

On the Optimal Precoding for MIMO Gaussian Wire-Tap Channels

Arash Khabbazi¹, Maksym A. Girnyk², Sergiy A. Vorobyov^{1,3}, Mikko Vehkaperä^{2,3}, Lars K. Rasmussen²

¹University of Alberta, Department of Electrical and Computer Engineering, Alberta, Canada

email: {khabbazi, svorobyov}@ualberta.ca

²KTH Royal Institute of Technology, School of Electrical Engineering and the ACCESS Linnaeus Center, Sweden

email: {mgyr, lkra}@kth.se

³Aalto University, School of Electrical Engineering, Department of Signal Processing and Acoustics, Finland

email: {mikko.vehkaperä, sergiy.vorobyov}@aalto.fi

Abstract—We consider the problem of finding secrecy rate of a multiple-input multiple-output (MIMO) wire-tap channel. A transmitter, a legitimate receiver, and an eavesdropper are all equipped with multiple antennas. The channel states from the transmitter to the legitimate user and to the eavesdropper are assumed to be known at the transmitter. In this contribution, we address the problem of finding the optimal precoder/transmit covariance matrix maximizing the secrecy rate of the given wire-tap channel. The problem formulation is shown to be equivalent to a difference of convex functions programming problem and an efficient algorithm for addressing this problem is developed.

I. INTRODUCTION

In the recent years, the field of wireless physical layer security has received considerable attention. Due to its inherent randomness, wireless channels can be used for enhancing the secrecy of communication. On the other hand, broadcast nature of a wireless medium creates possibility for illegitimate parties to eavesdrop the transmission. The corresponding basic setup (see Fig. 1), referred to as the *wire-tap channel*, was firstly introduced by Wyner in [1]. The *secrecy rate*, introduced as a performance metric for this setup, reflects the amount of information per channel use that a source can reliably transmit to a destination, provided that an eavesdropper does not get any information. To achieve strictly positive secrecy rates, legitimate parties must have statistically better channel than that of the eavesdroppers'. Therefore, many techniques have been proposed to enhance the performance of secure communication systems, e.g., multiple-input multiple-output (MIMO) communications [2], [3]. As an extension, a scenario with multiple eavesdroppers has been studied in [4], [5].

The *secrecy capacity* is defined as the maximum achievable secrecy rate, and its evaluation is, in general, problematic due to the non-convex nature of the corresponding optimization problem. In the context of MIMO communications, the attempts of finding the secrecy-capacity achieving precoding/covariance matrices have been made already in [4], [5]. Later, in [6], the secrecy capacity has been characterized for

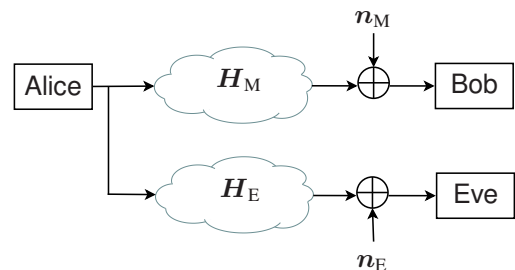


Fig. 1. MIMO wire-tap channel.

the case of channel matrices with certain rank properties. The special case with two-antenna legitimate parties and a single-antenna eavesdropper has been addressed in [2].

Recently, several computationally efficient approaches have been proposed for tackling the problem. For instance, in [7], a beamforming method based on the *generalized singular value decomposition* (GSVD) was proposed. Later, in [8], two other sub-optimal approaches have been studied. A *zero-forcing* (ZF) precoder nulls out the leakage to the eavesdropper, thereby increasing the secrecy rate, whereas the other approach maximizes the *signal-to-leakage-plus-noise* ratio (SLNR). The three approaches outperform each other in various particular settings. Nevertheless, there yet exists no unified approach allowing to compute the secrecy capacity achieving transmit covariance matrix for the general MIMO setting.

