AN ALTERNATING OPTIMIZATION ALGORITHM FOR THE MIMO SECRECY CAPACITY PROBLEM UNDER SUM POWER AND PER-ANTENNA POWER CONSTRAINTS

Qiang Li[†]*, Mingyi Hong*[‡]*, Hoi-To Wai*[†]*, Wing-Kin Ma*[†]*, Ya-Feng Liu*[§] *and Zhi-Quan Luo*[‡]

[†]Dept. of Electronic Engineering, The Chinese University of Hong Kong, Shatin N.T., Hong Kong
[‡]Dept. of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA
[§]State Key Lab. of Sci. and Eng. Computing, Chinese Academy of Sciences, Beijing 100190, China

ABSTRACT

This paper considers transmit covariance optimization for a multi-input multi-output (MIMO) Gaussian wiretap channel. Specifically, we aim to maximize the MIMO secrecy capacity by judiciously designing the transmit covariance under the sum power and per-antenna power constraints. The MIMO secrecy capacity maximization (SCM) problem is nonconvex, and so far there is no tractable solution available. We propose an alternating optimization (AO) approach to handle the SCM problem. In particular, our development consists of two steps: First, we show that the SCM problem can be reexpressed to a form that can be conveniently processed by AO. Second, we develop a custom-designed fast algorithm for each AO iteration. Interestingly, with this fast implementation, the overall AO algorithm can be viewed as performing iterative reweighting and water-filling. Finally, the convergence of the proposed algorithm to a stationary solution of SCM is shown, and numerical results are provided to demonstrate its efficacy.

Index Terms— secrecy capacity, alternating optimization, water-filling, Per-antenna power constraints

1. INTRODUCTION

Recently, *physical-layer secrecy*, a means of providing confidentiality at the physical layer, has received considerable attention. In contrast to the cryptographic approach, physical-layer secrecy has at least two advantages, namely, provably perfect security and no need for encryption keys. The latter makes physical-layer secrecy very attractive for wireless applications, since the open nature of the wireless medium makes the encryption key more vulnerable to eavesdropping and impersonation attack [1,2]. Among the physical-layer secrecy studies, the multi-input multi-output (MIMO) Gaussian wiretap channel is of particular interest [3–5], because this fundamental wiretap model has its importance in understanding the use of multiple antennas for enhancing secrecy. Several concurrent works [3–5] have shown that the secrecy capacity of the MIMO Gaussian wiretap channel under the *sum power constraint* is given by

$$C_s^\star = \max_{\mathbf{W}\succeq\mathbf{0},\mathrm{Tr}(\mathbf{W})\le P_0} \ \ln|\mathbf{I}+\mathbf{H}^H\mathbf{W}\mathbf{H}| - \ln|\mathbf{I}+\mathbf{G}^H\mathbf{W}\mathbf{G}|, \quad (1)$$

where P_0 is a given total transmit budget; $\mathbf{W}$ is the covariance matrix of the transmit signal, or simply the transmit covariance; $\mathbf{H}/\mathbf{G}$ represents the MIMO legitimate / eavesdropping channel, respectively (resp.). A more detailed model description will be provided in the next section. In the sequel, we will call problem (1) the MIMO *secrecy capacity maximization* (SCM) problem. The SCM problem (1) is nonconvex, and so far there is no efficient, tractable solution to it for general $\mathbf{H}$ and $\mathbf{G}$. As a compromise, some suboptimal designs have been proposed, e.g., the fix-point iteration design [6], and generalized singular value decomposition (GSVD) precoding [4, 7].

In this paper, we consider a more general form of SCM, where *per-antenna power constraints* (PAPC) are incorporated into the SCM problem (1). This SCM-PAPC design is motivated by the fact that each antenna is often equipped with its own power amplifier (PA). In order to operate within the linear region of each PA, one may want to limit the per-antenna peak power [8]. Owing to the per-antenna power constraints, the SCM-PAPC problem turns out to be more difficult than SCM. In this paper, we propose an efficient optimization algorithm for SCM-PAPC. The proposed algorithm is based on an equivalent reformulation of SCM-PAPC and an alternating optimization (AO) methodology. In particular, we show that SCM-PAPC can be equivalently expressed to a form that can be conveniently processed by AO. Moreover, a custom-derived water-filling-like solution is developed for each AO iteration. In addition, we also prove that the proposed algorithm is guaranteed to converge to a stationary solution of SCM-PAPC.

