf{B}^\dagger$ needs to be calculated once per symbol transmission.

B. The Case of Strong Transmitter ($N_e < N_t$, Relaxed Phase)

The phases of the received signals in (13) are fixed, which decreases the degrees of freedom in designing $\mathbf{w}$, and consequently the power efficiency. To improve the power efficiency in the transmitter side, we can consider a region instead of a line for the phase of the received signal on each receiving antenna. In the M-PSK modulation, each symbol has a detection region within $\pm\frac{\pi}{M}$ degrees of its phase. The detection and relaxed phase regions for a reference symbol s_0 with the angle $\varphi_{s_0} = arg(s_0)$ are shown in Fig. 4 [42]. According to the characterization in Fig. 4, the relaxed phase design problem is defined as [13], [16], [42]

$$\begin{aligned} \min_{\mathbf{w}} \ & \|\mathbf{w}\|^2 \\ \text{s.t.} \ & \operatorname{Im}\left(\mathbf{h}_{n_r}^T \mathbf{w} e^{i\varphi_{n_r}}\right) \geq b_1 \operatorname{Re}\left(\mathbf{h}_{n_r}^T \mathbf{w} e^{i\varphi_{n_r}}\right) + a_1, \quad (33a) \\ & \operatorname{Im}\left(\mathbf{h}_{n_r}^T \mathbf{w} e^{i\varphi_{n_r}}\right) \leq b_2 \operatorname{Re}\left(\mathbf{h}_{n_r}^T \mathbf{w} e^{i\varphi_{n_r}}\right) + a_2, \quad (33b) \end{aligned}$$

for $r = 1, ..., R$ and $n = 1, ..., N$, where

$$\begin{aligned} a_1 &= c_1 - \sqrt{\left(\cos^{-2}\left(\varphi_{s_0} - \frac{\pi}{M}\right) - 1\right) c_2^2}, \\ a_2 &= -\tan\left(\varphi_{s_0} + \frac{\pi}{M}\right)\left[c_2 - \sqrt{\left(\sin^{-2}\left(\varphi_{s_0} + \frac{\pi}{M}\right) - 1\right) c_1^2}\right], \\ b_1 &= \tan\left(\varphi_{s_0} - \frac{\pi}{M}\right), \ b_2 = \tan\left(\varphi_{s_0} + \frac{\pi}{M}\right), \\ c_1 &= \sqrt{\gamma}\sin\left(\arg\left(s_0\right)\right), \ c_2 = \sqrt{\gamma}\cos\left(\arg\left(s_0\right)\right), \end{aligned} \tag{34}$$

and $\varphi_{n_r} = \arg(s_{n_r})$. The value of φ_{n_r} can be absorbed in the channel to rewrite (33) as

$$\begin{aligned} \min_{\mathbf{w}} \ & \|\mathbf{w}\|^2 \\ \text{s.t.} \ & \operatorname{Im}\left(\widetilde{\mathbf{h}}_{n_r}^T \mathbf{w}\right) \geq b_1 \operatorname{Re}\left(\widetilde{\mathbf{h}}_{n_r}^T \mathbf{w}\right) + a_1, \quad (35a) \\ & \operatorname{Im}\left(\widetilde{\mathbf{h}}_{n_r}^T \mathbf{w}\right) \leq b_2 \operatorname{Re}\left(\widetilde{\mathbf{h}}_{n_r}^T \mathbf{w}\right) + a_2. \quad (35b) \end{aligned}$$

By stacking the constraints, we can encapsulate (35) as

$$\begin{aligned} \min_{\mathbf{w}} \ & \|\mathbf{w}\|^2 \\ \text{s.t.} \ & \operatorname{Im}\left(\widetilde{\mathbf{H}}_U \mathbf{w}\right) \geq b_1 \operatorname{Re}\left(\widetilde{\mathbf{H}}_U \mathbf{w}\right) + a_1 \mathbf{1}, \quad (36a) \\ & \operatorname{Im}\left(\widetilde{\mathbf{H}}_U \mathbf{w}\right) \leq b_2 \operatorname{Re}\left(\widetilde{\mathbf{H}}_U \mathbf{w}\right) + a_2 \mathbf{1}, \quad (36b) \end{aligned}$$

where $\mathbf{1}$ is an $N_U \times 1$ unit vector. We can use the relations developed in (19) to transform (36) into

$$\begin{aligned} \min_{\mathbf{w}} \ & \|\mathbf{w}\|^2 \\ \text{s.t.} \ & \mathbf{B}_1 \mathbf{w} \geq \mathbf{a}, \end{aligned} \tag{37}$$

where

$$\mathbf{B}_1 = \begin{bmatrix} \widetilde{\mathbf{H}}_{U_2} - b_1 \widetilde{\mathbf{H}}_{U_1} \\ b_2 \widetilde{\mathbf{H}}_{U_1} - \widetilde{\mathbf{H}}_{U_2} \end{bmatrix}, \ \mathbf{a} = \begin{bmatrix} a_1 \mathbf{1} \\ -a_2 \mathbf{1} \end{bmatrix}. \tag{38}$$

Using a similar approach as in Section IV-A, (37) can be efficiently solved using the proposed iterative approach or the non-negative least squares formulation.

C. The Case of Strong Eavesdropper ($N_e \geq N_t$)

In this case, as the results in Section VI show, E can get a lower SER compared to the $N_e < N_t$ case. This capability of E comes from the fact that it has more antennas than T and owns global CSI knowledge, which puts E in a superior position compared to T from hardware and CSI knowledge point of view. Nevertheless, there is still one possible way to enhance the security. Focusing on the signal part and ignoring the noise, we can see from (5), for ZF estimator, or (7), for MMSE estimator, that $\hat{\mathbf{w}} = \mathbf{w}$. This means that the estimated symbols by E are equal to those induced on receiver antennas, $\mathbf{H}_U \mathbf{w}$, for the noiseless case, therefore, we can design the precoder such that the SNR of the received signal becomes

equal to the required level for successful decoding, which is defined by ACM.

As the results of the case $N_e < N_t$ in Section VI shows, the SNR level at E is lower than that of the users, which may prevent successful decoding of the M-PSK symbol at E. Based on this, we can minimize the sum power of the received signals at the users, $\|\mathbf{H}_U\mathbf{w}\|^2$, which is the same as the sum power of the estimated signals at E. In this frame, minimizing the sum power of the received signals is equivalent to minimizing the power of received signal on each receiving antenna. Since the power of the received signal on each receiving antenna is constrained, minimizing the sum power results in the minimum possible power on each receiving antenna. This results in a sort of *"security fairness"* among the users. The precoder design problem for the signal level minimization precoder can be defined as

$$\begin{aligned} \min_{\mathbf{w}} \ & \|\mathbf{H}_U\mathbf{w}\|^2 \\ \text{s.t.} \ & \arg\left(\mathbf{h}_{n_r}^T\mathbf{w}\right) = \arg\left(s_{n_r}\right), && \text{(39a)} \\ & \mathrm{Re}\left(s_{n_r}\right)\mathrm{Re}\left(\mathbf{h}_{n_r}^T\mathbf{w}\right) \geq \sqrt{\gamma}\mathrm{Re}^2\left(s_{n_r}\right), && \text{(39b)} \end{aligned}$$

for $r = 1, ..., R$ and $n = 1, ..., N$. Similar as in (13), the phase of the received signal on each receiving antenna in (39) is fixed, hence, we need to consider the signal level constraint on the real or imaginary part of the received signal. Following a similar procedure as in Section IV-A, (39) can be transformed to

$$\begin{aligned} \min_{\mathbf{w}} \ & \|\mathbf{H}_U\mathbf{w}\|^2 \\ \text{s.t.} \ & \mathbf{A}\mathrm{Re}\left(\mathbf{H}_U\mathbf{w}\right) - \mathrm{Im}\left(\mathbf{H}_U\mathbf{w}\right) = \mathbf{0}, \\ & \mathrm{Re}\left(\mathbf{S}\right)\mathrm{Re}\left(\mathbf{H}_U\mathbf{w}\right) \geq \sqrt{\gamma}\,\mathbf{s}_r, \end{aligned} \tag{40}$$