In this paper, we study the non-convex problem of secrecy rate maximization under the total power constraint in its most general formulation. A widely used assumption that the channel states from the transmitter to both the legitimate user and eavesdropper are known at the transmitter [4], [7] is adopted here and the corresponding problem is formulated as a difference of convex functions (*DC*) *programming* problem. We then develop an efficient algorithm for addressing the problem. The proposed algorithm is based on the eigenvalue decomposition of the transmit covariance matrix and a subsequent alternate optimization of the eigenvectors and eigenvalues of the matrix. The numerical results show that the proposed method outperforms the other existing alternatives.

II. SYSTEM MODEL AND PROBLEM FORMULATION

The scenario of interest, depicted in Fig. 1, consists of the following two channels: the *main channel* between transmitter (Alice) and legitimate receiver (Bob), and *eavesdropper's channel* between Alice and the eavesdropper (Eve). The corresponding input-output relations are given by

$$\mathbf{y}_M = \mathbf{H}_M \mathbf{x} + \mathbf{n}_M, \quad (1a)$$

$$\mathbf{y}_E = \mathbf{H}_E \mathbf{x} + \mathbf{n}_E, \quad (1b)$$

where $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_M, \mathbf{P})$ is the channel input vector, $\mathbf{n}_M \sim \mathcal{CN}(\mathbf{0}_{N_M}, \mathbf{I}_{N_M})$ and $\mathbf{n}_E \sim \mathcal{CN}(\mathbf{0}_{N_E}, \mathbf{I}_{N_E})$ are additive noise vectors at the receivers of the legitimate user and eavesdropper, respectively. Here \mathbf{I} denotes the identity matrix of the size given by its subscript. The entries of the channel matrices $\mathbf{H}_M \in \mathbb{C}^{N_M \times M}$ and $\mathbf{H}_E \in \mathbb{C}^{N_E \times M}$ are assumed to be independent identically distributed according to $\mathcal{CN}(0, \rho_M/M)$ and $\mathcal{CN}(0, \rho_E/M)$, respectively, where ρ_M and ρ_E represent the corresponding signal-to-noise ratios (SNRs). The channel state information (CSI), consisting of the two above matrices, is assumed to be perfectly known at the transmitter. Furthermore, the total power constraint $\mathbb{E}\{\text{tr}\{\mathbf{x}\mathbf{x}^H\}\} \leq M$ is employed at the transmitter. Here $\mathbb{E}\{\cdot\}$, $\text{tr}\{\cdot\}$ and $(\cdot)^H$ stand for the mathematical expectation, trace of a matrix and Hermitian transpose, respectively.

Based on the available CSI, it is possible to determine the optimal covariance matrix \mathbf{P} that maximizes the achievable secrecy rate, formally given as follows. Let w be a confidential message with entropy $H(w)$, which Alice wants to communicate to Bob, and let $p_{e,n}$ be the probability of error at Bob's receiver. Then a (weak) secrecy rate R_s is achievable if there exists a sequence of $(2^{nR_s}, n)$ codes, such that $p_{e,n} \rightarrow 0$ and $\frac{1}{n}H(w|\mathbf{y}_E) \leq \frac{1}{n}H(w) - \varepsilon_n$, for some ε_n that tends to zero as $n \rightarrow \infty$.

Given CSI $\{\mathbf{H}_M, \mathbf{H}_E\}$ and transmit covariance matrix \mathbf{P} , the achievable secrecy rate is given by [3]

$$R_s = I(\mathbf{x}; \mathbf{y}_M | \mathbf{H}_M) - I(\mathbf{x}; \mathbf{y}_E | \mathbf{H}_E) \quad (2a)$$

$$= \ln \det \left(\mathbf{I}_{N_M} + \mathbf{H}_M \mathbf{P} \mathbf{H}_M^H \right) - \ln \det \left(\mathbf{I}_{N_E} + \mathbf{H}_E \mathbf{P} \mathbf{H}_E^H \right). \quad (2b)$$

The corresponding secrecy capacity is then obtained by

$$C_s = \max_{\substack{\mathbf{P}: \text{tr}\{\mathbf{P}\} \leq M \\ \mathbf{P} \succeq \mathbf{0}_M}} R_s, \quad (3)$$

where $\mathbf{P} \succeq \mathbf{0}_M$ means that \mathbf{P} is a positive semi-definite matrix.

