# Energy Efficient Artificial Noise–Aided Precoding Designs for Secured Visible Light Communication Systems

SCHOLARONE™
Manuscripts

1

# Energy Efficient Artificial Noise-Aided Precoding Designs for Secured Visible Light Communication Systems

Thanh V. Pham *Member, IEEE,* and Anh T. Pham *Senior Member, IEEE*

## Abstract

Physical layer security (PLS) has recently gained a lot of attention in the research and development of visible light communication (VLC). In this paper, we study the designs of PLS in VLC systems in the presence of multiple unauthorized users (i.e. eavesdroppers) using artificial noise (AN)-aided precoding. The design objective focuses on minimizing the total transmit power subject to specific constraints on the signal-to-interference-plus-noise ratios (SINRs) of the legitimate and unauthorized users. In particular, two design approaches are investigated considering the availability of unauthorized users' channel state information (CSI) at the transmitter. In the case of unknown CSI, the AN is constructed to lie on the null-space of the legitimate user's channel. The design problem is convex, thus, can be effectively solved. When the CSI is available, the design additionally imposes constraints on the maximum allowable unauthorized users' SINRs. The design problem, in this case, is, nevertheless, non-convex. Therefore, instead of finding the optimal solution, we examine two different sub-optimal yet low-complexity approaches to solve the problem, namely: Concave-Convex Procedure (CCP) and Semidefinite Relaxation (SDR). Additionally, robust designs that take into account channel uncertainty are also investigated. Extensive numerical results are shown to demonstrate the feasibility and performance of each design with practical parameters.

## Index Terms

Visible light communication, physical layer security, artificial noise, precoding.

2

# I. INTRODUCTION

The past decade has witnessed an explosive growth of mobile devices and an increasing number of data-intensive multimedia applications. According to Cisco System Inc., the global data traffic will increase from 7 exabytes to 49 exabytes per month during the period of 2016 to 2021 [1]. This tremendous demand for data traffic has been posing a serious burden on present wireless technologies and urges the use of a new electromagnetic spectrum for communication purposes due to the already congested radio frequency (RF) spectrum. Visible light communication (VLC) that uses the visible spectrum for data transmission is gaining prominence as an alternative or complementary solution to the existing RF technologies. Possessing several unique advantages, such as huge unlicensed spectrum and immunity to RF waves, the research and development of VLC are also motived by the recent popularity of light-emitting diodes (LEDs) in illumination [2], [3].

As an emerging technology, the majority of research works have been concentrated on improving the performance and practicality of VLC systems [4]–[9]. In addition to these aspects, security and privacy are also crucial concerns in the design of VLC. This is because of the broadcast nature of visible light, which makes the VLC channel vulnerable to eavesdropping by malicious users. Currently, security measures heavily rely on traditional key-based cryptographic techniques performing at the upper layers of the OSI model. The secrecy of those techniques mainly comes from the complexity of deriving secret keys. With the current computational power, there is a belief that this key derivation problem is infeasible to be solved in a reasonable time. Nevertheless, it is expected that rapid advances in hardware (e.g., the development of quantum computing) will threaten the security of several cryptographic algorithms in the near future. This has motivated numerous research on security at the physical layer (also known as physical layer security (PLS)), which utilizes the inherent uncertainty of the transmission media to deal with eavesdropping by unauthorized users. The foundational theory of PLS was initiated by A. D. Wyner [10] as he introduced the concept of the wiretap channel and the notation of secrecy capacity as the performance measure. Since then, many research efforts have been devoted to characterizing the secrecy capacity for various channel models and systems configurations [11]–[16].

While PLS has been extensively studied in the context of RF systems, only recently, its adoption to VLC has received attention. First, in the scenario of a single transmitter, the authors in

3

[17] examined the conventional wiretap channel (i.e., one legitimate user, and one eavesdropper) and comprehensively analyzed the lower and upper bound secrecy capacity of single-input single-output (SISO) VLC channels. Also, in the presence of multiple legitimate users, the use of the non-orthogonal multiple access (NOMA) scheme to enhance the secrecy outage probability (SOP) performance was investigated in [18]. As practical VLC systems need to guarantee a uniform and certain illumination for indoor purposes, the scenarios with multiple LED luminaries (i.e., transmitters) is a logical setting. Modeling the locations of LED transmitters, legitimate users, and eavesdroppers as 2-D independent, homogenous Poisson point processes, closed-form expressions for outage probability and ergodic secrecy rate were derived in [19] using mathematical tools from stochastic geometry. Furthermore, the configuration of multi-transmitter also enables the use of precoding and artificial noise (AN) techniques that exploit the spatial degrees of freedom to improve secrecy performance.

For the case of precoding with a single legitimate user and single eavesdropper, the study in [20] characterized a lower bound on the secrecy capacity for the multiple-input single-output (MISO) VLC channels, where the obtained bound was represented as a fractional function of the precoder. The optimal solution to the problem of maximizing the lower bound secrecy capacity was given for the specific case of zero-forcing (ZF) precoding. Then, the optimal precoding designs for the generic precoder structure were given in [21] for two scenarios: perfect and imperfect (i.e., uncertainty) channel state information (CSI). In the case of multi-eavesdropper, precoding designs to minimize transmitted power and maximize the minimum secrecy rate were investigated in [22]. The case of multiple legitimate users was reported in [23] that aimed to maximize users' secrecy sum-rate using ZF precoding for two scenarios of known and unknown eavesdropper's CSI at the transmitter. In [24], a special configuration of two legitimate users was examined, where the transmission from a single LED transmitter was assisted by a number of trusted relay luminaries. Achievable secrecy rate regions were then obtained for there relaying schemes: cooperative jamming, decode-and-forward, and amplify-and-forward. The authors in [25] studied a multi-user VLC configuration considering that users' messages must be kept confidential to each other. Under this assumption, the optimal precoding to maximize other secrecy capacity-related measures, namely: max-min fairness, harmonic mean, proportional fairness, and weighted fairness, was designed.

Regarding the AN approach, initial studies focused on the friendly jamming method, where a specific set of LED transmitters is designated to transmit AN [26], [27]. A different method

4

where the AN is added to the information-bearing signal through precoding (also known as AN-aided precoding) was also reported in [28]–[31]. In particular, [28] examined the AN designs for a single-user multiple-eavesdropper configuration with the availabilities of user's and eavesdroppers' CSI at the transmitters. The designs focused on maximizing the user's signal-to-interference-plus-noise ratio (SINR) while eavesdroppers' SINRs were limited to a certain threshold. With the same design objective, the authors in [30] attempted to design AN under the assumption that the eavesdropper is randomly distributed. Also using SINR as the performance measure, [29] investigated AN designs to maximize the fairness of users' SINRs for a multi-user wiretap channel. On the other hand, optimal AN-aided precoding designs to maximize user's secrecy rate were reported in [31], in which the secrecy rate expression was derived for the discrete input distributions of information-bearing and jamming signal.

Aside from the conference version of this paper [32], it should be noted that the issue of energy efficiency has not been taken into account in previous AN-aided precoding design studies. As a matter of fact, indoor VLC systems consume energy for both illumination and data transmission. While the energy for illumination (i.e., the energy of the direct current (DC) bias) is set independently and usually fixed for specific usage, the combined energy for transmission of the information-bearing signals and AN generation should be optimized with respect to a certain targeted secrecy performance. Considering the legitimate user's and unauthorized users' SINRs as the characterization of the secrecy performance, the aim of this paper is to study AN-aided precoding designs from the perspective of energy efficiency. The general design purpose targets at the minimization of the total transmit power while guaranteeing a minimum achievable SINR by the legitimate user.

Two different approaches are considered for the two scenarios of known and unknown unauthorized users' CSI at the transmitter. In the case of unknown unauthorized users' CSI, the AN is generated to lie on the null-space of the user's channel. By doing so, it does not interfere with the information-bearing signal intended to the legitimate user, yet potentially degrades the quality of unauthorized users' channels. In this design, we are interested in the impact of the AN size (defined as the size of AN symbol vector) and the AN adjusting power parameter on the total transmit power and the resulting unauthorized users' SINRs. Numerical results show that either increasing the AN size or adjusting the power parameter can lead to the reduction of unauthorized users' SINRs (i.e., better secrecy performance), and that it is more beneficial to increase AN size for the sake of power saving. When unauthorized users' CSI is available at

the transmitter, a more proper design is investigated so that specific constraints on unauthorized users' SINRs can be imposed. The design problem is, nevertheless, shown to be non-convex. Therefore, rather than finding the optimal solution, which can be computationally intensive, we focus on the use of two low-complexity sub-optimal approaches, namely: concave-convex procedure (CCP) and semidefinite relaxation (SDR), in solving the problem. The CCP technique involves an iterative procedure, which, depending on the required accuracy, might need several iterations to arrive at a local optimum. On the other hand, the use of SDR leads to a simpler upper bound solution, it is however not always applicable. The two approaches, thus, can be considered as complementary to each other. It is also proved that the AN size of one is sufficient for the optimal design. Furthermore, the assumption that users' CSI is perfectly known by the transmitter is, in practice, rather unrealistic. Therefore, taking into account the uncertainty of channel estimation, we explore robust designs based on the use of SDR.

