Phase-Only Zero Forcing for Secure Communication with Multiple Antennas

Wanyao Zhao, Si-Hyeon Lee, and Ashish Khisti

***Abstract*—Artificial noise (AN) transmission can enhance the secrecy in multi-antenna wireless channels by superimposing judiciously constructed synthetic noise signals over information signals during transmission. Motivated by large-scale MIMO systems, we study AN transmission under a constraint that all elements (weights) in each beamforming vector must have a constant magnitude, but can have arbitrary phases.**

For the special case of one AN beamforming vector and one legitimate receiver, we derive a necessary and sufficient condition for finding a beamforming vector in the null space of the legitimate receiver's channel and provide a geometric interpretation. For the i.i.d. Rayleigh fading channel, we derive an approximate expression for the probability of failing to find such a beamforming vector, and show that it decreases exponentially in the square of the number of transmit antennas.

For the general case, we propose a numerical algorithm for obtaining a set of mutually orthogonal AN beamforming vectors in the null space of all the receivers. Our approach involves reducing the problem to an unconstrained non-linear programming problem, which is then solved using the Gauss-Newton method. We show numerically that our proposed algorithm performs significantly better than a heuristic relaxation approach. Finally, for a multi-antenna system with multiple RF chains, we show numerically that the secrecy rate achieved by our proposed approach is close to that achieved by AN transmission with unconstrained beamforming, when the number of transmit antennas is sufficiently large.

I. Introduction

Artificial noise (AN) transmission [1] is a natural mechanism for enhancing the secrecy in multi-antenna transmission systems. In this approach, a transmitter transmits synthetic noise symbols in the null-space of the legitimate user's channel. These signals are super-imposed with information symbols from either the same transmitter or a separate transmitter. Through such a mechanism, the legitimate user's signal is not interfered by the artificial noise, whereas an eavesdropping receiver observes some additional interference.

In recent years, there has been a growing interest in deploying large scale antenna systems, often referred to as massive MIMO [2]. Massive MIMO can be built with inexpensive, low-power antenna elements with a limited number of RF chains. Each RF chain is associated with an *analog beamforming* vector – all elements (weights) must have constant magnitudes, but can have arbitrary phases – see, e.g., [3], [4]. This is unlike the classical MIMO systems, where each element of the the beamforming vectors can have an arbitrary amplitude and phase value. In the context of AN transmission, two interesting questions arise: (1) with such a constraint, is it still possible to find beamforming vectors that lie in the null space of the legitimate user's channel? and (2) given that the constraint is non-convex, can we devise an effective algorithm for finding the beamforming vectors? In this work we will address these questions.

We first consider the case of a single RF chain and a single receiver to examine how the constant-magnitude constraint affects the success of finding a beamforming vector in the null-space of the legitimate receiver's channel. We will refer to such a vector as a phase-only zero forcing (PZF) vector. We provide a geometric interpretation for PZF and establish a necessary and sufficient condition for its existence. We show that the probability of failing to find such a vector decreases exponentially in the square of the number of transmitting antennas for the case of i.i.d. Rayleigh-fading channels. For some other channel models we present simulation results that demonstrate that these channels also exhibit a similar performance.

For the general case, we develop algorithms to construct PZF vectors. We first consider the case of single RF chain and multiple receivers and propose an algorithm to construct a PZF vector that lies in the common null-space of all the legitimate users. Our approach involves reformulating the problem as an unconstrained non-linear programming (NLP) problem, where the objective function can be expressed as a sum-of-squares. We numerically compare our approach with a heuristic relaxation approach based on semi-definite programming (SDP) and show that it performs significantly better. We further extend the NLP algorithm for the general case of multiple RF chains and multiple receivers.

The remainder of this paper is organized as follows. The system model is given is Section II. In Section III, we present analytical results for the special case of single RF chain and single receiver. In Section IV, we develop algorithms for finding PZF vectors for the general case. In Section V, we provide simulation results. To illustrate the performance of our algorithm, the secrecy rate achieved by our algorithm is derived and compared with other schemes in Section VI. Section VII concludes the paper.

A. Related Work

AN transmission was originally proposed in [1] under the assumption of perfect channel state information. It has since been studied in a number of works [5]–[21], which we briefly

W. Zhao was with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada when this work was done. He is now with Bank of China Limited (e-mail: wanyao.zhao@mail.utoronto.ca). S.-H. Lee and A. Khisti are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada (e-mail: sihyeon.lee@utoronto.ca; akhisti@comm.utoronto.ca).

survey. Reference [5] showed that the AN transmission is near-optimal at high SNR for the multi-input, single-output, multi-eavesdropper (MISOME) wiretap channel. In [6], an optimal power allocation between the message and the AN transmissions was studied for fading MIMO channels. AN transmission in cellular networks such as the single-cell and multi-cell MIMO downlink systems is treated in [7]–[9]. In [11]–[14], the effect of channel estimation on AN transmission was studied. In particular, [11], [12] considered the imperfect legitimate receiver's CSI at the transmitter and [13], [14] considered adversarial jamming during training. For practical implementations of AN transmission, readers are referred to [17], [18]. Other approaches involving AN transmission are studied in [19]–[21]. We note that our approach of AN transmission with PZF beamforming has not been studied previously.

With the advent of massive MIMO, there is a growing body of literature on the design of beamforming matrix under the constraint on the number of RF chains. The work [22] designed an optimal analog preprocessing network (APN) beamforming matrix that has the phase-only constraint. The aim of this work is to minimize the interference at the input of the analog to digital converter (ADC) at the receiver. The works [3], [4], [23] studied a combination of digital and analog beamforming techniques for systems with a limited number of RF chains. The design criteria of these works is however different from ours because they do not consider a secrecy constraint.

We next discuss two lines of research that incorporate the constraint on the number of RF chains for secure communication with multiple antennas. Villiapan, Lozano and Heath [24] proposed a technique known as antenna subset modulation (ASM), which generates message symbols using different subset of antenna for each symbol transmission. By providing a simple inter-antenna phase shift and driving a different subset of antennas at each symbol interval, the authors show that it is possible to create a direction-dependent modulated signal. This allows the transmitter to introduce additional randomness in the constellations viewed at angles other than the target direction. The subset of antennas is chosen randomly at each time to confuse the eavesdropper. Another line of research is on directional modulation (DM) [25]–[29], where the transmitter sends intentionally distorted signals along all spatial directions other than a pre-selected secure communication direction. In [27], it was shown that using only four antennas the bit error rate at the eavesdropper is considerably higher than the legitimate receiver. There are two main drawbacks of these approaches. First, they require switching/antenna-modulation techniques at the symbol rate. In contrast, the proposed PZF scheme uses a constant beamforming direction. Second these schemes do not appear to easily generalize when there are multiple legitimate users. The PZF scheme naturally extends by transmitting noise symbols in the common null-space of all the legitimate users.

II. System Model

We consider a network that consists of one transmitter equipped with N antennas, K legitimate receivers each with a single antenna, and one eavesdropper with N_e antennas. The transmitter generates an input vector $\mathbf{x} \in \mathbb{C}^N$ under the average power constraint of P, i.e., $\mathrm{E}[\|\mathbf{x}\|^2] \leq P$. The channel outputs at the k-th legitimate receiver and the eavesdropper are given as[1]

$$y_k = \mathbf{h}_k^T \mathbf{x} + n_k, \tag{1}$$

$$\mathbf{z} = \mathbf{G}\mathbf{x} + \mathbf{n}_e, \tag{2}$$

respectively, where $n_k \sim \mathcal{CN}(0,1)$, $\mathbf{n}_e \sim \mathcal{CN}(0, \mathbf{I})$, $\mathbf{I}$ is the identity matrix of size N_e, and $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$ and $\mathbf{G} \in \mathbb{C}^{N_e \times N}$ denote the associated channel transfer matrices. For notational convenience, let us define $\mathbf{H} \triangleq \begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2 & \dots & \mathbf{h}_K\end{bmatrix}^T \in \mathbb{C}^{K\times N}$. We mainly consider the following two models for $\mathbf{H}$ and $\mathbf{G}$.

1) i.i.d. Rayleigh fading model: $h_{i,j} \sim \mathcal{CN}(0, \sigma^2)$ for $i \in [1:K]$ and $j \in [1:N]$,[2] and $g_{i,j} \sim \mathcal{CN}(0,\sigma^2)$ for $i \in [1:N_e]$ and $j \in [1:N]$ for some $\sigma^2 > 0$.
2) Geometric model with L paths:
$$\mathbf{h}_k = \sqrt{\frac{N}{L}} \sum_{l=1}^{L} \alpha_l \mathbf{a}_l^k(\phi_l^k) \tag{3}$$
for $k = [1:K]$, where $\alpha_l \sim \mathcal{CN}(0,1)$ denotes the channel gain for the l-th path and $\mathbf{a}_l^k(\phi_l^k)$ denotes the antenna array response vector given as follows assuming that the transmitter's antennas form a uniform linear array (ULA):
$$\frac{1}{\sqrt{N}}[1, e^{j\frac{2\pi}{\lambda}d\sin\phi_l^k}, \cdots, e^{j\frac{2\pi}{\lambda}(N-1)d\sin\phi_l^k}]^T. \tag{4}$$
In the above, ϕ_l^k denotes the random azimuth angle of departure for the l-th path to the k-th receiver, which is drawn independently and uniformly over $[0, 2\pi]$; d denotes the antenna spacing; and λ denotes the wavelength. Each row in $\mathbf{G}$ is assumed to follow the same model (3).