Using (18) to (19), we expand $\|\mathbf{H}_U\mathbf{w}\|^2$ as

$$\begin{aligned} \|\mathbf{H}_U\mathbf{w}\|^2 &= \widetilde{\mathbf{w}}^T\mathbf{H}_{U_1}^T\mathbf{H}_{U_1}\widetilde{\mathbf{w}} + \widetilde{\mathbf{w}}^T\mathbf{H}_{U_2}^T\mathbf{H}_{U_2}\widetilde{\mathbf{w}} \\ &= \widetilde{\mathbf{w}}^T\left(\mathbf{H}_{U_1}^T\mathbf{H}_{U_1} + \mathbf{H}_{U_2}^T\mathbf{H}_{U_2}\right)\widetilde{\mathbf{w}}, \end{aligned} \tag{41}$$

which along with (19) helps us convert (40) into

$$\begin{aligned} \min_{\widetilde{\mathbf{w}}} \ & \widetilde{\mathbf{w}}^T\left(\mathbf{H}_{U_1}^T\mathbf{H}_{U_1} + \mathbf{H}_{U_2}^T\mathbf{H}_{U_2}\right)\widetilde{\mathbf{w}} \\ \text{s.t.} \ & \left(\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}\right)\widetilde{\mathbf{w}} = \mathbf{0}, \\ & \mathrm{Re}\left(\mathbf{S}\right)\mathbf{H}_{U_1}\widetilde{\mathbf{w}} \geq \sqrt{\gamma}\,\mathbf{s}_r. \end{aligned} \tag{42}$$

For (42) to be feasible, $\widetilde{\mathbf{w}}$ has to be in the null space of $\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}$. Hence, we can write $\widetilde{\mathbf{w}}$ as a linear combination of the null space basis of $\mathbf{A}\mathbf{H}_{U_1} - \mathbf{H}_{U_2}$ yielding $\widetilde{\mathbf{w}} = \mathbf{E}\boldsymbol{\lambda}$, where $\mathbf{E}$ and $\boldsymbol{\lambda}$ are as in (21). This way, (42) boils down to[5]

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \ & \boldsymbol{\lambda}^T\mathbf{E}^T\left(\mathbf{H}_{U_1}^T\mathbf{H}_{U_1} + \mathbf{H}_{U_2}^T\mathbf{H}_{U_2}\right)\mathbf{E}\boldsymbol{\lambda} \\ \text{s.t.} \ & \mathbf{B}\boldsymbol{\lambda} \geq \sqrt{\gamma}\mathbf{s}_r, \end{aligned} \tag{43}$$

where $\mathbf{B} = \mathrm{Re}\left(\mathbf{S}\right)\mathbf{H}_{U_1}\mathbf{E}$. Similar as in Section IV-A, in the following, we propose an iterative algorithm and non-negative least squares formulation to solve (43).

[5]The design problem (43) can be extended to M-QAM modulation by changing the constraint into equality. A detailed derivation falls beyond the scope of this paper.

1) Iterative solution: By introducing the new variable $\mathbf{u}$, we can rewrite (43) as

$$\begin{aligned} \min_{\boldsymbol{\lambda},\mathbf{u}} \ & \boldsymbol{\lambda}^T\mathbf{E}^T\left(\mathbf{H}_{U_1}^T\mathbf{H}_{U_1} + \mathbf{H}_{U_2}^T\mathbf{H}_{U_2}\right)\mathbf{E}\boldsymbol{\lambda} \\ \text{s.t.} \ & \mathbf{B}\boldsymbol{\lambda} = \sqrt{\gamma}\mathbf{s}_r + \mathbf{u}. \end{aligned} \tag{44}$$

We can adapt Algorithm 1 to solve (43) by replacing the solution to $\boldsymbol{\lambda}^\star$ as

$$\boldsymbol{\lambda}^\star = \left(\frac{\mathbf{E}^T\left(\mathbf{H}_{U_1}^T\mathbf{H}_{U_1} + \mathbf{H}_{U_2}^T\mathbf{H}_{U_2}\right)\mathbf{E}}{\eta} + \mathbf{B}^T\mathbf{B}\right)^{-1}\mathbf{B}^T\left(\mathbf{a} + \mathbf{u}\right), \tag{45}$$

which is derived using a similar procedure as in Section IV-A1. Similar as in (29), the matrix inversion in (45) needs to be calculated only once per symbol transmission.

2) Non-negative least squares: Assuming that $\mathbf{H}_{U_1}$ and $\mathbf{H}_{U_2}$ are non-singular, the matrix $\mathbf{E}^T\left(\mathbf{H}_{U_1}^T\mathbf{H}_{U_1} + \mathbf{H}_{U_2}^T\mathbf{H}_{U_2}\right)\mathbf{E}$ is positive definite, hence, its Cholesky decomposition $\mathbf{E}^T\left(\mathbf{H}_{U_1}^T\mathbf{H}_{U_1} + \mathbf{H}_{U_2}^T\mathbf{H}_{U_2}\right)\mathbf{E} = \mathbf{L}\mathbf{L}^T$ exists and can be used in order to rewrite (44) as

$$\begin{aligned} \min_{\boldsymbol{\lambda},\mathbf{u}} \ & \left\|\mathbf{L}^T\boldsymbol{\lambda}\right\|^2 \\ \text{s.t.} \ & \mathbf{B}\boldsymbol{\lambda} = \sqrt{\gamma}\mathbf{s}_r + \mathbf{u}. \end{aligned} \tag{46}$$

We can derive $\boldsymbol{\lambda}$ using the constraint of (46) as $\boldsymbol{\lambda} = \mathbf{B}^\dagger\left(\sqrt{\gamma}\mathbf{s}_r + \mathbf{u}\right)$ and replace it back into the objective of (46) to get

$$\begin{aligned} \min_{\mathbf{u}} \ & \left\|\mathbf{L}^T\mathbf{B}^\dagger\mathbf{u} + \mathbf{L}^T\mathbf{B}^\dagger\sqrt{\gamma}\mathbf{s}_r\right\|^2 \\ \text{s.t.} \ & \mathbf{u} \geq \mathbf{0}, \end{aligned} \tag{47}$$

which is a non-negative least squares optimization problem. Since $\mathbf{L}^T\mathbf{B}^\dagger$ and $\mathbf{L}^T\mathbf{B}^\dagger\sqrt{\gamma}\mathbf{s}_r$ are real valued, we can use [40], [41] to solve (47) in an efficient way.