1.1. Relation to Prior Work

This paper considers MIMO secrecy capacity maximization under the sum power and per-antenna power constraints. Similar problems have been investigated in [4–7, 9], but the works [4, 6, 7, 9] focus on the sum power constraint only, while [5] considers a different matrix covariance constraint, i.e., by replacing $\mathrm{Tr}(\mathbf{W}) \le P_0$ in (1) with $\mathbf{W} \preceq \mathbf{M}$ for some given positive semidefinite matrix $\mathbf{M}$. In such a case, the SCM problem has a closed-form optimal solution [5]. However, this closed-form design does not apply to the sum power constraint or the per-antenna power constraints. Another notable difference from [4–7] is the way that we handle the SCM-PAPC problem. Specifically, in [6] Li and Petropulu developed a fixed-point iteration to SCM by exploiting the Karush-Kuhn-Tucker (KKT) optimality conditions. In [4, 7], a generalized singular value decomposition (GSVD) precoding is applied to SCM by prefixing the transmit covariance structure through GSVD, and then performing power allocation on each parallelized wiretap channel. Here, we take a different approach. We first derive an equivalent formulation of SCM-PAPC. Based on this new formulation, an alternating optimization approach is developed to obtain an efficient design for SCM-PAPC in an iterative water-filling-like manner.

2. SYSTEM MODEL AND PROBLEM STATEMENT

Consider the aforementioned three-terminal MIMO Gaussian wiretap model, which consists of a transmitter, a legitimate receiver and

This work is supported by a Direct Grant by the Chinese University of Hong Kong (Project ID: 2050489).

an eavesdropper. All the terminals are equipped with multiple antennas. For ease of the subsequent description, we will call the transmitter, the legitimate receiver and the eavesdropper *Alice*, *Bob* and *Eve*, resp. Assuming quasi-static frequency-flat fading channels for all the communication links, the received signals at Bob and Eve may be modeled as

$$\mathbf{y}_b(t) = \mathbf{H}^H \mathbf{s}(t) + \mathbf{n}_b(t), \tag{2a}$$

$$\mathbf{y}_e(t) = \mathbf{G}^H \mathbf{s}(t) + \mathbf{n}_e(t), \tag{2b}$$

resp., where $\mathbf{H} \in \mathbb{C}^{N_t \times N_b}/\mathbf{G} \in \mathbb{C}^{N_t \times N_e}$ represent the MIMO channels from Alice to Bob/Eve, resp.; N_t, N_b and N_e are the number of antennas employed by Alice, Bob and Eve, resp.; $\mathbf{n}_b(t)$ and $\mathbf{n}_e(t)$ are i.i.d. complex Gaussian noise with zero mean and unit variance; $\mathbf{s}(t) \in \mathbb{C}^{N_t}$ is the coded confidential information intended for Bob.

The problem of interest here is the MIMO secrecy capacity maximization under the sum power and per-antenna power constraints (SCM-PAPC), that is [5]

$$\begin{aligned} C_s^\star = \max_{\mathbf{W} \succeq \mathbf{0}} \; & C_s(\mathbf{W}) && \text{(3a)} \\ (\mathsf{SCM-PAPC}) \quad \text{s.t. } & \mathrm{Tr}(\mathbf{W}) \le P_0, && \text{(3b)} \\ & [\mathbf{W}]_{ii} \le P_i,\ i = 1, \ldots, N_t, && \text{(3c)} \end{aligned}$$

where $P_0 > 0$ and $P_i > 0$ for all i are given sum power and per-antenna power limits, resp.; $\mathbf{W} = \mathrm{E}\{\mathbf{s}(t)\mathbf{s}(t)^H\}$ denotes the covariance of $\mathbf{s}(t)$; $[\mathbf{W}]_{ii}$ is the ith diagonal element of $\mathbf{W}$; and

$$C_s(\mathbf{W}) = \ln|\mathbf{I} + \mathbf{H}^H\mathbf{W}\mathbf{H}| - \ln|\mathbf{I} + \mathbf{G}^H\mathbf{W}\mathbf{G}|,$$

which is the mutual information difference between the Alice-to-Bob and Alice-to-Eve channels. Readers are referred to [1, 10] for more details about the notion of MIMO physical-layer secrecy. Here (3c) represents the per-antenna power constraints by imposing an upper bound on each diagonal entry of the transmit covariance [8].

The SCM-PAPC problem (3) is a nonconvex problem. Under the sum power constraint only, there exist special cases where problem (3) is solvable, namely, when $N_b = 1$ [11], $N_e = 1$ [12], or $\mathbf{H}\mathbf{H}^H \succ \mathbf{G}\mathbf{G}^H$ [9]. However, to the best of our knowledge, there is no tractable solution for (3) in general. As mentioned in Introduction, several concurrent endeavors consider suboptimal, but easy-to-implement, solutions for (3) [4, 6, 7]. In the next section, we will propose a different solution to problem (3) using an alternating optimization (AO) approach. The advantages of the proposed approach will become clear in the subsequent development.