Remark 1. *We note in advance that our proposed techniques for finding PZF vectors will apply to arbitrary channel matrices. The statistical models discussed above are used to evaluate the performance of our schemes. Also note that the i.i.d. Rayleigh fading model is commonly used in rich scattering environments, while the L-path geometric model is used when the number of scatterers is relatively small [3].*

The transmitter uses M RF chains for AN transmission. Throughout this paper unless otherwise stated, we restrict the input $\mathbf{x}$ to take the following form:

$$\mathbf{x} = \sqrt{\frac{P}{MN}} \sum_{i=1}^{M} s_i \cdot \mathbf{b}_i \tag{5}$$

where $s_i \sim \mathcal{CN}(0,1)$ is the i-th noise symbol and $\mathbf{b}_i$ is the beamforming vector for s_i and is associated with the i-th RF chain. We note that $\mathbf{b}_i$ for $i \in [1:M]$ is implemented by using analog phase shifters [3], [4] and must be of the form:

$$\mathbf{b}_i = [e^{j\phi_{i,1}}, \cdots, e^{j\phi_{i,N}}]^T, \tag{6}$$

i.e., all the entries have unit amplitudes but arbitrary phases. See Figure 1 for an illustration of the system model.

[1]Throughout this paper we use the symbol ^{T} for ordinary transpose and the symbol † for the hermitian/conjugate transpose. We also use * to denote conjugate.

[2]For two integers i and j, $[i:j]$ denotes the set $\{i, i+1, \cdots, j\}$.

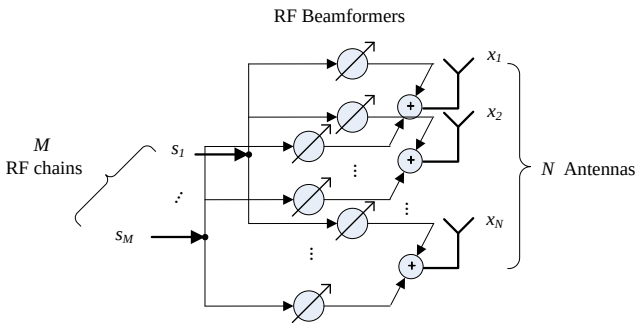

Figure 1. System model for the AN transmission with M RF chains

For AN transmission, we seek to find the beamforming vectors $\mathbf{b}_1, \cdots, \mathbf{b}_M$ belonging to the null-space of the legitimate receivers' channel i.e.,

$$\mathbf{H}\mathbf{b}_i = \mathbf{0}, \qquad i = 1, \ldots, M. \tag{7}$$

Furthermore we will design the beamforming vectors $\mathbf{b}_i$ to be mutually orthogonal. This choice follows naturally when the eavesdropper's channel is not known, since it ensures that the noise signal is distributed as uniformly as possible in the subspace of the eavesdropper. Furthermore from a practical point of view, our algorithm for constructing a single PZF vector can be naturally extended to generate multiple, mutually orthogonal PZF vectors using successive projections. It is interesting to note that orthogonal linear beamforming has been widely studied in downlink-MIMO when there is limited CSI available involving the receiver's channel, see e.g., [30], [31]. We refer to each $\mathbf{b}_i$ satisfying (6) and (7) as a phase-only zero forcing (PZF) vector.

Remark 2. *The motivation for transmitting $M > 1$ noise symbols is that it guarantees secrecy from an eavesdropper with $N_e \leq M$ antennas, see e.g., [5].*

Remark 3. *In our proposed model, the transmission of message symbols to the legitimate receivers is not discussed. The information symbols could be transmitted using additional RF chains by the same transmitter or by a separate transmitter. While we briefly discuss this in Section VI, the primary focus of this paper is on constructing the PZF vectors.*

III. Phase-Only Zero Forcing for the Special Case of $M = 1$ and $K = 1$

In this section, we give analytical results for the special case of $M = 1$ and $K = 1$, i.e., one RF chain and one legitimate receiver. For brevity, we denote $\mathbf{h}_1$ and $\mathbf{b}_1$ by $\mathbf{h} = [h_1, \cdots, h_N]^T$ and $\mathbf{b} = [e^{j\phi_1}, \cdots, e^{j\phi_N}]^T$, respectively, in this section. Note that $\mathbf{b}$ must satisfy the following:

$$\mathbf{h}^T\mathbf{b} = \sum_{i=1}^{N} h_i e^{j\phi_i} = 0. \tag{8}$$

The following proposition gives a necessary and sufficient condition for such a beamforming vector $\mathbf{b}$ to exist. The proof is provided in Appendix A.

Proposition 1. *Given any $\mathbf{h} \in \mathbb{C}^N$ with $N \geq 3$, there exists a PZF vector $\mathbf{b}$ satisfying (8) if and only if the following condition holds:*

$$a_{\max} \triangleq \max\left(|h_1|, \ldots, |h_N|\right) \leq \frac{1}{2}\sum_{j=1}^{N} |h_j|, \tag{9}$$

where $|h_j|$ denotes the magnitude of the complex number h_j.

Proof: See Appendix A. ∎

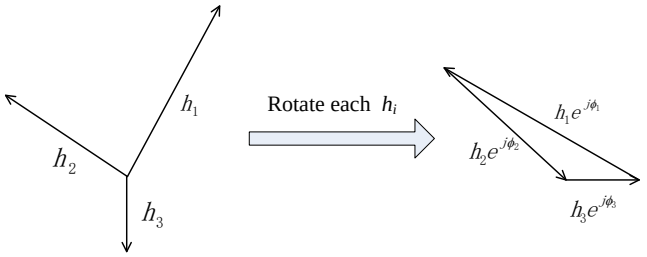

Figure 2. Geometric interpretation of equation (8) when $N = 3$: rotating each h_i such that a triangle can completed.

We make a few remarks regarding Proposition 1. Note that if $\mathbf{b}$ does not have the unit magnitude constraint, the solution to $\mathbf{h}^T\mathbf{b} = 0$ is obtained by selecting any vector in the null-space of $\mathbf{h}$. In this case, a non-trivial zero-forcing vector exists for $N \geq 2$.

Under the unit magnitude constraint, we must have $N \geq 3$. If $N = 2$, it is easy to verify that the equation $h_1 e^{j\phi_1} + h_2 e^{j\phi_2} = 0$ does not have a solution unless $|h_1| = |h_2|$ holds. For $N = 3$, the condition in (9) reduces to the following condition (assuming without loss of generality $|h_1| \geq |h_2| \geq |h_3|$):

$$|h_1| \leq |h_2| + |h_3|. \tag{10}$$

This condition has a natural geometric interpretation. We can view each complex number h_i as a 2-dimensional vector in the real-imaginary plane. Then, $h_i e^{j\phi_i}$ is geometrically a rotated version of this vector. The condition $h_1 e^{j\phi_1} + h_2 e^{j\phi_2} + h_3 e^{j\phi_3} = 0$ is equivalent to "completing a triangle" whose sides are of lengths $|h_i|$, as illustrated in Figure 2. Then the well-known law of cosines guarantees that condition (10) is a necessary and sufficient condition. Eq. (9) is a generalization of this condition where N sides of lengths $|h_i|$ must complete a polygon.

The following proposition shows that the probability of failing to meet the condition in Proposition 1 decreases exponentially in the square of the number of transmit antennas for the i.i.d. Rayleigh fading channel.

Proposition 2. *Let $\mathcal{E}$ denote the set of channel vectors for which a PZF vector does not exist i.e.,*

$$\mathcal{E} \triangleq \left\{ \mathbf{h} \in \mathbb{C}^N : \max\left(|h_1|, \ldots, |h_N|\right) > \frac{1}{2}\sum_{j=1}^{N} |h_j| \right\}. \tag{11}$$

If the entries of $\mathbf{h}$ are i.i.d. $\mathcal{CN}(0, \sigma^2)$, it follows that

$$\Pr(\mathcal{E}) \approx N e^{-\frac{N^2\pi}{16}}, \qquad N \gg 1. \tag{12}$$

Proof: When $h_i \sim \mathcal{CN}(0, \sigma^2)$, its magnitude $|h_i|$ follows the Rayleigh distribution (i.e., $f(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$). Let us define $a_{\max} \triangleq \max\{|h_1|, \cdots, |h_N|\}$ and recall that its cumulative distribution function (cdf) is given by:

$$F_{a_{\max}}(x) = (1 - e^{-\frac{x^2}{2\sigma^2}})^N. \tag{13}$$

Furthermore the cdf of its scaled version $a' = \frac{2}{N} a_{\max}$ is given

by $F_{a'}(x) = (1 - e^{-\frac{N^2x^2}{8\sigma^2}})^N$.