D. Benchmark Scheme

We consider the ZF at the transmitter [17] as the benchmark scheme since both our design and the benchmark scheme use the CSI knowledge at the transmitter to design the precoder.

In the benchmark scheme, ZF precoder is applied at the transmitter to remove the interference among the symbol streams. The received signals at users and E in the benchmark scheme are

$$\mathbf{y}_U = \mathbf{H}_U\mathbf{W}\mathbf{s}\beta + \mathbf{n}_\mathbf{U}, \tag{48}$$

$$\mathbf{y}_E = \mathbf{H}_E\mathbf{W}\mathbf{s}\beta + \mathbf{n}_\mathbf{E}, \tag{49}$$

where $\mathbf{W} = \mathbf{H}_U^H\left(\mathbf{H}_U\mathbf{H}_U^H\right)^{-1}$ is the precoding vector, $\mathbf{s}$ contains the symbols, and β is the amplification factor for the symbols which acts similar as $\sqrt{\gamma}$ in the directional modulation scheme. For a fair comparison, we pick up the same values for $\sqrt{\gamma}$ and β in the simulations.

When using the benchmark, E can use ZF and MMSE as two possible ways to estimate the symbols. In contrast to our method E can use the knowledge of $\mathbf{H}_U$ to calculate $\mathbf{W}$ in the benchmark scheme.

In the ZF approach, given that $N_e \geq N_t$, E can estimate $\mathbf{s}\beta$ as

$$\begin{aligned}\widehat{\mathbf{s}\beta} &= \left[(\mathbf{H}_E\mathbf{W})^H\mathbf{H}_E\mathbf{W}\right]^{-1}(\mathbf{H}_E\mathbf{W})^H\mathbf{y}_E \\ &= \mathbf{s}\beta + \left[(\mathbf{H}_E\mathbf{W})^H\mathbf{H}_E\mathbf{W}\right]^{-1}(\mathbf{H}_E\mathbf{W})^H\mathbf{n}_E \end{aligned} \quad (50)$$

where $\widehat{\mathbf{s}\beta}$ is the estimated $\mathbf{s}\beta$ at E. Since $\mathbf{H}_E\mathbf{W}$ is $N_e \times N_U$, $\left[(\mathbf{H}_E\mathbf{W})^H\mathbf{H}_E\mathbf{W}\right]^{-1}(\mathbf{H}_E\mathbf{W})^H\mathbf{H}_E\mathbf{W} = \mathbf{I}$ for $N_e \geq N_U$. Hence, in the benchmark scheme, E can derive the precoder and estimate the symbols using the ZF method when $N_e \geq N_U$. On the other hand, since our designed precoder depends on both the channels and symbols, E cannot derive the precoder and estimate the symbols using the ZF method when $N_e \geq N_U$.

In the MMSE approach, E can estimate $\mathbf{s}\beta$ as

$$\widehat{\mathbf{s}\beta} = \mathbf{G}_3\mathbf{y}_E, \quad (51)$$

where

$$\mathbf{G}_3 = \left[(\mathbf{H}_E\mathbf{W})^H\mathbf{C}_{\mathbf{w}}^{-1}\mathbf{H}_E\mathbf{W} + \mathbf{C}_{\mathbf{N}_E}^{-1}\right]^{-1}(\mathbf{H}_E\mathbf{W})^H\mathbf{C}_{\mathbf{w}}^{-1}. \quad (52)$$

When using the benchmark method, we will see in Section VI that SER at E when using the MMSE method depends on the difference between N_e and N_U, while the SER at E depends on the difference between N_e and N_t in our method. Broadly speaking, the base station has usually more antennas than the users, hence, it is more likely to have a higher difference between N_e and N_t rather than N_e and N_U, especially with a large-scale array. Therefore, it is more probable to preserve the security in our design compared to the benchmark scheme. Furthermore, by comparing (5) and (7) with (50), we see that E has to multiply $\widehat{\mathbf{W}}$ by $\mathbf{H}_U$ in our design whereas E does need to do this in the benchmark scheme.

V. Remarks on Computational Complexity

In this part, we analyze the computational complexity of our method and the benchmark scheme assuming that we pick up the non-negative formulation approach to design our precoder. The computational complexity of the non-negative least squares approach when using the interior point, (53), and fast projected gradient algorithms, (54), are, respectively, as [43]

$$O\left(N_t^3 \ln \varepsilon^{-1}\right), \quad (53)$$

$$O\left(\lambda_0^{\frac{1}{2}} \left\|\mathbf{w}_0 - \mathbf{w}^\star\right\| N_t^2 \varepsilon^{-\frac{1}{2}}\right), \quad (54)$$

where ε is the upper bound on the difference between the current, $f(\mathbf{w}_{itr})$, and the optimal value, $f(\mathbf{w}^\star)$, of the objective function as $f(\mathbf{w}_{itr}) - f(\mathbf{w}^\star) \leq \varepsilon$, $\lambda_0 = \lambda_{\max}\left(\mathbf{D}^T\mathbf{D}\right)$ with $\mathbf{D} = \mathbf{B}^\dagger$ for (32), $\mathbf{D} = \mathbf{B}_1^\dagger$ for (37), and $\mathbf{D} = \mathbf{L}^T\mathbf{B}^\dagger$ for (47).

Next, we derive the computational complexity of the benchmark scheme. Considering the structure of $\mathbf{W}$, the complexity of the benchmark scheme is derived as

$$2O\left(N_t N_U^2\right) + O\left(N_U^3\right) + O\left(N_t N_U\right). \quad (55)$$

Each of the problems in (32), (37), and (47), need to be solved once per group of symbols communications. In other words, N_U symbols can be communicated for each designed precoder. Therefore, a higher N_U means that more symbols can be communicated to the users for each designed precoder. On the other hand, the designed precoder in the benchmark scheme can be used as far as the channel is fixed. Hence, the computational complexity comparison between our scheme and the benchmark method depends on the channel changing rate, the total number of users' antennas, and the required accuracy in the non-negative least squares solution in (53) or (54).

VI. Simulation Results

In this part, we present different simulation scenarios to analyze the security and the performance of the directional modulation scheme for different precoding designs, and compare them with a benchmark scheme. In all simulations, channels are considered to be quasi static block Rayleigh which are generated using i.i.d. complex Gaussian random variables with distribution $\mathcal{CN}(0,1)$ and remain fixed during the interval that the M-PSK symbols are being induced at the receiver. Also, the noise is generated using i.i.d. complex Gaussian random variables with distribution $\mathcal{CN}(0,\sigma^2)$, and the modulation order used in all of the scenarios is 8-PSK modulation. Here, we simulate each precoder for both strong transmitter, $N_e < N_t$, and strong eavesdropper, $N_e \geq N_t$, cases. This way, we show the benefit of the power minimizer precoder in the strong transmitter case and the signal level minimizer precoder in the strong eavesdropper case. We use the acronym "min" instead of minimization in the legend of the figures. Unless otherwise mentioned, the power minimization precoder used in the scenario is the one with fixed phase. Here, the SER at E is derived by assuming that E decodes the symbols of all users.

