A Rotation-Based Method for Precoding in Gaussian MIMOME Channels

Xinliang Zhang, *Graduate Student Member, IEEE*, Yue Qi, *Graduate Student Member, IEEE*, and Mojtaba Vaezi, *Senior Member, IEEE*

Abstract—The problem of maximizing secrecy rate of multiple-input multiple-output multiple-eavesdropper (MIMOME) channels with arbitrary numbers of antennas at each node is studied in this paper. First, the optimization problem corresponding to the secrecy capacity of the MIMOME channel is converted to an equivalent optimization based on Givens rotations and eigenvalue decomposition of the covariance matrix. In this new formulation, precoder is a rotation matrix which results in a positive semi-definite (PSD) covariance matrix by construction. This removes the PSD matrix constraint and makes the problem easier to tackle. Next, a Broyden-Fletcher-Goldfarb-Shanno (BFGS)-based algorithm is developed to find the rotation and power allocation parameters. Further, the generalized singular value decomposition (GSVD)-based precoding is used to initialize this algorithm. The proposed rotation-BFGS method provides an efficient approach to find a near-optimal transmit strategy for the MIMOME channel and outperforms various state-of-the-art analytical and numerical methods. In particular, the rotation-BFGS precoding achieves higher secrecy rates than the celebrated GSVD precoding, with a reasonably higher computational complexity. Extensive numerical results elaborate on the effectiveness of the rotation-BFGS precoding. The new framework developed in this paper can be applied to a variety of similar problems in the context of multi-antenna channels with and without secrecy.

Index Terms—Physical layer security, MIMO wiretap channel, secrecy capacity, beamforming, precoding, covariance, rotation.

I. INTRODUCTION

AS A complement to higher-layer security measures, *physical layer security* has emerged as a significant technique for security in the lowest layer of communication, i.e., the physical layer. Founded on information-theoretic security, which is built on classical Shannon's notion of perfect secrecy, physical layer security can offer unbreakable security, unlike conventional secret-key-based cryptosystems. Physical layer security was laid in the 1970s in Wyner's seminal work on the *wiretap channel* [2] where the idea of secure communication based on the communication channel itself without using encryption keys was first introduced. In this work, Wyner proved that in a wiretap channel (a channel

in which a transmitter conveys information to a legitimate receiver in the presence of an eavesdropper) communication can be both robust to transmission errors (*reliable*) and confidential (*secure*), to a certain degree, provided that the legitimate user's channel is better than the eavesdropper's channel.[1] He established the capacity of the *degraded* wiretap channel. Later, Csiszar and Korner [4] generalized this result to arbitrary, not necessarily degraded, wiretap channels.

In the past decades, physical layer security has been applied to enhance the classical wiretap channels (e.g., by including more realistic assumptions) and to study advanced wiretap channels (e.g., quantum communication [5], [6]). Particularly, as multiple-input multiple-output (MIMO) networks continue to flourish worldwide, a significant effort has been made to study the MIMO wiretap channel which allows for the exploitation of space/time/user dimensions of wireless channels for secure communications. Specifically, secrecy capacity of Gaussian multiple-input multiple-output multiple-eavesdropper (MIMOME) channels under an average total power constraint was established independently in [7]–[9]. The capacity result is abstracted as an optimization problem over input covariance matrix. This problem is non-convex and its optimal solution is known only for limited settings [10]–[14].

Among notable sub-optimal solutions that can be applied to the MIMOME channel is the *generalized singular value decomposition* (GSVD)-based precoding [15]. GSVD-based precoding decomposes transmitted channel matrices into several parallel subchannels and confidential information is transmitted over subchannels where the legitimate user is stronger than the eavesdropper. This method gives a closed-form solution for achievable secrecy rate which is relatively fast and is asymptotically optimal at high signal-to-noise ratios (SNRs). However, its performance is not good at certain settings, e.g., when the eavesdropper has a single antenna while other nodes have multiple antennas [14]. Another important sub-optimal solution is Li *et al.'s* alternating optimization and water filling (AOWF) algorithm [16] which alternates the original optimization problem to a convex problem and finds the corresponding Lagrange multipliers in an iterative manner. AOWF is more computationally expensive than GSVD-based precoding but it can provide a better secrecy rate in some settings. The performance of this method also varies depending

[1]Later in the 1990s, Maurer proved that secret key generation through public communication over an insecure yet authenticated channel is possible even when a legitimate user has a worse channel than an eavesdropper [3].

on the number of antennas at different nodes. For example, its performance is not as good as the GSVD-based precoding when the number of antennas at the eavesdropper is greater than that of the transmitter. There are also other numerical solutions for this optimization problem [17]–[19]. Specifically, in [17], a barrier method based iterative algorithm is tailored to obtain global optimal with guaranteed convergence. However, it is not straightforward to obtain the barrier parameter and this involves a challenging optimization problem. In addition, this solution is devised based on an upper bound custom-made for the MIMOME channel and its extension to other related problems is not straightforward.

Recently, based on a trigonometric parameterization of the covariance matrix, a closed-form solution for optimal precoding and power allocation of the MIMOME channel with two transmit antennas was obtained in [14], [20]. This approach in finding the optimal covariance matrix is completely different from existing linear beamforming methods. It does not require degradedness condition of [12] and [21], and thus provides the optimal solution for both full-rank and rank-deficient cases in one shot. The above beamforming and power allocation schemes are, however, limited to two transmit antenna cases, and the optimal transmit covariance matrix is still open in general. Givens rotation has been previously investigated for quantizer design in MIMO broadcast channels [22]. Later, Givens rotation is used as a beamformer to maximize sum-rate under average signal-to-noise ratio (SNR) with limited feedback in multiuser multiple-input single-output (MISO) channels [23]. However, Givens rotation theory is the first time to be applied on MIMO wiretap channel to the authors' knowledge.

In this paper, we use Givens rotation matrices to generalize the approach of [20] to arbitrary numbers of antennas at each node, and introduce a new method for precoding and power allocation in the MIMOME channel. In this approach, without loss of generality, the precoding matrix is formed using a rotation matrix. Then, the covariance matrix can be seen as an operator to appropriately stretch (by power allocation) and rotate (by precoding) the input symbols to form a transmit signal that best fits the channels of the legitimate user and eavesdropper. The capacity expression is then transformed into optimizing rotation angles and power allocation parameters. This problem is very different from the original problem, which requires exploring new techniques to solve it. One advantage of the new problem formulation is that it converts the matrix covariance constraint (symmetric and positive semi-definiteness (PSD)) into a set of linear constraints and thus simplifies the optimization problem. We then provide a numerical solution for the new optimization problem and show that it outperforms the existing methods in various antenna settings. The method in [14] can be seen as using two-dimension Givens rotation for precoding.

A. Contributions

The main contributions of this paper are listed below:
- We use a rotation modeling method for the parameterization of the covariance matrix. This parameterization gives a new representation of the capacity expression for the MIMOME channel for which more efficient solutions can be exploited. Particularly, the PSD constraint on the covariance matrix is removed because in the proposed method the covariance matrix is PSD by construction.
- For $n_t = 1$ and $n_t = 2$, the proposed scheme reduces to those in [11] and [14], respectively, for which closed-form solutions are known. For $n_t \geq 3$, finding a closed-form solution is still challenging. In such a case, we introduce a Broyden-Fletcher-Goldfarb-Shanno (BFGS)-based method to iteratively solve the rotation and power allocation parameters. We name this approach rotation-BFGS method in which a rectifier is designed to remove the constraints. Numerical results in different antenna settings confirm that the proposed scheme works better than the well-known GSVD and AOWF. Specifically, the proposed approach outperforms GSVD when $n_e < n_t$, and AOWF approach when $n_e \geq n_t$, where n_e is the number of antennas at the eavesdropper. Particularly, the gap between the proposed and GSVD-based methods is remarkably high when the eavesdropper has a single antenna.
- To improve the computational complexity of the proposed algorithm (by reducing the number of iterations), we develop an algorithm to exploit GSVD as an initialization for our rotation-BFGS method. This initialization improves the results and reduces the computational complexity as it reduces the number of iterations in the optimization problem.

B. Other Related Works

An interesting aspect of the proposed approach is its generality and its great potential for extension to other related problems. The MIMOME channel has turned out as a fundamental tool for the study of physical-layer security in many other related problems throughout the past decade. Many solutions developed for the MIMOME has appeared to be instrumental in designing transmit strategies that maximize the secrecy rate of extensions of this basic channel model to MIMO channels with multiple eavesdroppers [16], secure relaying [24], [25], ergodic secrecy rate [26], finite alphabet signaling [27]–[29], artificial noise [30] and cooperative jammer [31], among others. The rotation method is also applicable to any problem that can be cast as an optimization problem over a covariance matrix. Therefore, it is worth studying the optimal covariance matrix of the MIMOME channel as a general tool for physical-layer security in various MIMO settings.

C. Organizations and Notations

The remainder of this paper is organized as follows. Section II describes the system model and related works. Section III introduces and elaborates on the Givens rotation and reformulates the secrecy capacity for the MIMOME channel. Section IV, details a rotation-BFGS based algorithm to optimize the achievable secrecy rate. In Section V, numerous simulation results are carried out to demonstrate the effectiveness of the proposed method. Section VI draws the conclusion.
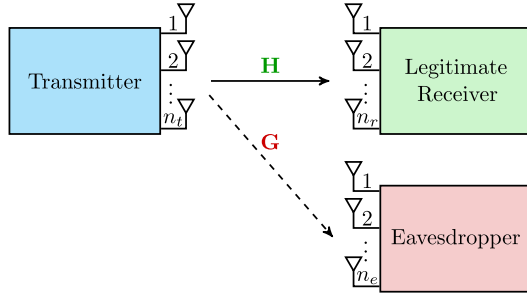
Fig. 1. The MIMOME channel with n_t, n_r, and n_e antennas at the transmitter, legitimate receiver, and eavesdropper.

Fig. 2. The structure of linear precoding and power allocation. $s_1, \ldots, s_{n_t}$ are the independent input symbols, $\lambda_1, \ldots, \lambda_{n_t}$ are their corresponding allocated powers, and $x_1, \ldots, x_{n_t}$ are the transmitted signals.