By the weak law of large numbers, $\frac{1}{N}\sum_{i=1}^{N}|h_i|$ converges to the expectation of $|h_1|$, i.e. $\sigma\sqrt{\frac{\pi}{2}}$, as N becomes large. Thus we have that:

$$\Pr\left\{a_{\max} \le \frac{1}{2}\sum_{i=1}^{N}|h_i|\right\}$$

$$= \Pr\left\{\frac{2}{N}a_{\max} \le \frac{1}{N}\sum_{i=1}^{N}|h_i|\right\} \tag{14}$$

$$\approx \Pr\left\{\frac{2}{N}a_{\max} \le \sigma\sqrt{\frac{\pi}{2}}\right\} = F_{a'}\left(\sigma\sqrt{\frac{\pi}{2}}\right) \tag{15}$$

$$= (1 - e^{-\frac{N^2\pi}{16}})^N. \tag{16}$$

Thus it follows that:

$$\Pr(\mathcal{E}) = 1 - \Pr\left(a_{\max} \le \frac{1}{2}\sum_{i=1}^{N}a_i\right) \tag{17}$$

$$\approx 1 - \left(1 - e^{-\frac{N^2\pi}{16}}\right)^N \tag{18}$$

$$\approx N \cdot e^{-\frac{N^2\pi}{16}}. \tag{19}$$

■

In Section V-A, we provide a plot that compares the analytical result in Proposition 2 with a simulation result for the i.i.d. Rayleigh fading channel. Furthermore, in the same section, we also plot simulation results for the Kronecker model (see e.g., [32, eq. (3)]) as well as the geometric model with L paths. For these two models, it seems difficult to obtain analytical results like Proposition 2, but the simulation results indicate that the outage probabilities are even smaller for such channel models. In the case when $K > 1$, we present some simulations in Section V-B, where correlation between channel gains also seems to be advantageous.

IV. Phase-Only Zero Forcing: General Case

For the special case of $M = 1$ and $K = 1$ in Section III, the condition for phase-only zero forcing is interpreted as forming a N-sided polygon in a 2-dimensional space. For the general case, we require each PZF vector to simultaneously form K polygons, each with N sides. In this case, there does not seem to be a simple way to generalize the analysis.

In this section, we provide algorithms to construct PZF vectors for the general case. We first consider the case of $M = 1$ and $K \ge 1$. For this case, we discuss two approaches: the SDP relaxation and the NLP approach (see e.g., [33], [34]). The former approach is based on a heuristic relaxation, while the latter involves transforming the original problem into an objective function involving sum of squares, which can be then evaluated using the iterative Gauss-Newton method. It turns out that the latter approach not only outperforms the SDP relaxation approach but also seems to suggest that a solution exists with high probability when $N \ge 2K + 3$. Thus, for the case when $M \ge 1$, we focus on the second approach and generalize it.

A. Case of $M = 1$ and $K \ge 1$

In this subsecion, we consider the case of $M = 1$ and $K \ge 1$ and provide two algorithms for finding the PZF vector $\mathbf{b}$.

1) SDP Relaxation Approach: The problem can be written as

$$\min\ \mathbf{b}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{b} \tag{20}$$

$$\text{s.t. } b_i b_i^* = 1,\ i = 1, \cdots, N. \tag{21}$$

By introducing $\mathbf{U} = \mathbf{b}\mathbf{b}^\dagger$ and $\mathbf{C} = \mathbf{H}^\dagger\mathbf{H}$, the aforementioned problem can be rewritten as follows:

$$\min\ \text{trace}(\mathbf{C}\mathbf{U}) \tag{22}$$

$$\text{s.t. } \mathbf{U}_{i,i} = 1,\ i = 1, \cdots, N, \tag{23a}$$

$$\text{rank}(\mathbf{U}) = 1, \tag{23b}$$

$$\mathbf{U} \succeq 0, \tag{23c}$$

where trace$(\cdot)$ and rank$(\cdot)$ denote the trace and the rank of a matrix, respectively.

By getting rid of the rank constraint, we have a relaxed SDP problem. The solution $\mathbf{U}^\star$ in general does not have unit rank. A heuristic approach is to consider only the eigenvector of $\mathbf{U}^\star$ corresponding to the largest eigenvalue, and then normalize each element of this eigenvector to have unit-amplitude.

2) NLP Approach: We first denote the null space of the channel matrix $\mathbf{H}$ as

$$\mathbf{H}^\perp = \begin{bmatrix}\mathbf{h}_1' & \mathbf{h}_2' & \cdots & \mathbf{h}_{N-K}'\end{bmatrix}, \tag{24}$$

where $\mathbf{h}_i' \in \mathbb{C}^N$ are column vectors satisfying $\mathbf{h}_j^T\mathbf{h}_i' = 0$. Solving equation $\mathbf{H}\mathbf{b} = \mathbf{0}$ is equivalent to finding a complex vector $\mathbf{m} = [m_1, m_2, \cdots, m_{N-K}]^T$ such that

$$\mathbf{H}^\perp \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_{N-K}\end{bmatrix} = \begin{bmatrix} e^{j\phi_1} \\ e^{j\phi_2} \\ \vdots \\ e^{j\phi_N}\end{bmatrix}. \tag{25}$$

With the i-th row of $\mathbf{H}^\perp$ denoted as $\mathbf{r}_i^T \in \mathbb{C}^{1\times(N-K)}$, it is required that $\mathbf{r}_i$ satisfies $\mathbf{r}_i^T\mathbf{m} = e^{j\phi_i}$ or $|\mathbf{r}_i^T\mathbf{m}| = 1$.

To simplify further, we represent $\mathbf{r}_i$ and $\mathbf{m}$ as

$$\mathbf{r}_i = \mathbf{c}_i + j\mathbf{d}_i, \qquad \mathbf{m} = \mathbf{p} + j\mathbf{q}$$

where $\mathbf{c}_i$ (or $\mathbf{p}$) and $\mathbf{d}_i$ (or $\mathbf{q}$) denote the real and imaginary part of $\mathbf{r}_i$ (or $\mathbf{m}$) respectively. Then, the condition $|\mathbf{r}_i^T\mathbf{m}| = 1$ simplifies to:

$$[\mathbf{p}^T, \mathbf{q}^T]\begin{bmatrix}\mathbf{c}_i \\ -\mathbf{d}_i\end{bmatrix}[\mathbf{c}_i^T, -\mathbf{d}_i^T]\begin{bmatrix}\mathbf{p} \\ \mathbf{q}\end{bmatrix} +$$

$$[\mathbf{p}^T, \mathbf{q}^T]\begin{bmatrix}\mathbf{d}_i \\ \mathbf{c}_i\end{bmatrix}[\mathbf{d}_i^T, \mathbf{c}_i^T]\begin{bmatrix}\mathbf{p} \\ \mathbf{q}\end{bmatrix} = 1 \tag{26}$$

for $i = 1, \cdots, N$. Eq. (26) can be further simplified to

$$[\mathbf{p}^T, \mathbf{q}^T]\begin{bmatrix}\mathbf{c}_i\mathbf{c}_i^T + \mathbf{d}_i\mathbf{d}_i^T & \mathbf{d}_i\mathbf{c}_i^T - \mathbf{c}_i\mathbf{d}_i^T \\ \mathbf{c}_i\mathbf{d}_i^T - \mathbf{d}_i\mathbf{c}_i^T & \mathbf{c}_i\mathbf{c}_i^T + \mathbf{d}_i\mathbf{d}_i^T\end{bmatrix}\begin{bmatrix}\mathbf{p} \\ \mathbf{q}\end{bmatrix} = 1 \tag{27}$$

for $i = 1, \cdots, N$. Now, define

$$\mathbf{w} \triangleq \begin{bmatrix}\mathbf{p} \\ \mathbf{q}\end{bmatrix},\ \mathbf{Q}_i \triangleq \begin{bmatrix}\mathbf{c}_i\mathbf{c}_i^T + \mathbf{d}_i\mathbf{d}_i^T & \mathbf{d}_i\mathbf{c}_i^T - \mathbf{c}_i\mathbf{d}_i^T \\ \mathbf{c}_i\mathbf{d}_i^T - \mathbf{d}_i\mathbf{c}_i^T & \mathbf{c}_i\mathbf{c}_i^T + \mathbf{d}_i\mathbf{d}_i^T\end{bmatrix}, \tag{28}$$

then we need to solve the system of quadratic equations

$$\mathbf{w}^T\mathbf{Q}_i\mathbf{w} = 1,\ i = 1, \cdots, N. \tag{29}$$

To be more compact, let $F(\mathbf{w}) = [f_1(\mathbf{w}), \cdots, f_N(\mathbf{w})]^T$ where

$$f_i(\mathbf{w}) \triangleq \mathbf{w}^T\mathbf{Q}_i\mathbf{w} - 1.$$

To solve the system of equations $F(\mathbf{w}) = 0$, we transform it into the following sum-of-squares optimization problem without constraints:

$$\min g(\mathbf{w}), \tag{30}$$

where $g(\mathbf{w}) \triangleq \sum_{i=1}^{N} f_i(\mathbf{w})^2$. This can be solved using Gauss-Newton method to find a locally optimal solution $\mathbf{w}_{opti}$. If the value of objective function in (30) can be minimized to zero, $\mathbf{w}_{opti}$ satisfies (29) obviously. In this case the solution to the PZF vector is given by:

$$\mathbf{b}_{opti} = \mathbf{H}^{\perp}(\mathbf{w}_{opti,[1:N-K]} + j\mathbf{w}_{opti,[N-K+1:2(N-K)]}). \tag{31}$$

Also note that when there is no solution to the original problem $\mathbf{Hb} = \mathbf{0}$, $\mathbf{b}_{opti}$ returned by the algorithm will not satisfy the unit-amplitude constraint. To find a nearest valid beamforming vector, we normalize each of $\mathbf{b}_{opti}$'s component to have unit-amplitude.