3. AN ALTERNATING OPTIMIZATION APPROACH TO SCM-PAPC

3.1. An Equivalent Formulation of Problem (3) for AO

To describe our approach, we need to reexpress problem (3) to a form that can be conveniently processed by alternating optimization (AO). The following lemma will serve our purpose.

Lemma 1 ([13]). *Let $\mathbf{E} \in \mathbb{C}^{N \times N}$ be any matrix such that $\mathbf{E} \succ \mathbf{0}$. Consider the function $f(\mathbf{S}) = -\mathrm{Tr}(\mathbf{S}\mathbf{E}) + \ln|\mathbf{S}| + N$. Then,*

$$\ln|\mathbf{E}^{-1}| = \max_{\mathbf{S} \in \mathbb{C}^{N \times N}, \mathbf{S} \succeq \mathbf{0}} f(\mathbf{S}), \tag{4}$$

and the optimal solution to the right-hand side of (4) *is $\mathbf{S}^\star = \mathbf{E}^{-1}$.*

By applying Lemma 1 to problem (3) via setting $\mathbf{E} = \mathbf{I} + \mathbf{G}^H\mathbf{W}\mathbf{G}$, we obtain an equivalent formulation of problem (3) as follows

$$\begin{aligned} \max_{\mathbf{W},\mathbf{S}} \; & \ln|\mathbf{I} + \mathbf{H}^H\mathbf{W}\mathbf{H}| - \mathrm{Tr}(\mathbf{S}(\mathbf{I} + \mathbf{G}^H\mathbf{W}\mathbf{G})) + \ln|\mathbf{S}| \\ \text{s.t. } & \mathrm{Tr}(\mathbf{A}_i\mathbf{W}) \le P_i,\ \forall i \in \mathcal{I},\ \mathbf{W} \succeq \mathbf{0},\ \mathbf{S} \succeq \mathbf{0}, \end{aligned} \tag{5}$$

where $\mathcal{I} \triangleq \{0, 1, \ldots, N_t\}$, $\mathbf{A}_0 = \mathbf{I}$, $\mathbf{A}_i = \boldsymbol{e}_i\boldsymbol{e}_i^H, \forall i \in \mathcal{I}\backslash\{0\}$ and $\boldsymbol{e}_i$ is a unit vector with the ith entry being one. While the SCM-PAPC equivalent problem (5) is still nonconvex with respect to (w.r.t.) both $\mathbf{W}$ and $\mathbf{S}$, the advantage of the reformulation is that *fixing either $\mathbf{W}$ or $\mathbf{S}$, problem* (5) *is convex w.r.t. the other decision variable.* This coordinate-wise convexity property naturally leads to an alternating optimization for problem (5). Specifically, let $(\mathbf{W}^n, \mathbf{S}^n)$ be the AO iterate at the nth iteration. We alternatingly solve the following two optimization problems to obtain $(\mathbf{W}^n, \mathbf{S}^n)$ for $n = 1, 2, \ldots$

$$\mathbf{S}^n = \arg\max_{\mathbf{S} \succeq \mathbf{0}} \; \ln|\mathbf{S}| - \mathrm{Tr}(\mathbf{S}(\mathbf{I} + \mathbf{G}^H\mathbf{W}^{n-1}\mathbf{G})), \tag{6a}$$

$$\begin{aligned} \mathbf{W}^n = \arg\max_{\mathbf{W} \succeq \mathbf{0}} \; & \ln|\mathbf{I} + \mathbf{H}^H\mathbf{W}\mathbf{H}| - \mathrm{Tr}(\mathbf{G}\mathbf{S}^n\mathbf{G}^H\mathbf{W}) \\ \text{s.t. } & \mathrm{Tr}(\mathbf{A}_i\mathbf{W}) \le P_i,\ \forall i \in \mathcal{I}. \end{aligned} \tag{6b}$$

For problem (6a), the optimal solution can be computed in closed form by Lemma 1, that is

$$\mathbf{S}^n = (\mathbf{I} + \mathbf{G}^H\mathbf{W}^{n-1}\mathbf{G})^{-1}. \tag{7}$$

Moreover, in the next subsection, we will describe an efficient way to compute the optimal solution of (6b). At this point, we should point out the insight of the above AO iteration: The problem (6b) is reminiscent of the MIMO capacity maximization problem, except for an additional Eve-induced penalty term $\mathrm{Tr}(\mathbf{G}\mathbf{S}^n\mathbf{G}^H\mathbf{W})$. Intuitively, this penalty term plays a role in degrading Eve's reception, thereby achieving a balance between maximizing Bob's channel capacity and suppressing signal leakage to Eve. Moreover, as seen from (6b) and (7), this penalty term will be adaptively updated according to the previous transmit covariance $\mathbf{W}^{n-1}$.