In all the experiments, the computation times of the iterative method and non-negative least squares were considerably lower than the computation time of CVX. For example, in the case $N_t = 20$ and $N_U = 20$, while the average required time for the iterative method and non-negative least squares was 173.4 and 10.5 milliseconds, respectively, the same task was accomplished by CVX in 999.3 milliseconds.

In the first scenario, the effect of the number of transmitter antennas, N_t, on transmitter's consumed power and the SER at users and E are investigated for power minimization, fixed and relaxed phase, and signal level minimization precoders in (13), (37), and (39), and the benchmark scheme. The average consumed power, $\|\mathbf{w}\|^2$, with respect to N_t is shown in Fig. 5 for $N_U = 8, 10$. As N_t increases, the power consumption of our design with power minimization precoders, fixed and relaxed phase, converge to that of other two schemes. The power consumed by power minimization precoders with fixed and relaxed phase have the largest difference with the other two schemes, almost 6 and 8 dB, for $N_t = N_U$. We see that power minimization precoder with relaxed phase has 2.5 dB less power consumption compared to the power minimization precoder with fixed phase. The signal level minimization precoder

has almost the same power consumption as the benchmark scheme for $N_t = N_U = 10$. When the difference between N_t and N_U increases, all four schemes consume considerably less power. When N_t is larger than N_U, the degrees of freedom of the signal level minimization design increases and the power consumed by the signal level minimization precoder approaches that of the power minimization precoder.

The average total SER at users and the average SER at E with respect to N_t are presented in Figures 6 and 7 where the eavesdropper uses ZF and MMSE to estimate the symbols. Our designed precoders, power and signal level minimization, cause considerably more SER at E compared to the benchmark scheme for a long range of N_t. Furthermore, as N_e increases, there are cases, e.g., $N_t = 16$, that the error caused at E by the benchmark scheme decreases while the error caused by our designed precoders remains almost fixed when E uses the ZF estimator and reduces slightly when E uses the MMSE estimator. As Fig. 8 shows, our design with signal level minimization precoder and the benchmark scheme keep users' signal level norm constant. This leads into a constant SNR at E.

We see in Figures 6 and 7 that the MMSE estimator results in a less SER at the eavesdropper compared to the ZF estimator when the difference between N_t and N_U increases. On the other hand, for close values of N_t and N_U, the MMSE approach leads into the same SER as the ZF approach. Although the MMSE estimator reduces the SER at the eavesdropper, the error at the eavesdropper is still much higher than the users. For example, in Fig. 6, the SER at the eavesdropper is 0.2 while the SER at the users is 10^{-3}. We see in Fig. 7 that for $N_t = N_U = 10$, the eavesdropper can reduce the SER more in the benchmark scheme compared to our method. Since the directional modulation with signal level minimization imposes more error on E and consumes the same power as the benchmark scheme, it is the preferable choice for secure communication when $N_e \geq N_t$. Comparing Fig. 5 with Figures 6 and 7 shows that when the difference between N_t and N_U goes above a specific amount, the power and signal level minimization precoders converge in both power consumption and the SER at E and users.

The instantaneous power of the induced symbols to average noise power is shown in Figures 9 and 10 for power, fixed and relaxed phase, and signal level minimization precoders when $N_e > N_t$. As we see, even with E being able to estimate the symbols, the SNR at E is lower than the users. This shows that the processes carried out at E to perform ZF and MMSE estimations of $\mathbf{w}$ cause the SNR to be less than that of the users. As Fig. 10 shows, the signal level minimization precoder keeps the SNR at the users and E at the lowest possible level. The SNR at the users is on the required threshold for decoding while the SNR at E is much lower than that of the users and below the required threshold for successful decoding, which imposes the maximum SER on E.

In the second scenario, T's average power consumption, total average SER at the users, and average SER at E are plotted with respect to total receiving antennas, N_U. Fig. 11 shows the average consumed power with respect to N_U. Increasing N_U decreases the degrees of freedom and increases the power

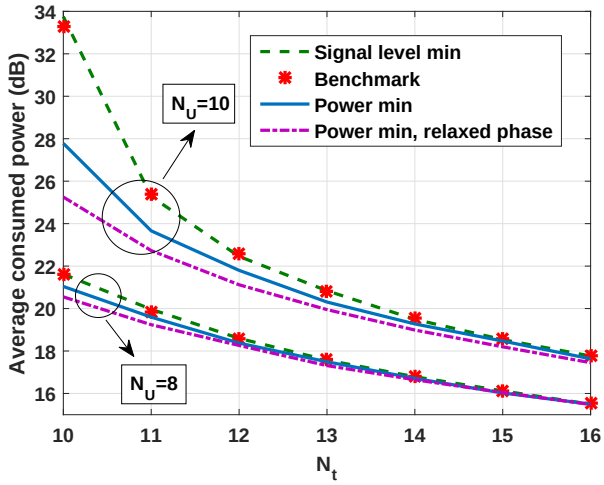

Fig. 5. Average consumed power with respect to N_t for our designed precoders and the benchmark scheme when $\gamma = 15.56$ dB and $\beta^2 = 15.56$ dB.

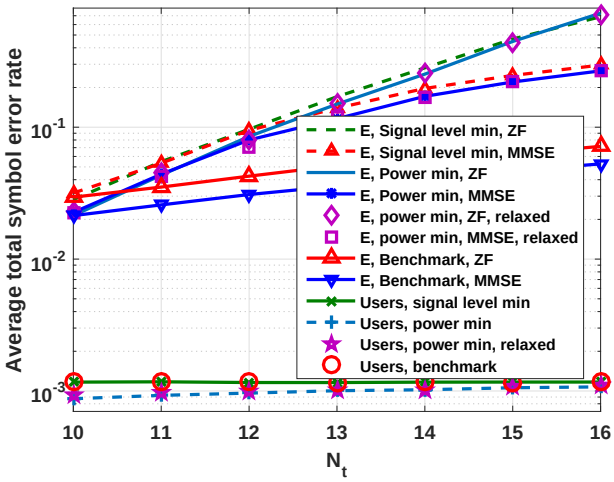

Fig. 6. Average total SER at the users and average SER at E with respect to N_t for our designed precoders and the benchmark scheme when $N_U = 10$, $N_e = 15$, $\gamma = 15.56$ dB, and $\beta^2 = 15.56$ dB.

consumption. As N_U approaches N_t, the difference between the power consumed by the power minimization precoder and the other two schemes increases.