We conclude by summarizing the main steps of the Gauss-Newton algorithm [34]:

Algorithm 1.

1) **Initialization:** *Choose* $\mathbf{w}_0$ *as the initial starting point. Select* ϵ *to be a small positive number and let* $k = 0$.
2) **Update:** *Update the solution as follows:*
$$\mathbf{w}_{k+1} = \mathbf{w}_k - (J_F^T(\mathbf{w}_k)J_F(\mathbf{w}_k))^{-1}J_F^T(\mathbf{w}_k)F(\mathbf{w}_k).$$
3) **Termination:** *When* $|g(\mathbf{w}_{k+1}) - g(\mathbf{w}_k)| < \epsilon$*, stop and return* $\mathbf{w}_{k+1}$ *as the optimal solution. Otherwise, go to Step 2 with* $k = k + 1$.

In Algorithm 1, $J_F(\mathbf{w}) \in \mathbb{R}^{N\times 2(N-K)}$ denotes the Jacobian matrix of the mapping $F(\mathbf{w})$. For the case $2K < N$ when $J_F^T(\mathbf{w}_k)J_F(\mathbf{w}_k)$ is non-invertible, a pseudo-inverse can be adopted [34, p. 270, eq. (13)].[3] The computational complexity of the Gauss-Newton algorithm is determined by the number of iterations and the complexity in each iteration which is of the order of $O(N^4)$. The latter is dominated by the complexity of computing the Jacobian and the inverse of the matrix. The number of iterations of the algorithm is influenced by the selection of the initial point and the tolerable error ϵ. The initial point $\mathbf{w}_0$ we choose for the algorithm is the all-zero vector. With ϵ chosen to be 10^{-6}, the algorithm converges mostly within 10 to 20 iterations. The SDP approach, which works faster than the NLP approach, uses interior point method to find the optimal solution, and its computation complexity analysis can be found in standard references [33], [37]. We omit the details as the SDP relaxation approach does not generally lead to a satisfactory solution.

B. General case of $M \geq 1$ and $K \geq 1$

In this subsection, we extend the NLP approach for the general case of $M \geq 1$ and $K \geq 1$. We are required to compute M beamforming vectors $\mathbf{b}_1, \cdots, \mathbf{b}_M$, each of which satisfies (6) and (7). Our algorithm uses Gram-Schmidt procedure to find a subspace that is orthogonal to all the previous beamforming vectors. In the following, the inner product is denoted as $\langle \cdot, \cdot \rangle$ and the l_2 norm is denoted as $||\cdot||$.

Algorithm 2.

1) **Initialization:** *Start at* $i = 1$ *and let* $\mathbf{H}_1^{\perp}$ *contain the columns in the null-space of* $\mathbf{H}$*, defined as in* (24).
2) **Compute** $\mathbf{b}_i$**:** *Given* $\mathbf{H}_i^{\perp}$ *apply Algorithm 1 in Section IV-A2 and* (31) *to find* $\mathbf{m}_i$ *and* $\mathbf{b}_i = [e^{j\phi_{i,1}}, \cdots, e^{j\phi_{i,N}}]^T$ *such that* $\mathbf{H}_i^{\perp}\mathbf{m}_i = \mathbf{b}_i$ *as in* (25).
3) **Projection:** *If* $i > M$*, stop. Otherwise, let*
$$\mathbf{h}'_{i+1,j} = \mathbf{h}'_{i,j} - c^*_{i,j}\mathbf{b}_i$$
where $c_{i,j} = \langle \mathbf{h}'_{i,j}, \mathbf{b}_i \rangle / ||\mathbf{b}_i||^2$*. Let*
$$\mathbf{H}_{i+1}^{\perp} = [\mathbf{h}'_{i+1,1}, \mathbf{h}'_{i+1,2}, \cdots, \mathbf{h}'_{i+1,N-K}].$$
4) *Let* $i = i + 1$*, go to 2).*

We note that since $\mathbf{b}_i$ is in the subspace spanned by the columns of $\mathbf{H}_i^{\perp}$, the dimension of $\mathbf{H}_{i+1}^{\perp}$ is reduced by one in step 3), and still in the null space of the channel matrix $\mathbf{H}$. The Gram-Schmidt procedure also guarantees that the PZF vectors satisfy the orthogonality constraint.

V. Simulation Results

In this section we complement the results in the previous sections by presenting simulation results for statistical channel models. In Section V-A we consider the case of single RF chain and single receiver and compute the probability of not finding a PZF vector as a function of the number of transmit antennas for various statistical models. We find that the outage probability is largest for the Rayleigh fading channel and is in close agreement with the analytical bound in Proposition 2. In Section V-B we consider the case of single RF chain and multiple receivers and compare the performance of the NLP approach and the SDP relaxation, which were discussed in Section IV-A. We find that the NLP approach not only outperforms SDP relaxation, but also that it exhibits an interesting threshold phenomenon. In particular it appears that the algorithm succeeds with high probability whenever $N \geq 2K + 3$ i.e., the number of transmit antennas is approximately twice the number of users. This is in contrast to the case without amplitude constraint, where it suffices that $N > K$. Finally we treat the fully general case of multiple RF chains and multiple receivers in Section V-C and report a similar threshold effect in Fig. 8.

A. Case of $M = 1$ and $K = 1$

Figure 3 illustrates the analytical approximation in Proposition 2 and the simulation result for the i.i.d. Rayleigh fading channel. We also plot the simulation results for the geometric channel model with L paths when $L = 5, 10$ and the Kronecker channel model. The Kronecker model is given as $\mathbf{H} = \mathbf{\Theta}_R^{1/2}\mathbf{H}_w(\mathbf{\Theta}_T^{1/2})^{\dagger}$ (similar for $\mathbf{G}$), where we set $\mathbf{H}_w$ to be i.i.d Rayleigh-faded with unit variance for each entry, $\mathbf{\Theta}_R$ to be the identity matrix, and $(\mathbf{\Theta}_T)_{i,j} = e^{-0.05d^2(i-j)^2}$, in which d is the antenna spacing normalized by the wavelength. We pick $d = 1/2$ for these two classes of channel models as suggested in [38]. The simulation results are obtained by examining whether each channel realization satisfies the condition in Proposition 1. For each point we run the simulation until approximately 50 failures are recorded. We can see that the probability of outage decreases rapidly as N

[3] An alternative approach is the Levenberg-Marquardt algorithm, in which $(J^TJ)^{-1}$ is replaced by $(J^TJ + \lambda I)^{-1}$ where λ is a non-negative damping factor and I is the identity matrix [35], [36].

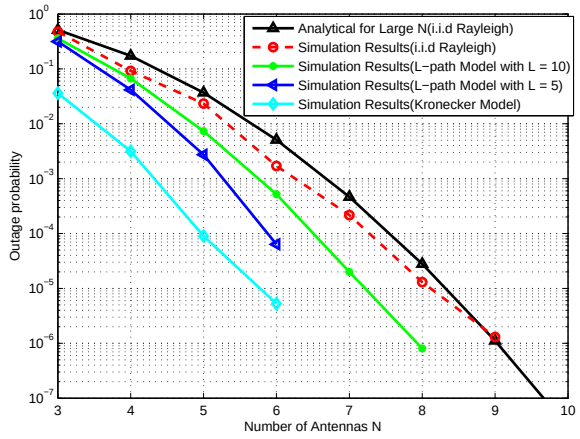

Figure 3. Simulation results and the analytic approximation using Proposition 2 for the probability of failing to find a PZF vector for the i.i.d Rayleigh fading channel, as well as the simulation results for the geometric channel model with L paths when $L = 5, 10$ and the Kronecker channel model.

increases. In fact for the i.i.d Rayleigh fading channel, we have $\Pr(\mathcal{E}) \approx 0.02$ for $N = 5$ and $\Pr(\mathcal{E}) \approx 0.002$ for $N = 6$. For the other channel models, the outage probabilities are even smaller, possibly due to the correlation among the channel gains from different transmit antennas to one receiver.