The remainder of the paper is organized as follows. The AN-aided precoding model for the considered VLC system is described in Section II. Section III presents the AN-aided precoding designs for two scenarios of unknown and known unauthorized users' CSI with the assumption that the transmitter has perfect CSI knowledge. In Section IV, robust designs taking into account the channel uncertainty are examined. Representative numerical results are given in Section IV. Finally, Section V concludes the paper.

*Notation*: The following notations are used throughout the paper. $\mathbb{R}$ is the sets of real-valued numbers. Bold upper case letters denote matrices, e.g., $\mathbf{A}$ whereas bold lower case letters indicate vectors, e.g., $\mathbf{a}$. The transpose of $\mathbf{A}$ is written as $\mathbf{A}^T$ while $[\mathbf{A}]_{i,j}$ is the element at the $i-$th row and the $j-$th column of $\mathbf{A}$. The $i-$th row vector of matrix $\mathbf{A}$ and the $i-$th element of vector $\mathbf{a}$ are denoted as $[\mathbf{A}]_{i,:}$ and $[\mathbf{a}]_i$, respectively. $\|\cdot\|_1$, $\|\cdot\|_2$, and $|\cdot|$ are the norm-1, Euclidean norm, and absolute value operator. $\mathbf{I}_N$ is the identity matrix of size $N$, $\mathbf{0}_N$ is the all-zeros vector of size $N$, and $\mathbf{e}_n$ is the all-zero vector with the $n-$th element being 1. Finally $\text{Tr}(\cdot)$ denotes the trace of a square function.

## II. AN-AIDED PRECODING MODEL FOR VLC CHANNEL

We consider a VLC broadcast system as illustrated in Fig. 1, where a legitimate sender (Alice) equipped with $N_t$ ($N_t \geq 2$) LED luminaries serves a legitimate user (Bob) in the presence of $K$ non-colluding unauthorized users (Eve 1, Eve 2, ..., Eve $K$). The purpose of using AN is to increase confusion at Eves while keeping its impact on Bob as small as possible. Precisely, let $d$
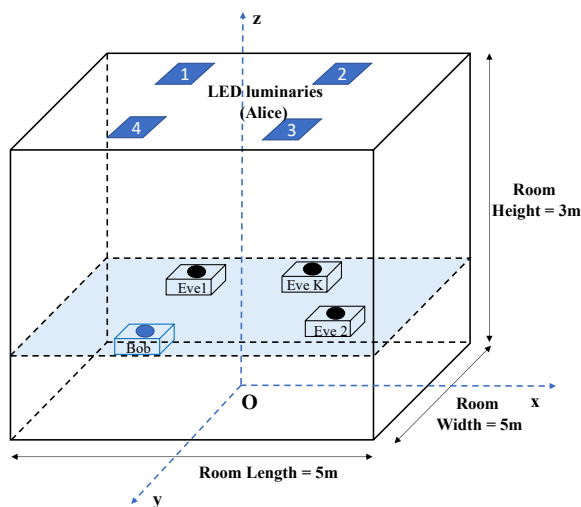
Figure 1: Geometrical configuration of a VLC system with a legitimate user and multiple unauthorized users.

and $\mathbf{z}$ respectively be the pulse amplitude modulation (PAM) data symbols intended for Bob and the AN symbol vector. To simplify the analysis, $d$ and $\mathbf{z} \in \mathbb{R}^{N_s}$ are assumed to be zero-mean and uniformly distributed over the normalized range $[-1, 1]$. Note that one can improve the secrecy rate at the expense of more complicated SINR expressions by adopting more general input distributions, such as truncated Gaussian distribution [27] and truncated generalized normal distribution [33]. In this paper, we refer $N_s$ as the AN size. At the $n-$th LED luminary, the data symbol and AN vector are multiplied with their corresponding precoding weights $v_n$ and $\mathbf{w}_n \in \mathbb{R}^{N_s}$ as

$$s_n = v_n d + \mathbf{w}_n^T \mathbf{z}. \tag{1}$$

Notice that the AN-added information-bearing signal $s_n$ can take negative values, it is thus can not be used directly as the drive current for the LEDs. As a result, a DC bias $I_n^{\mathrm{DC}}$ is added to $s_n$ as

$$x_n = s_n + I_n^{\mathrm{DC}}. \tag{2}$$

To guarantee a minimum dimming level and ensure the drive current operating within its linear range, $x_n$ should be constrained to a certain range, e.g.,

$$I_{\min} \leq x_n \leq I_{\max}, \tag{3}$$

where $I_{\min}$ and $I_{\max}$ are the minimum and maximum allowable values for the drive current. From (1) with the assumption that $d$ and $\mathbf{z}$ are uniformly distributed over $[-1, 1]$, the above constraints imply

$$|v_n| + \left\| \mathbf{w}_n^T \right\|_1 \le \Delta_n, \tag{4}$$

where $\Delta_n = \min\left(I_n^{\mathrm{DC}} - I_{\min}, I_{\max} - I_n^{\mathrm{DC}}\right)$. Denote $\mathbf{p}^t = \begin{bmatrix} p_1^t & p_2^t & \cdots & p_{N_t}^t \end{bmatrix}^T$ as the LED luminaries' emitted optical power vector where each element is given by $p_n^t = \eta x_n$ with $\eta$ being the LED conversion factor. The received signal at the $k-$th user[1] (either Bob or Eves) after the optical-electrical conversion can be expressed as

$$y_k = \gamma \mathbf{h}_k^T \mathbf{p}^t + n_k = \gamma \eta \left( \underbrace{\mathbf{h}_k^T \mathbf{v} d}_{\text{information-bearing signal}} + \underbrace{\mathbf{h}_k^T \mathbf{W} \mathbf{z}}_{\text{AN}} + \underbrace{\mathbf{h}_k^T \mathbf{I}^{\mathrm{DC}}}_{\text{DC current}} \right) + \underbrace{n_k}_{\text{receiver noise}}, \tag{5}$$

where $\mathbf{h}_k = \begin{bmatrix} h_{1,k} & h_{2,k} & \cdots & h_{N_t,k} \end{bmatrix}^T$ is the user's light-of-sight (LoS) channel vector[2], $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_{N_t} \end{bmatrix}^T$ and $\mathbf{W} = \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_{N_t} \end{bmatrix}^T$ are precoders of the data symbol and the AN vector, respectively. $\mathbf{I}^{\mathrm{DC}} = \begin{bmatrix} I_1^{\mathrm{DC}} & I_2^{\mathrm{DC}} & \cdots & I_{N_t}^{\mathrm{DC}} \end{bmatrix}^T$ is the DC-bias vector and $n_k$ is the receiver noise, which can be well modeled as additive white Gaussian noise with zero mean and variance $\sigma_k^2$ given by

$$\sigma_k^2 = 2\gamma e \overline{p_k^r} B + 4\pi e A_r \gamma \chi_{\mathrm{amb}} (1 - \cos(\Psi)) B + i_{\mathrm{amp}}^2 B, \tag{6}$$

where $\overline{p_k^r} = \mathbb{E}[p_k^r] = \eta \mathbf{h}_k^T \mathbf{I}^{\mathrm{DC}}$ is the average received optical power of the $k-$th user, $e$ is the elementary charge, and $B$ is the systems bandwidth. $\chi_{\mathrm{amb}}$ and $i_{\mathrm{amp}}$ are the ambient light photocurrent and the pre-amplifier noise current density, respectively.

The DC current term $\mathbf{h}_k^T \mathbf{I}^{\mathrm{DC}}$, which carries no information, can be effectively removed by AC coupling, resulting in

$$y_k = \gamma \eta \left( \mathbf{h}_k^T \mathbf{v} d + \mathbf{h}_k^T \mathbf{W} \mathbf{z} \right) + n_k. \tag{7}$$

The average SINR of the $k-$th user is then written as

$$\mathrm{SINR}_k = \frac{\sigma_d^2 (\gamma \eta)^2 \left| \mathbf{h}_k^T \mathbf{v} \right|^2}{\sigma_z^2 (\gamma \eta)^2 \left\| \mathbf{h}_k^T \mathbf{W} \right\|_2^2 + \sigma_k^2}, \tag{8}$$

---

[1]For notational convenience, the subscript $k = 0$ is used to indicate Bob while $k = 1, 2, ..., K$ refers to Eves.

[2]For the VLC channel model, refer to [34] for the details.

8

where $\sigma_d^2$ and $\sigma_z^2$ are the variances of $d$ and $z$, respectively. Since $d$ and $z$ are uniformly distributed over $[-1, 1]$, $\sigma_d^2 = \sigma_z^2 = \frac{1}{3}$. Therefore, for the sake of mathematical analyses in later parts of the paper, we rewrite the above expression as

$$\text{SINR}_k = \frac{\left|\mathbf{h}_k^T \mathbf{v}\right|^2}{\left\|\mathbf{h}_k^T \mathbf{W}\right\|_2^2 + \tilde{\sigma}_k^2}, \tag{9}$$

where $\tilde{\sigma}_k^2 = \frac{\sigma_k^2}{\frac{1}{3}(\gamma\eta)^2}$.