We investigate the effect of N_U on average total SER at the users and the average SER at E in Figures 12 and 13. As N_U increases, the SNR provided by the power minimization precoder goes more above the threshold. This reduces the average SER at both users and E. On the other hand, regardless of difference between N_t and N_U, our design with signal level minimization precoder always preserves the SER at E in the maximum value. Compared to the ZF estimator, when our precoders are used, the MMSE approach reduces the SER at E for close values of N_U and N_t. As N_U approaches N_t,

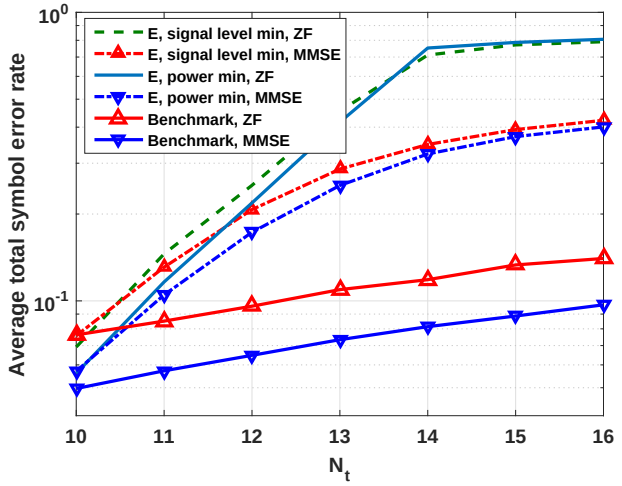

Fig. 7. Average SER at E with respect to N_t for our designed precoders and the benchmark scheme when $N_U = 10$, $N_e = 13$, $\gamma = 15.56$ dB, and $\beta^2 = 15.56$ dB.

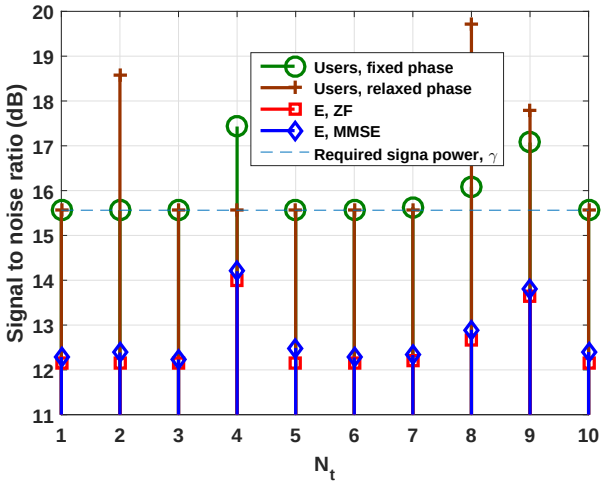

Fig. 9. Instantaneous symbol power to average noise power for power minimization precoder with fixed and relaxed phase designs when $N_t = 11$, $N_U = 10$, $N_e = 16$ and $\gamma = 15.56$ dB.

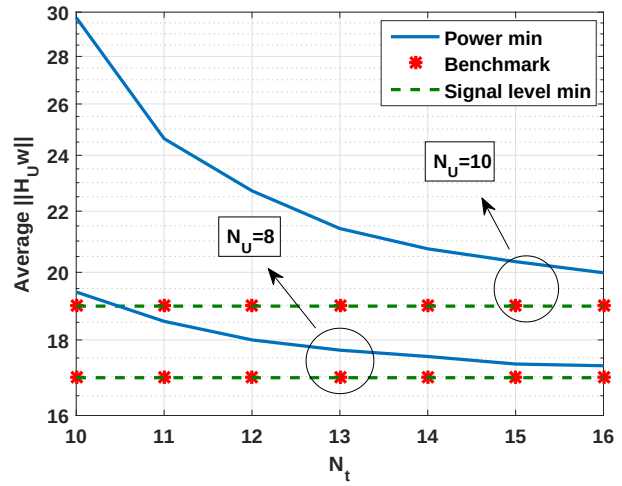

Fig. 8. Average $\|\mathbf{H}_U\mathbf{w}\|$ for our designed precoders and the benchmark scheme when $\gamma = 15.56$ dB, and $\beta^2 = 15.56$ dB.

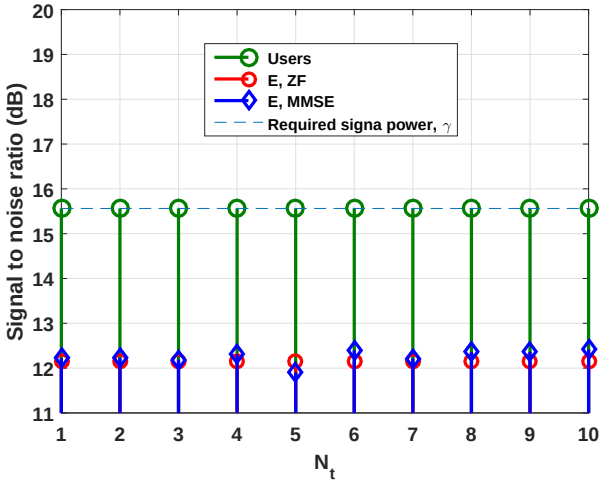

Fig. 10. Instantaneous symbol power to average noise power for signal level minimization precoder when $N_t = 11$, $N_U = 10$, $N_e = 16$ and $\gamma = 15.56$ dB.

the performance of ZF and MMSE techniques get closer. As shown in Fig. 13, the MMSE estimator at E reduces the SER more compared to the ZF estimator when the signal level min precoder is used. When $N_e > N_U$, our design imposes more SER at E compared to the benchmark scheme since $N_e \geq N_U$ is required for E to estimate the symbols in the benchmark scheme. As N_U approaches N_t, the SER imposed on E by the signal level minimization precoder and the benchmark scheme get closer.

The next scenario inspects the effect of the required SNR for the received signals, γ, on T's consumed power and the SER at users and E. Fig. 14 shows the average consumed power with respect to γ for our design and the benchmark scheme. The difference between the power consumed by the power minimization precoder and the other two schemes in low SNRs is more than that of high SNRs. The average total SER at users and the average SER at E with respect to γ is shown in Fig. 15. As SNR increases, the SER imposed on E by our design becomes more than that of the benchmark scheme. Furthermore, the performance of ZF and MMSE get closer as the SNR increases. The difference between the average total SER at the users for power and signal level minimization precoders remains almost constant as γ increases. The effect of low-density parity-check (LDPC) codes on the average total bit error rate (BER) at the users and the average BER at E is shown in Fig. 16 when the signal level minimization precoder is used for the case $N_e \geq N_t$.

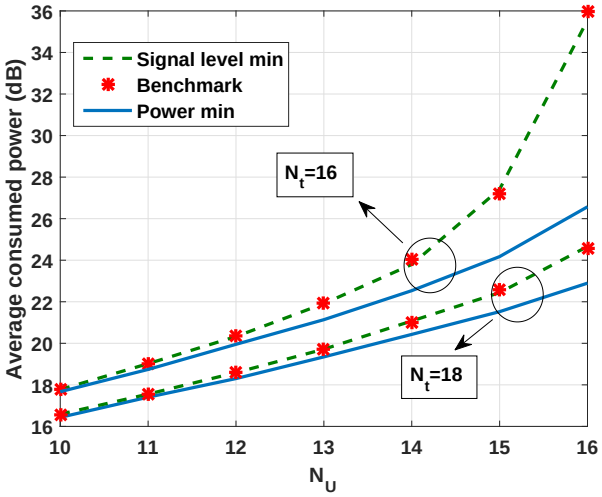

Fig. 11. Average consumed power with respect to N_U for our designed precoders and the benchmark scheme when $\gamma = 15.56$ dB, and $\beta^2 = 15.56$ dB.