Since the objective function of the above problem is *not concave*, the problem cannot be solved directly with aid of convex optimization tools. The GSVD approach is known to be efficient for addressing the class of DC programming problems that (3) belongs to [7]. In [9], however, we have developed a more general approach, which allows for deriving analytic results about the global optimality. The corresponding algorithm is able to solve a class of DC programming problems in polynomial time and is therefore referred to as *POLynomial*

Time DC (POTDC) method. In [9], the algorithm is applied to the problem of optimal amplification matrix design for a two-way amplify-and-forward network. The terminals therein are equipped with a single antenna and only the relay has multiple antennas, while in (3) all terminals are equipped with multiple antennas, thereby making the problem more difficult. In what follows, we develop a novel approach for addressing problems of type (3).

III. MAIN RESULT

Let $\mathbf{P} = \mathbf{U}^H \mathbf{\Lambda} \mathbf{U}$ be the eigenvalue decomposition of the covariance matrix \mathbf{P} , where $\mathbf{U}_{M \times M}$ is a unitary matrix and $\mathbf{\Lambda}_{M \times M} \triangleq \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_M)$ is a diagonal matrix whose elements are non-negative and correspond to the eigenvalues of the matrix \mathbf{P} . Substituting $\mathbf{P} = \mathbf{U}^H \mathbf{\Lambda} \mathbf{U}$ into (3) results in the following optimization problem

$$\begin{aligned} \max_{\mathbf{U}, \mathbf{\Lambda}} \quad & \ln \det(\mathbf{I}_{N_M} + \mathbf{H}_M \mathbf{U}^H \mathbf{\Lambda} \mathbf{U} \mathbf{H}_M^H) \\ & - \ln \det(\mathbf{I}_{N_E} + \mathbf{H}_E \mathbf{U}^H \mathbf{\Lambda} \mathbf{U} \mathbf{H}_E^H) \\ \text{s.t.} \quad & \text{tr}\{\mathbf{\Lambda}\} \leq M, \mathbf{U}^H \mathbf{U} = \mathbf{I}_M, \lambda_i \geq 0, i = 1, \dots, M \end{aligned} \quad (4)$$

Let us now use Sylvester's determinant identity, which implies that for any arbitrary matrices $\mathbf{A}_{M \times N}$ and $\mathbf{B}_{N \times M}$, the following equality holds [10]

$$\det(\mathbf{I}_M + \mathbf{A}\mathbf{B}) = \det(\mathbf{I}_N + \mathbf{B}\mathbf{A}). \quad (5)$$

Thus, (4) can be equivalently written as

$$\begin{aligned} \max_{\mathbf{U}, \mathbf{\Lambda}} \quad & \ln \det(\mathbf{I}_{N_M} + \mathbf{H}_M \mathbf{U}^H \mathbf{\Lambda} \mathbf{U} \mathbf{H}_M^H) - \ln \det(\mathbf{I}_M + \mathbf{D}(\mathbf{U})\mathbf{\Lambda}) \\ \text{s.t.} \quad & \text{tr}\{\mathbf{\Lambda}\} \leq M, \mathbf{U}^H \mathbf{U} = \mathbf{I}_M, \lambda_i \geq 0, i = 1, \dots, M \end{aligned} \quad (6)$$

where $\mathbf{D}(\mathbf{U}) \triangleq \mathbf{U} \mathbf{H}_E^H \mathbf{H}_E \mathbf{U}^H$ is defined for notation simplicity. By a careful inspection of problem (6), one can readily observe that it is a DC programming problem over the matrix $\mathbf{\Lambda}$ for a fixed value of \mathbf{U} . It can be addressed using the POTDC algorithm [9]. Moreover, for a fixed $\mathbf{\Lambda}$, (6) is an optimization problem over unitary matrices that has been comprehensively studied in the literature (see, e.g., [11]). Based on the latter fact, (6) can be addressed via alternating optimization, by first optimizing with respect to $\mathbf{\Lambda}$ for a fixed value of \mathbf{U} , and then further optimizing with respect to \mathbf{U} with $\mathbf{\Lambda}$ being set to the optimal value from the previous iteration. These alternations continue until convergence as it is explained later in the paper.