## III. AN-AIDED PRECODING DESIGNS WITH PERFECT CSI

As mentioned in the previous section, our AN-aided precoding design strategy is to minimize the average total data transmit power subject to certain SINR thresholds on Bob and Eves. The design also needs to take the dimming constraint in (4) into consideration. In the following, we present two AN-aided precoder designs in accordance with two scenarios: unknown and known Eves' CSI at the LED transmitters.

### A. Unknown Eves' CSI

In most practical cases, Eves are passive malicious users who do not feed their CSI back to the transmitters. As a result, it is impossible to impose specific constraints on the instantaneous Eves' SINRs. The use of AN, however, is still beneficial since it can possibly degrade Eves' channel quality, while is designed to interfere with Bob's transmission as little as possible. Straightforwardly, the design problem can mathematically be formulated as[3]

$$\mathcal{P}1 : \underset{\mathbf{v},\mathbf{W}}{\text{minimize}} \quad \|\mathbf{v}\|_2^2 + \|\mathbf{W}\|_2^2 \tag{10a}$$

$$\text{subject to } \text{SINR}_0 \geq \gamma_0, \tag{10b}$$

$$\left|[\mathbf{v}]_n\right| + \left\|[\mathbf{W}]_{n,:}\right\|_1 \leq \Delta_n, \ n = 1, 2, ..., N_t. \tag{10c}$$

As can be seen from the problem formulation, the precoders are designed in such a way that Bob's SINR meets a minimum predefined threshold of $\gamma_0$. Intuitively, since Eves' SINRs are uncertain from this AN-aided precoder design perspective, $\gamma_0$ should be chosen sufficiently large to guarantee a good secrecy performance (i.e., large gaps between Bob's and Eves' SINRs). We can notice that $\text{SINR}_0$ is bounded due to the boundedness of the channel gain $\mathbf{h}_0$ and the

---

[3]The actual average total transmit power is $\sigma_d^2\|\mathbf{v}\|_2^2 + \sigma_z^2\|\mathbf{W}\|_2^2 = \frac{1}{3}\left(\|\mathbf{v}\|_2^2 + \|\mathbf{W}\|_2^2\right)$. For the sake conciseness, we omit the constant term $\frac{1}{3}$ of the objective function from the problem formulation.

elements of $\mathbf{v}$ and $\mathbf{W}$. Therefore, $\mathcal{P}1$ may not be feasible if $\gamma_0$ is set too demanding. The following proposition gives the necessary and sufficient conditions on the feasibility of $\mathcal{P}1$.

**Proposition 1:** $\mathcal{P}1$ is feasible if and only if $\gamma_0 \leq \frac{|\mathbf{h}_0^T \boldsymbol{\Delta}|^2}{\tilde{\sigma}_0^2}$, where $\boldsymbol{\Delta} = \begin{bmatrix} \Delta_1 & \Delta_2 & \dots & \Delta_{N_t} \end{bmatrix}^T$.

*Proof*: It suffices to prove that $\mathcal{P}1$ is feasible if and only if $\gamma_0$ is not strictly greater than the maximum attainable value of $\mathrm{SINR}_0$. Then, the proof is immediately implied from the following observation.

$$\mathrm{SINR}_0 \leq \frac{|\mathbf{h}_0 \mathbf{v}|^2}{\tilde{\sigma}_0^2} \leq \frac{|\mathbf{h}_0 \boldsymbol{\Delta}|^2}{\tilde{\sigma}_0^2}. \tag{11}$$

Note that, the right-hand side inequality is due to the non-negativeness of the elements of $\mathbf{h}_0$.

**Proposition 1** is meaningful in the sense that it points out the maximum allowable $\gamma_0$ for $\mathcal{P}1$ being feasible given specific values of $\mathbf{h}_0$, $\boldsymbol{\Delta}$, and $\tilde{\sigma}_0^2$. In practice, it is often more preferable to characterize the feasibility of $\mathcal{P}1$ over the possible ranges of these parameters, especially the position of users (in other words, $\mathbf{h}_0$) because of user movement. Nonetheless, even if $\boldsymbol{\Delta}$ is fixed, an analytical investigation for this could be rather complicated due to the dependency of $\tilde{\sigma}_0^2$ on $\mathbf{h}_0$. Instead, the probability of $\mathcal{P}1$ being feasible is numerically evaluated in Section V of the paper. In the following, for the sake of analysis, we assume that $\gamma_0$ is chosen to ensure the feasibility of $\mathcal{P}1$ for at least one realization of $\mathbf{h}_0$.

With this assumption, it can be seen that the design in $\mathcal{P}1$ is problematic as the optimal solution to $\mathbf{W}$ is $\mathbf{W}^* = \mathbf{0}$. To see this, let $\mathbf{v}^*$ and $\mathbf{W}^*$ be the optimal solution to $\mathbf{v}$ and $\mathbf{W}$, respectively, and assume that $\mathbf{W}^* \neq \mathbf{0}$. Since $(\mathbf{v}^*, \mathbf{W}^*)$ is feasible to $\mathcal{P}1$, $(\mathbf{v}_{\mathrm{opt}}, \mathbf{0})$ is also feasible and offer a smaller objective value. This contradicts with the assumption that $\mathbf{v}^*$ and $\mathbf{W}^*$ are optimal. Hence, $\mathbf{W}^* = \mathbf{0}$. This precoding design (referred to as no-AN design) statistically offers no physical security in the sense that the average Bob's and Eves' SINRs are the same.

To exploit the benefit of AN, one can choose $\mathbf{W}$ to be a non-zero matrix whose columns lie on the null-space of $\mathbf{h}_0^T$. Assume that $\mathbf{h}_0^T$ is full rank[4], then according to the rank-nullity theorem, the dimension of the null-space of $\mathbf{h}_0^T$, denoted as $\mathrm{Nullity}(\mathbf{h}_0^T) = N_t - \mathrm{rank}(\mathbf{h}_0^T) = N_t - 1 \geq 1$ since $N_t \geq 2$. Hence, assume that $\mathbf{W}$ is full rank, the maximum allowable AN size is $N_s = N_t - 1$. It is expected that larger AN size would result in lower Eves' SINRs at the expense of increased total transmit power. Let $\overline{\mathbf{W}}$ be a $\mathbb{R}^{N_t \times N_s}$ matrix whose $N_s$ columns are chosen from the $(N_t - 1)$

---

[4]Since Bob and Eves are equipped with one photodiode resulting in $N_T \times 1$ channel matrices, the full rank assumption is always valid as long as $\mathbf{h}_0 \neq \mathbf{0}$ and $\mathbf{h}_k \neq \mathbf{0}$ (i.e. there is at least one link between the user and LED luminaries).

vectors of an orthonormal basis for the null-space of $\mathbf{h}_0^T$. It is seen that if $\mathbf{W}$ is set to be $\overline{\mathbf{W}}$, the constraints (10c) might not always be guaranteed. To ensure that (10c) holds, one can set $\mathbf{W}$ as a scaled version of $\overline{\mathbf{W}}$ as

$$\mathbf{W} = \rho \frac{\min_{n} \Delta_n}{\max_{n} \left\| \left[\overline{\mathbf{W}}\right]_{n,:} \right\|_1} \overline{\mathbf{W}}, \tag{12}$$

where $\rho \in (0,1]$ is referred to as the AN power adjusting parameter, which controls the magnitude of $\mathbf{W}$. As a result, increasing $\rho$ not only decreases Eves' SINRs but also reduces the probability that Bob's SINR meets the threshold requirement since less power is allocated for the transmission of data symbols. In deed, finding $\rho$ to balance a trade-off between the design feasibility and Eves' SINRs is of importance. This, however, is a challenging issue as the instantaneous Eves' SINRs are unknown to the transmitter. For the sake brevity, we leave this question to future investigation. With the AN being specified as in (12), we might consider the following design problem

$$\mathcal{P}2 : \underset{\mathbf{v},\rho}{\text{minimize}} \quad \|\mathbf{v}\|_2^2 + \|\mathbf{W}\|_2^2 \tag{13a}$$

$$\text{subject to} \quad \frac{\left|\mathbf{h}_0^T \mathbf{v}\right|^2}{\tilde{\sigma}_0^2} \geq \gamma_0, \tag{13b}$$

$$\left|[\mathbf{v}]_n\right| + \left\| [\mathbf{W}]_{n,:} \right\|_1 \leq \Delta_n, \ \ n = 1, 2, ..., N_t, \tag{13c}$$

$$0 < \rho \leq 1. \tag{13d}$$

Nonetheless, the solution to the above problem is not definite as it can be seen that the smaller $\rho$ is the better. To fix it, we turn to the design problem with $\rho$ being chosen as a positive constant. With this in mind, $\mathcal{P}2$ reduces to the problem of minimizing the data-bearing signal power

$$\mathcal{P}3(\rho) : \underset{\mathbf{v}}{\text{minimize}} \quad \|\mathbf{v}\|_2^2 \tag{14a}$$

$$\text{subject to} \quad \frac{\left|\mathbf{h}_0^T \mathbf{v}\right|^2}{\tilde{\sigma}_0^2} \geq \gamma_0, \tag{14b}$$

$$\left|[\mathbf{v}]_n\right| \leq \Delta_n - \rho \frac{\min_{n} \Delta_n}{\max_{n} \left\| \left[\overline{\mathbf{W}}\right]_{n,:} \right\|_1} \left\| \left[\overline{\mathbf{W}}\right]_{n,:} \right\|_1, \ \ n = 1, 2, ..., N_t. \tag{14c}$$

**Proposition 2:** The optimal solution $\mathbf{v}^*$ to $\mathcal{P}3(\rho)$ must satisfy $\frac{\left|\mathbf{h}_0^T \mathbf{v}^*\right|^2}{\tilde{\sigma}_0^2} = \gamma_0$.