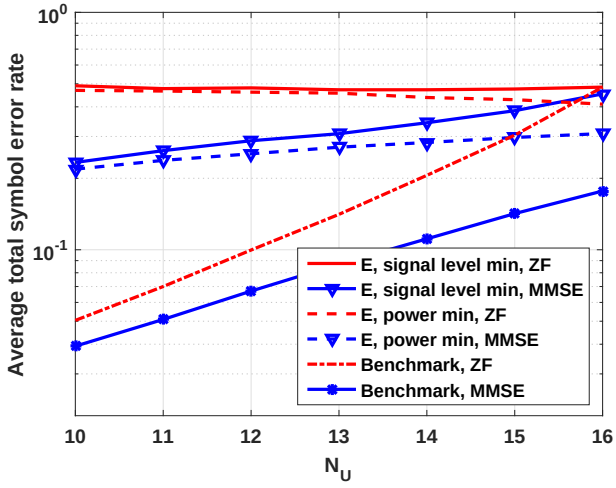

Fig. 13. Average SER versus N_U for our designed precoders and the benchmark scheme when $N_t = 16$, $N_e = 16$, $\gamma = 15.56$ dB, and $\beta^2 = 15.56$ dB.

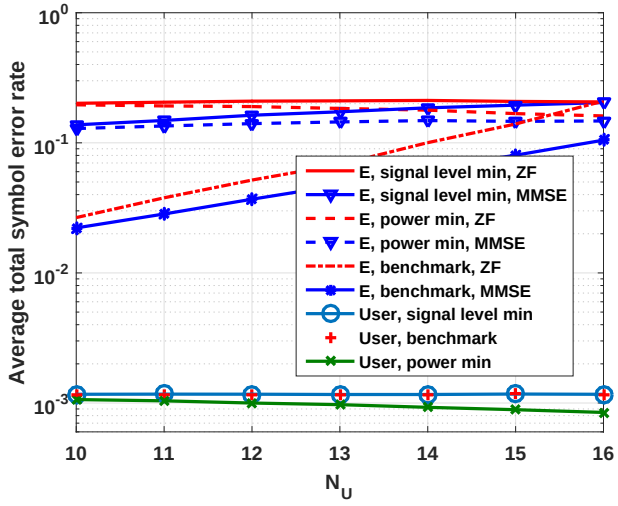

Fig. 12. Average SER versus N_U for our designed precoders and the benchmark scheme when $N_t = 16$, $N_e = 18$, $\gamma = 15.56$ dB, and $\beta^2 = 15.56$ dB.

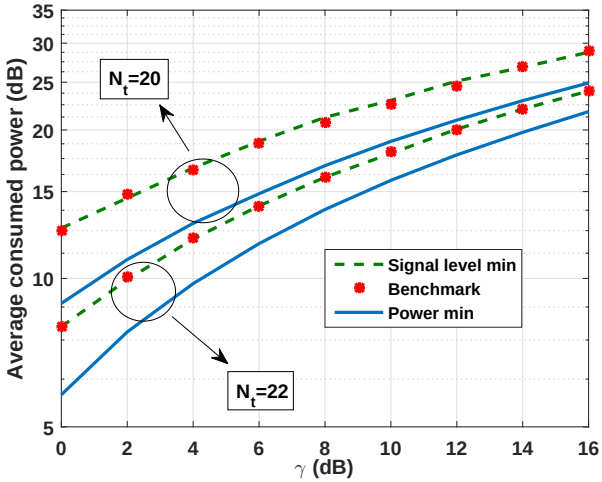

Fig. 14. Average consumed power with respect to required SNR for our designed precoders and the benchmark scheme when $N_U = 19$.

Next, similar to [8], we consider a LOS channel and use a uniform linear array (ULA). In this scenario, five single-antenna users are located on the circumference of a circle with radius 4 meters in the angles $10°, 50°, 110°, 260°, 310°$. The SER with respect to direction of transmission, θ, is shown in Fig. 17. As we see, the SER sharply decreases to 2×10^{-3} from 0.6 in the locations that users are present.

In the last scenario, we quantify the required time at E to perform brute-force method mentioned in Section III-D over all the possible communicated symbols between the transmitter and the receiver. The average brute-force time at E for the proposed precoders using an ordinary computer is shown in Fig. 18 for different modulation orders. As we see, increasing the system dimension or the modulation order increases the brute-force consumed time enormously.

VII. Conclusions

We used the directional modulation technology and followed a signal processing approach to enhance the security over multiuser MIMO channels in the presence of a multi-antenna eavesdropper. We studied the feasibility of different MIMO receiving algorithms at the eavesdropper and showed that the eavesdropper is able to use the ZF and MMSE algorithms to estimate the users' symbols. The legitimate users can directly decode the received signal via the conventional detectors, e.g., ML, while the results show that the usage of ZF or MMSE causes much more SER at the eavesdropper compared to the users. In addition, we derived the necessary condition

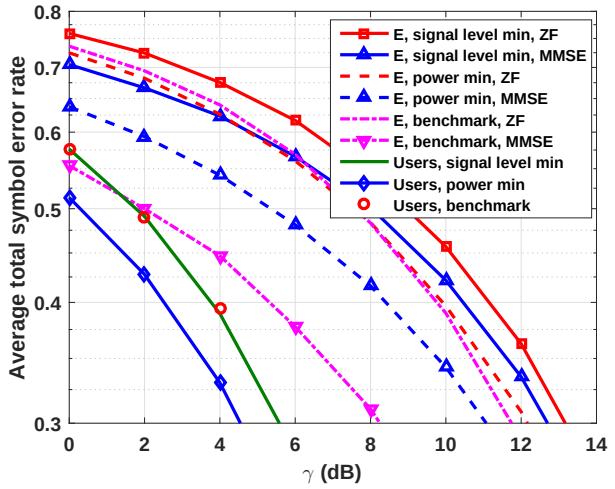

Fig. 15. Average SER versus required SNR for our designed precoders and the benchmark scheme when $N_t = 15$, $N_e = 17$, and $N_U = 14$.

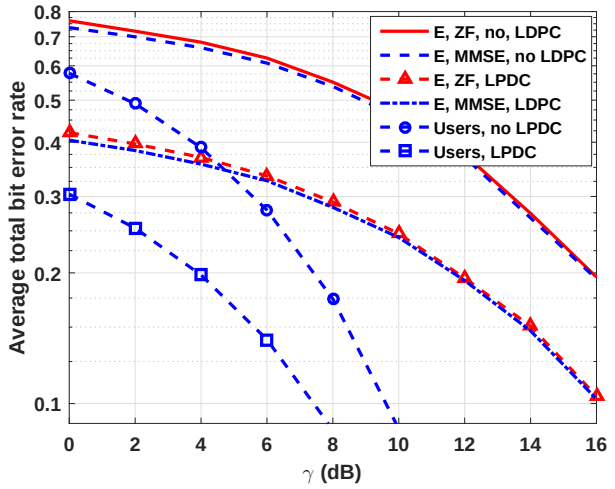

Fig. 16. Average BER versus required SNR for the signal level min precoder when $N_t = 10$, $N_e = 11$, and $N_U = 10$ with the code rate $5/6$.