As a basic property of AO, the AO iterations yield a nondecreasing sequence of the objective values of SCM-PAPC; i.e., $C_s(\mathbf{W}^n) \ge C_s(\mathbf{W}^{n-1}) \ge \ldots \ge C_s(\mathbf{W}^0)$. Moreover, we have the following convergence result on the iterate $\mathbf{W}^n$:

Proposition 1. *Every limit point $\bar{\mathbf{W}}$ of the iterates $\{\mathbf{W}^n\}$ generated by the AO process in* (6a)-(6b) *is a stationary point of the SCM-PAPC problem* (3)*.*

The proof of Proposition 1 is given in the Appendix. A key ingredient in proving Proposition 1 is to apply a specific block coordinate descent (BCD) convergence result for the two blocks case; see [14, Corollary 2].

3.2. An Iterative Water-filling-like Algorithm for Problem (6b)

In this subsection, we focus on developing an efficiently computable solution to (6b). Our idea is to solve the dual of problem (6b). There are two reasons for doing so. First, one can easily verify that strong duality holds for problem (6b), and thus it suffices to consider the dual of problem (6b). Second, as we will show shortly, the specific dual problem structure allows us to compute a primal-dual optimal pair in an iterative water-filling-like manner, thereby possessing low per-iteration complexity.

Consider the Lagrangian dual of problem (6b), that is

$$\begin{aligned} &\min_{\boldsymbol{\lambda}} f(\boldsymbol{\lambda}) \\ &\text{s.t. } \lambda_i \geq 0,\ \forall i \in \mathcal{I} \end{aligned} \tag{8}$$

where $\lambda_i \geq 0$ is the dual variable associated with $\mathrm{Tr}(\mathbf{A}_i\mathbf{W}) \leq P_i$ for all $i \in \mathcal{I}$, and

$$f(\boldsymbol{\lambda}) \triangleq \max_{\mathbf{W}\succeq\mathbf{0}} \mathcal{L}(\boldsymbol{\lambda}, \mathbf{W}), \tag{9}$$

$$\begin{aligned} &\mathcal{L}(\boldsymbol{\lambda}, \mathbf{W}) \\ &= \ln|\mathbf{I} + \mathbf{H}^H\mathbf{W}\mathbf{H}| + \sum_{i\in\mathcal{I}} \lambda_i P_i - \mathrm{Tr}((\mathbf{G}\mathbf{S}^n\mathbf{G}^H + \sum_{i\in\mathcal{I}} \lambda_i\mathbf{A}_i)\mathbf{W}). \end{aligned}$$

We apply a standard dual descent method (or more commonly known as the dual ascent method when the primal problem is in the minimization form [15]) to problem (8). Specifically, let $\boldsymbol{\lambda}^{n,l}$ denote the lth iterate of the dual descent searches for (8). Then the iterates $\{\boldsymbol{\lambda}^{n,l}\}$ are generated by the following projected subgradient (PSG) update formula

$$\boldsymbol{\lambda}^{n,l+1} = \left[\boldsymbol{\lambda}^{n,l} - \alpha^{n,l}\boldsymbol{\delta}_{\boldsymbol{\lambda}^{n,l}}\right]^+, \quad l = 1, 2, \ldots \tag{10}$$

where $[\cdot]^+$ denotes an elementwise projection onto the set of nonnegative numbers; $\{\alpha^{n,l}\}$ is a step-size sequence; and $\boldsymbol{\delta}_{\boldsymbol{\lambda}^{n,l}}$ denotes a subgradient of $f(\boldsymbol{\lambda})$ at the point $\boldsymbol{\lambda}^{n,l}$. Readers are referred to the optimization literature [15] regarding other general operational details of the dual descent method. Here, we are concerned with one crucial component of the iteration— calculation of the subgradient $\boldsymbol{\delta}_{\boldsymbol{\lambda}^{n,l}}$. According to a standard result of the dual descent method, the subgradient $\boldsymbol{\delta}_{\boldsymbol{\lambda}^{n,l}}$ takes the following form (cf. [15, Section 6.1])

$$\boldsymbol{\delta}_{\boldsymbol{\lambda}^{n,l}} = [P_0 - \mathrm{Tr}(\mathbf{W}^\star_{\boldsymbol{\lambda}^{n,l}}\mathbf{A}_0), \ldots, P_{N_t} - \mathrm{Tr}(\mathbf{W}^\star_{\boldsymbol{\lambda}^{n,l}}\mathbf{A}_{N_t})]^T, \tag{11}$$

where $\mathbf{W}^\star_{\boldsymbol{\lambda}^{n,l}}$ denotes an optimal solution of problem (9) when fixing $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{n,l}$. Apparently, efficiently computing $\mathbf{W}^\star_{\boldsymbol{\lambda}^{n,l}}$ is the key to achieving low-complexity iterations in (10). According to the results in [16,17], the optimal solution of problem (9) can be computed in closed form, as summarized in the following lemma:

Lemma 2 ([16, 17])**.** *The optimal solution* $\mathbf{W}^\star_{\boldsymbol{\lambda}}$ *of problem* (9) *is given by*

$$\mathbf{W}^\star_{\boldsymbol{\lambda}} = \mathbf{R}^{-H}\mathbf{U}\mathbf{D}\mathbf{U}^H\mathbf{R}^{-1} \tag{12}$$

where $\mathbf{R}$ *is a square-root factorization of* $\mathbf{G}\mathbf{S}^n\mathbf{G}^H + \sum_{i\in\mathcal{I}}\lambda_i\mathbf{A}_i$ [1]*, i.e.,* $\mathbf{R}\mathbf{R}^H = \mathbf{G}\mathbf{S}^n\mathbf{G}^H + \sum_{i\in\mathcal{I}}\lambda_i\mathbf{A}_i$*;* $\mathbf{U} \in \mathbb{C}^{N_t\times N_t}$ *is the left singular vectors of* $\mathbf{R}^{-1}\mathbf{H}$*;* $\mathbf{D} = \mathrm{Diag}\,(d_1, \ldots, d_r, 0, \ldots, 0) \in \mathbb{R}^{N_t\times N_t}$ *with*

$$d_i = \left[1 - 1/\sigma_i^2\right]^+,\ i = 1, \ldots, r,$$

r being the rank of $\mathbf{R}^{-1}\mathbf{H}$*, and* $\sigma_i > 0, i = 1, \ldots, r$*, being the positive singular values of* $\mathbf{R}^{-1}\mathbf{H}$*.*

Note that $\mathbf{W}^\star_{\boldsymbol{\lambda}}$ in (12) exhibits a similar form as the classic waterfilling solution, except that an additional prewhitening of $\mathbf{G}\mathbf{S}^n\mathbf{G}^H + \sum_{i\in\mathcal{I}}\lambda_i\mathbf{A}_i$ is needed. We summarize the PSG iteration in (10), together with the whole AO process in Algorithm 1.

[1]Without loss of generality, we can assume $\mathbf{G}\mathbf{S}^n\mathbf{G}^H + \sum_{i\in\mathcal{I}}\lambda_i\mathbf{A}_i \succ \mathbf{0}$, for otherwise $\mathbf{G}\mathbf{S}^n\mathbf{G}^H + \sum_{i\in\mathcal{I}}\lambda_i\mathbf{A}_i$ must be rank-deficient, and we can always find a vector $\boldsymbol{w} \neq \mathbf{0}$ such that $(\mathbf{G}\mathbf{S}^n\mathbf{G}^H + \sum_{i\in\mathcal{I}}\lambda_i\mathbf{A}_i)\boldsymbol{w} = \mathbf{0}$. On the other hand, if $\mathbf{H}$ is a randomly generated channel matrix, it follows that $\mathbf{H}^H\boldsymbol{w} \neq \mathbf{0}$ with probability 1. Therefore, $\mathcal{L}(0, \alpha\boldsymbol{w}\boldsymbol{w}^H)$ will be unbounded above as $\alpha \to \infty$, which is apparently infeasible for the dual problem (8).

Algorithm 1 AO Algorithm for SCM-PAPC (3)

1: Initialize $n = 1$, $\epsilon > 0$, $\boldsymbol{\lambda}^0 \geq \mathbf{0}$, and $\mathbf{W}^0 \succeq \mathbf{0}$ such that $\mathrm{Tr}(\mathbf{A}_i\mathbf{W}^0) \leq P_i, \forall i \in \mathcal{I}$ and a maximum number of PSG iterations $L > 0$;
2: **while** $|C_s(\mathbf{W}^n)) - C_s(\mathbf{W}^{n-1})| > \epsilon$ **do**
3: Update $\mathbf{S}^n$ according to (7);
4: Set $l = 0$ and $\boldsymbol{\lambda}^{n,0} = \boldsymbol{\lambda}^{n-1}$;
5: **while** $l \leq L$ **do**
6: Calculate $\mathbf{W}^\star_{\boldsymbol{\lambda}^{n,l}}$ and $\boldsymbol{\delta}_{\boldsymbol{\lambda}^{n,l}}$ according to (12) and (11);
7: $\boldsymbol{\lambda}^{n,l+1} = \left[\boldsymbol{\lambda}^{n,l} - \alpha^{n,l}\boldsymbol{\delta}_{\boldsymbol{\lambda}^{n,l}}\right]^+$;
8: $l = l + 1$;
9: **end while**
10: $(\mathbf{W}^n, \boldsymbol{\lambda}^n) = (\mathbf{W}^{n,l^\star}, \boldsymbol{\lambda}^{n,l^\star})$, where $l^\star = \arg\min_{l=1,\ldots,L} f(\boldsymbol{\lambda}^{n,l})$;
11: $n = n + 1$;
12: **end while**
13: Output $\mathbf{W}^n$.