*Proof*: Firstly, one can realize that there exists an optimal solution $\mathbf{v}^*$ such that $\mathbf{h}_0^T \mathbf{v}^* \geq 0$. As proof, assume that $\mathbf{v}_{\text{opt}}$ is an optimal solution. Then, it is easy to see that $\mathbf{v}^* = \text{sign}(\mathbf{h}_0^T \mathbf{v}_{\text{opt}}) \mathbf{v}_{\text{opt}}$, which satisfies $\mathbf{h}_0^T \mathbf{v}^* \geq 0$, is also feasible and offers the same objective value.

Now, we proof the proposition by contradiction. Assume that there is an optimal solution $\mathbf{v}^*$ so that $\frac{\left|\mathbf{h}_0^T \mathbf{v}^*\right|^2}{\tilde{\sigma}_0^2} > \gamma_0$. According to the above observation, we can write $\mathbf{h}_0^T \mathbf{v}^* > \tilde{\sigma}_0 \sqrt{\gamma_0}$. This implies that there is at least a positive element, say $[\mathbf{v}^*]_p$, in $\mathbf{v}^*$. Then, there exists $\bar{\mathbf{v}}^*$, where $[\bar{\mathbf{v}}^*]_i = [\mathbf{v}^*]_i$ for $i \neq p$ and $[\bar{\mathbf{v}}^*]_p = [\mathbf{v}^*]_p - \beta$ (for some $\beta > 0$ sufficiently small), such that $[\bar{\mathbf{v}}^*]_p > 0$ and $\mathbf{h}_0^T \bar{\mathbf{v}}^* \geq \tilde{\sigma}_0 \sqrt{\gamma_0}$. Obviously, $\bar{\mathbf{v}}^*$ is feasible to $\mathcal{P}3(\rho)$ and offers a smaller objective value. This contradicts to the assumption that $\mathbf{v}^*$ is optimal. Hence, $\frac{\left|\mathbf{h}_0^T \mathbf{v}^*\right|^2}{\tilde{\sigma}_0^2} = \gamma_0$. The proof is completed.

As a result, the constraint in (13b) can now be equivalently expressed by $\mathbf{h}_0^T \mathbf{v} = \tilde{\sigma}_0^2 \sqrt{\gamma_0}$, which is convex with respect to $\mathbf{v}$. Therefore, $\mathcal{P}3(\rho)$ is a convex optimization problem, which can be solved efficiently using off-the-shelf optimization packages [37], [38]. The impact of choosing $\rho$ is shown in Fig. 4, where it is clearly observed that increasing $\rho$ considerably lowers Eves' SINRs at the expense of higher power consumption.

### B. Known Eves' CSI

In several scenarios where Eves are active (i.e., the transmitters know their CSI), specific constraints on Eves' SINRs can be taken into account in the design, resulting in the following problem

$$\mathcal{P}4: \underset{\mathbf{v}, \mathbf{w}}{\text{minimize}} \quad \|\mathbf{v}\|_2^2 + \|\mathbf{W}\|_2^2 \tag{15a}$$

$$\text{subject to} \quad \text{SINR}_0 \geq \gamma_0, \tag{15b}$$

$$\text{SINR}_k \leq \gamma_k, \quad k = 1, 2, ..., K, \tag{15c}$$

$$\left|[\mathbf{v}]_n\right| + \left\|[\mathbf{W}]_{n,:}\right\|_1 \leq \Delta_n, \quad n = 1, 2, ..., N_t. \tag{15d}$$

Different from $\mathcal{P}1$, this design strategy additionally imposes maximum allowable values on Eves' SINRs (i.e., $\gamma_k$ for the $k-$th Eve). These constraints make characterizing necessary and sufficient conditions for the feasibility of $\mathcal{P}4$ cumbersome. Perhaps, we can only infer from **Proposition 1** that the problem is not feasible when $\gamma_0 > \frac{\left|\mathbf{h}_0^T \boldsymbol{\Delta}\right|^2}{\tilde{\sigma}_0^2}$. Similar to the case of unknown Eves' CSI, we assume that $(\gamma_0, \gamma_k\text{'s})$ are chosen to ensure the feasibility of $\mathcal{P}4$ for at least one realization of $(\mathbf{h}_0, \mathbf{h}_k\text{'s})$.

Though the Bob's SINR constraint can be transformed to be convex as

$$\frac{1}{\gamma_0} \mathbf{h}_0^T \mathbf{v} \geq \sqrt{\|\mathbf{h}_0^T \mathbf{W}\|_2^2 + \tilde{\sigma}_0^2}, \tag{16}$$

it is not the case for Eves' SINR constraints. As a result, $\mathcal{P}4$ is a non-convex optimization problem, which is generally hard to be optimally solved. In this paper, rather than finding its optimal solution, which may be computationally expensive, we aim to investigate two different sub-optimal yet low-complexity approaches to solve the problem. Particularly, we focus on the use of CCP and SDR technique in solving $\mathcal{P}4$.

*1) CCP Approach:* The CCP is a heuristic method for finding local optimal solutions of non-convex optimization problems [35], [36]. The technique involves an iterative procedure, which solves a sequence of surrogate convex optimization problems with guaranteed convergence. In this case, since the objective function and Bob's SINR constraint are convex, we focus on modifying the constraints on Eves' SINRs to be convex. These constraints can be rewritten as

$$\frac{1}{\gamma_k}\left(\mathbf{h}_k^T\mathbf{v}\right)^2 \leq \left\|\mathbf{h}_k^T\mathbf{W}\right\|_2^2 + \tilde{\sigma}_k^2, \ \ k=1,2,...,K, \tag{17}$$

where the left-hand side is convex but the right-hand side $\left\|\mathbf{h}_k^T\mathbf{W}\right\|_2^2$ term is a quadratic form, hence not convex. The CCP relies on using a linear lower bound of the quadratic form obtained from its first-order Taylor expansion. Specifically, at the $i-$th iteration of the procedure, the following lower bound is used

$$\left\|\mathbf{h}_k^T\mathbf{W}^{(i-1)}\right\|_2^2 + 2\left[\mathbf{W}^{(i-1)}\right]^T\mathbf{h}_k\mathbf{h}_k^T\left(\mathbf{W}^{(i)}-\mathbf{W}^{(i-1)}\right) + \tilde{\sigma}_k^2 \leq \left\|\mathbf{h}_k^T\mathbf{W}^{(i)}\right\|_2^2 + \tilde{\sigma}_k^2, \ \ k=1,2,...,K, \tag{18}$$

where $\mathbf{W}^{(i-1)}$ is the optimal solution obtained from the previous iteration. Then, a local solution to $\mathcal{P}4$ can be found via solving a sequence of the following convex optimization problems

$$\mathcal{P}5 : \underset{\mathbf{v},\mathbf{W}^{(i)}}{\text{minimize}} \ \|\mathbf{v}\|_2^2 + \left\|\mathbf{W}^{(i)}\right\|_2^2 \tag{19a}$$

$$\text{subject to}$$

$$\frac{1}{\gamma_0}\mathbf{h}_0^T\mathbf{v} \geq \sqrt{\left\|\mathbf{h}_0^T\mathbf{W}^{(i)}\right\|_2^2 + \tilde{\sigma}_0^2}, \tag{19b}$$

$$\frac{1}{\gamma_k}\left(\mathbf{h}_k^T\mathbf{v}\right)^2 \leq \left\|\mathbf{h}_k^T\mathbf{W}^{(i-1)}\right\|_2^2 + 2\left[\mathbf{W}^{(i-1)}\right]^T\mathbf{h}_k\mathbf{h}_k^T\left(\mathbf{W}^{(i)}-\mathbf{W}^{(i-1)}\right) + \tilde{\sigma}_k^2, \ \ k=1,2,...,K, \tag{19c}$$

$$\left|[\mathbf{v}]_n\right| + \left\|\left[\mathbf{W}^{(i)}\right]_{n,:}\right\|_1 \leq \Delta_n, \ \ n=1,2,...,N_t. \tag{19d}$$

A CCP algorithm for solving $\mathcal{P}4$ is described as follows

---

**Algorithm 1** CCP algorithm for solving problem $\mathcal{P}4$

---

1: **Initialization**

   1) Estimate channel vectors $\mathbf{h}_k$'s and noise variance $\sigma_k^2$'s of Bob and Eves.