Notations: Bold lowercase letters denote column vectors and bold uppercase letters denote matrices. $|x|$ and $\log_2(x)$ denote the absolute value and the binary logarithm of the scalar x. $\mathbf{A}(i,j)$ denotes the entry (i,j) of matrix $\mathbf{A}$. Besides, $(\mathbf{A})^T$, $\mathrm{tr}(\mathbf{A})$, $|\mathbf{A}|$, and $\mathrm{diag}(\mathbf{A})$ are the transpose, trace, determinate, and diagonal of the matrix $\mathbf{A}$. $[\mathbf{A}]^+$ replaces the negative elements with zeros. $\mathbf{I}_n$ is a $n \times n$ identity matrix. And $E\{\cdot\}$ is the expectation of random variables.

II. SYSTEM MODEL AND RELATED WORKS

A. System Model

We consider a MIMOME channel with n_t antennas at the transmitter, n_r antennas at the receiver, and n_e antennas at the eavesdropper, as depicted in Fig. 1. The transmitter knows the perfect channel state information (CSI) of both users.[2] The received signals at the legitimate receiver and the eavesdropper can be, respectively, expressed as

$$\mathbf{y}_r = \mathbf{H}\mathbf{x} + \mathbf{w}_r, \tag{1a}$$

$$\mathbf{y}_e = \mathbf{G}\mathbf{x} + \mathbf{w}_e, \tag{1b}$$

in which $\mathbf{H} \in \mathbb{R}^{n_r \times n_t}$ and $\mathbf{G} \in \mathbb{R}^{n_e \times n_t}$ are the channels corresponding to the receiver and eavesdropper, $\mathbf{x} \in \mathbb{R}^{n_t}$ is the transmitted signal, and $\mathbf{w}_r \in \mathbb{R}^{n_r}$ and $\mathbf{w}_e \in \mathbb{R}^{n_e}$ are independent and identically distributed (i.i.d) Gaussian noises with zero means and identity covariance matrices. A representation of secrecy capacity is given by [9]

$$\text{(P1)} \quad \mathcal{C}_s = \max_{\mathbf{Q} \succeq \mathbf{0}, \mathrm{tr}(\mathbf{Q}) \leq P_t} \frac{1}{2} \log_2 \frac{|\mathbf{I}_{n_r} + \mathbf{H}\mathbf{Q}\mathbf{H}^T|}{|\mathbf{I}_{n_e} + \mathbf{G}\mathbf{Q}\mathbf{G}^T|}. \tag{2}$$

Based on Slyvester's determinant theorem, i.e., $\det(\mathbf{I}+\mathbf{X}\mathbf{Y}) = \det(\mathbf{I} + \mathbf{Y}\mathbf{X})$, (2) can also be rewritten as

$$\mathcal{C}_s = \max_{\mathbf{Q} \succeq \mathbf{0}, \mathrm{tr}(\mathbf{Q}) \leq P_t} \frac{1}{2} \log_2 \frac{|\mathbf{I}_{n_t} + \mathbf{H}^T\mathbf{H}\mathbf{Q}|}{|\mathbf{I}_{n_t} + \mathbf{G}^T\mathbf{G}\mathbf{Q}|}, \tag{3}$$

where $\mathbf{Q} = E\{\mathbf{x}\mathbf{x}^T\} \in \mathbb{R}^{n_t \times n_t}$ is the covariance matrix of the channel input $\mathbf{x}$ and P_t is the total transmit power. $\mathbf{Q}$ is symmetric and positive semi-definite (PSD) by definition.

---

[2] A perfect CSI is assumed, as we are deriving the theoretical limits. This may provide an upper bound in terms of achievable secrecy rates. The method we are developing in this paper is, however, applicable to the case with imperfect CSI. In our future works, we will relax this idealized assumption and consider the practical scenarios with imperfect CSI.
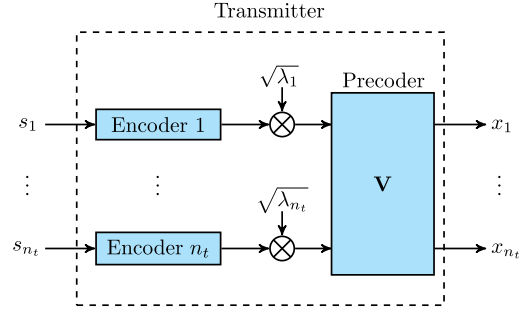
The architecture of the linear precoding and power allocation is depicted in Fig. 2. In this figure, $s_1, \ldots, s_{n_t}$ are input symbols which are i.i.d Gaussian random variables with zero means and unit variances, λ_is are power allocation coefficients, $\mathbf{V}$ is the precoding matrix, and $\mathbf{x} = [x_1, \ldots, x_{n_t}]^T$ is the transmit vector whose covariance is $\mathbf{Q}$.

B. Existing Results

The optimal transmission over the MIMOME channel is still an open problem in general. However, there are a number of notable analytical results for special numbers of antennas as well as numerical results as listed below.

1) Analytical Solutions: An analytical capacity-achieving covariance matrix is known only for special cases. These are limited to:

- $n_t = 1$: this is the single-input multiple-output (SIMO) case in which $\mathbf{Q}$ is a scalar and the optimal solution is either P_t or 0 [11].
- $n_r = 1$: the so-called multiple-input single-output multiple-eavesdropper (MISOME) channel in which *generalized eigenvalue decomposition* of $\mathbf{H}$ and $\mathbf{G}$ achieves the capacity [32].
- $n_t = 2$, $n_r = 2$, and $n_e = 1$: the optimization problem is shown to be the Rayleigh quotient and optimal signaling, which is the maximum eigenvalue of this problem, is unit-rank [10].
- $n_t = 2$: in which the secrecy capacity is obtained by modeling the covariance matrix as a 2×2 rotation matrix [14], [20].
- $\mathbf{Q}$ is full-rank (which implies $\mathbf{H}^T\mathbf{H} - \mathbf{G}^T\mathbf{G} \succ \mathbf{0}$) and also P_t is greater than a certain threshold [12], [21]: in this case the problem is convex and Karush–Kuhn–Tucker (KKT) conditions are used to find the optimal $\mathbf{Q}$.

  It is worth noting that, as n_t grows, few channel realizations satisfy the above conditions. For $n_t = 3$, $n_r = 3$, and $n_e = 1$, for example, the probability of having a full-rank solution is less than 18.2%.[3] This value decreases when n_e goes up. Therefore, it can be said that an analytical solution for the MIMOME channel is still an open problem in many practical cases.

---

[3] This is obtained by Monte Carlo experiments with 10^6 trails where $\mathbf{H}$ and $\mathbf{G}$ have the same distributions.

2) Suboptimal Analytical Solution: For a general MIMOME channel, a sub-optimal solution can be obtained using GSVD-based beamforming [15]. By applying GSVD on $\mathbf{H}$ and $\mathbf{G}$, the optimization problem (3) is simplified to a set of parallel non-interfering channels whose optimal power allocation can be obtained using Lagrange multiplier and KKT conditions. Since parallelization using GSVD does not necessarily convert this problem into an equivalent one, GSVD-based beamforming is not the optimal solution in general. It is not even close enough to the capacity in some cases. For example, GSVD-based beamforming achieves less than 70% of the capacity when $n_t = 3$, $n_r = 2$, and $n_e = 1$.

3) Numerical Solutions: There are still important cases of the MIMOME for which optimal $\mathbf{Q}$ is unknown. Due to the intractability of the problem in an analytical form, numerical solutions have been developed to tackle this problem. The AOWF [16], which is computationally efficient to implement, is one of them. Despite its effectiveness in many cases, AOWF experiences problems when n_e is greater than n_t, for example, which is caused by a failure in finding an optimal Lagrange multiplier. We modify this issue in this paper, as we will see later in Section V. The price is a higher time consumption in the modified approach.

In the next section, we apply a rotation-based model for the covariance matrix $\mathbf{Q}$ which is a generalization of the solution in [14], [16], from $n_t = 2$ to any arbitrary n_t. This model is then used to find transmit signaling that can be used to achieve secrecy capacity of the MIMOME channel regardless of the number of antennas at different nodes.

III. A ROTATION MODELING OF THE PROBLEM

In the following subsections, we further model matrix $\mathbf{V}$ using the rotation matrix related method by reviewing $n_t = 2$ [20] first and then generalizing it to an arbitrary n_t with proof.

The covariance matrix $\mathbf{Q}$ can be eigendecomposed as

$$\mathbf{Q} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T, \tag{4}$$

in which $\boldsymbol{\Lambda}$ is a diagonal matrix and its diagonal elements are the eigenvalues of $\mathbf{Q}$, which are real and non-negative, i.e.,

$$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_i), \ \lambda_i \geq 0, \quad i = 1, 2, \ldots, n_t, \tag{5}$$

and the total power constraint $\mathrm{tr}(\mathbf{Q}) \leq P_t$ will be equivalent to

$$\sum_{i=1}^{n_t} \lambda_i \leq P_t. \tag{6}$$

Also, $\mathbf{V} \in \mathbb{R}^{n_t \times n_t}$ is the matrix composed of n_t corresponding eigenvectors of $\mathbf{Q}$ in $\mathbb{R}^{n_t}$ vector space. Since $\mathbf{Q}$ is symmetric, the matrix $\mathbf{V}$ is *orthonormal*.

With this, an immediate implication of the above decomposition is that linear precoding can achieve the capacity of the MIMOME channel.