Optimization problem (6), when \mathbf{U} is fixed to \mathbf{U}_0 , can be expressed as

$$\begin{aligned} \max_{\mathbf{\Lambda}} \quad & \ln \det(\mathbf{I}_{N_M} + \mathbf{H}_M \mathbf{U}_0^H \mathbf{\Lambda} \mathbf{U}_0 \mathbf{H}_M^H) \\ & - \ln \det(\mathbf{I}_M + \mathbf{D}(\mathbf{U}_0)\mathbf{\Lambda}) \\ \text{s.t.} \quad & \text{tr}\{\mathbf{\Lambda}\} \leq M, \lambda_i \geq 0, i = 1, \dots, M. \end{aligned} \quad (7)$$

The resulting problem (7) is addressed via the POTDC algorithm, which is efficiently used for the DC programming problems, whose non-concave part is a function of a single variable. Since the non-concave part of problem (7) is not a function of a single variable, we instead maximize a lower bound on the objective function of (7), whose non-concave

parts depend on a single variable each. For this goal, we utilize the Hadamard's inequality [12], which implies that the determinant of a Hermitian positive semidefinite matrix is upper-bounded by the product of its diagonal elements. By applying this inequality, the optimization problem of maximizing the lower bound of the objective function of (7) can be recast as

$$\begin{aligned} \max_{\mathbf{\Lambda}} \quad & \ln \det(\mathbf{I}_{N_M} + \mathbf{H}_M \mathbf{U}_0^H \mathbf{\Lambda} \mathbf{U}_0 \mathbf{H}_M^H) \\ & - \sum_{i=1}^M \ln(1 + [\mathbf{D}(\mathbf{U}_0)]_{(i,i)} \lambda_i) \\ \text{s.t.} \quad & \text{tr}\{\mathbf{\Lambda}\} \leq M, \lambda_i \geq 0, i = 1, \dots, M, \end{aligned} \quad (8)$$

where $[\mathbf{D}(\mathbf{U}_0)]_{(i,j)}$ denotes the element of the matrix $\mathbf{D}(\mathbf{U}_0)$ in i th row and j th column. All the terms of the objective function and the constraint functions of (8) are concave with respect to $\mathbf{\Lambda}$ except for the terms $-\ln(1 + [\mathbf{D}(\mathbf{U}_0)]_{(i,i)} \lambda_i)$, $i = 1, \dots, M$, which are convex, and each of which depends only on a single variable. These convex constraints can be handled iteratively in terms of their linear approximation around suitably selected points. Algorithm 1 shows how the POTDC approach can be used for addressing (8).

Algorithm 1 The iterative POTDC algorithm

Require: Choose the initialization point $\lambda_{i,c} \in [0, M]$, $i = 1, \dots, M$ such that $\sum_{i=1}^M \lambda_{i,c} = M$.
Select the termination threshold ζ_1 ,
and set i equal to 1.