B. Case of $M = 1$ and $K \geq 1$

We numerically compare the performance of the SDP relaxation approach and the NLP approach studied in Section IV-A. We sample the matrix $\mathbf{H} \in \mathbb{C}^{K \times N}$ from the i.i.d Rayleigh fading channel and the geometric channel and use these two algorithms to find the beamforming vector $\mathbf{b}_{opti}$. To compare, we use the average noise variance at the receivers and the eavesdropper. The average noise leakage at the intended receivers is given by:

$$\frac{1}{NK} E_{\mathbf{H}}[||\mathbf{H}\mathbf{b}_{opti}||^2]P. \tag{32}$$

The average noise variance observed by the eavesdropper whose channel vector $\mathbf{g}$ is independent of $\mathbf{H}$ is

$$\frac{1}{N} E_{\mathbf{H},\mathbf{g}}[|\mathbf{g}^T\mathbf{b}_{opti}|^2]P. \tag{33}$$

We show in Figures 4 and 5 the SDP/NLP noise variances at the receivers/eavesdropper when $N = 2K + 3$ for the i.i.d Rayleigh fading and geometric model with $L = 10$ paths, respectively. This scaling of N is numerically found to yield a feasible PZF vector with high probability. Each point is simulated for 1000 times approximately. The blue line marked with triangles, which is very close to zero, is the noise power at the legitimate receiver when using the NLP approach. This shows that the algorithm succeeds in finding a PZF vector for most channel realizations. The blue line marked with circles is the noise power at the legitimate receiver from the SDP solver. We note that it increases with K, which indicates that this heuristic relaxation scheme is not effective as we increase the number of receivers. The red lines are the noise variance at the eavesdroppers. For both the schemes, the noise remains almost the constant on average at the eavesdropper as the number of antennas increases.

The probabilities of failing to find a PZF vector using these two approaches for the i.i.d Rayleigh fading channel are shown

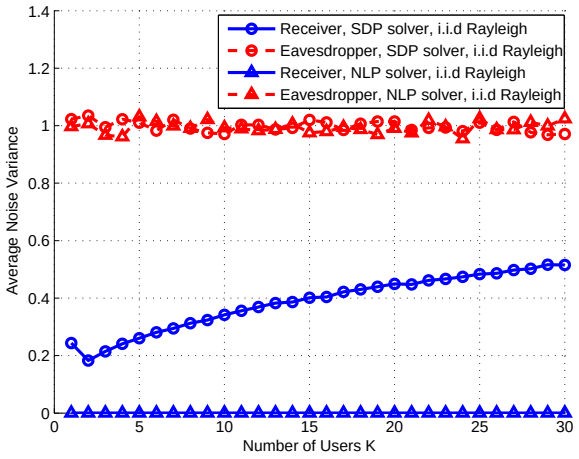

Figure 4. The average noise variance observed at the receiver and the eavesdropper for $N = 2K+3$ when $P = 1$ for i.i.d Rayleigh fading channel.

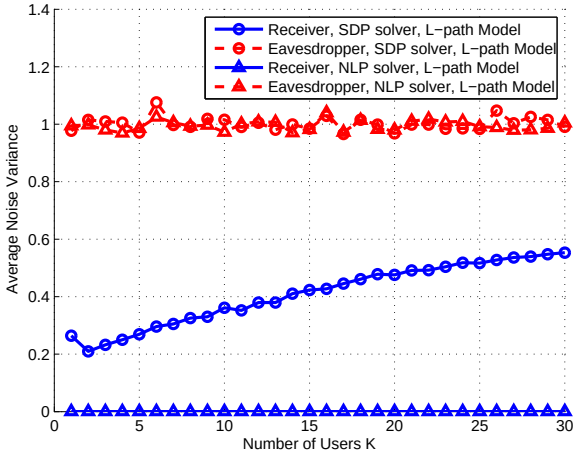

Figure 5. The average noise variance observed at the receiver and the eavesdropper for $N = 2K + 3$ when $P = 1$ for geometric channel with L paths when $L = 10$.

in Figure 6, for the cases when $N = 2K + 3$ as well as when $N = \lceil 2.05K \rceil$. It clearly shows that the SDP relaxation solver fails to find a solution to $\mathbf{H}\mathbf{b} = \mathbf{0}$ almost surely, while NLP solver maintains an outage probability of 10^{-2} for $N = 2K + 3$. The staircase-like shape of the outage probability for $N = \lceil 2.05K \rceil$ reveals that the level of outage probability is likely to be determined by the difference between N and $2K$. For relation $N = \lceil \gamma K \rceil$, the outage probability decreases as K increases when $\gamma > 2$. Ten thousand simulation trials has been done to get each outage probability in Figure 6.

Although we observe that the SDP solver works faster than the NLP solver, especially when K is large, its noise leakage to the legitimate receivers is unsatisfactory.

In Figure 7, we compare the outage probability using NLP approach for different channel models. In particular we show the performance over the i.i.d. Rayleigh fading channel model and the geometric channel model with $L = 5$ and $L = 10$ when $N = 2K + 2$ and $N = 2K + 3$. Interestingly even as we increase K, it appears that channel correlation might be beneficial in finding a PZF vector that simultaneously belongs

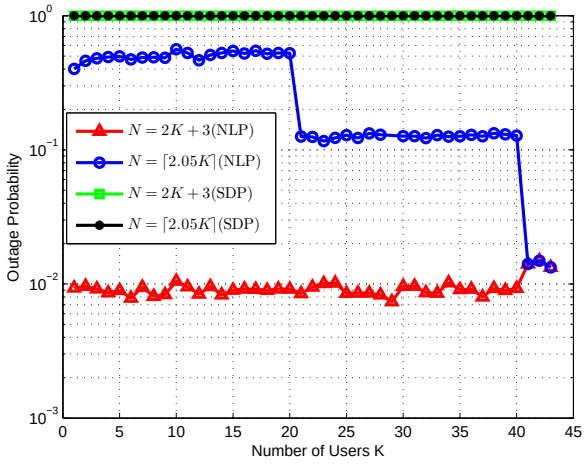

Figure 6. Outage probability for SDP relaxation and NLP approaches when $N = 2K + 3$ and $N = \lceil 2.05K \rceil$.

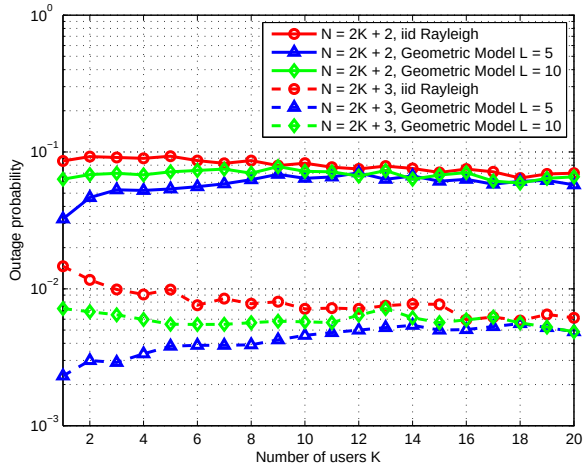

Figure 7. Outage probability for $N = 2K + 2$ and $N = 2K + 3$ for i.i.d Rayleigh fading channel and geometric channel with $L = 5$ and $L = 10$.

to null space of all channel vectors.

C. Case of $M \geq 1$ and $K \geq 1$

To measure how good beamforming vectors $\mathbf{b}_1, \cdots, \mathbf{b}_M$ are, let us introduce two criteria: *noise leakage* and *orthogonality loss*. Similar to (32), the average noise leakage at the intended receivers can be computed as

$$E_{\mathbf{H}}\Big[\frac{1}{NMK}\sum_{i=1}^{M}||\mathbf{H}\mathbf{b}_i||^2 P\Big]. \tag{34}$$

The orthogonality loss is defined as

$$E_{\mathbf{H}}\Big[\frac{2}{M(M-1)}\sum_{i=1}^{M}\sum_{j=i+1}^{M}|\langle \mathbf{b}_i, \mathbf{b}_j\rangle|\Big] \tag{35}$$

which checks the orthogonality of $\{\mathbf{b}_1, \cdots, \mathbf{b}_M\}$.