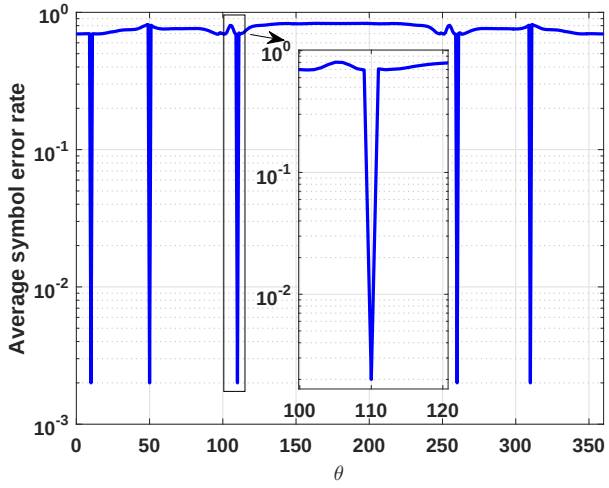

Fig. 17. Average SER versus the location in degrees the signal level min precoder when $N_t = 5$ and T communicates with five single-antenna users, i.e., $N_1 = N_2 = N_3 = N_4 = N_5 = 1$.

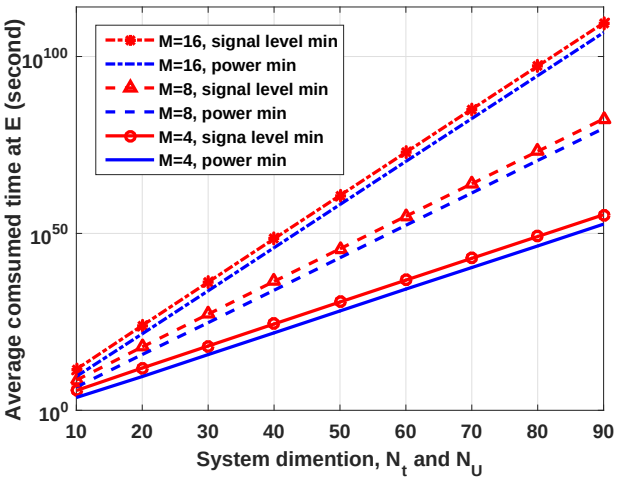

Fig. 18. Average consumed time at the eavesdropper with respect to the system dimension when using the proposed precoders with different modulation orders, $N_e = N_t$.

for the feasibility of the optimal precoder for the directional modulation. We proposed an iterative algorithm and non-negative least squares formulation to reduce the design time of the optimal precoders. The results showed that in most of the cases, our designed directional modulation precoders impose a considerable amount of SER on the eavesdropper compared to the conventional precoding. This is due to the fact that our precoders depend on both the CSI knowledge and the symbols while the conventional precoder only depends on the CSI knowledge and the eavesdropper can calculate it. The simulations showed that regardless of the number of antennas, the signal level minimization precoder keeps the SER at the eavesdropper on the maximum value, and it consumes the same power as the power minimization precoder when the difference between the number of transmit and receive antennas is above a specific value. In addition, the numerical examples showed that both the power and signal level minimization precoders outperform the benchmark scheme in terms of the power consumption and/or the imposed SER at the eavesdropper.

REFERENCES

[1] A. Kalantari, M. Soltanalian, S. Maleki, S. Chatzinotas, and B. Ottersten, "Secure M-PSK communication via directional modulation," in *IEEE Int. Conf. on Acoustics, Speech and Signal Proces. (ICASSP)*, Shanghai, China, Mar. 2016, pp. 3481–3485.

[2] A. D. Wyner, "The wire-tap channel," *Bell Systems Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Jan. 1975.

[3] M. Baldi, F. Chiaraluce, N. Laurenti, S. Tomasin, and F. Renna, "Secrecy transmission on parallel channels: Theoretical limits and performance of practical codes," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 11, pp. 1765–1779, Nov. 2014.

[4] E. Tekin and A. Yener, "The general gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008.

[5] A. Kalantari, S. Maleki, G. Zheng, S. Chatzinotas, and B. Ottersten, "Joint power control in wiretap interference channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3810–3823, Jul. 2015.

[6] A. Kalantari, G. Zheng, Z. Gao, Z. Han, and B. Ottersten, "Secrecy analysis on network coding in bidirectional multibeam satellite communications," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 9, pp. 1862–1874, Sep. 2015.

[7] A. Babakhani, D. Rutledge, and A. Hajimiri, "Transmitter architectures based on near-field direct antenna modulation," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 2674–2692, Dec. 2008.

[8] M. Daly and J. Bernhard, "Directional modulation technique for phased arrays," *IEEE Trans. Antennas Propag.*, vol. 57, no. 9, pp. 2633–2640, Sep. 2009.

[9] Y. Ding and V. Fusco, "MIMO inspired synthesis of directional modulation systems," *IEEE Antennas Wireless Propag. Lett.*, vol. PP, no. 99, 2015.

[10] ——, "Orthogonal vector approach for synthesis of multi-beam directional modulation transmitters," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 1330–1333, Jun. 2015.

[11] M. Hafez and H. Arslan, "On directional modulation: An analysis of transmission scheme with multiple directions," in *IEEE International Conference on Communication Workshop (ICCW)*, London, UK, Jun. 2015, pp. 459–463.

[12] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive multiuser interference in symbol level precoding for the MISO downlink channel," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2239–2252, May 2015.

[13] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3628–3640, Jul. 2015.

[14] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive interference through symbol level precoding for multi-level modulation," in *IEEE Global Commun. Conf. (GLOBECOM)*, CA, San Diego, Dec. 2015.

[15] ——, "Symbol-level multiuser MISO precoding for multi-level adaptive modulation: A multicast view," 2016. [Online]. Available: http://arxiv.org/abs/1601.02788

[16] ——, "Energy-efficient symbol-level precoding in multiuser MISO based on relaxed detection region," *IEEE Trans. Wireless Commun.*, 2016.

[17] L.-U. Choi and R. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 20–24, Jan. 2004.

[18] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[19] K. Ren, H. Su, and Q. Wang, "Secret key generation exploiting channel characteristics in wireless communications," *IEEE Wireless Commun.*, vol. 18, no. 4, pp. 6–12, August 2011.

[20] E. J. Baghdady, "Directional signal modulation by means of switched spaced antennas," *IEEE Trans. Commun.*, vol. 38, no. 4, pp. 399–403, Apr. 1990.

[21] M. Daly and J. Bernhard, "Beamsteering in pattern reconfigurable arrays using directional modulation," *IEEE Trans. Antennas Propag.*, vol. 58, no. 7, pp. 2259–2265, Jul. 2010.

[22] J. Lavaei, A. Babakhani, A. Hajimiri, and J. Doyle, "A study of near-field direct antenna modulation systems using convex optimization," in *American Control Conference (ACC)*, Baltimore, MD, Jun. 2010, pp. 1065–1072.

[23] M. Daly, E. Daly, and J. Bernhard, "Demonstration of directional modulation using a phased array," *IEEE Trans. Antennas Propag.*, vol. 58, no. 5, pp. 1545–1550, May 2010.

[24] M. Daly and J. Bernhard, "Directional modulation and coding in arrays," in *IEEE International Symposium on Antennas and Propagation (APSURSI)*, Spokane, WA, Jul. 2011, pp. 1984–1987.