3.3. Secrecy Capacity Maximization with the Sum Power Constraint Only

In this subsection, let us particularize the above AO algorithm to SCM with the sum power constraint only. In such a case, there is a more convenient and efficient way to search for the dual optimal solution, rather than using PSG iterations in (10). To see this, note that $(\mathbf{W}^\star_{\lambda_0^\star}, \lambda_0^\star)$ is a primal-dual optimal pair to problems (6b) and (8) (with $\mathcal{I} = \{0\}$) if the following condition holds [15]

$$\lambda_0^\star(\mathrm{Tr}(\mathbf{W}^\star_{\lambda_0^\star}) - P_0) = 0,\ \lambda_0^\star \geq 0, \tag{13}$$

where $\mathbf{W}^\star_{\lambda_0^\star}$ is given by (12) with $\lambda = \lambda_0^\star$. The following fact (see [18, Sec. 13.1, Lemma 1]) sheds lights on the search for such $\lambda_0^\star$.

Fact 1. *The function* $\mathrm{Tr}(\mathbf{W}^\star_{\lambda_0})$ *is nonincreasing w.r.t.* λ_0*.*

This monotonicity means that we can adopt an efficient bisection approach to find $\lambda_0^\star$. Specifically, for a given λ_0, we compute $\mathbf{W}^\star_{\lambda_0}$ and $\mathrm{Tr}(\mathbf{W}^\star_{\lambda_0})$ from (12). If $\mathrm{Tr}(\mathbf{W}^\star_{\lambda_0}) > P_0$ (resp. $\mathrm{Tr}(\mathbf{W}^\star_{\lambda_0}) < P_0$), we increase λ_0 (resp. decrease λ_0). After a number of searches, we have either $\lambda_0 = 0$ and $\mathrm{Tr}(\mathbf{W}^\star_{\lambda_0}) \leq P_0$ or $\lambda_0 > 0$ and $\mathrm{Tr}(\mathbf{W}^\star_{\lambda_0}) = P_0$, thereby satisfying the optimality condition (13).

4. SIMULATION RESULTS AND CONCLUSIONS

We provide two simulation examples to demonstrate the performance gains of the proposed AO algorithm. The simulation settings are as follows: The number of antennas at Alice, Bob and Eve are $N_t = 5$, $N_b = N_e = 4$, resp. In each simulation trial, Bob and Eve's channels are randomly generated following an i.i.d. complex Gaussian distribution with zero mean and unit variance.

In the first example, we consider SCM with the sum power constraint only, and compare the proposed AO algorithm (cf. Section 3.3) with the fixed-point method (FPM) [6], the GSVD method [4] and the projected SVD (P-SVD) method [12]. Fig. 1 shows the secrecy rates of the various methods w.r.t. the transmit power. From the figure, we can see that the proposed AO algorithm yields performance identical to FPM, and outperforms the other methods over the whole range of powers tested. Table 1 shows the average running times of AO and FPM under the same setting as Fig. 1. As seen, AO is much faster than FPM, especially for large powers.

In the second example, we compare the proposed AO algorithm with the closed-form design in [5] under the per-antenna power constraints. Note that [5] deals with a matrix-covariance constrained SCM problem (cf. Introduction), and we provide the result of [5] here as a reference. To facilitate the comparison, we consider the per-antenna power constraints (3c) only, and the sum power constraint (3b) is omitted in the following simulation. The implementation details are as follows: We set $P_1 = \ldots = P_{N_t} = P_{\rm ant}$, $\lambda_1^0 = \ldots = \lambda_{N_t}^0 = 1$, $\mathbf{W}^0 = P_{\rm ant}\mathbf{I}$, $L = 1$ and ϵ=1e-3. As for the setting of [5], the covariance-matrix constraint $\mathbf{W} \preceq P_{\rm ant}\mathbf{I}$ is adopted for simplicity. Fig. 2 shows the secrecy rate behaviors of the two designs when we increase $P_{\rm ant}$ from 0 dB to 18 dB. It is evident that the AO design outperforms the closed-form design in [5], and there is about 1 bit/channel use constant rate gap between the two designs.