In Figure 8, we illustrate the simulation results for the cases of $(N, K) = (80, 20), (80, 25)$ and $(40, 10)$, where noise leakage and orthogonality loss are plotted in vesus M. The expectation over $\mathbf{H}$ in (34) and (35) is computed by averaging over sufficiently large number of i.i.d Rayleigh fading channel realizations in our simulation. The largest values of valid M

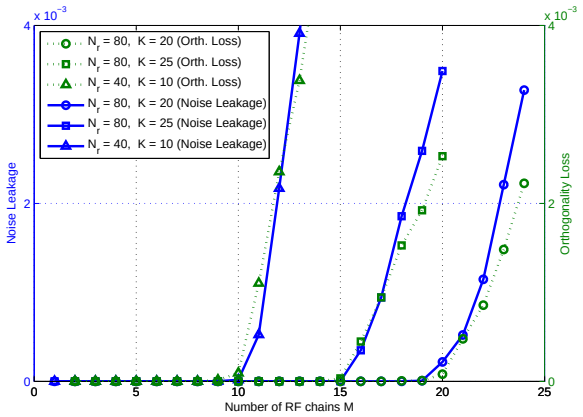

Figure 8. Average noise leakage and orthogonality loss of the beamforming vectors vs number of RF chains M for $(N, K) = (80, 20), (80, 25)$ and $(40, 10)$ when $P = 1$.

are shown to be 19, 15 and 10 respectively for these three cases.

Our simulation results indicates that the maximum M such that (7) and the orthogonality constraint hold is approximately $N/2 - K$. An intuitive explanation is as follows: for the M-th iteration in Algorithm 2, $\mathbf{H}_M^{\perp}$ has rank $N-K-(M-1)$, hence $2(N - K - (M - 1))$ real degrees of freedom, while having N quadratic constraints. To have extra degrees of freedom, we need $N > 2(N - K - (M - 1))$, which yields $M < N/2 - K + 1$.

VI. Secrecy Rate Analysis

In this section, we analyze the secrecy rate achieved by our algorithm. For brevity of results, we consider the single receiver case, i.e., $K = 1$. We assume one RF chain for message transmission and M RF chains for noise transmission. Accordingly, the input $\mathbf{x}$ has the following form:

$$\mathbf{x} = \sqrt{\frac{P}{N(M+1)}}\left(w\mathbf{u} + \sum_{k=1}^{M} s_k \mathbf{b}_k\right) \tag{36}$$

where w is the message symbol such that $\mathrm{E}[|w|^2] = 1$, $s_k \sim \mathcal{CN}(0, 1)$ is the k-th noise symbol, and $\mathbf{u}$ and $\mathbf{v}_k$ are the phase-only beamforming vectors for w and s_k, respectively. See Figure 9 for a system diagram corresponding to (36).

To evaluate the secrecy rate achieved by our algorithm, we first present a general secrecy rate expression in the following, which is for a general setting of J noise symbols and holds with or without phase-only constraints. The proof is provided in Appendix B.

Proposition 3 (Lower bound)**.** *Suppose that the transmitter sends one message symbol using beamforming vector* $\mathbf{v}$ *and* J *noise symbols using beamforming vectors* $\mathbf{v}_1, \cdots, \mathbf{v}_J$ *such that* $\|\mathbf{v}\|^2 + \sum_{k=1}^{J}\|\mathbf{v}_k\|^2 \leq P$*. Then, the following secrecy rate is achievable:*

$$R = \mathrm{E}_{\mathbf{h},\mathbf{G}}\left[\log\left(1 + \frac{|\mathbf{h}^T\mathbf{v}|^2}{1 + \sum_{k=1}^{J}|\mathbf{h}^T\mathbf{v}_k|^2}\right)\right.$$

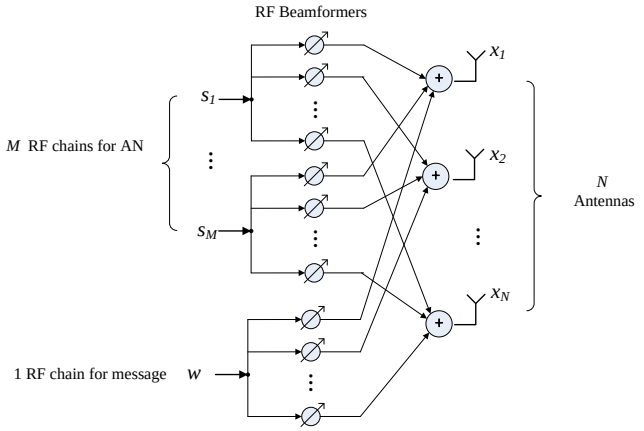

Figure 9. System model for the complete transmission scheme, with one RF chain for message transmission and M RF chains for AN transmission.

$$- \log \left(\frac{\det \left(I + \mathbf{G}\mathbf{v}\mathbf{v}^\dagger \mathbf{G}^\dagger + \sum_{k=1}^{J} \mathbf{G}\mathbf{v}_k \mathbf{v}_k^\dagger \mathbf{G}^\dagger \right)}{\det \left(I + \sum_{k=1}^{J} \mathbf{G}\mathbf{v}_k \mathbf{v}_k^\dagger \mathbf{G}^\dagger \right)} \right) \Bigg]. \quad (37)$$

The secrecy rate achieved by our algorithm is obtained by substituting $J \Leftarrow M, \mathbf{v} \Leftarrow \sqrt{\frac{P}{N(M+1)}}\mathbf{u}, \mathbf{v}_k \Leftarrow \sqrt{\frac{P}{N(M+1)}}\mathbf{b}_k$ in (37) of Proposition 3, where each element of $\mathbf{u}$ has the unit magnitude and the phase that compensates the phase of corresponding legitimate channel gain and $\mathbf{b}_1, \cdots, \mathbf{b}_M$ are constructed according to our algorithm in Section IV-B.

For comparison, we describe two other schemes in the following, which are also summarized in Table I. In each of the following schemes, the beamforming vector for message transmission is set as the same as in our scheme.

- Scheme based on [3]: It is shown in [3] that arbitrary J beamforming vectors can be synthesized by a hybrid Analog/Digital BF system with $2J$ RF chains. Inspired by this observation, in our proposed model, one can transmit $\lfloor M/2 \rfloor$ AN symbols using arbitrary beamforming vectors $\mathbf{v}_1, \cdots, \mathbf{v}_{\lfloor M/2 \rfloor}$ as long as the power constraint is satisfied. Hence, in this scheme, $\lfloor M/2 \rfloor$ AN symbols are transmitted by using orthogonal beamforming vectors $\mathbf{v}_1, \cdots, \mathbf{v}_{\lfloor M/2 \rfloor}$ in the null space of $\mathbf{h}$ (normalized to satisfy the power constraint) and its achievable secrecy rate can be evaluated from Proposition 3.
- Unconstrained scheme: As an upper bound, we consider the case without constraint on the number of RF chains. Hence, in this scheme, M AN symbols are transmitted using orthogonal beamforming vectors $\mathbf{v}_1, \cdots, \mathbf{v}_M$ in the null space of $\mathbf{h}$ (normalized to satisfy the power constraint) and its achievable secrecy rate can be evaluated from Proposition 3.

In the following subsection, we numerically plot the secrecy rates achieved by our scheme and the aforementioned two schemes.

A. Numerical Results

Numerical results are shown in Figures 10, 11, and 12 for the cases $(M, N, N_e) = (10, 20, 10)$, $(10, 30, 10)$ and $(10, 40, 10)$ respectively for the i.i.d Rayleigh fading channel.

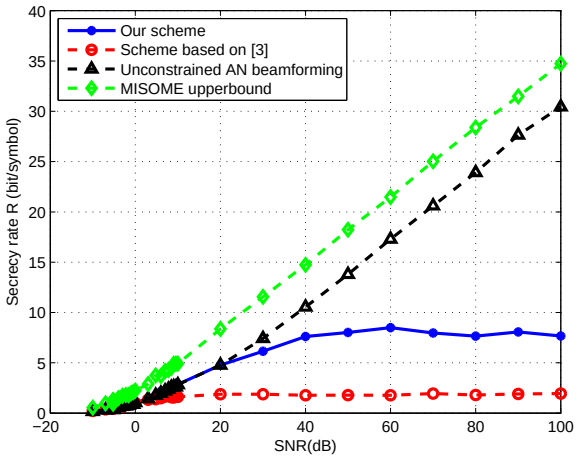

Figure 10. Secrecy rates when $M = 10$, $N = 20$ and $N_e = 10$ for i.i.d Rayleigh fading channel.

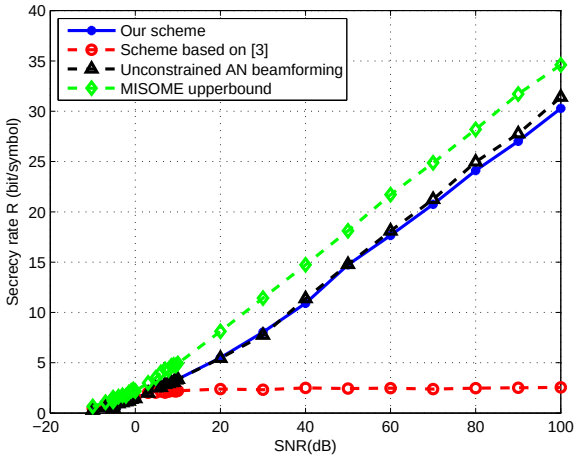

Figure 11. Secrecy rates when $M = 10$, $N = 30$ and $N_e = 10$ for i.i.d Rayleigh fading channel.