A. Rotation Modeling for $n_t = 2$

The capacity region of MIMOME channel with two transmit antennas has been established in [14], [20], using a rotation

matrix in two-dimensional (2D) space. In this case, the eigenvalue matrix $\boldsymbol{\Lambda}$ can be written as

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \tag{7}$$

and, without loss of generality, the eigenvectors can be written as the following orthonormal (rotation) matrix

$$\mathbf{V} = \mathbf{V}_{12} \triangleq \begin{bmatrix} \cos\theta_{12} & -\sin\theta_{12} \\ \sin\theta_{12} & \cos\theta_{12} \end{bmatrix}. \tag{8}$$

The rotation angle θ_{12} in rotation matrix $\mathbf{V}_{12}$ corresponds to the rotation from the direction of the standard basis (unit vector) $\mathbf{e}_1 = (1,0)^T$ to the standard basis $\mathbf{e}_2 = (0,1)^T$ in $\mathbb{R}^2$ vector space. This rotation is achieved on the plane defined by $\mathbf{e}_1$ and $\mathbf{e}_2$. Then, the covariance matrix $\mathbf{Q}$ can be built using three parameters: the two non-negative eigenvalues λ_1 and λ_2, and the rotation angle θ_{12}. These parameters are complete to represent any arbitrary 2×2 covariance matrix as in (4). Then the original optimization problem (3) can be equivalently converted to

$$\mathcal{C}_s = \max_{\lambda_1, \lambda_2, \theta_{12}} \frac{1}{2} \log_2 \frac{|\mathbf{I}_{n_t} + \mathbf{H}^T\mathbf{H}\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T|}{|\mathbf{I}_{n_t} + \mathbf{G}^T\mathbf{G}\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T|}, \tag{9a}$$
$$\mathrm{s.\,t.} \quad \lambda_1 \geq 0, \ \lambda_2 \geq 0, \ \lambda_1 + \lambda_2 \leq P_t. \tag{9b}$$

In light of this modeling, an analytical solution for optimal precoding matrix and power allocation scheme are obtained in [20] by finding θ_{12}, λ_1 and λ_2. In the next subsection, we extend this method to the cases for an arbitrary n_t.

B. Generalization to an Arbitrary n_t

To generalize the rotation modeling method to an arbitrary $n_t \times n_t$ covariance matrix, $\boldsymbol{\Lambda} \in \mathbb{R}^{n_t \times n_t}$ is a diagonal matrix with non-negative elements $\boldsymbol{\Lambda}(i,i) \triangleq \lambda_i$. $\mathbf{V}$ is a rotation matrix in $\mathbb{R}^{n_t \times n_t}$ vector space which can be obtained by

$$\mathbf{V} = \prod_{i=1}^{n_t-1} \prod_{j=i+1}^{n_t} \mathbf{V}_{ij}, \tag{10}$$

in which the basic rotation matrix $\mathbf{V}_{ij}$ is the Givens matrix [33] defined as

$$\mathbf{V}_{ij} = \begin{bmatrix} 1 & \cdots & & & & \cdots & 0 \\ \vdots & \ddots & & & & & \vdots \\ & & v_{ii} & \cdots & v_{ij} & & \\ & & \vdots & \ddots & \vdots & & \\ & & v_{ji} & \cdots & v_{jj} & & \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & & & & \cdots & 1 \end{bmatrix}, \tag{11}$$

and

$$\begin{bmatrix} v_{ii} & v_{ij} \\ v_{ji} & v_{jj} \end{bmatrix} = \begin{bmatrix} \cos\theta_{ij} & -\sin\theta_{ij} \\ \sin\theta_{ij} & \cos\theta_{ij} \end{bmatrix}. \tag{12}$$

$\mathbf{V}_{ij}$ represents a rotation from the ith standard basis to the jth standard basis in $\mathbb{R}^{n_t}$ vector space with a rotation angle θ_{ij}. That is, we show that an arbitrary orthogonal matrix $\mathbf{V}$ can be represented by (10). Further, an arbitrary covariance

matrix $\mathbf{Q} \in \mathbb{R}^{n_t \times n_t}$ can be represented by n_t non-negative eigenvalues and $\frac{1}{2}n_t(n_t - 1)$ rotation angles.

It should be noted that the order of multiplication in (10) is not unique, and a different order will lead to different rotation angles θ_{ij}. In this paper, without loss of generality, we use the order definition in (10).

Lemma 1: To reach the secrecy capacity of the MIMOME with $n_t \geq 2$, it is sufficient to use a PSD diagonal $\mathbf{\Lambda}$ and the rotation matrix $\mathbf{V}$ given in (10) to generate the input covariance matrix $\mathbf{Q} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$.

Proof: First, we prove that $\mathbf{Q} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ is a covariance matrix. It is straightforward to check that $\mathbf{V}$ in (10) is an orthonormal matrix, i.e., $\mathbf{V}\mathbf{V}^T = \mathbf{I}$. Since diagonal elements of $\mathbf{\Lambda}$ are non-negative, $\mathbf{Q}$ is symmetric and PSD, i.e., a covariance matrix.

Next, we prove that an arbitrary covariance matrix $\mathbf{Q}$ can be written as $\mathbf{Q} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ while $\mathbf{V}$ is defined as (10). It suffices to find θ_{ij} such that (13) holds for a given orthonormal $\mathbf{V}$

$$\left(\prod_{i=1}^{n_t-1} \prod_{j=i+1}^{n_t} \mathbf{V}_{ij} \right)^T \mathbf{V} = \mathbf{I}. \tag{13}$$

This process is carried out in Algorithm 1. In fact, (13) applies a series of Givens rotation on $\mathbf{V}$. They keep the relative orthogonality and rotates $\mathbf{V}$ to identity matrix $\mathbf{I}$.

To better appreciate this, we note that if we expand the product of matrices on the left side of $\mathbf{V}$ in (13), we see that $\mathbf{V}$ is initially multiplied with $\mathbf{V}_{12}^T$. Then, the corresponding rotation angle θ_{12} can be chosen to set the entry $(2, 1)$ of $(\mathbf{V}_{12}\mathbf{V})$ to zero. Next, $\mathbf{V}_{13}^T$ is multiplied to the new $\mathbf{V}$ and θ_{13} can be chosen to set the entry $(3, 1)$ to zero. This process continues until the last element under the main diagonal of $\mathbf{V}$, i.e., the entry (n_t, n_t-1), becomes zero. We note that, since $\mathbf{V}$ is orthonormal, the upper triangle also will be zero throughout this process. That is, the left side of (13) becomes an identity matrix. This proves Lemma 1. The rotation angles θ_{ij} can be obtained by Algorithm 1 which is a generalization of Algorithm 5.1.3 of [33] for vectors in $\mathbb{R}^{n_t}$, where $\mathrm{atan2}(\cdot, \cdot)$ is the four-quadrant inverse tangent denoted by `atan2` in MATLAB.

In certain cases, the eigenvalue decomposition of $\mathbf{Q}$ may give an improper rotation matrix [34] whose determinate is -1 and cannot be converted to an identical matrix by rotation. As a result, $\mathbf{V}_{ij}$ for (13) does not exist. To deal with this issue, the improper rotation matrix can be converted to a proper rotation matrix by exchanging arbitrary two eigenvectors of $\mathbf{V}$ and the corresponding eigenvalues of $\mathbf{\Lambda}$. As an example, we can define $\mathbf{V}' = \mathbf{V}\mathbf{I}'$ and $\mathbf{\Lambda}' = \mathbf{I}'^T \mathbf{\Lambda}\mathbf{I}'$, where

$$\mathbf{I}' = \begin{bmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n_t \times n_t}. \tag{14}$$

This specific $\mathbf{I}$ exchanges the last two eigenvectors and eigenvalues of $\mathbf{Q}$. Since $\mathbf{I}'\mathbf{I}'^T = \mathbf{I}$, it is clear that

$$\mathbf{V}'\mathbf{\Lambda}'\mathbf{V}'^T = (\mathbf{V}\mathbf{I}')(\mathbf{I}'^T\mathbf{\Lambda}\mathbf{I}')(\mathbf{V}\mathbf{I}')^T = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = \mathbf{Q}. \tag{15}$$

Algorithm 1 Rotation Angles Solution

1: Initialize $[\mathbf{V}, \mathbf{\Lambda}] = \mathrm{eig}(\mathbf{Q})$, $i = 1$;
2: **if** $\det(\mathbf{V}) = -1$ **then**
3:    Exchange first two columns of $\mathbf{V}$;
4:    Exchange first two values on diagonal of $\mathbf{\Lambda}$;
5: **end if**
6: **while** $i \leq (n_t - 1)$ **do**
7:    $j = i + 1$;
8:    **while** $j \leq n_t$ **do**
9:       $\theta_{ij} = -\mathrm{atan2}(-\mathbf{V}(j, i), \mathbf{V}(i, i))$;
10:      $\mathbf{V}_{\mathrm{rot}} = \mathbf{I}_{n_t}$;
11:      $\mathbf{V}_{\mathrm{rot}}(i, i) = \mathbf{V}_{\mathrm{rot}}(j, j) = \cos\theta_{ij}$;
12:      $\mathbf{V}_{\mathrm{rot}}(j, i) = -\mathbf{V}_{\mathrm{rot}}(i, j) = \sin\theta_{ij}$;
13:      $\mathbf{V} = \mathbf{V}_{\mathrm{rot}}\mathbf{V}$;
14:      $j = j + 1$;
15:    **end while**
16:    $i = i + 1$;
17: **end while**
18: Output θ_{ij}, $\forall 1 \leq i < j \leq n_t$.

In such a case, $\mathbf{V}'$ is a rotation matrix and the $\mathbf{Q}$ will remain the same. This completes the proof of Lemma 1. $\square$

Lemma 1 shows that any covariance matrix can be formed using a rotation matrix and a diagonal power allocation matrix. This is useful in many optimization problems, including that of the MIMOME channel in (3), as it removes the PSD constraint of the covariance matrix (i.e., $\mathbf{Q} \succeq \mathbf{0}$). Instead, we will have a set of linear constraints to make sure that the diagonal elements of the power allocation matrix are non-negative and their sum is not greater than P_t.