[25] Y. Ding and V. Fusco, "Directional modulation transmitter radiation pattern considerations," *IET Microwaves, Antennas & Propagation*, vol. 7, no. 15, pp. 1201–1206, Dec. 2013.

[26] ——, "Constraining directional modulation transmitter radiation patterns," *IET Microwaves, Antennas & Propagation*, vol. 8, no. 15, pp. 1408–1415, Jul. 2014.

[27] ——, "A vector approach for the analysis and synthesis of directional modulation transmitters," *IEEE Trans. Antennas Propag.*, vol. 62, no. 1, pp. 361–370, Jan. 2014.

[28] ——, "Directional modulation far-field pattern separation synthesis approach," *IET Microwaves, Antennas & Propagation*, vol. 9, no. 1, pp. 41–48, Aug. 2014.

[29] ——, "Directional modulation-enhanced retrodirective array," *Electronics Lett.*, vol. 51, no. 1, pp. 118–120, Jan. 2015.

[30] H. Shi and A. Tennant, "Simultaneous, multichannel, spatially directive data transmission using direct antenna modulation," *IEEE Trans. Antennas Propag.*, vol. 62, no. 1, pp. 403–10, Jan. 2014.

[31] S. M. Kay, *Fundamentals of statistical signal processing*, ser. Prentice Hall signal processing series. Upper Saddle River (N.J.): Prentice Hall, 1993, vol. I, Estimation theory.

[32] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *URSI International Symposium on Signals, Systems, and Electronics*, Danvers, MA, Sep. 1998, pp. 295–300.

[33] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639–1642, Jul. 1999.

[34] B. Tao and H. Wu, "Improving the biclique cryptanalysis of AES," in *Australasian Conference Information Security and Privacy (ACISP)*, vol. 9144, Brisbane, Australia, Jun. 2015, pp. 39–56.

[35] A. Goldsmith and S.-G. Chua, "Adaptive coded modulation for fading channels," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 595–602, May 1998.

[36] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.

[37] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jul. 2014.

[38] G. Strang, *Introduction to Linear Algebra*, 4th ed. Wellesley-Cambridge Press and SIAM, 2009.

[39] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY: Cambridge University Press, 2004.

[40] C. L. Lawson and R. J. Hanson, *Solving least squares problems*, ser. Classics in applied mathematics. Philadelphia (Pa.): Society for Industrial and Applied Mathematics (SIAM), 1995.

[41] R. Bro and S. De Jong, "A fast non-negativity-constrained least squares algorithm," *Journal of Chemometrics*, vol. 11, no. 5, pp. 393–401, Sep. 1997.

[42] C. Masouros and E. Alsusa, "Soft linear precoding for the downlink of DS/CDMA communication systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 203–215, Jan. 2010.

[43] R. A. Polyak, *Projected Gradient Method for Non-Negative Least Square*. Contemporary Mathematics, 2015, vol. 636.

Ashkan Kalantari was born in Yazd, Iran. He received his BSc and MSc degrees from K. N. Toosi University of Technology, Tehran, Iran in 2009 and 2012, and his Ph.D from from University of Luxembourg, Luxembourg in April 2016. Since May 2016 he has been with the SIGCOM research group in the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, where he is working on caching for satellite networks within the ESA project. His research interests include optimization in wireless communications and caching for satellite communications.

Mojtaba Soltanalian (S08-M14) received the Ph.D. degree in electrical engineering (with specialization in signal processing) at the Department of Information Technology, Uppsala University, Sweden, in 2014. He is currently with the faculty of the Electrical and Computer Engineering Department, University of Illinois at Chicago (UIC).

Before joining UIC, he held research positions at the Interdisciplinary Centre for Security, Reliability and Trust (SnT, University of Luxembourg), and California Institute of Technology (Caltech). His research interests include different aspects of signal design and optimization for active sensing, communications, biology and medicine. He has been a recipient of -or supported in part by- different research grants from the European Research Council (ERC), the Swedish Research Council (VR), and Ericsson.

Sina Maleki received his PhD degree from Delft University of Technology, Delft, The Netherlands, in 2013, and his MSc from the same university in 2009. From July 2008 to April 2009, he was an intern student at the Philips Research Center, Eindhoven, The Netherlands, working on spectrum sensing for cognitive radio networks. Since August 2013, he has been working at the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, where he is working on cognitive radio for satellite communications within the EU FP7 CoRaSat project, as well as Luxembourgish national projects CO2SAT, SeMIGod, and SATSENT.

Symeon Chatzinotas (S06M09SM13) is currently the Deputy Head of the SIGCOM Research Group, Interdisciplinary Centre for Security, Reliability, and Trust, University of Luxembourg, Luxembourg. He has received the M.Eng. degree in telecommunications from Aristotle University of Thessaloniki, Thessaloniki, Greece, and the M.Sc. and Ph.D. degrees in electronic engineering from the University of Surrey, Surrey, U.K., in 2003, 2006, and 2009, respectively. In the past, he has worked on numerous RD projects for the Institute of Informatics Telecommunications, National Center for Scientific Research Demokritos, Institute of Telematics and Informatics, Center of Research and Technology Hellas, and Mobile Communications Research Group, Center of Communication Systems Research, University of Surrey, Surrey, U.K. Since 2004, he has authored more than 200 technical papers in refereed international journals, conferences and scientific books. He is the corecipient of the 2014 Distinguished Contributions to Satellite Communications Award, and Satellite and Space Communications Technical Committee, IEEE Communications Society, and CROWNCOM 2015 Best Paper Award. His research interests include multiuser information theory, co-operative/cognitive communications and wireless networks optimization.

Björn Ottersten (S87M89SM99F04) was born in Stockholm, Sweden, in 1961. He received the M.S. degree in electrical engineering and applied physics from Linköping University, Linköping, Sweden, in 1986, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1989. He has held research positions at the Department of Electrical Engineering, Linköping University, the Information Systems Laboratory, Stanford University, the Katholieke Universiteit Leuven, Leuven, Belgium, and the University of Luxembourg, Luxembourg. From 1996 to 1997, he was the Director of Research at ArrayComm Inc, a start-up in San Jose, CA, based on his patented technology. In 1991, he was appointed a Professor of Signal Processing with the Royal Institute of Technology (KTH), Stockholm, Sweden. From 1992 to 2004, he was the Head of the Department for Signals, Sensors, and Systems, KTH, and from 2004 to 2008, he was the Dean of the School of Electrical Engineering, KTH. Currently, he is the Director for the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg. As Digital Champion of Luxembourg, he acts as an Adviser to the European Commission. His research interests include security and trust, reliable wireless communications, and statistical signal processing. He is a Fellow of the EURASIP and a Member of the IEEE Signal Processing Society Board of Governors. He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and on the Editorial Board of IEEE Signal Processing Magazine. He is currently Editor-in-Chief of EURASIP Signal Processing Journal and a Member of the Editorial Boards of EURASIP Journal of Applied Signal Processing and Foundations and Trends in Signal Processing. He has coauthored journal papers that received the IEEE Signal Processing Society Best Paper Award in 1993, 2001, 2006, and 2013 and three IEEE conference papers receiving Best Paper Awards. He was the recipient of the IEEE Signal Processing Society Technical Achievement Award in 2011. He was the first recipient of the European Research Council Advanced Research Grant.