In addition to the achievable secrecy rates using the three schemes, we also plot the MISOME upper bound [5] for comparison. In contrast to the scheme based on [3] where only $\lfloor M/2 \rfloor$ AN symbols are sent, our proposed scheme achieves a considerably higher secrecy rate as it can transmit M AN symbols. The unconstrained AN beamforming scheme is shown by the dashed black lines with triangles. We remind that in Section V-C, it is numerically shown that $M < \frac{N}{2} - K + 1$ is needed to find the PZF vectors. For the current simulation setup, this condition becomes $N > 2M = 20$. Indeed, we can see that our PZF scheme has a non-negligible performance gap compared to the unconstrained AN beamforming scheme for $N = 20$, but the gap becomes negligible for $N = 30, 40$.

In addition, for the geometric channel model with $L = 10$ paths, the cases $(M, N, N_e) = (10, 20, 10)$ and $(10, 30, 10)$ are shown in Figures 13 and 14 respectively. Similarly, we can see that our PZF scheme has some loss for $N = 20$, but performs well for $N = 30$.

Table I
COMPARISON OF THREE AN SCHEMES WHERE ONE RF CHAIN IS USED FOR MESSAGE TRANSMISSION AND M RF CHAINS ARE USED FOR AN TRANSMISSION.

Scheme	Our scheme	Scheme based on [3]	Unconstrained scheme
Number of AN symbols	M (Number of RF chains minus one)	$\lfloor \frac{M}{2} \rfloor$	M
Property of AN beamforming vectors	Not guaranteed to be orthogonal and zero-forcing	Guaranteed to be orthogonal and zero-forcing	Guaranteed to be orthogonal and zero-forcing

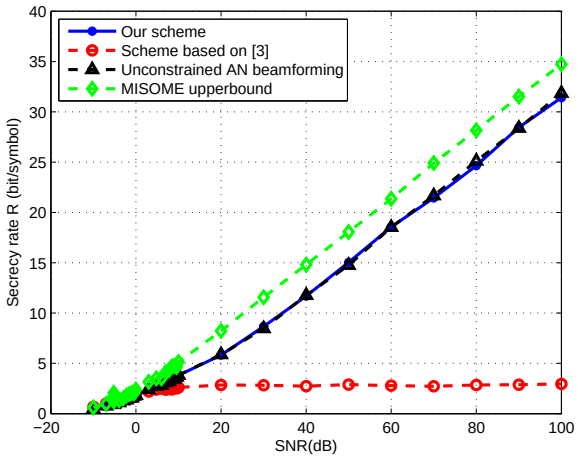

Figure 12. Secrecy rates when $M = 10$, $N = 40$ and $N_e = 10$ for i.i.d Rayleigh fading channel.

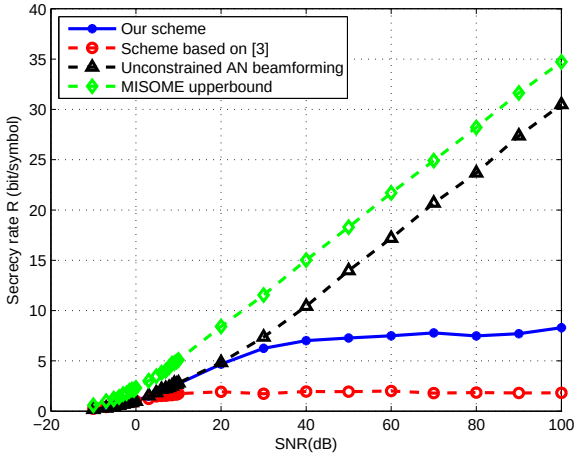

Figure 13. Secrecy rates when $M = 10$, $N = 20$ and $N_e = 10$ for geometric model with $L = 10$ paths.

VII. CONCLUSIONS

We study artificial noise transmission when each beamforming vector in the null space of the legitimate users' channel must satisfy a constant amplitude constraint on each element (weight). Our work is motivated by applications to emerging massive MIMO systems with a fixed number of RF chains and we refer to such beamforming vectors as PZF vectors. For the case of single RF chain and single receiver, we establish an analytical condition for the existence of such a PZF vector. For the general case we propose a numerical algorithm that uses the Gauss-Newton method to find multiple PZF vectors. Using simulations we note that for a system with single RF

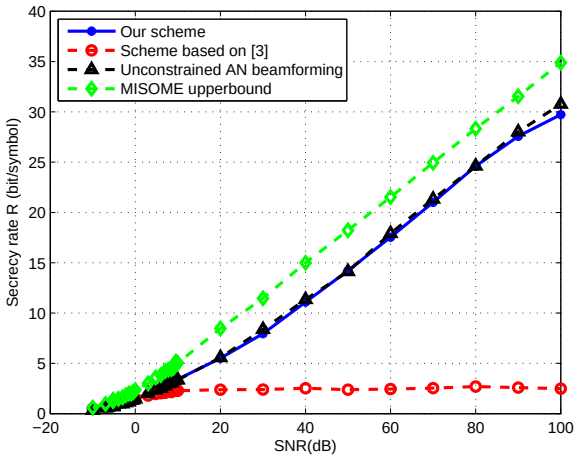

Figure 14. Secrecy rates when $M = 10$, $N = 30$ and $N_e = 10$ for geometric model with $L = 10$ paths.

chain, N antennas, and K users, a PZF vector exists with high probability if $N \geq 2K+3$. This is in sharp contrast to the case of unconstrained beamforming, where it suffices that $N > K$.

In terms of future work, it will be interesting to analytically prove the above threshold in the multiuser case. In another direction, it might be interesting to incorporate digital beamforming across AN symbols in the case of multiple RF chains, although off-the-shelf approaches do not appear to be sufficient in this regard.

APPENDIX A
PROOF OF PROPOSITION 1

Following the geometric interpretation in Section III it follows that one can view each $h_i \in \mathbb{C}$ as a 2-D vector in the real-imaginary plane. The condition

$$\sum_{i=1}^{N} h_i e^{j\phi_i} = 0 \tag{38}$$

is equivalent to rotating N sides of lengths $a_i = |h_i|$ to complete a N-sided polygon in 2-dimensions. We must show that a necessary and sufficient condition for such rotations $\{\phi_i\}$ to exist is that

$$a_i \leq \frac{1}{2}\sum_{j=1}^{N} a_j, \quad \forall i = 1, \ldots, N. \tag{39}$$

The necessity is immediate. If there exists a side, say h_1 that violates (39) then h_1 cannot be cancelled even when $h_2, \ldots, h_N$ are perfectly aligned in the direction opposite to h_1. Thus (38) cannot be satisfied.

The sufficiency part can be proved using mathematical induction. As discussed in Section III, the condition for $N = 3$ follows from the law of cosines.

Assume that the statement is true for $N-1$ and we are required to prove it for N. Given $a_i \leq \frac{1}{2}\sum_{j=1}^{N} a_j$, we can get that

$$a_{N-1} + a_N \geq 2a_i - \sum_{j=1}^{N-2} a_j, \quad \forall i = 1, \cdots, N-2$$

as well as

$$|a_N - a_{N-1}| \leq \sum_{j=1}^{N-2} a_j.$$

Combining these with the trivial inequalities:

$$a_{N-1} + a_N \leq |a_N - a_{N-1}|$$

and

$$a_i \leq \sum_{j=1}^{N-2} a_j, \quad \forall i = 1, \ldots, N-2$$

which hold since all a_j are non-negative, we have that for each $i \in \{1, \ldots, N-2\}$:

$$\min\{a_{N-1} + a_N, \sum_{j=1}^{N-2} a_j\}$$
$$\geq \max\{2a_i - \sum_{j=1}^{N-2} a_j, |a_N - a_{N-1}|\}.$$

Hence there exists d such that

$$d \leq \min\{a_{N-1} + a_N, \sum_{j=1}^{N-2} a_j\}$$
$$d \geq \max\{2a_i - \sum_{j=1}^{N-2} a_j, |a_N - a_{N-1}|\},$$

equivalently, we have

$$\sum_{j=1}^{N-2} a_j \geq d \geq 2a_i - \sum_{j=1}^{N-2} a_j, \ i = 1, ..., N-2 \quad (40)$$

$$a_{N-1} + a_N \geq d \geq |a_N - a_{N-1}|. \quad (41)$$

Based on the $N-1$ case, (40) infers that $(a_1, \cdots, a_{N-2}, d)$ can form a polygon. (41) infers that a triangle can be formed using (d, a_{N-1}, a_N). Sharing a common edge, the polygon with $N-1$ edges and the triangle can form a polygon with N edges. This completes the proof. ∎