Specifically, for any n_t, the optimization problem (3) is equivalently reformulated as

$$\text{(P2)} \quad \mathcal{C}_s = \max_{\mathbf{\lambda}, \mathbf{\theta}} \frac{1}{2}\log_2 \frac{|\mathbf{I}_{n_t} + \mathbf{H}^T\mathbf{H}\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T|}{|\mathbf{I}_{n_t} + \mathbf{G}^T\mathbf{G}\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T|}, \tag{16a}$$

$$\text{s.t.} \quad \sum_{i=1}^{n_t} \lambda_i \leq P_t, \tag{16b}$$

$$\lambda_i \geq 0, i \in \{1, \ldots, n_t\}, \tag{16c}$$

in which we have defined

$$\mathbf{\lambda} \triangleq \{\lambda_i\}, \ 1 \leq i \leq n_t, \tag{17a}$$

$$\mathbf{\theta} \triangleq \{\theta_{ij}\}, \ 1 \leq i < j \leq n_t. \tag{17b}$$

as the compact form of the parameters of (10)-(12).[4]

Different from (3), this new representation replaces the constraint that $\mathbf{Q}$ is symmetric and PSD by linear constraints on $\mathbf{\lambda}$ only while rotation angles can take any number, i.e., $\mathbf{\theta} \in \mathbb{R}$. Then, numerical methods can be applied to optimize the parameters $\mathbf{\theta}$ and $\mathbf{\lambda}$ to obtain the optimal secrecy rate. For $\mathbf{Q}$

---

[4]It is worth mentioning that Givens rotation method applied in many problems can finally boil down to finding the optimal rotation parameters. However, the application of the rotation method is different. In [23], only the optimal unitary beamforming without power allocation is evaluated because it constrained the precoding matrices to be unitary for limited feedback. On the other hand, the optimization approach is in different manners. In their iterative algorithm, the rotation angles of each covariance matrix are obtained one by one. In our paper, these parameters can be updated vector-wised once.
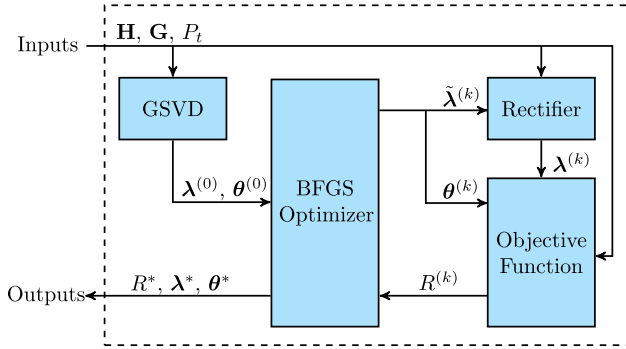
Fig. 3.   The system design of rotation-BFGS method.

with dimension n_t, the required number of eigenvalues is n_t, whilst this number is $\frac{1}{2}n_t(n_t-1)$ for the rotation angle. The total number of parameters is $\frac{1}{2}n_t(n_t+1)$, which is equal to the number of the elements of the upper triangular of $\mathbf{Q}$. Theoretically, the rotation modeling method can provide a systematic approach to traverse $\mathbf{Q}$ by traversing $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$ in finite regions. The feasible region of $\boldsymbol{\lambda}$ is given in (16b)-(16c) and for each rotation angle the regions can be $[0, 2\pi)$.[5] The rotation method can be applied to many other problems, some listed in Section I-B.

As mentioned earlier, for $n_t = 2$ a closed-form solution for (P2) is known in [14]. However, finding closed-form solutions of $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$ is challenging for $n_t \geq 3$. Next, we introduce a structure to solve this problem effectively.

IV. SOLVING THE NEW OPTIMIZATION PROBLEM

In this section, we develop a novel method to solve (P2). Specifically, we convert (P2) to an unconstrained problem and resort to BFGS [35] to solve it. We should highlight that various iterative optimization methods, such as gradient descent, Newton's method, and quasi-Newton algorithms with interior-point method (that is, MATLAB convex optimization tool `fmincon`), may be used to optimize (P2). Among them, Newton's method is the fastest, but it requires the local Hessian matrix and its inverse, which may not exist at some points. Quasi-Newton methods keep the convergence advantage of Newton's method without requiring the local Hessian matrix and its inverse. BFGS algorithm is an outstanding variety of quasi-Newton methods that can converge faster for non-convex problems [36], [37]. The proposed method is called rotation-BFGS. The block diagram of this method is given in Fig. 3 in which the inputs are channel matrices $\mathbf{H}$ and $\mathbf{G}$ and average power P_t, and the outputs are the parameters of the rotation model (i.e., $\boldsymbol{\lambda}^*$ and $\boldsymbol{\theta}^*$) as well as the corresponding secrecy rate R^*. As shown in Fig. 3, there are four blocks in the BFGS-based optimizer: GSVD for initialization, BFGS for optimization, eigenvalues rectifier, and objective function (16a) evaluator. In the following subsections, we introduce the functionality of each block and the optimization algorithm.

---

[5]It can be proved that in some cases the optimal θ_{ij} is in $[0, \pi)$. See, for example, the case for $n_t = 2$ [14].

A. Functionality of Each Block

1) Initialization Using GSVD: While initial values of $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$ can be chosen randomly, efficient initial values can save time by reducing the number of iterations. For this reason, and knowing that GSVD-based beamforming is a good solution for this problem, we use GSVD to find the initial values ($\boldsymbol{\lambda}^{(0)}$ and $\boldsymbol{\theta}^{(0)}$) for our rotation-BFGS algorithm. The solution given by GSVD-based beamforming provides a precoding matrix $\mathbf{E}$ which satisfies:

$$\mathbf{HE} = \boldsymbol{\Psi}_r \mathbf{C}, \tag{18a}$$

$$\mathbf{GE} = \boldsymbol{\Psi}_e \mathbf{D}, \tag{18b}$$

$$\mathbf{C}^T\mathbf{C} + \mathbf{D}^T\mathbf{D} = \mathbf{I}, \tag{18c}$$

where $\mathbf{E} \in \mathbb{R}^{n_t \times q}$, $q = \min(n_t, n_r + n_e)$, $\mathbf{C}^T\mathbf{C} = \mathrm{diag}(c_i)$ and $\mathbf{D}^T\mathbf{D} = \mathrm{diag}(d_i)$, $i \in \{1, \ldots, q\}$, are diagonal matrices, and $\boldsymbol{\Psi}_r \in \mathbb{R}^{n_r \times n_r}$ and $\boldsymbol{\Psi}_e \in \mathbb{R}^{n_e \times n_e}$ are orthonormal matrices. Besides, the power allocation matrix $\mathbf{P} = \mathrm{diag}(p_i)$ is determined by [15]

$$p_i = \begin{cases} \max(0, \dfrac{2(c_i-d_i)/(\mu e_i)-2}{1+\sqrt{1-4c_id_i+4(c_i-d_i)c_id_i/(\mu e_i)}}), & \text{if } c_i > d_i, \\ 0, & \text{otherwise,} \end{cases} \tag{19}$$

in which p_i and e_i are the ith diagonal element of $\mathbf{P}$ and $\mathbf{E}^T\mathbf{E}$ respectively, and μ is the Lagrange multiplier to ensure

$$\mathrm{tr}(\mathbf{EPE}^T) = P_t. \tag{20}$$

GSVD-based $\mathbf{Q}$ is then $\mathbf{EPE}^T$. Thus, we can determine $\mathbf{V}^{(0)}$ and $\boldsymbol{\Lambda}^{(0)}$ from eigenvalue decomposition of $\mathbf{EPE}^T$, i.e., from

$$\mathbf{Q}^{(0)} = \mathbf{EPE}^T = \mathbf{V}^{(0)}\boldsymbol{\Lambda}^{(0)}\mathbf{V}^{(0)T}. \tag{21}$$

Then, $\boldsymbol{\lambda}^{(0)} = \mathrm{diag}(\boldsymbol{\Lambda}^{(0)})$ and $\boldsymbol{\theta}^{(0)}$ can be obtained from $\mathbf{V}^{(0)}$ using Algorithm 1.

2) BFGS Optimizer: Different iterative optimization methods, such as gradient descent, Newton's method, and quasi-Newton algorithms may be used to optimize (P2). Among the above methods, Newton's method is the fastest method, but it requires a local Hessian matrix and its inverse, which may not exist or hard to obtain. Quasi-Newton methods keep the convergence advantage of Newton's method without requiring the local Hessian matrix and its inverse. An outstanding variety of quasi-Newton methods is the BFGS algorithm [35]. We give a brief introduction to this method in the following.

Given an unconstrained optimization problem

$$\arg\min_{\mathbf{x}} f(\mathbf{x}), \tag{22}$$

with argument vector $\mathbf{x}$ and objective function $f(\mathbf{x})$, BFGS algorithm updates the vector $\mathbf{x}$ iteratively according to

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha^{(k)}\mathbf{M}^{(k)}\mathbf{g}^{(k)}, \quad k \geq 0, \tag{23}$$

in which

- $\alpha^{(k)}$ is the step size to minimize $f(\mathbf{x}^{(k+1)})$, which can be obtained via a line search.

- $\mathbf{g}^{(k)}$ is the gradient of $f(\cdot)$ at $\mathbf{x}^{(k)}$.
- $\mathbf{M}^{(k)}$ is an approximation of the inverse of the Hessian matrix.

The matrix $\mathbf{M}$ is initialized by a unity matrix, i.e., $\mathbf{M}^{(0)} = \mathbf{I}$, and is then updated as [35]

$$\mathbf{M}^{(k+1)} = \left(\mathbf{I} - \frac{\boldsymbol{\delta}_\mathbf{x}^{(k)} \boldsymbol{\delta}_\mathbf{g}^{(k)T}}{\boldsymbol{\delta}_\mathbf{g}^{(k)T} \boldsymbol{\delta}_\mathbf{x}^{(k)}} \right) \mathbf{M}^{(k)} \left(\mathbf{I} - \frac{\boldsymbol{\delta}_\mathbf{g}^{(k)} \boldsymbol{\delta}_\mathbf{x}^{(k)T}}{\boldsymbol{\delta}_\mathbf{g}^{(k)T} \boldsymbol{\delta}_\mathbf{x}^{(k)}} \right) + \frac{\boldsymbol{\delta}_\mathbf{x}^{(k)} \boldsymbol{\delta}_\mathbf{x}^{(k)T}}{\boldsymbol{\delta}_\mathbf{g}^{(k)T} \boldsymbol{\delta}_\mathbf{x}^{(k)}}, \quad (24)$$

in which

$$\boldsymbol{\delta}_\mathbf{x}^{(k)} \triangleq \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}, \quad (25a)$$

$$\boldsymbol{\delta}_\mathbf{g}^{(k)} \triangleq \mathbf{g}^{(k+1)} - \mathbf{g}^{(k)}, \quad (25b)$$

respectively, represent the difference between the arguments and the gradients in two successive iterations. The gradient can be obtained analytically or numerically.