   2) Initialize the starting point $\mathbf{W}^{(0)}$ to be sufficiently small to ensure that $\mathcal{P}5$ is feasible.

2: **Iteration**: At the $i-$th iteration

   1) Given $\mathbf{W}^{(i-1)}$ obtained from the previous iteration, solve $\mathcal{P}5$ using CVX toolbox [38].

   2) $i = i + 1$.

3: **Termination**: terminate the iteration when one of the following two conditions is met

   1) $\left\| \mathbf{W}^{(i)} - \mathbf{W}^{(i-1)} \right\|_2^2 \leq \epsilon$, where $\epsilon$ is a predefined convergence error.

   2) $i = L$, where $L$ is a predefined maximum number of iterations.

---

It should be noted that the CCP algorithm guarantees a convergence to a local minimum, which may depend on the chosen initial point $\mathbf{W}^{(0)}$. In oder to obtain a high-quality sub-optimal solution, one can run the algorithm multiple times with different initial points.

*2) SDR Approach:* A drawback of the presented CCP approach is its iterative nature, which may require a number of iterations to get to convergence point. To address this issue, we investigate the use of SDR in this section. SDR is an efficient approximation approach to handle non-convex quadratic programming with or without quadratic constraints [39]. To make use of the technique, let us rewrite $\mathcal{P}3$ as

$$\mathcal{P}6: \quad \underset{\mathbf{v},\mathbf{W}}{\text{minimize}} \quad \text{Tr}\left(\mathbf{v}\mathbf{v}^T\right) + \text{Tr}\left(\mathbf{W}\mathbf{W}^T\right) \tag{20a}$$

$$\text{subject to} \quad \frac{\mathbf{h}_0^T \mathbf{v}\mathbf{v}^T \mathbf{h}_0}{\mathbf{h}_0^T \mathbf{W}\mathbf{W}^T \mathbf{h}_0 + \tilde{\sigma}_0^2} \geq \gamma_0, \tag{20b}$$

$$\frac{\mathbf{h}_k^T \mathbf{v}\mathbf{v}^T \mathbf{h}_k}{\mathbf{h}_k^T \mathbf{W}\mathbf{W}^T \mathbf{h}_k + \tilde{\sigma}_k^2} \leq \gamma_k, \quad k = 1, 2, ..., K, \tag{20c}$$

$$\left| [\mathbf{v}]_n \right| + \left\| [\mathbf{W}]_{n,:} \right\|_1 \leq \Delta_n, \quad n = 1, 2, ..., N_t. \tag{20d}$$

SDR works by introducing the new variables $\mathbf{V} = \mathbf{v}\mathbf{v}^T$ and $\widetilde{\mathbf{W}} = \mathbf{W}\mathbf{W}^T$. These variable transformations are equivalent to rank$(\mathbf{V}) = 1$, rank$(\widetilde{\mathbf{W}}) = $ rank$(\mathbf{W})$, and $\mathbf{V}, \widetilde{\mathbf{W}}$ being symmetric positive semidefinite matrices. Then (20a), (20b), and (20c) are functions of $\mathbf{V}$ and

$\widetilde{\mathbf{W}}$. However, it is seen that (20d) can not be equivalently expressed in terms of $\mathbf{V}$ and $\widetilde{\mathbf{W}}$. To overcome this, one can upper bound $|[\mathbf{v}]_n| + \left\| [\mathbf{W}]_{n,:} \right\|_1$ by

$$\frac{\left( |[\mathbf{v}]_n| + \left\| [\mathbf{W}]_{n,:} \right\|_1 \right)^2}{2} \leq [\mathbf{V}]_{n,n} + \left[ \widetilde{\mathbf{W}} \right]_{n,n}, \quad n = 1, 2, ..., N_t. \tag{21}$$

Now, the following problem generally gives an upper bound solution to $\mathcal{P}5$

$$\mathcal{P}7: \quad \underset{\mathbf{V}, \widetilde{\mathbf{W}}}{\text{minimize}} \ \text{Tr}\left(\mathbf{V}\right) + \text{Tr}\left(\widetilde{\mathbf{W}}\right) \tag{22a}$$

$$\text{subject to} \ \frac{\mathbf{h}_0^T \mathbf{V} \mathbf{h}_0}{\mathbf{h}_0^T \widetilde{\mathbf{W}} \mathbf{h}_0 + \tilde{\sigma}_0^2} \geq \gamma_0, \tag{22b}$$

$$\frac{\mathbf{h}_k^T \mathbf{V} \mathbf{h}_k}{\mathbf{h}_k^T \widetilde{\mathbf{W}} \mathbf{h}_k + \tilde{\sigma}_k^2} \leq \gamma_k, \quad k = 1, 2, .., K, \tag{22c}$$

$$\left[ \mathbf{V} \right]_{n,n} + \left[ \widetilde{\mathbf{W}} \right]_{n,n} \leq \frac{\Delta_n^2}{2}, \quad n = 1, 2, .., N_t, \tag{22d}$$

$$\mathbf{V} \succeq \mathbf{0}, \ \widetilde{\mathbf{W}} \succeq \mathbf{0}, \tag{22e}$$

$$\text{rank}\left(\mathbf{V}\right) = 1, \tag{22f}$$

$$\text{rank}\left(\widetilde{\mathbf{W}}\right) = \text{rank}\left(\mathbf{W}\right), \tag{22g}$$

where (20d) is replaced by a tighter constraint in (22d). However, due to this replacement, a feasible point to $\mathcal{P}6$ might not be feasible to $\mathcal{P}7$. Therefore, it is only applicable to use $\mathcal{P}7$ as an upper bound solution to $\mathcal{P}6$ when the two problems are both feasible. Let us assume this condition to be satisfied. It is seen that $\mathcal{P}7$ is still not a convex optimization problem due to the non-convexity of (22f) and (22g). The next step of SDR technique is to omit the rank constraints, resulting in an approximation to $\mathcal{P}7$ as

$$\mathcal{P}8: \quad \underset{\mathbf{V}, \widetilde{\mathbf{W}}}{\text{minimize}} \ \text{Tr}\left(\mathbf{V}\right) + \text{Tr}\left(\widetilde{\mathbf{W}}\right) \tag{23a}$$

$$\text{subject to} \ (22b) - (22e). $$

The above problem is a semidefinite programing (SDP), which is convex and can be solved efficiently. The question now is how good of the solution to $\mathcal{P}8$ in comparison with that to $\mathcal{P}7$. Interestingly, the following theorem proves that the optimal $\mathbf{V}^*$ and $\widetilde{\mathbf{W}}^*$ to $\mathcal{P}8$ always satisfy $\text{rank}(\mathbf{V}^*) = 1$ and $\text{rank}\left(\widetilde{\mathbf{W}}^*\right) \leq 1$.

**Theorem 1:** Assume that $\mathcal{P}8$ is feasible, then its optimal solutions $\mathbf{V}^*$ and $\widetilde{\mathbf{W}}^*$ satisfy $\text{rank}\left(\mathbf{V}^*\right) = 1$ and $\text{rank}\left(\widetilde{\mathbf{W}}^*\right) \leq 1$.

*Proof*: The proof is presented in Appendix A.

If rank $\left(\widetilde{\mathbf{W}}^*\right) = 0$, then $\widetilde{\mathbf{W}}^* = \mathbf{0}$ and equivalently $\mathbf{W}^* = \mathbf{0}$. Hence, the optimal solution is the same as the no-AN design (i.e. $\mathcal{P}6$ without $\mathbf{W}$). When rank $\left(\widetilde{\mathbf{W}}^*\right) = 1$, the optimal precoders $\mathbf{v}^*$ and $\mathbf{W}^*$ are obtained as $\mathbf{v}^* = \sqrt{\lambda_v}\mathbf{q}_v$ and $\mathbf{W}^* = \sqrt{\lambda_w}\mathbf{q}_w$, where $\mathbf{q}_v$ and $\mathbf{q}_w$ are the eigenvectors of $\mathbf{V}^*$ and $\widetilde{\mathbf{W}}^*$, which associate with the non-zero eigenvalues $\lambda_v$ and $\lambda_w$, respectively. This implies that it suffices to choose the AN size $N_s = 1$ for the optimal design in the case of known Eve's CSI, resulting in rank $\left(\widetilde{\mathbf{W}}\right) = \text{rank}\left(\mathbf{W}\right) = 1$. Thus, rank $\left(\widetilde{\mathbf{W}}\right) = \text{rank}\left(\mathbf{W}\right)$ always hold, which proves that $\mathcal{P}7$ and $\mathcal{P}8$ are equivalent.

*3) Complexity Analysis:* In this section, we analyze the computational complexities of the presented CCP and SDR approach. Following the same arguments in [25], [31], the number of Newton steps, denotes as $N_{step}$, is used as the complexity measure. Assume that the interior point algorithm is used to solve the convex optimization problem under consideration, $N_{step}$ is the number of recursive iterations required for the algorithm to find a local solution. Then, for a non-linear convex problem, the worst-case $N_{step}$ to reach a local solution is given by

$$N_{step} \sim \sqrt{\text{problem size}}, \tag{24}$$

where the problem size is the number of optimization scalar variables.