repeat

Solve the following convex optimization problem using $\lambda_{i,c}$, $i = 1, \dots, M$ to obtain $\lambda_{i,\text{opt}}$, $i = 1, \dots, M$

$$\begin{aligned} \max_{\mathbf{\Lambda}} \quad & \ln \det(\mathbf{I}_{N_M} + \mathbf{H}_M \mathbf{U}_0^H \mathbf{\Lambda} \mathbf{U}_0 \mathbf{H}_M^H) \\ & - \sum_{i=1}^M \frac{[\mathbf{D}(\mathbf{U}_0)]_{(i,i)} (\lambda_i - \lambda_{i,c})}{1 + [\mathbf{D}(\mathbf{U}_0)]_{(i,i)} \lambda_{i,c}} \\ \text{s.t.} \quad & \text{tr}\{\mathbf{\Lambda}\} \leq M, \lambda_i \geq 0, i = 1, \dots, M \end{aligned} \quad (9)$$

and set

$$\begin{aligned} \mathbf{\Lambda}_{\text{opt}} & \leftarrow \mathbf{\Lambda}_{\text{opt},i} = \text{diag}(\lambda_{1,\text{opt}}, \dots, \lambda_{M,\text{opt}}), \\ \lambda_{i,c} & \leftarrow \lambda_{i,\text{opt}}, \quad i \leftarrow i + 1 \end{aligned}$$

until the difference between two objective values in consecutive iterations is less than or equal to the termination threshold ζ_1 .

The POTDC algorithm, as applied to (8), is guaranteed to convergence to a point which satisfies the Karush-Kuhn-Tucker (KKT) optimality conditions. Moreover, the value of the objective function is guaranteed to be non-decreasing over the iterations of the algorithm [9].

In the next step, we address (7) with respect to \mathbf{U} , when $\mathbf{\Lambda}$ is fixed to $\mathbf{\Lambda}_0 = \mathbf{\Lambda}_{\text{opt}}$ obtained from Algorithm 1. The corresponding optimization problem can be written as

$$\begin{aligned} \max_{\mathbf{U}} \quad & \ln \frac{\det(\mathbf{I}_{N_M} + \mathbf{H}_M \mathbf{U}^H \mathbf{\Lambda}_0 \mathbf{U} \mathbf{H}_M^H)}{\det(\mathbf{I}_{N_E} + \mathbf{H}_E \mathbf{U}^H \mathbf{\Lambda}_0 \mathbf{U} \mathbf{H}_E^H)} \\ \text{s.t.} \quad & \mathbf{U}^H \mathbf{U} = \mathbf{I}_M. \end{aligned} \quad (10)$$

This is an optimization problem over the complex-valued matrix \mathbf{U} under the constraint that \mathbf{U} is a unitary matrix. In order to address this problem, we adopt the steepest descent algorithm on the Lie group of $M \times M$ unitary matrices developed in [11]. This algorithm is shown to move towards the optimal point over the iterations. To apply this method, the gradient of the objective function of (10) with respect to the complex-valued matrix \mathbf{U} is required. This gradient can be easily derived as

$$\begin{aligned} \nabla_{\mathbf{U}} \ln \frac{\det(\mathbf{I}_{N_M} + \mathbf{H}_M \mathbf{U}^H \mathbf{\Lambda}_0 \mathbf{U} \mathbf{H}_M^H)}{\det(\mathbf{I}_{N_E} + \mathbf{H}_E \mathbf{U}^H \mathbf{\Lambda}_0 \mathbf{U} \mathbf{H}_E^H)} \\ = \mathbf{\Lambda}_0 \mathbf{U} \mathbf{H}_M^H (\mathbf{I}_{N_M} + \mathbf{H}_M \mathbf{U}^H \mathbf{\Lambda}_0 \mathbf{U} \mathbf{H}_M^H)^{-1} \mathbf{H}_M \\ - \mathbf{\Lambda}_0 \mathbf{U} \mathbf{H}_E^H (\mathbf{I}_{N_E} + \mathbf{H}_E \mathbf{U}^H \mathbf{\Lambda}_0 \mathbf{U} \mathbf{H}_E^H)^{-1} \mathbf{H}_E, \end{aligned} \quad (11)$$

where $\nabla(\cdot)$ stands for the gradient operator.

The overall algorithm for addressing the precoding matrix design problem for MIMO Gaussian wire-tap channels can be then described in terms of Algorithm 2.