Due to its fast convergence speed and self-correcting property [36], BFGS is widely used in unconstrained optimization problems. However, the problem in this paper is a constrained optimization problem due to (16b) and (16c). To overcome this limitation, we relax the constraints of (P2) when using the BFGS optimizer, but we add a new block called a rectifier, as shown in Fig. 3. The rectifying block is used to ensure that the constraint on λ_is are satisfied as elaborated on in what follows.

3) Rectifying Eigenvalues: To use BFGS, the optimization problem should be unconstrained. However, the eigenvalues in (P2) need to be constrained. To overcome this issue, we rectify the eigenvalues to ensure that the constraints in (16b)-(16c) are satisfied.

Before talking about the rectification process, we highlight that we only optimize the first $n_t - 1$ eigenvalues because the last eigenvalue will be obtained by $\sum_{i=1}^{n_t} \lambda_i = P_t$. That is, we use $P_t - \sum_{i=1}^{n_t-1} \lambda_i$ as the value of the last eigenvalue (λ_{n_t}) since in the MIMOME channel it is known that optimal solution uses the total power [7], [13], [14].

Suppose $\tilde{\boldsymbol{\lambda}} \in \mathbb{R}^{n_t-1}$ is the vector of first $n_t - 1$ unconstrained eigenvalues obtained from the BFGS algorithm. We obtain $\boldsymbol{\lambda} \in \mathbb{R}^{n_t}$ (the rectified eigenvalue vector) from

$$\boldsymbol{\lambda} = \mathrm{r}(\tilde{\boldsymbol{\lambda}}, P_t), \quad (26)$$

where the rectifying function $\mathrm{r}(\cdot, \cdot)$ is defined by the following successive processes:

$$\boldsymbol{\lambda}^+ = [\tilde{\boldsymbol{\lambda}}]^+, \quad (27a)$$

$$\bar{\boldsymbol{\lambda}} = \begin{cases} \boldsymbol{\lambda}^+ \cdot \frac{P_t}{\sum_{i=1}^{n_t-1} \lambda_i^+}, & \sum_{i=1}^{n_t-1} \lambda_i^+ > P_t, \\ \boldsymbol{\lambda}^+, & \text{otherwise,} \end{cases} \quad (27b)$$

$$\boldsymbol{\lambda} = \left[\bar{\boldsymbol{\lambda}}, P_t - \sum_{i=1}^{n_t-1} \bar{\boldsymbol{\lambda}} \right]. \quad (27c)$$

In (27a), $[\cdot]^+$ is an element-wise operation which forces all negative elements of $\tilde{\boldsymbol{\lambda}}$ to 0 while keeping non-negative values unaltered. This will take care of the constraints in (16c). But, the elements of the new vector $\boldsymbol{\lambda}^+$ may not still satisfy (16b).
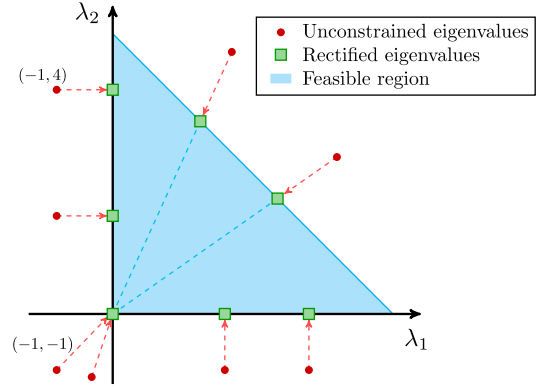

Fig. 4.   Illustration of eigenvalues rectification for $n_t = 3$.

In such a case, in (27b), we scale the new vector such that the sum of the eigenvalues does not exceed P_t. Finally, in the (27c), the last eigenvalue is added. Thus, the problem (P2) can be solved using an unconstrained optimization method, namely, the BFGS.

This process illustrated in Fig. 4 for $n_t = 3$ and $P_t = 5$, as an example. In this figure, the blue (shaded) area denotes the feasible region of the eigenvalues, red points outside the region are unconstrained eigenvalues $\tilde{\boldsymbol{\lambda}} = [\lambda_1, \lambda_2, \cdots, \lambda_{n_t-1}]$ (the output of the BFGS block), and green squares denote their corresponding rectified eigenvalues. With the proposed rectification in (27a)-(27b), unrestricted inputs will be forced to new points on the boundary of the feasible region. For example, for the point $(-1, 4)$ we have $\tilde{\boldsymbol{\lambda}} = [-1, 4]$, $\bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^+ = [0, 4]$ (the second case in (27b)) and $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, P - \lambda_1 - \lambda_2] = [0, 4, 1]$, following (27a)-(27b). Similarly, the red point $(-1, -1)$ will result in $\boldsymbol{\lambda} = [0, 0, 5]$.

4) The Objective Function: In this paper, our goal is to maximize the secrecy rate R defined by

$$R(\boldsymbol{\lambda}, \boldsymbol{\theta}) \triangleq \tfrac{1}{2} \log_2 \frac{|\mathbf{I}_{n_t} + \mathbf{V}^T \mathbf{H}^T \mathbf{H} \mathbf{V} \boldsymbol{\Lambda}|}{|\mathbf{I}_{n_t} + \mathbf{V}^T \mathbf{G}^T \mathbf{G} \mathbf{V} \boldsymbol{\Lambda}|}, \quad (28)$$

which is a function of $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$ as shown in (P2). Noting that our problem is a maximization rather than a minimization problem, in comparison to (22), we define the objective function $f(\mathbf{x}) = -R$. With this, we can link BFGS and rotation method together.

B. Rotation-BFGS Algorithm

As shown in Fig. 3, GSVD provides an initial value for the argument $\mathbf{x}$ in the BFGS block. The rectifier and the objective function can be considered together which require the current value of $\mathbf{x}$ and provide corresponding achievable rate back to the BFGS. Then, BFGS will update $\mathbf{x}$ and check the termination condition.

To link the BFGS to the rotation model, we need to define the relation between $\mathbf{x}$, the arguments inside BFGS, and the parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$. This is given by

$$\mathbf{x} \triangleq [\tilde{\boldsymbol{\lambda}}, \boldsymbol{\theta}], \quad (29)$$

where $\tilde{\boldsymbol{\lambda}}$ is the first $n_t - 1$ elements of $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$ is the same as we defined in (17b). In this way, we reduce one

argument (eigenvalue) for efficiency. Then, in the kth iteration ($k \geq 0$) of the rotation-BFGS algorithm, the relation between argument $\mathbf{x}^{(k)}$ in the BFGS and the eigenvalues $\boldsymbol{\lambda}^{(k)}$ in precoding is

$$\mathbf{x}^{(k)} \triangleq [\tilde{\boldsymbol{\lambda}}^{(k)}, \boldsymbol{\theta}^{(k)}], \tag{30a}$$

$$\boldsymbol{\lambda}^{(k)} \triangleq \mathrm{r}(\tilde{\boldsymbol{\lambda}}^{(k)}, \mathrm{P_t}), \tag{30b}$$

in which $\tilde{\boldsymbol{\lambda}}^{(k)}$ denotes the first $n_t - 1$ unconstrained eigenvalues and $\boldsymbol{\lambda}^{(k)}$ is the constrained (rectified) eigenvalues which are obtained by (26)-(27c). Besides, the value of the function is denoted as

$$f^{(k)} \triangleq f(\mathbf{x}^{(k)}) = -R(\mathrm{r}(\tilde{\boldsymbol{\lambda}}^{(k)}, P_t), \boldsymbol{\theta}^{(k)}), \tag{31}$$

in which $\boldsymbol{\lambda}^{(k)}$ and $\boldsymbol{\theta}^{(k)}$ can be obtained from $\mathbf{x}^{(k)}$ according to (30a)-(30b), and $f(\cdot)$ denotes the mapping from $\mathbf{x}^{(k)}$ to $f^{(k)}$. Finally, the gradient vector $\mathbf{g}^{(k)}$ with respect to $\mathbf{x}^{(k)}$ is obtained numerically. Specifically, the ith element of $\mathbf{g}$ in the kth iteration is given by

$$g_i^{(k)} = \frac{f(\mathbf{x}^{(k)} + \boldsymbol{\epsilon}_i) - f^{(k)}}{|\boldsymbol{\epsilon}_i|}, \tag{32}$$

where $\boldsymbol{\epsilon}_i$ has the same length as $\mathbf{x}$ and its all components are zero except for the ith element which is a constant ϵ_1.

In our proposed rotation-BFGS method, as shown in Fig. 3, GSVD-based beamforming is first applied to find initial values of $\boldsymbol{\lambda}^{(0)}$, $\boldsymbol{\theta}^{(0)}$ which are obtained by GSVD decomposition as described in Section IV-A.1. Thus $\mathbf{x}^{(0)} = [\tilde{\boldsymbol{\lambda}}^{(0)}, \boldsymbol{\theta}^{(0)}]$. We should also highlight that $\tilde{\boldsymbol{\lambda}}^{(0)}$ contains the first $n_t - 1$ elements of $\boldsymbol{\lambda}^{(0)}$.

Algorithm 2 Rotation-BFGS Method

1: Initialize: $\epsilon_1 = \epsilon_2 = 10^{-4}$, $k = 0$, and $\mathbf{M}^{(0)} = \mathbf{I}$;
2: Find $\mathbf{V}^{(0)}$ and $\boldsymbol{\Lambda}^{(0)}$ using GSVD in (21);
3: Find $\boldsymbol{\lambda}^{(0)}$ which is the diagonal of $\boldsymbol{\Lambda}^{(0)}$;
4: Find $\tilde{\boldsymbol{\lambda}}^{(0)}$ which is the first $(n_t - 1)$ elements in $\boldsymbol{\lambda}^{(0)}$;
5: Find $\boldsymbol{\theta}^{(0)}$ using Algorithm 1;
6: Find $R^{(0)}$ using (28);
7: Find $\mathbf{x}^{(0)}$, $\mathbf{g}^{(0)}$, and $f^{(0)}$ using (30a)-(32);
8: **while** 1 **do**
9:     Find $\alpha^{(k)} = \alpha^*$ by line search using Algorithm 3;
10:    Find $\mathbf{x}^{(k+1)}$ using (23);
11:    Find $\boldsymbol{\lambda}^{(k+1)}$ and $\boldsymbol{\theta}^{(k+1)}$ from $\mathbf{x}^{(k)}$ using (30a)-(30b);
12:    Find $R(\boldsymbol{\lambda}^{(k+1)}, \boldsymbol{\theta}^{(k+1)})$ using (28);
13:    Find $\mathbf{g}^{(k+1)}$ and $f^{(k+1)}$ using (30a)-(32);
14:    Find $\mathbf{M}^{(k+1)}$ using BFGS by (24)-(25b);
15:    **if** $|f^{(k+1)} - f^{(k)}| < \epsilon_2$ **then**
16:      $R^* = -f^{(k+1)}$, $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^{(k+1)}$, and $\boldsymbol{\theta}^* = \boldsymbol{\theta}^{(k+1)}$;
17:      Break;
18:    **end if**
19:    $k = k + 1$;
20: **end while**
21: Output: R^*, $\boldsymbol{\lambda}^*$, and $\boldsymbol{\theta}^*$.