In the case of CCP approach, the number of optimization scalar variables of the surrogate problem $\mathcal{P}5$ is $2N_T$. Since the **Algorithm 1** involves solving $\mathcal{P}5$ at most $L$ times, the worst-case complexity of the CCP approach is $N_{step}^{\text{CCP}} \sim \sqrt{2N_T}L$. For the SDR approach, the number of optimization scalar variables is $2N_T^2$. Its complexity is thus $N_{step}^{\text{SDR}} \sim \sqrt{2}N_T$. We have

$$\frac{N_{step}^{\text{SDR}}}{N_{step}^{\text{CCP}}} \sim \frac{\sqrt{N_T}}{L}, \tag{25}$$

which indicates that the SDR might exhibit a higher complexity when there is a large number of LED luminaries. On the other hand, the complexity of the CCP approach could be higher if the predefined convergence error $\epsilon$ is set too small and/or the number of Eves is large as this results in more iterations needed for the **Algorithm 1** to converge.

## IV. ROBUST DESIGNS WITH CHANNEL UNCERTAINTY

The assumption that Bob and Eves' CSI are perfectly known at the transmitter is rather unrealistic, especially in the case of moving users. Unlike RF communications where the CSI estimation can usually be done at the transmitter using uplink-downlink reciprocity, the CSI in

the case of VLC should be estimated at the receiver then fed back to the transmitter using an RF or infrared uplink. As a consequence, outdated CSI estimations are inevitable due to user movement. In this section, we hence investigate robust AN-aided precoding designs taking into account uncertainties in the channel model.

## A. Channel Uncertainty Model

We consider an additive uncertainty in the channel model described by

$$\mathbf{h}_k = \hat{\mathbf{h}}_k + \mathbf{u}_k, \tag{26}$$

$\hat{\mathbf{h}}_k$ is the estimation of the actual channel gain $\mathbf{h}_k$ and $\mathbf{u}_k$ represents the estimation error vector. In case of outdated CSI caused by user movement, the estimation error can be bounded by

$$\|\mathbf{u}_k\|_2 \leq \delta_k, \tag{27}$$

where $\delta_k$ depends on maximal changes in channel gains between CSI estimation and feedback. Therefore, this error bound is a function of user velocity, feedback rate, and $\hat{\mathbf{h}}_k$. Since an explicit characterization of $\delta_k$ could be cumbersome, we adopt this simplified model $\delta_k = \alpha \|\hat{\mathbf{h}}_k\|_2$, where $\alpha \in [0, 1)$ measures the magnitude of CSI uncertainty [40].

## B. Unknown Eves' CSI

In this scenario, the robust design aims to guarantee a certain threshold of the minimal (i.e., worst-case) Bob's SINR. With the above-defined uncertainty model and the AN design given in (12), the worst-case Bob's SINR is given by

$$\text{SINR}_0^{\text{worst-case}} = \min_{\|\mathbf{e}_0\|_2 \leq \delta_0} \frac{\left| \left( \hat{\mathbf{h}}_0^T + \mathbf{e}_0^T \right) \mathbf{v} \right|^2}{\left( \rho \frac{\min_n \Delta_n}{\max_n \left\| [\overline{\mathbf{W}}]_{n,:} \right\|_1} \right)^2 \left\| \mathbf{e}_0^T \overline{\mathbf{W}} \right\|_2^2 + \widetilde{\sigma}_0^2} \geq \frac{\min_{\|\mathbf{e}_0\|_2 \leq \delta_0} \left| \left( \hat{\mathbf{h}}_0^T + \mathbf{e}_0^T \right) \mathbf{v} \right|^2}{\left( \rho \frac{\min_n \Delta_n}{\max_n \left\| [\overline{\mathbf{W}}]_{n,:} \right\|_1} \right)^2 \max_{\|\mathbf{e}_0\|_2 \leq \delta_0} \left\| \mathbf{e}_0^T \overline{\mathbf{W}} \right\|_2^2 + \widetilde{\sigma}_0^2} \tag{28}$$

The robust AN-aided precoding design can then be formulated as

$$\mathcal{P}9 : \underset{\mathbf{v}}{\text{minimize}} \quad \|\mathbf{v}\|_2^2 \tag{29a}$$

$$\text{subject to} \quad \frac{\min_{\|\mathbf{u}_0\|_2 \leq \delta_0} \left| \left( \hat{\mathbf{h}}_0^T + \mathbf{u}_0^T \right) \mathbf{v} \right|^2}{\left( \rho \frac{\min_n \Delta_n}{\max_n \left\| [\overline{\mathbf{W}}]_{n,:} \right\|_1} \right)^2 \max_{\|\mathbf{u}_0\|_2 \leq \delta_0} \left\| \mathbf{u}_0^T \overline{\mathbf{W}} \right\|_2^2 + \widetilde{\sigma}_0^2} \geq \gamma_0, \tag{29b}$$

$$\left|[\mathbf{v}]_n\right| \leq \Delta_n - \rho \frac{\min\limits_{n}\Delta_n}{\max\limits_{n}\left\|\left[\overline{\mathbf{W}}\right]_{n,:}\right\|_1} \left\|\left[\overline{\mathbf{W}}\right]_{n,:}\right\|_1, \quad n = 1, 2, ..., N_t. \tag{29c}$$

It is seen that $\mathcal{P}9$ is not a convex optimization problem due to the non-convex constraint (29b). To convexify it, we make use of the following lemma.

**Lemma 1 (S-Procedure [41]):** Let $\mathbf{A}_1$, $\mathbf{A}_2 \in \mathbb{R}^{n \times n}$ be symmetric matrices, $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{x} \in \mathbb{R}^n$, and $c_1$, $c_2 \in \mathbb{R}$. The implication

$$\mathbf{x}^T \mathbf{A}_1 \mathbf{x} + 2\mathbf{g}_1^T \mathbf{x} + c_1 \leq 0 \implies \mathbf{x}^T \mathbf{A}_2 \mathbf{x} + 2\mathbf{g}_2^T \mathbf{x} + c_2 \leq 0 \tag{30}$$

holds if and only if there exists a non-negative $\lambda$ such that

$$\lambda \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^T & c_1 \end{bmatrix} - \begin{bmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^T & c_2 \end{bmatrix} \succeq 0 \tag{31}$$

provided that there exists a point $\hat{\mathbf{x}}$ with $\hat{\mathbf{x}}^T \mathbf{A}_1 \hat{\mathbf{x}} + 2\mathbf{g}_1^T \hat{\mathbf{x}} + c_1 < 0$.

Next, we introduce the following slack variables $\tau_0 = \min\limits_{\|\mathbf{u}_0\|_2 \leq \delta_0} \left|\left(\hat{\mathbf{h}}_0^T + \mathbf{u}_0^T\right)\mathbf{v}\right|^2$ and $\omega_0 = \max\limits_{\|\mathbf{u}_0\|_2 \leq \delta_0} \left\|\mathbf{u}_0^T \overline{\mathbf{W}}\right\|_2^2$. Now, we obtain the following implications

$$\|\mathbf{u}_0\|_2^2 - \delta_0^2 \leq 0 \implies \tau_0 - \left|\left(\hat{\mathbf{h}}_0^T + \mathbf{u}_0^T\right)\mathbf{v}\right|^2 \leq 0, \tag{32}$$

$$\|\mathbf{u}_0\|_2^2 - \delta_0^2 \leq 0 \implies \left\|\mathbf{u}_0^T \overline{\mathbf{W}}\right\|_2^2 - \omega_0 \leq 0. \tag{33}$$

Applying **Lemma 1** for the implication in (32) with $\mathbf{A}_1 = \mathbf{I}_{N_t}$, $\mathbf{A}_2 = -\mathbf{v}\mathbf{v}^T$, $\mathbf{b}_1 = \mathbf{0}_{N_t}$, $\mathbf{b}_2 = -(\mathbf{v}\mathbf{v}^T)\hat{\mathbf{h}}_0$, $c_1 = -\delta_0^2$, $c_2 = \tau_0 - \hat{\mathbf{h}}_0^T\mathbf{v}\mathbf{v}^T\hat{\mathbf{h}}_0$ yields

$$\begin{bmatrix} \lambda_0 \mathbf{I}_{N_t} + \mathbf{v}\mathbf{v}^T & \mathbf{v}\mathbf{v}^T\hat{\mathbf{h}}_0 \\ \hat{\mathbf{h}}_0^T\mathbf{v}\mathbf{v}^T & -\lambda_0\delta_0^2 + \hat{\mathbf{h}}_0^T\mathbf{v}\mathbf{v}^T\hat{\mathbf{h}}_0 - \tau_0 \end{bmatrix} \succeq 0, \lambda_0 \geq 0. \tag{34}$$

Similarly, the implication in (33) leads to

$$\begin{bmatrix} \xi_0 \mathbf{I}_{N_t} - \overline{\mathbf{W}}\,\overline{\mathbf{W}}^T & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t}^T & -\xi_0\delta_0^2 + \omega_0 \end{bmatrix} \succeq 0, \xi_0 \geq 0. \tag{35}$$