Algorithm 2 The algorithm for precoding design for MIMO Gaussian wire-tap channels

Require: Choose arbitrary $\mathbf{P}_c = \mathbf{U}_c \mathbf{\Lambda}_c \mathbf{U}_c^H \succeq \mathbf{0}$ such that $\text{tr}\{\mathbf{P}_c\} = M$.
Choose the termination threshold ζ_2 , and
set k equal to 1.

repeat

- i) Solve (8) using Algorithm 1 for $\mathbf{U}_0 = \mathbf{U}_c$ and $\mathbf{\Lambda}_c$ as the initialization point.
 $\mathbf{\Lambda}_c \leftarrow \mathbf{\Lambda}_{\text{opt}}$
- ii) Solve (10) for $\mathbf{\Lambda}_0 = \mathbf{\Lambda}_c$ to obtain \mathbf{U}_{opt}
 $\mathbf{U}_c \leftarrow \mathbf{U}_{\text{opt}}$, $k \leftarrow k + 1$

until the difference between two objective values in consecutive iterations is less than or equal to the termination threshold ζ_2 .

It is noteworthy to mention that both steps in Algorithm 2, corresponding to optimization problems (8) and (10), result in a non-decreasing value of the objective function. A detailed convergence analysis and optimality of the proposed algorithm is left for the future work.

IV. NUMERICAL EXAMPLES

In this section, we present the numerical results illustrating our theoretical findings. The channel matrices \mathbf{H}_M and \mathbf{H}_E are assumed to be quasi-static flat Rayleigh fading. The displayed results are averaged over 500 independent channel realizations. The initialization point for the proposed algorithm, \mathbf{P}_c , (see Algorithm 2) is chosen randomly.

For the first scenario, we set $\rho_M = \rho_E = \rho$ with $M = N_E = 2$ and $N_M = 1$. The corresponding wire-tap channel is referred to as multiple-input single-output multi-eavesdropper (MISOME) channel and its secrecy capacity has been characterized analytically in [4], setting a benchmark for the approach proposed in this paper. Fig. 2 compares the

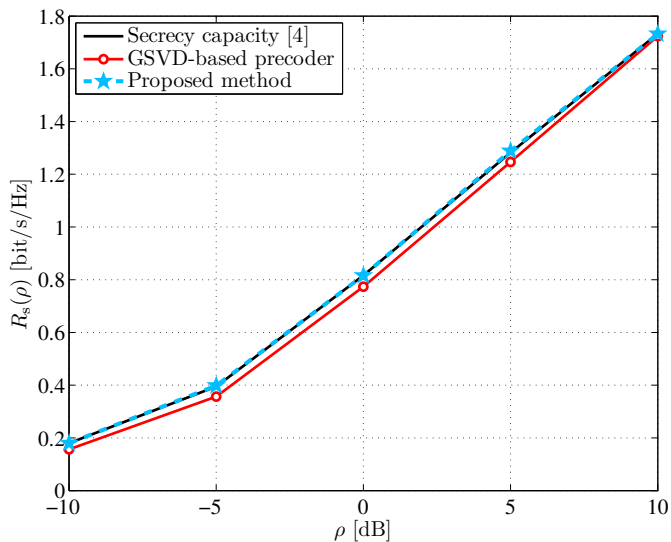


Fig. 2. Achievable secrecy rates and the secrecy capacity as functions of SNR $\rho_M = \rho_E = \rho$ for $M = N_E = 2$ and $N_M = 1$.