Algorithm 2 illustrates the details of the proposed optimization process. Within this algorithm, we require a line

TABLE I

ACHIEVABLE RATE (IN BPS/Hz) OF EACH METHOD FOR $n_t = 3$ AND $P_t = 30W$

Rotation-BFGS						
$n_t = 3$		n_e				
		1	2	3	4	5	6
n_r	1	2.58	1.99	1.14	0.63	0.39	0.23
	2	3.91	2.99	1.75	1.09	0.71	0.43
	3	4.84	3.59	2.24	1.47	1.01	0.70
	4	5.48	4.16	2.70	1.73	1.24	0.89
	5	5.97	4.52	3.01	2.03	1.45	1.09
	6	6.46	4.88	3.30	2.35	1.76	1.28

Rotation-Fmin						
$n_t = 3$		n_e				
		1	2	3	4	5	6
n_r	1	2.58	1.99	1.14	0.63	0.39	0.23
	2	3.91	2.99	1.75	1.09	0.71	0.43
	3	4.84	3.59	2.24	1.47	1.01	0.70
	4	5.48	4.16	2.70	1.73	1.24	0.89
	5	5.97	4.52	3.01	2.03	1.45	1.09
	6	6.46	4.88	3.30	2.35	1.76	1.28

GSVD [15]						
$n_t = 3$		n_e				
		1	2	3	4	5	6
n_r	1	2.56	1.82	1.12	0.63	0.39	0.23
	2	2.97	2.87	1.74	1.09	0.70	0.43
	3	4.41	3.50	2.23	1.46	1.01	0.70
	4	5.17	4.10	2.69	1.73	1.24	0.89
	5	5.76	4.46	3.00	2.02	1.45	1.09
	6	6.28	4.83	3.29	2.35	1.76	1.28

AOWF [16]						
$n_t = 3$		n_e				
		1	2	3	4	5	6
n_r	1	2.58	1.99	1.13	0.63	0.38	0.23
	2	3.92	2.98	1.74	1.08	0.69	0.42
	3	4.86	3.58	2.23	1.45	0.99	0.68
	4	5.49	4.15	2.69	1.70	1.20	0.85
	5	5.99	4.51	2.98	1.97	1.39	1.02
	6	6.47	4.87	3.26	2.28	1.67	1.18

search. The details of the line search are given in Algorithm 3 in the Appendix. Algorithm 2 will terminate if the secrecy rate at two successive iterations are very close, i.e., when their difference is smaller than a tolerance (ϵ_2). Algorithm 2 guarantees a global convergence based on the global convergence theorem [35, Chapter 7, pp. 196–204], because the optimized function (31) (the objective function together with the rectifier in Fig. 3) is continuous, and the iteration in (23) can lead to a decreasing sequence, i.e., $f(\mathbf{x}^{(k+1)}) \leq f(\mathbf{x}^{(k)})$.

We analyze the complexity of the different algorithms here. The computation of matrix multiplications and matrix inverse yields the complexity of $\mathcal{O}(L^3)$ where $L = \max(n_t, n_r, n_e)$. GSVD-based precoding [15], used as the initial point generator, has the complexity of $\mathcal{O}(L^3 + L\log(1/\epsilon))$ [38] in which ϵ is the convergence tolerance of the algorithm, while bisection search requires $\mathcal{O}(\log(1/\epsilon))$ iterations [39]. Besides, the BFGS algorithm has the complexity of $\mathcal{O}(n^2)$ [37], where n is the size of input variables which is the total number of optimized rotation parameters, i.e., $\frac{1}{2}n_t(n_t + 1)$. Thus, the overall complexity of Algorithm 2 is $\mathcal{O}(n_t^4 + L^3 + L\log(1/\epsilon))$. AOWF [16, Algorithm 1] yields $\mathcal{O}(\frac{L^3}{\epsilon}\log(1/\epsilon))$, in which $\mathcal{O}(\frac{1}{\epsilon})$ and $\mathcal{O}(\log(1/\epsilon))$ are the outer layer loop and inner bisection search, respectively. It should also be mentioned that rotation-Fmin has the same asymptotic complexity as rotation-BFGS. However, for small values of n_t,

TABLE II

ACHIEVABLE RATE (IN BPS/HZ) OF EACH
METHOD FOR $n_t = 4$ AND $P_t = 30W$

Rotation-BFGS						
$n_t = 4$		n_e				
		1	2	3	4	5	6
n_r	1	3.04	2.63	2.06	1.25	0.77	0.44
	2	4.72	4.01	3.12	2.02	1.32	0.87
	3	5.91	4.97	3.78	2.60	1.72	1.21
	4	6.81	5.68	4.40	3.05	2.08	1.55
	5	7.63	6.28	4.82	3.46	2.50	1.83
	6	8.24	6.76	5.22	3.78	2.74	2.10

Rotation-Fmin						
$n_t = 4$		n_e				
		1	2	3	4	5	6
n_r	1	3.04	2.63	2.06	1.25	0.77	0.44
	2	4.73	4.02	3.12	2.02	1.32	0.87
	3	5.91	4.98	3.78	2.60	1.72	1.21
	4	6.81	5.69	4.40	3.05	2.08	1.55
	5	7.63	6.29	4.82	3.46	2.50	1.83
	6	8.24	6.77	5.22	3.78	2.74	2.10

GSVD [15]						
$n_t = 4$		n_e				
		1	2	3	4	5	6
n_r	1	3.03	2.59	1.82	1.22	0.76	0.44
	2	4.27	3.02	2.97	1.99	1.31	0.87
	3	3.87	4.42	3.67	2.57	1.72	1.21
	4	5.80	5.33	4.31	3.03	2.08	1.55
	5	6.91	6.03	4.75	3.44	2.50	1.82
	6	7.70	6.56	5.16	3.76	2.73	2.09

AOWF [16]						
$n_t = 4$		n_e				
		1	2	3	4	5	6
n_r	1	3.04	2.63	2.06	1.24	0.76	0.44
	2	4.74	4.04	3.12	2.01	1.31	0.86
	3	5.94	4.98	3.77	2.58	1.70	1.18
	4	6.82	5.68	4.39	3.04	2.06	1.52
	5	7.63	6.29	4.81	3.44	2.46	1.78
	6	8.24	6.75	5.21	3.74	2.69	2.03

(a) Relative rate improvement (%) of the proposed method to GSVD.

(b) Relative rate improvement (%) of the proposed method to AOWF.

Fig. 5. Comparisons between the secrecy transmission rates of the proposed method with GSVD and AOWF.

rotation-BFGS has about an order of magnitude smaller complexity than rotation-Fmin, as we show in Section V.
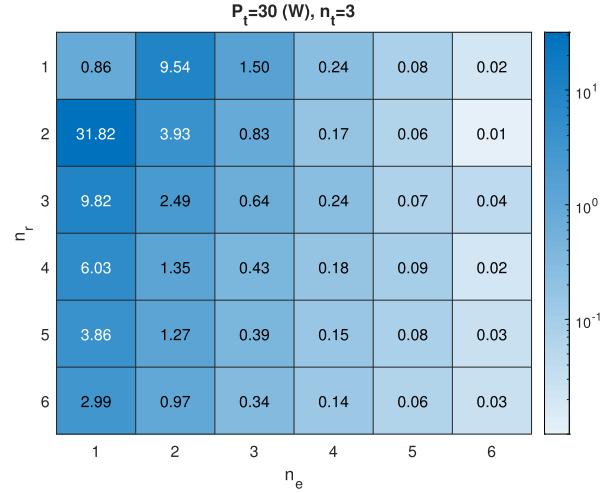
V. NUMERICAL RESULT

In this section, extensive numerical results are provided to illustrate the performance of the proposed rotation-BFGS method. Four methods are taken into account:
- **Rotation-BFGS**: the proposed rotation-BFGS parameterization solved by the BFGS method.
- **Rotation-Fmin**: the proposed rotation modeling solved using MATLAB convex optimization tool `fmincon`.
- **GSVD**: GSVD-based beamforming with optimal power allocation [15], as described in Section IV-A.1.
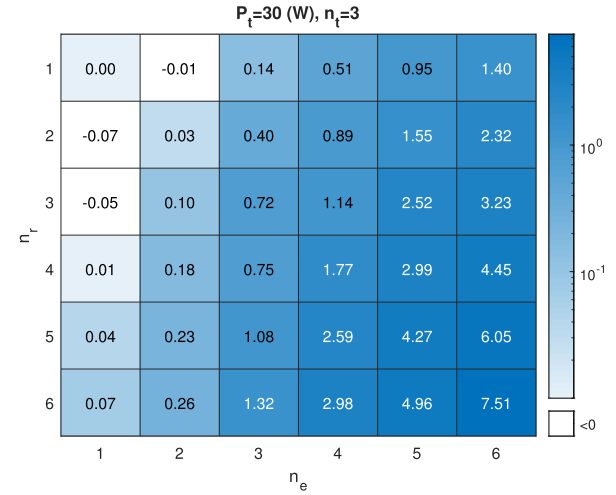- **AOWF**: alternating optimization and water-filling [16].

All results are based on averaging over 1000 realizations of independent **H** and **G**. These entries of **H** and **G** are generated based on the standard Gaussian distribution, i.e., $\mathcal{N}(0, 1)$. Two performance metrics are considered. In the first part, we focus on the achievable secrecy rate. In the second part, the time consumption is analyzed for different methods in various antenna configurations.