It is obvious that (35) is convex yet (34) it not due to the non-linear tem $\mathbf{v}\mathbf{v}^T$. To deal with this, we again use the SDR technique described in the previous section. By defining $\mathbf{V}_{\text{rb}} = \mathbf{v}\mathbf{v}^T$ and omitting the rank constraint $\text{rank}(\mathbf{V}_{\text{rb}}) = 1$, $\mathcal{P}9$ can then be reformulated to be a convex optimization problem as

$$\mathcal{P}10 : \underset{\mathbf{V}_{\text{rb}}, \tau_0, \omega_0,}{\text{minimize}} \quad \text{Tr}(\mathbf{V}_{\text{rb}}) \tag{36a}$$
$$\lambda_0, \xi_0$$

18

subject to

$$\begin{bmatrix} \lambda_0 \mathbf{I}_{N_t} + \mathbf{V}_{\text{rb}} & \mathbf{V}_{\text{rb}} \hat{\mathbf{h}}_0 \\ \hat{\mathbf{h}}_0^T \mathbf{V}_{\text{rb}} & -\lambda_0 \delta_0^2 + \hat{\mathbf{h}}_0^T \mathbf{V}_{\text{rb}} \hat{\mathbf{h}}_0 - \tau_0 \end{bmatrix} \succeq 0, \tag{36b}$$

$$\begin{bmatrix} \xi_0 \mathbf{I}_{N_t} - \overline{\mathbf{W}}\overline{\mathbf{W}}^T & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t}^T & -\xi_0 \delta_0^2 + \omega_0 \end{bmatrix} \succeq 0, \tag{36c}$$

$$\mathbf{V}_{\text{rb}} \succeq 0 \tag{36d}$$

$$[\mathbf{V}_{\text{rb}}]_{n,n} \leq \left( \Delta_n - \rho \frac{\min\limits_{n} \Delta_n}{\max\limits_{n} \left\| \left[\overline{\mathbf{W}}\right]_{n,:} \right\|_1} \left\| \left[\overline{\mathbf{W}}\right]_{n,:} \right\|_1 \right)^2, \quad n = 1, 2, ..., N_t, \tag{36e}$$

$$\tau_0 - \gamma_0 \left( \left( \rho \frac{\min\limits_{n} \Delta_n}{\max\limits_{n} \left\| \left[\overline{\mathbf{W}}\right]_{n,:} \right\|_1} \right)^2 \omega_0 + \tilde{\sigma}_0^2 \right) \geq 0, \tag{36f}$$

$$\lambda_0 \geq 0, \xi_0 \geq 0. \tag{36g}$$

Similar to **Theorem 1**, the following theorem proves the equivalence of $\mathcal{P}9$ and $\mathcal{P}10$.

**Theorem 2:** Assume that $\mathcal{P}9$ is feasible, then its optimal $\mathbf{V}_{\text{rb}}^*$ solution satisfies $\text{rank}\,(\mathbf{V}_{\text{rb}}^*) = 1$.

*Proof*: The proof is presented in Appendix B.

## C. Known Eves' CSI

In this case, the worst-case Bob's and Eves' SINRs are given by

$$\text{SINR}_0^{\text{worst-case}} = \min_{\|\mathbf{u}_0\|_2 \leq \delta_0} \frac{\left| \left( \hat{\mathbf{h}}_0^T + \mathbf{u}_0^T \right) \mathbf{v} \right|^2}{\left\| \left( \hat{\mathbf{h}}_0^T + \mathbf{u}_0^T \right) \mathbf{W} \right\|_2^2 + \tilde{\sigma}_0^2} \geq \frac{\min\limits_{\|\mathbf{u}_0\|_2 \leq \delta_0} \left| \left( \hat{\mathbf{h}}_0^T + \mathbf{u}_0^T \right) \mathbf{v} \right|^2}{\max\limits_{\|\mathbf{u}_0\|_2 \leq \delta_0} \left\| \left( \hat{\mathbf{h}}_0^T + \mathbf{u}_0^T \right) \mathbf{W} \right\|_2^2 + \tilde{\sigma}_0^2}, \tag{37}$$

$$\text{SINR}_k^{\text{worst-case}} = \max_{\|\mathbf{u}_k\|_2 \leq \delta_k} \frac{\left| \left( \hat{\mathbf{h}}_k^T + \mathbf{u}_k^T \right) \mathbf{v} \right|^2}{\left\| \left( \hat{\mathbf{h}}_k^T + \mathbf{u}_k^T \right) \mathbf{W} \right\|_2^2 + \tilde{\sigma}_k^2} \leq \frac{\max\limits_{\|\mathbf{u}_k\|_2 \leq \delta_k} \left| \left( \hat{\mathbf{h}}_k^T + \mathbf{u}_k^T \right) \mathbf{v} \right|^2}{\min\limits_{\|\mathbf{u}_k\|_2 \leq \delta_k} \left\| \left( \hat{\mathbf{h}}_k^T + \mathbf{u}_k^T \right) \mathbf{W} \right\|_2^2 + \tilde{\sigma}_k^2}, \quad k = 1, ..., K. \tag{38}$$

Let us define $\kappa_0 = \max\limits_{\|\mathbf{u}_0\|_2 \leq \delta_0} \left\| \left( \hat{\mathbf{h}}_0^T + \mathbf{u}_0^T \right) \mathbf{W} \right\|_2^2$, $\upsilon_k = \min\limits_{\|\mathbf{u}_k\|_2 \leq \delta_k} \left\| \left( \hat{\mathbf{h}}_k^T + \mathbf{u}_k^T \right) \mathbf{W} \right\|_2^2$, and $\varphi_k = \max\limits_{\|\mathbf{u}_k\|_2 \leq \delta_k} \left| \left( \hat{\mathbf{h}}_k^T + \mathbf{u}_k^T \right) \mathbf{v} \right|^2$. Using the same procedures described in the previous section, we obtain

the following constraints

$$\begin{bmatrix} \zeta_0 \mathbf{I}_{N_t} - \mathbf{W}\mathbf{W}^T & -\mathbf{W}\mathbf{W}^T \hat{\mathbf{h}}_0 \\ -\hat{\mathbf{h}}_0^T \mathbf{W}\mathbf{W}^T & -\zeta_0 \delta_0^2 - \hat{\mathbf{h}}_0^T \mathbf{W}\mathbf{W}^T \hat{\mathbf{h}}_0 + \kappa_0 \end{bmatrix} \succeq 0, \zeta_0 \geq 0, \tag{39}$$

$$\begin{bmatrix} \varrho_k \mathbf{I}_{N_t} + \mathbf{W}\mathbf{W}^T & \mathbf{W}\mathbf{W}^T \hat{\mathbf{h}}_k \\ \hat{\mathbf{h}}_k^T \mathbf{W}\mathbf{W}^T & -\varrho_k \delta_k^2 + \hat{\mathbf{h}}_k^T \mathbf{W}\mathbf{W}^T \hat{\mathbf{h}}_k - \upsilon_k \end{bmatrix} \succeq 0, \varrho_k \geq 0, \ k = 1, ..., K, \tag{40}$$

$$\begin{bmatrix} \varsigma_k \mathbf{I}_{N_t} - \mathbf{v}\mathbf{v}^T & -\mathbf{v}\mathbf{v}^T \hat{\mathbf{h}}_k \\ -\hat{\mathbf{h}}_k^T \mathbf{v}\mathbf{v}^T & -\varsigma_k \delta_k^2 - \hat{\mathbf{h}}_k^T \mathbf{v}\mathbf{v}^T \hat{\mathbf{h}}_k + \varphi_k \end{bmatrix} \succeq 0, \varsigma_k \geq 0, \ k = 1, ..., K. \tag{41}$$

Finally, by using the transformation $\widetilde{\mathbf{W}}_{\mathrm{rb}} = \mathbf{W}\mathbf{W}^T$ and omitting the rank constraint $\mathrm{rank}\left(\widetilde{\mathbf{W}}_{\mathrm{rb}}\right) = \mathrm{rank}\left(\mathbf{W}\right)$, the robust AN-aided design problem for this case is given as

$$\mathcal{P}11: \underset{\substack{\mathbf{V}_{\mathrm{rb}}, \widetilde{\mathbf{W}}_{\mathrm{rb}}, \\ \tau_0, \kappa_0, \lambda_0, \zeta_0 \\ \upsilon_k, \varphi_k, \varrho_k, \varsigma_k}}{\mathrm{minimize}} \quad \mathrm{Tr}\left(\mathbf{V}_{\mathrm{rb}}\right) + \mathrm{Tr}\left(\widetilde{\mathbf{W}}_{\mathrm{rb}}\right) \tag{42a}$$

subject to

$$\begin{bmatrix} \lambda_0 \mathbf{I}_{N_t} + \mathbf{V}_{\mathrm{rb}} & \mathbf{V}_{\mathrm{rb}} \hat{\mathbf{h}}_0 \\ \hat{\mathbf{h}}_0^T \mathbf{V}_{\mathrm{rb}} & -\lambda_0 \delta_0^2 + \hat{\mathbf{h}}_0^T \mathbf{V}_{\mathrm{rb}} \hat{\mathbf{h}}_0 - \tau_0 \end{bmatrix} \succeq 0, \tag{42b}$$