To conclude, this paper has developed an alternating optimization (AO) approach to the MIMO secrecy capacity maximization problem under the sum power and per-antenna power constraints. The proposed AO algorithm can be efficiently implemented in an iterative water-filling-like manner, and is guaranteed to converge to a stationary point of the secrecy capacity maximization problem (3). The efficacy of the proposed approach has been demonstrated by simulations.

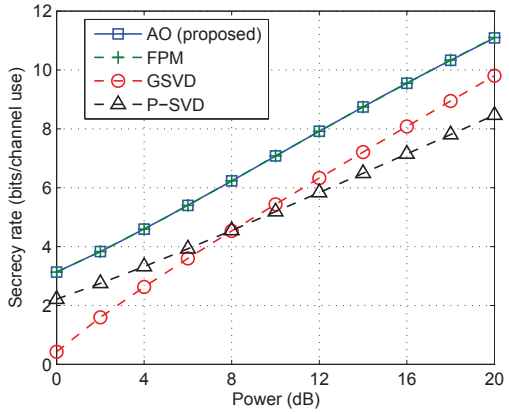

Fig. 1: Secrecy rates versus the sum power.

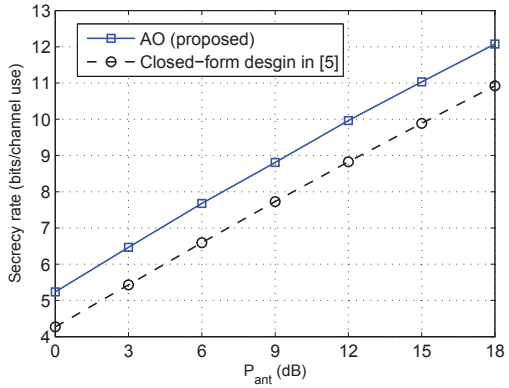

Fig. 2: Secrecy rates versus the per-antenna power.

Table 1: Average running times (in secs.) of AO and FPM

Algorithm	Power (dB)				
	4	8	12	16	20
AO	**0.0172**	**0.0264**	**0.0414**	**0.0646**	**0.0930**
FPM	0.1254	0.1805	0.3237	0.7672	13.5714

5. APPENDIX

We divide the proof into two steps: First, we show that every limit point of the iterates generated by (6a)-(6b) is a stationary point of (5); secondly, we show that every stationary point of (5) is also a stationary point of SCM-PAPC (3), thereby establishing our claim in Proposition 1.

Step 1: We need the following convergence result.

Lemma 3 ([14, Corollary 2]). *Consider the problem:*

$$\min_{\mathbf{x}} \; f(\mathbf{x_1}, \mathbf{x_2}) \quad \text{s.t.}\, \mathbf{x} \in \mathcal{X} \triangleq \mathcal{X}_1 \times \mathcal{X}_2 \tag{14}$$

where $f : \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \to \mathbb{R}$ is a continuously differentiable function; $\mathcal{X}_i \subseteq \mathbb{R}^{m_i}$, $i = 1, 2$ is a closed, nonempty and convex subset. Suppose that the sequence $\{\mathbf{x}^n\}$ generated by optimizing $\mathbf{x}_1$ and $\mathbf{x}_2$ alternatingly has limit points. Then every limit point of $\{\mathbf{x}^n\}$ is a stationary point of (14).

It can be verified that the objective function of (5) is continuously differentiable, and the feasible set is closed, nonempty and convex. Moreover, the iterates $(\mathbf{W}^n, \mathbf{S}^n)$ must be bounded, owing to the total power constraint in (5). Then, by Bolzano-Weierstrass theorem, we know that $(\mathbf{W}^n, \mathbf{S}^n)$ must have limit points. Therefore, invoking Lemma 3, we conclude that every limit point of $(\mathbf{W}^n, \mathbf{S}^n)$ generated by (6a)-(6b) is a stationary point of (5).