APPENDIX B
PROOF OF PROPOSITION 3

The following secrecy rate is achievable from [39]

$$R = I(A; Y|\mathbf{h}) - I(A; \mathbf{Z}|\mathbf{G}) \quad (42)$$

for any distribution $f(a)f(\mathbf{x}|a, \mathbf{h})$ such that $\mathrm{E}[\|\mathbf{X}\|^2] \leq P$, by considering $f(\mathbf{x}|a, \mathbf{h})$ as a prefix channel. Then, if we substitute $A \sim \mathcal{CN}(0,1)$ and $\mathbf{X} = A\mathbf{v} + \sum_{k=1}^{J} s_k \mathbf{v}_k$ with $s_k \sim \mathcal{CN}(0,1)$ in (42), we have

$$\begin{aligned} R &= I(A; Y|\mathbf{h}) - I(A; \mathbf{Z}|\mathbf{G}) \quad (43) \\ &= I(A; A\mathbf{h}^T\mathbf{v} + \sum_{k=1}^{J} s_k \mathbf{h}^T\mathbf{v}_k + n|\mathbf{h}) \\ &\quad - I(A; A\mathbf{G}\mathbf{v} + \sum_{k=1}^{J} s_k \mathbf{G}\mathbf{v_k} + \mathbf{n}_e|\mathbf{G}) \quad (44) \\ &= h(A\mathbf{h}^T\mathbf{v} + \sum_{k=1}^{J} s_k \mathbf{h}^T\mathbf{v}_k + n|\mathbf{h}) \\ &\quad - h(\sum_{k=1}^{J} s_k \mathbf{h}^T\mathbf{v}_k + n|\mathbf{h}) \\ &\quad - h(A\mathbf{G}\mathbf{v} + \sum_{k=1}^{J} s_k \mathbf{G}\mathbf{v_k} + \mathbf{n}_e|\mathbf{G}) \\ &\quad + h(\sum_{k=1}^{J} s_k \mathbf{G}\mathbf{v_k} + \mathbf{n}_e|\mathbf{G}) \quad (45) \\ &= \mathrm{E}_{\mathbf{h}}\left[\log\left(1 + |\mathbf{h}^T\mathbf{v}|^2 + \sum_{k=1}^{J} |\mathbf{h}^T\mathbf{v}_k|^2\right)\right. \\ &\quad \left. - \log\left(1 + \sum_{k=1}^{J} |\mathbf{h}^T\mathbf{v}_k|^2\right)\right] \\ &\quad - \mathrm{E}_{\mathbf{G}}\left[\log\det\left(I + \mathbf{G}\mathbf{v}\mathbf{v}^\dagger\mathbf{G}^\dagger + \sum_{k=1}^{J} \mathbf{G}\mathbf{v}_k\mathbf{v}_k^\dagger\mathbf{G}^\dagger\right)\right. \\ &\quad \left. - \log\det\left(I + \sum_{k=1}^{J} \mathbf{G}\mathbf{v}_k\mathbf{v}_k^\dagger\mathbf{G}^\dagger\right)\right], \quad (46) \end{aligned}$$

which completes the proof. ∎

REFERENCES

[1] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 6, pp. 2180–2189, 2008.

[2] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 186–195, 2014.

[3] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale MIMO systems," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 2929–2933.

[4] A. Alkhateeb and R. W. Heath Jr, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *arXiv preprint arXiv:1510.00609*, 2015.

[5] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *Information Theory, IEEE Transactions on*, vol. 56, no. 7, pp. 3088–3104, 2010.

[6] X. Zhou and M. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 8, pp. 3831–3842, 2010.

[7] D. Ng, E. Lo, and R. Schober, "Energy-efficient resource allocation for secure OFDMA systems," *Vehicular Technology, IEEE Transactions on*, vol. 61, no. 6, pp. 2572–2585, 2012.

[8] M. Pei, J. Wei, K. Wong, and X. Wang, "Masked beamforming for multiuser MIMO wiretap channels with imperfect CSI," *IEEE Transactions on Wireless Communications,*, vol. 11, no. 2, pp. 544–549, 2012.

[9] J. Zhu, R. Schober, and V. Bhargava, "Secure transmission in multicell massive MIMO systems," *IEEE Transactions on Wireless Communications,*, vol. 13, no. 9, pp. 4766–4781, 2014.

[10] Q. Li and W. Ma, "Spatially selective artificial-noise aided transmit optimization for MISO multi-eves secrecy rate maximization," *Signal Processing, IEEE Transactions on*, vol. 61, no. 10, pp. 2704–2717, 2013.

[11] X. Zhang, M. McKay, X. Zhou, and R. Heath, "Artificial-noise-aided secure multi-antenna transmission with limited feedback," *Wireless Communications, IEEE Transactions on*, vol. 14, no. 5, pp. 2742–2754, 2015.

[12] H.-M. Wang, C. Wang, and D. W. K. Ng, "Artificial noise assisted secure transmission under training and feedback," *IEEE Trans. Signal Process.*, vol. 63, pp. 6285–6298, Dec. 2015.

[13] Y. O. Basciftci, C. E. Koksal, and A. Ashikhmin, "Securing massive MIMO at the physical layer," in *Communications and Network Security (CNS), 2015 IEEE Conference on*, Sept 2015, pp. 272–280.

[14] S. Lin, K. Huang, W. Luo, and Y. Zou, "Analysis of pilot contamination on the security performance of artificial noise in MIMO systems," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5.

[15] N. Yang, S. Yan, J. Yuan, R. Malaney, R. Subramanian, and I. Land, "Artificial noise: Transmission optimization in multi-input single-output wiretap channels," *IEEE Transactions on Communications*, vol. 63, no. 5, pp. 1771–1783, May 2015.

[16] A. Khina, Y. Kochman, and A. Khisti, "The gap to practical MIMO wiretap codes," in *Proc. International Zurich Seminar on Communications (IZS)*, Mar. 2016, p. 115.

[17] W. Shen, P. Ning, X. He, and H. Dai, "Ally friendly jamming: How to jam your enemy and maintain your own wireless connectivity at the same time," in *Security and Privacy (SP), 2013 IEEE Symposium on*. IEEE, 2013, pp. 174–188.

[18] S. Gollakota, H. Hassanieh, B. Ransford, D. Katabi, and K. Fu, "They can hear your heartbeats: non-invasive security for implantable medical devices," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 4, pp. 2–13, 2011.

[19] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "Interference assisted secret communication," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3153–3167, May 2011.

[20] L. Lai and H. E. Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4005–4019, Sept 2008.

[21] A. Khisti and D. Zhang, "Artificial-noise alignment for secure multicast using multiple antennas," *Communications Letters, IEEE*, vol. 17, no. 8, pp. 1568–1571, 2013.

[22] V. Venkateswaran and A. J. van der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4131–4143, Aug 2010.

[23] R. W. Heath Jr, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *arXiv preprint arXiv:1512.03007*, 2015.

[24] N. Valliappan, A. Lozano, and R. W. Heath, "Antenna subset modulation for secure millimeter-wave wireless communication," *Communications, IEEE Transactions on*, vol. 61, no. 8, pp. 3231–3245, 2013.

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

[26] ——, "Near-field direct antenna modulation," *IEEE Microwave Magazine*, vol. 10, no. 1, pp. 36–46, February 2009.

[27] M. P. Daly and J. T. Bernhard, "Directional modulation technique for phased arrays," *Antennas and Propagation, IEEE Transactions on*, vol. 57, no. 9, pp. 2633–2640, 2009.

[28] ——, "Beamsteering in pattern reconfigurable arrays using directional modulation," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 7, pp. 2259–2265, July 2010.

[29] Y. Ding and V. F. Fusco, "A vector approach for the analysis and synthesis of directional modulation transmitters," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 1, pp. 361–370, Jan 2014.

[30] R. De Francisco and D. Slock, "An optimized unitary beamforming technique for MIMO broadcast channels," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 3, pp. 990–1000, 2010.

[31] K. Huang, R. W. Heath Jr, and J. G. Andrews, "Space division multiple access with a sum feedback rate constraint," *Signal Processing, IEEE Transactions on*, vol. 55, no. 7, pp. 3879–3891, 2007.

[32] A. M. Tulino, A. Lozano, and S. Verdű, "Impact of antenna correlation on the capacity of multiantenna channels," *Information Theory, IEEE Transactions on*, vol. 51, no. 7, pp. 2491–2509, 2005.

[33] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[34] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*. Siam, 1970, vol. 30.

[35] K. Levenberg, "A method for the solution of certain non–linear problems in least squares," 1944.

[36] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[37] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Methods in Convex Programming*. Society for Industrial and Applied Mathematics, 1994.

[38] T.-S. Chu and L. J. Greenstein, "A semi-empirical representation of antenna diversity gain at cellular and pcs base stations," *Communications, IEEE Transactions on*, vol. 45, no. 6, pp. 644–646, 1997.

[39] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *Information Theory, IEEE Transactions on*, vol. 24, no. 3, pp. 339–348, 1978.