A. Achievable Secrecy Rate

We evaluate the performance of the Rotation-BFGS, Rotation-Fmin, GSVD, and AOWF approaches with respect to the variation of the number of antennas. The average secrecy

rates are listed in Tables I and II respectively for $n_t = 3$ and $n_t = 4$. As expected, increasing the number of antennas at the eavesdropper decreases secrecy rate. On the contrary, the secrecy rate will increase when the legitimate receiver or transmitter has higher number of antennas. To better appreciate the improvement due to our proposed numerical method, relative secrecy rate improvement between Rotation-BFGS and GSVD and Rotation-BFGS and AOWF is investigated in the following and illustrated in Fig. 5. Let us define relative rate improvement with respect to GSVD and AOWF, respectively, as

$$\eta_g = \frac{R_r - R_g}{R_g} \times 100\%, \tag{33a}$$

$$\eta_a = \frac{R_r - R_a}{R_a} \times 100\%, \tag{33b}$$

where R_r, R_g, and R_a represent the secrecy rate achieved by the proposed method, GSVD, and AOWF, respectively.[6] As shown in Fig. 5, the proposed approach is capable of

[6]Rotation-Fmin is similar to rotation-BFGS in terms of the achievable rate (details are in Table I and II), so relative rate improvements of that to GSVD and AOWF are the same as those in rotation-BFGS.
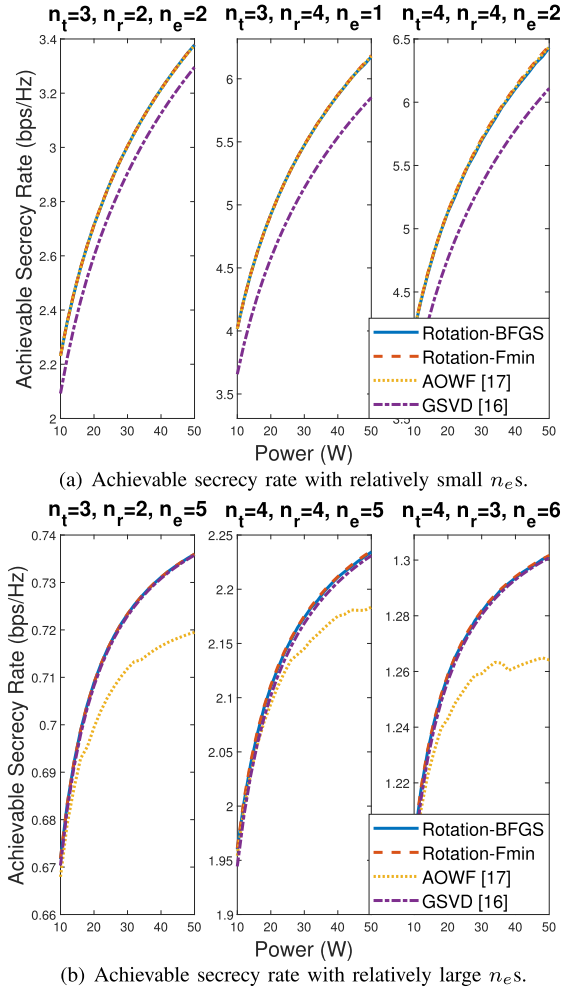
Fig. 6. Secrecy rate of the MIMOME channel versus the transmit power. The proposed algorithms (Rotation-BFGS and Rotation-Fmin) are compared with GSVD and AOWF.

Fig. 7. Time costs of the proposed methods, GSVD and AOWF, for (a) $n_t = 3$, (b) $n_t = 4$, with n_r and n_e from 1 to 6.

achieving the same or better secrecy rate in almost any antenna setting. The darker the cell color, the higher improvement is achieved by the proposed method.

There is a clear pattern that when the eavesdropper has a smaller number of antennas than the transmitter, the proposed method outperforms GSVD. The best case is $[n_r, n_e] = [2, 1]$ where the improvement is about 32%. On the other hand, for cases with larger n_e, the rotation-BFGS performs better than AOWF. The best case is $[n_r, n_e] = [6, 6]$ where the improvement is 7.5%. This pattern not only exist when $n_t = 3$ but also at higher n_ts, at least $n_t = 4$. Some more detailed comparisons are presented as follows:

1) The MIMOME With Small n_e: GSVD-based beamforming fails to get close to the secrecy capacity. This phenomenon has been verified by previous literature [14]. As can be seen in Fig. 6(a), rotation-BFGS and AOWF can achieve similar results. On the contrary, GSVD clearly has a gap with those methods which results from sub-optimality of GSVD.

2) The MIMOME With Large n_e: In this setting, AOWF does not perform very well because the Lagrange multiplier of AOWF cannot be obtained properly. As illustrated in Fig. 6(b) and Fig. 5, our proposed method outperforms AOWF in this regime.
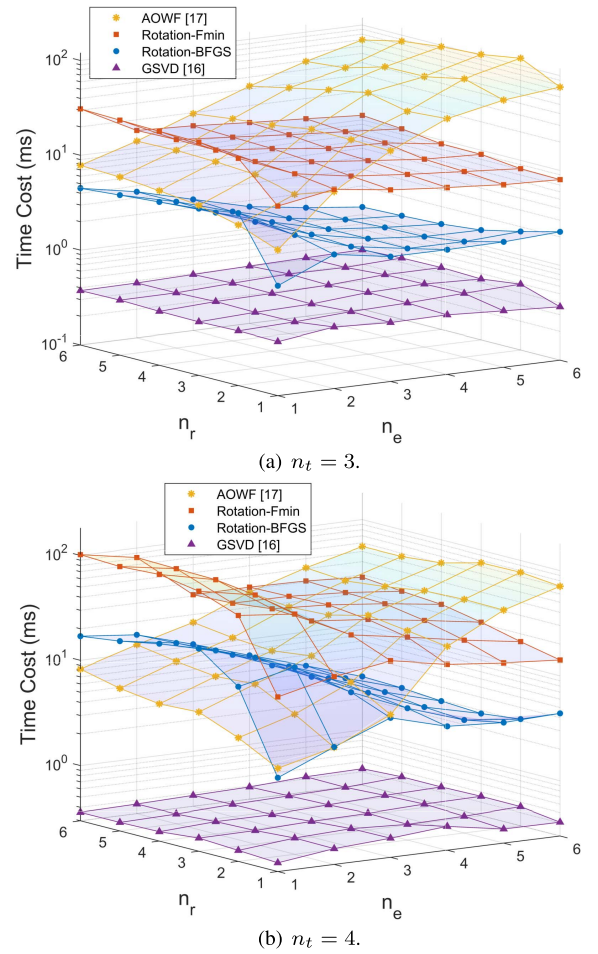
From the simulation results, it is seen that the proposed rotation-BFGS is robust in a wide range of practical antennas settings on each node. This is a big advantage as the robustness towards the variation of n_e is necessary to guarantee secure communication. We note that the eavesdropper can have any number of antennas, and the solution for wiretap channels should be robust to such variations. This makes the proposed precoding highly competitive the existing solutions of the MIMOME channel.

B. Time Consumption

Besides computational complexity, the execution time of each algorithm is also important and provides a means to evaluate the complexity. This evaluation is shown in Figs. 7(a)-7(b) for $n_t = 3$ and $n_t = 4$ respectively, which are averaged over 1000 channel realizations in each setting. GSVD has, by far, the best time cost since it transforms the problem into finding the best Lagrange multiplier. However, as we illustrated previously, the precoding provided by GSVD may be far from the optimal solution. The proposed rotation-BFGS method clearly outperforms AOWF, whose time consumption is very high, especially when $n_e \in [4, 5, 6]$. The gap becomes larger when n_e increases. In such cases, AOWF

Algorithm 3 Line Search (Golden Section [35]) for the BFGS

1: Requires from Algorithm 2: ϵ_2, $\mathbf{x}^{(k)}$, $\mathbf{M}^{(k)}$, $\mathbf{g}^{(k)}$, and $f^{(k)}$;
2: Define: $\mathbf{x} \triangleq \mathbf{x}^{(k)}$, $\mathbf{d} \triangleq \mathbf{M}^{(k)}\mathbf{g}^{(k)}$;
3: Initialize: $\epsilon_3 = 5 \times 10^{-4}$, $\tau_1 = 3$, $\tau_2 = 0.382$, and $\tau_3 = 0.618$;
4: Initialize: $\boldsymbol{\alpha} \triangleq [\alpha_1, \alpha_2, \alpha_3, \alpha_4] = [0, 0, 0, 0.1]$;
5: Initialize: $\boldsymbol{f} \triangleq [f_1, f_2, f_3, f_4] = [f^{(k)}, 0, 0, f(\mathbf{x} - \alpha_4\mathbf{d})]$;
6: **while** $\alpha_4 < 20$ and $f_1 < f_4$ **do**
7:    Let $\alpha_4 = \tau_1\alpha_4$ and $f_4 = f(\mathbf{x} - \alpha_4\mathbf{d})$;
8: **end while**
9: Let $\alpha_2 = \tau_2\alpha_4$ and $\alpha_3 = \tau_3\alpha_4$;
10: Let $f_2 = f(\mathbf{x} - \alpha_2\mathbf{d})$ and $f_3 = f(\mathbf{x} - \alpha_3\mathbf{d})$;
11: **while** $\alpha_4 - \alpha_1 > \epsilon_3$ and $\max(\boldsymbol{f}) - \min(\boldsymbol{f}) > \epsilon_2$ **do**
12:    Define: m is the index of the smallest element in $\boldsymbol{f}$;
13:    Define: $q_1 \triangleq \max(1, m - 1)$ and $q_2 \triangleq \min(4, m + 1)$;
14:    Let $\alpha_1 = \alpha_{q_1}$ and $\alpha_4 = \alpha_{q_2}$;
15:    Let $\alpha_2 = \alpha_1 + \tau_2(\alpha_4 - \alpha_1)$ and $\alpha_3 = \alpha_1 + \tau_3(\alpha_4 - \alpha_1)$;
16:    **if** $m = 2$ **then**
17:      Let $f_3 = f_2$ and $f_2 = f(\mathbf{x} - \alpha_2\mathbf{d})$;
18:    **else if** $m = 3$ **then**
19:      Let $f_2 = f_3$ and $f_3 = f(\mathbf{x} - \alpha_3\mathbf{d})$;
20:    **else**
21:      Let $f_2 = f(\mathbf{x} - \alpha_2\mathbf{d})$ and $f_3 = f(\mathbf{x} - \alpha_3\mathbf{d})$;
22:    **end if**
23: **end while**
24: Output: $\alpha^* = \alpha_m$.

does not perform well either in terms of achievable rate or time cost. On the other hand, our approach is robust in any antenna setting and more efficient in time cost compared with AOWF. Rotation-Fmin achieves almost the same performance, but its execution time is higher. This is because `fmincon` includes BFGS and interior-point method, and instead of the interior-point method, we use a rectifier which is more efficient in terms of programming.