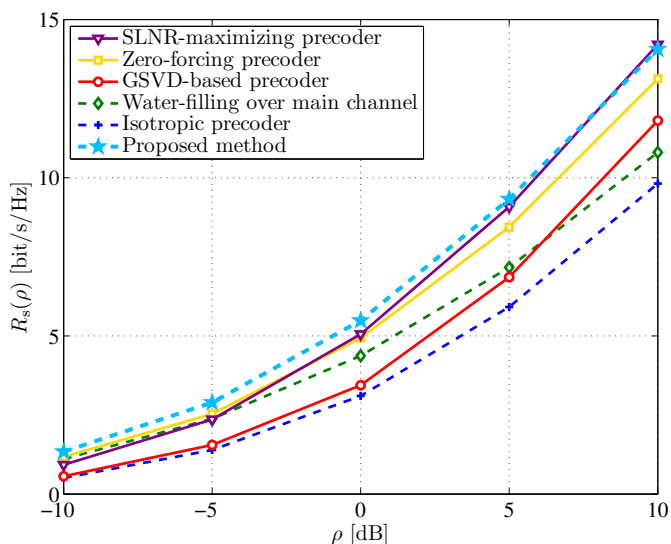


Fig. 3. Achievable secrecy rates as functions of SNR $\rho_M = \rho_E = \rho$ for $M = N_M = 6$ and $N_E = 2$.

secrecy capacity of the MISOME channel with the achievable secrecy rates obtained by the GSVD-based beamforming [7] and the proposed POTDC method. It can be seen from the figure that the proposed precoder outperforms the GSVD-based method and approaches the secrecy capacity derived in [4] for the given setup.

For the second scenario, we set $M = N_M = 6$ and $N_E = 2$. Since the capacity is not known for this setting, we compare our method to the existing alternatives. Thus, Fig. 3 compares the secrecy rates achievable by the proposed precoder with those, obtained by the ZF and SLNR-maximizing precoders

from [8]. For illustration, we also add conventional water-filling over the main channel (see, e.g., [13]), as well as an isotropic precoder. It can be seen from the figure that the latter two are poor strategies in a wire-tap setting. Then, most importantly, we observe that the proposed approach outperforms the rest of the strategies.

V. CONCLUSIONS

In this paper, we considered the problem of finding secrecy capacity of a MIMO wire-tap channel. It has been assumed that the transmitter, the legitimate receiver, and also the eavesdropper are all equipped with multiple antennas while the transmitter knows the channel states from itself to the legitimate receiver and eavesdropper perfectly. We have developed a novel method for maximizing the secrecy rate of the aforementioned wire-tap channel. The proposed algorithm optimizes the eigenvalues and eigenvectors of the precoding/transmit covariance matrix in alternate manner. Our numerical results confirm the superiority of the proposed method over the other state-of-the-art methods.

REFERENCES

- [1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1334-1387, 1975.
- [2] S. Shafiq, 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, Sept. 2009.
- [3] 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.
- [4] A. Khisti and G. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088-3104, Jul. 2010.
- [5] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961-4972, Aug. 2011.
- [6] J. Li and A. Petropulu, "Optimal input covariance for achieving secrecy capacity in Gaussian MIMO wiretap channels," in *Proc. IEEE Int. Conf. Acoust. Speech Sig. Process.*, Dallas, Texas, USA, Mar. 2010, pp. 3362-3365.
- [7] S. 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*, Cambridge, MA, USA, Jul. 2012, pp. 2321-2325.
- [8] K. Wang, X. Wang, and X. Zhang, "SLNR-based transmit beamforming for MIMO wiretap channel," *Wireless Pers. Commun.*, pp. 1-13, 2012.
- [9] A. Khabbazi-basmenj, F. Roemer, S. A. Vorobyov, and M. Haardt, "Sum-rate maximization in two-way AF MIMO relaying: Polynomial time solutions to a class of DC programming problems," *IEEE Trans. Sig. Process.*, vol. 60, no. 10, pp. 5478-5493, Oct. 2012.
- [10] A. G. Akritas, E. K. Akritas, and G. I. Malaschonok, "Various proofs of Sylvester's (determinant) identity," *Mathematics and Computers in Simulation*, vol. 42, no. 4, pp. 585-593, 1996.
- [11] T. E. Abruđan, J. Eriksson, and V. Koivunen, "Steepest descent algorithms for optimization under unitary matrix constraint," *IEEE Trans. Sig. Process.*, vol. 56, no. 3, pp. 1134-1147, Mar. 2008.
- [12] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1988.
- [13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.