Step 2: Let $\phi_1(\mathbf{W}, \mathbf{S})$ and $\phi_2(\mathbf{W})$ be the objectives of problems (5) and (3), resp., and $\mathcal{W} \triangleq \{\mathbf{W} \mid \mathbf{W} \succeq \mathbf{0},\ \mathrm{Tr}(\mathbf{A}_i\mathbf{W}) \le P_i,\ \forall i \in \mathcal{I}\}$. Suppose that $(\mathbf{W}^\star, \mathbf{S}^\star)$ is a stationary point of (5). Then, we have

$$\mathrm{Tr}\left(\nabla_{\mathbf{W}}\phi_1(\mathbf{W}^\star, \mathbf{S}^\star)^H(\mathbf{W} - \mathbf{W}^\star)\right) \le 0,\ \forall\, \mathbf{W} \in \mathcal{W} \tag{15a}$$

$$\mathrm{Tr}\left(\nabla_{\mathbf{S}}\phi_1(\mathbf{W}^\star, \mathbf{S}^\star)^H(\mathbf{S} - \mathbf{S}^\star)\right) \le 0,\ \forall\, \mathbf{S} \succeq \mathbf{0}. \tag{15b}$$

It follows from (15b) that

$$\mathbf{S}^\star = (\mathbf{I} + \mathbf{G}^H\mathbf{W}^\star\mathbf{G})^{-1}. \tag{16}$$

Notice that for a given $\mathbf{W}^\star$, the corresponding optimal $\mathbf{S}^\star$ is uniquely given by (16). Hence, by applying Danskin's theorem [15], we can substitute (16) into (15a), and have that

$$\mathrm{Tr}\big(\nabla_{\mathbf{W}}\phi_1\big(\mathbf{W}^\star, (\mathbf{I} + \mathbf{G}^H\mathbf{W}^\star\mathbf{G})^{-1}\big)^H(\mathbf{W} - \mathbf{W}^\star)\big) \le 0 \tag{17}$$

holds for all $\mathbf{W} \in \mathcal{W}$. On the other hand, it can be verified that

$$\begin{aligned} & \nabla_{\mathbf{W}}\phi_2(\mathbf{W}^\star) \\ = \;& \mathbf{H}(\mathbf{I} + \mathbf{H}^H\mathbf{W}^\star\mathbf{H})^{-1}\mathbf{H}^H - \mathbf{G}(\mathbf{I} + \mathbf{G}^H\mathbf{W}^\star\mathbf{G})^{-1}\mathbf{G}^H \\ = \;& \nabla_{\mathbf{W}}\phi_1\big(\mathbf{W}^\star, (\mathbf{I} + \mathbf{G}^H\mathbf{W}^\star\mathbf{G})^{-1}\big). \end{aligned} \tag{18}$$

Therefore, we conclude from (17) and (18) that

$$\mathrm{Tr}\left(\nabla_{\mathbf{W}}\phi_2(\mathbf{W}^\star)^H(\mathbf{W} - \mathbf{W}^\star)\right) \le 0, \quad \forall\, \mathbf{W} \in \mathcal{W},$$

i.e., $\mathbf{W}^\star$ is a stationary point of problem (3).

6. REFERENCES

[1] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4-5, pp. 355–580, 2008.

[2] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*. Boca Raton, FL, USA: CRC Press, 1996.

[3] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[4] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas II: The MIMOME wiretap channel," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[5] R. Bustin, R. Liu, H. V. Poor, and S. Shamai (Shitz), "An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP Journal on Wireless Communications and Networking*, 2009.

[6] J. Li and A. P. Petropulu, "Transmitter optimization for achieving secrecy capacity in Gaussian MIMO wiretap channels," available online at `http://arxiv.org/abs/0909.2622`.

[7] S. Fakoorian and A. L. Swindlehurst, "Optimal power allocation for GSVD-based beamforming in the MIMO Gaussian wiretap channel," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, July 2012, pp. 1 – 6.

[8] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, June 2007.

[9] S. Fakoorian and A. L. Swindlehurst, "Full rank solutions for the MIMO Gaussian wiretap channel with an average power constraint," available online at `http://arxiv.org/abs/1210.4795`.

[10] A. Mukherjee, S. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," available online at `http://arxiv.org/abs/1011.3754`.

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

[12] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and S. Cui, "On the relationship between the multi-antenna secrecy communications and cognitive radio communications," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1877–1886, Jun. 2010.

[13] J. Jose, N. Prasad, M. Khojastepour, and S. Rangarajan, "On robust weighted-sum rate maximization in MIMO interference networks," in *Proc. IEEE Int. Conf. Communications (ICC)*, June 2011.

[14] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss-Seidel method under convex constraints," *Operation research letter*, vol. 26, pp. 127–136, 2000.

[15] D. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1999.

[16] R. Zhang, "Cooperative multi-cell block diagonalization with per-base-station power constraints," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1435–1445, Dec. 2010.

[17] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Trans. Info. Theory*, vol. 57, no. 5, pp. 3117–3131, May 2011.

[18] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*. New York: Springer, 2008.