In summary, the proposed rotation-BFGS method gives robust precoding and power allocation for the MIMOME channel. By applying BFGS as an optimizer, the rotation method has a better performance compared to other numerical methods like AOWF in secrecy and time complexity. In addition, the framework we provided in this paper can be used to solve many other problems, some listed in Section I-B.

VI. Conclusion

In this paper, we have developed a rotation-based method, called rotation-BFGS, for precoding and power allocation of Gaussian MIMOME channels. In this method, the transmit covariance matrix is constructed using Givens rotation matrices. With this construction, the PSD constraint of the transmit covariance matrix is removed and the capacity optimization problem is simplified. The precoding (rotation) matrix and power allocation coefficients have been obtained iteratively using a modified BFGS algorithm. Compared to existing approaches, the proposed method is robust and performs well regardless of the number of antennas at each node. The proposed method outperforms the GSVD-based beamforming

when $n_e < n_t$ and AOWF particularly when $n_e \geq n_t$. This approach can also use the results of existing precoding and power allocation methods, such as the GSVD-based approach, as an initial point to expedite finding the solution.

In addition, the proposed rotation-BFGS method has a great potential for finding precoding and power allocation in various other applications, including in MIMO broadcast channel and MIMO channel with energy harvesting constraints, both with and without secrecy. Future works will focus on further improving the efficiency of solving parameters and extension of this approach to other related problems.

APPENDIX A
Line Search Algorithm for the BFGS Method

The line search method we use is the Golden section search [35] and summarized in Algorithm 3, where α_i and f_i for $1 \leq i \leq 4$ are temporary records for line search steps size and function values.

References

[1] X. Zhang, Y. Qi, and M. Vaezi, "A new precoding for the MIMO Gaussian channel with multi-antenna eavesdroppers," in *Proc. IEEE 31st Annu. Int. Symp. Pers., Indoor Mobile Radio Commun.*, Aug. 2020, pp. 1–6.
[2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
[3] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
[4] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.
[5] N. Cai, A. Winter, and R. W. Yeung, "Quantum privacy and quantum wiretap channels," *Problems Inf. Transmiss.*, vol. 40, no. 4, pp. 318–336, Oct. 2004.
[6] J. Wu, Z. Lin, L. Yin, and G. Long, "Security of quantum secure direct communication based on Wyner's wiretap channel theory," *Quantum Eng.*, vol. 1, no. 4, p. e26, Dec. 2019.
[7] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—Part II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.
[8] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.
[9] T. Liu and S. Shamai, "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.
[10] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4033–4039, Sep. 2009.
[11] P. Parada and R. Blahut, "Secrecy capacity of SIMO and slow fading channels," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Sep. 2005, pp. 2152–2155.
[12] S. A. A. Fakoorian and A. L. Swindlehurst, "Full rank solutions for the MIMO Gaussian wiretap channel with an average power constraint," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2620–2631, May 2013.
[13] S. Loyka and C. D. Charalambous, "Optimal signaling for secure communications over Gaussian MIMO wiretap channels," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7207–7215, Dec. 2016.
[14] M. Vaezi, W. Shin, and H. V. Poor, "Optimal beamforming for Gaussian MIMO wiretap channels with two transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6726–6735, Oct. 2017.
[15] S. A. A. Fakoorian and A. L. Swindlehurst, "Optimal power allocation for GSVD-based beamforming in the MIMO Gaussian wiretap channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 2321–2325.
[16] Q. Li, M. Hong, H.-T. Wai, Y.-F. Liu, W.-K. Ma, and Z.-Q. Luo, "Transmit solutions for MIMO wiretap channels using alternating optimization," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1714–1727, Sep. 2013.

[17] S. Loyka and C. D. Charalambous, "An algorithm for global maximization of secrecy rates in Gaussian MIMO wiretap channels," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2288–2299, Jun. 2015.

[18] J. Steinwandt, S. A. Vorobyov, and M. Haardt, "Secrecy rate maximization for MIMO Gaussian wiretap channels with multiple eavesdroppers via alternating matrix POTDC," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 5686–5690.

[19] J. Li and A. Petropulu, "Transmitter optimization for achieving secrecy capacity in Gaussian MIMO wiretap channels," 2009, *arXiv:0909.2622*. [Online]. Available: http://arxiv.org/abs/0909.2622

[20] M. Vaezi, W. Shin, H. V. Poor, and J. Lee, "MIMO Gaussian wiretap channels with two transmit antennas: Optimal precoding and power allocation," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 1708–1712.

[21] S. Loyka and C. D. Charalambous, "On optimal signaling over secure MIMO channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 443–447.

[22] M. Ansari Sadrabadi, A. K. Khandani, and F. Lahouti, "Channel feedback quantization for high data rate MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3335–3338, Dec. 2006.

[23] R. de Francisco and D. Slock, "An optimized unitary beamforming technique for MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 990–1000, Mar. 2010.

[24] Y. Yang, Q. Li, W.-K. Ma, J. Ge, and P. C. Ching, "Cooperative secure beamforming for AF relay networks with multiple eavesdroppers," *IEEE Signal Process. Lett.*, vol. 20, no. 1, pp. 35–38, Jan. 2013.

[25] J. Huang and A. L. Swindlehurst, "Cooperative jamming for secure communications in MIMO relay networks," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4871–4884, Oct. 2011.

[26] J. Li and A. P. Petropulu, "On ergodic secrecy rate for Gaussian MISO wiretap channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1176–1187, Apr. 2011.

[27] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, "Linear precoding for finite-alphabet signaling over MIMOME wiretap channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2599–2612, Jul. 2012.

[28] W. Zeng, C. Xiao, M. Wang, and J. Lu, "Linear precoding for finite-alphabet inputs over MIMO fading channels with statistical CSI," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3134–3148, Jun. 2012.

[29] S. Rezaei Aghdam, A. Nooraiepour, and T. M. Duman, "An overview of physical layer security with finite-alphabet signaling," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1829–1850, 2nd Quart., 2019.

[30] S. R. Aghdam and T. M. Duman, "Joint precoder and artificial noise design for MIMO wiretap channels with finite-alphabet inputs based on the cut-off rate," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3913–3923, Jun. 2017.

[31] X. Yang and A. L. Swindlehurst, "Limited rate feedback in a MIMO wiretap channel with a cooperative jammer," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4695–4706, Sep. 2016.

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

[33] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins Univ. Press, 2012.

[34] S. Matthies, "Orientations and rotations: computations in crystallographic textures," *J. Appl. Crystallogr.*, vol. 38, no. 6, pp. 1042–1043, 2005.

[35] D. G. Luenberger *et al.*, *Linear and Nonlinear Programming*, vol. 2. New York, NY, USA: Springer, 1984.

[36] R. H. Byrd and J. Nocedal, "A tool for the analysis of quasi-Newton methods with application to unconstrained minimization," *SIAM J. Numer. Anal.*, vol. 26, no. 3, pp. 727–739, Jun. 1989.

[37] J. Nocedal and S. Wright, *Numerical Optimization*. New York, NY, USA: Springer, 2006.

[38] D. Park, "Weighted sum rate maximization of MIMO broadcast and interference channels with confidential messages," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 1742–1753, Mar. 2016.

[39] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
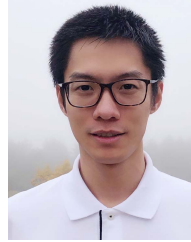
Xinliang Zhang (Graduate Student Member, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering from Xidian University, Xi'an, China, in 2015 and 2018, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Villanova University, Villanova, PA, USA. His research interests lie in physical layer security, machine learning for wireless communications, MIMO networks, and signal processing.

Yue Qi (Graduate Student Member, IEEE) received the B.Eng. and M.Eng. degrees in electronic engineering from Xidian University, Xi'an, China, in 2015 and 2018, respectively. She is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Villanova University, Villanova, PA, USA. Her research interests include NOMA, MIMO system, IoT heterogeneous networks, green communication, and signal processing. She has received the IEEE Communication Society Student Grant in ICC conference, 2020.

Mojtaba Vaezi (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees from the Amirkabir University of Technology (Tehran Polytechnic) and the Ph.D. degree from McGill University, all in electrical engineering. From 2015 to 2018, he was with Princeton University as a Postdoctoral Research Fellow and Associate Research Scholar. He is currently an Assistant Professor of ECE with Villanova University. Before joining Princeton, he was a Researcher with Ericsson Research, Montreal, Canada. His research interests include the broad areas of signal processing and machine learning for wireless communications with an emphasis on physical layer security and fifth generation (5G) and beyond radio access technologies. Among his publications in these areas is the book *Multiple Access Techniques for 5G Wireless Networks and Beyond* (Springer, 2019).

Dr. Vaezi is an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE COMMUNICATIONS LETTERS. He has co-organized six NOMA workshops at IEEE VTC 2017-Spring, Globecom'17, 18 and ICC'18, 19, 20. He is a recipient of several academic, leadership, and research awards, including McGill Engineering Doctoral Award, IEEE Larry K. Wilson Regional Student Activities Award in 2013, the Natural Sciences and Engineering Research Council of Canada (NSERC) Postdoctoral Fellowship in 2014, Ministry of Science and ICT of Korea's best paper award in 2017, IEEE Communications Letters Exemplary Editor Award in 2018, and the 2020 IEEE Communications Society Fred W. Ellersick Prize.