$$\begin{bmatrix} \zeta_0 \mathbf{I}_{N_t} - \widetilde{\mathbf{W}}_{\mathrm{rb}} & -\widetilde{\mathbf{W}}_{\mathrm{rb}} \hat{\mathbf{h}}_0 \\ -\hat{\mathbf{h}}_0^T \widetilde{\mathbf{W}}_{\mathrm{rb}} & -\zeta_0 \delta_0^2 - \hat{\mathbf{h}}_0^T \widetilde{\mathbf{W}}_{\mathrm{rb}} \hat{\mathbf{h}}_0 + \kappa_0 \end{bmatrix} \succeq 0, \tag{42c}$$

$$\begin{bmatrix} \varrho_k \mathbf{I}_{N_t} + \widetilde{\mathbf{W}}_{\mathrm{rb}} & \widetilde{\mathbf{W}}_{\mathrm{rb}} \hat{\mathbf{h}}_k \\ \hat{\mathbf{h}}_k^T \widetilde{\mathbf{W}}_{\mathrm{rb}} & -\varrho_k \delta_k^2 + \hat{\mathbf{h}}_k^T \widetilde{\mathbf{W}}_{\mathrm{rb}} \hat{\mathbf{h}}_k - \upsilon_k \end{bmatrix} \succeq 0, \ k = 1, 2, ..., K, \tag{42d}$$

$$\begin{bmatrix} \varsigma_k \mathbf{I}_{N_t} - \mathbf{V}_{\mathrm{rb}} & -\mathbf{V}_{\mathrm{rb}} \hat{\mathbf{h}}_k \\ -\hat{\mathbf{h}}_k^T \mathbf{V}_{\mathrm{rb}} & -\varsigma_k \delta_k^2 - \hat{\mathbf{h}}_k^T \mathbf{V}_{\mathrm{rb}} \hat{\mathbf{h}}_k + \varphi_k \end{bmatrix} \succeq 0, \ k = 1, 2, ..., K, \tag{42e}$$

$$\mathbf{V}_{\mathrm{rb}} \succeq 0, \widetilde{\mathbf{W}}_{\mathrm{rb}} \succeq 0, \tag{42f}$$

$$[\mathbf{V}_{\mathrm{rb}}]_{n,n} + \left[\widetilde{\mathbf{W}}_{\mathrm{rb}}\right]_{n,n} \leq \Delta_n, \ n = 1, 2, ..., N_t, \tag{42g}$$

$$\tau_0 - \gamma_0(\kappa_0 + \tilde{\sigma}_0^2) \geq 0, \tag{42h}$$

$$-\varphi_k + \gamma_k(\upsilon_k + \tilde{\sigma}_k^2) \geq 0, \ k = 1, 2, ..., K, \tag{42i}$$

$$\lambda_0 \geq 0, \zeta_0 \geq 0, \varrho_k \geq 0, \varsigma_k \geq 0. \tag{42j}$$

**Theorem 3:** Assume that $\mathcal{P}11$ is feasible, then its optimal $\mathbf{V}_{\text{rb}}^*$ and $\widetilde{\mathbf{W}}_{\text{rb}}^*$ solutions satisfy $\text{rank}\left(\mathbf{V}_{\text{rb}}^*\right) = 1$ and $\text{rank}\left(\widetilde{\mathbf{W}}_{\text{rb}}^*\right) \leq 1$.

The proof for the above theorem follows the same arguments described in the proof of **Theorem 2** in Appendix B. Therefore, we omit it for the sake of brevity.

## V. NUMERICAL RESULTS AND DISCUSSIONS

This section presents numerical results to illustrate the performance of the proposed design. The room configuration is shown in Fig. 1 with $N_t = 4$ and a Cartesian coordinate system whose the origin is the center of the floor is used to specify the positions of Bob, Eves and the LED luminaries. It is assumed that the legitimate and unauthorized users are located $0.5$m above the floor. The average optical power of each LED luminary is set to $\overline{p}_1^t = \overline{p}_2^t = ... = \overline{p}_{N_t}^t = 35$ dBm, which is approximately equivalent to $3.16$ Watts. For the sake of simulations, we also assume that $I_{\min} = 0$ and $I_{\max} \gg I_n^{\text{DC}}$, then $\Delta_n = I_n^{\text{DC}}$. In addition, all simulations are obtained by averaging the results from $10,000$ different channel realizations of Bob and Eves (according to $10,000$ randomly distributed positions). Unless otherwise specified, the system parameters are the same as those given in [34, Table I].

### A. Unknown Eves' CSI

Firstly, Fig. 2 shows the feasibility probability of $\mathcal{P}1$ in accordance with Bob's SINR threshold $\gamma_0$ for three different values of the AN adjusting power parameter $\rho = 0.2$, $0.5$, and $0.8$. The AN size is set to 1. At a given $\gamma_0$, as $\rho$ increases, less power is allocated to the information-bearing signal. Hence, the probability that $\mathcal{P}1$ is feasible, which depends on the feasibility of (10b), decreases. The results indicate that setting $\gamma_0$ below $35$ dB is reasonable to ensure $\mathcal{P}1$ is feasible for a sufficiently large number of Bob's positions.

Next, the feasibility probability of the robust design in $\mathcal{P}9$ is illustrated in Fig. 3 for different values of channel uncertainty level $\alpha = 0.01$, $0.03$, and $0.05$ when $\rho = 0.5$ and $N_s = 1$. Compared to the design with perfect CSI, the channel uncertainty causes significantly lower feasibility probabilities in the robust design. For example, when $\alpha = 0.05$, to achieve a probability close to 1, the threshold for Bob's SINR should be chosen below $28$ dB, which is $10$ dB smaller than that of the perfect CSI design in $\mathcal{P}1$.

In Fig. 4, the total transmit powers and Eve's SINR are illustrated as functions of $\rho$ for different AN sizes. The threshold for Bob's SINR is set to $\gamma_0 = 30$ dB. The results of perfect
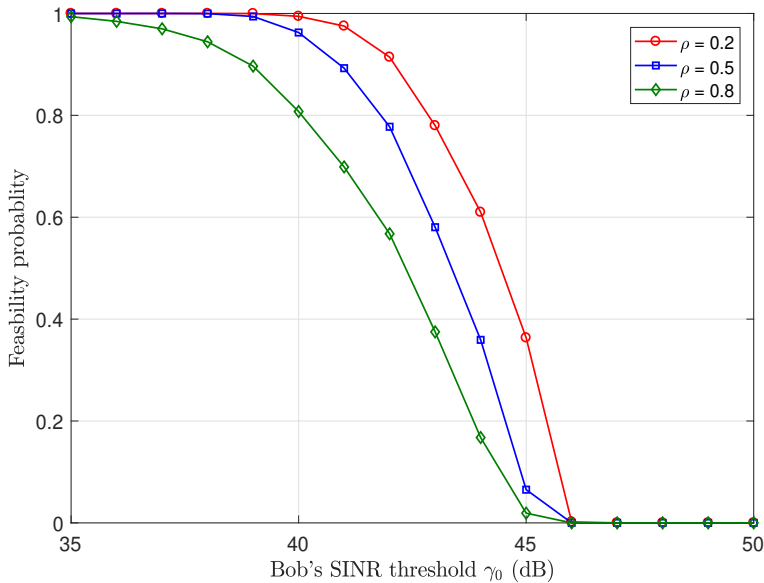
IEEE Transactions on Wireless Communications

Figure 2: Feasibility probability of $\mathcal{P}1$ versus Bob's SINR threshold $\gamma_0$ for different AN power factors $\rho$. AN size $N_s = 1$.

and uncertain CSI designs are compared, in which the magnitude of uncertainty $\alpha = 0.01$ is chosen. The claim of **Proposition 2** is also demonstrated in the figure, where the resulting Bob's SINRs are exactly equal to $\gamma_0$. This, however, does not hold in the case of robust design, where Bob's SINRs increase as $\rho$ increases. Those unwanted increases in Bob's SINRs result in the higher transmit powers in the case of uncertain CSI. It is observed that significant gaps between Bob's and Eve's SINRs can be achieved when AN is employed in both designs. Specifically, when $\rho = 0.1$, these are roughly 15, 20, and 21 dB in case of $N_s = 1$, 2, and 3, respectively. Moreover, these gaps increase following an increase of $\rho$ at the expense of more consumed transmit power. Though Eve's SINRs can be degraded by increasing either $N_s$ or $\rho$, we can notice that increasing AN size might be more beneficial. For example in the case of perfect CSI design, to target Eve's SINR at 0 dB, $\rho = 0.75$, 0.35, and 0.3 are required for $N_s = 1$, 2, and 3, respectively. These values of $\rho$ result in the transmit power of 41.14, 37, and 36.53 dBm. The reason for this is because the impact of $\rho$ on the AN amplitude is multiplicative while it is additive in the case of AN size (see (12)). As a result, the reduction of degrees of freedom for selecting the precoder $\mathbf{v}$ due to increased $\rho$ is more severe than that caused by increasing AN size.
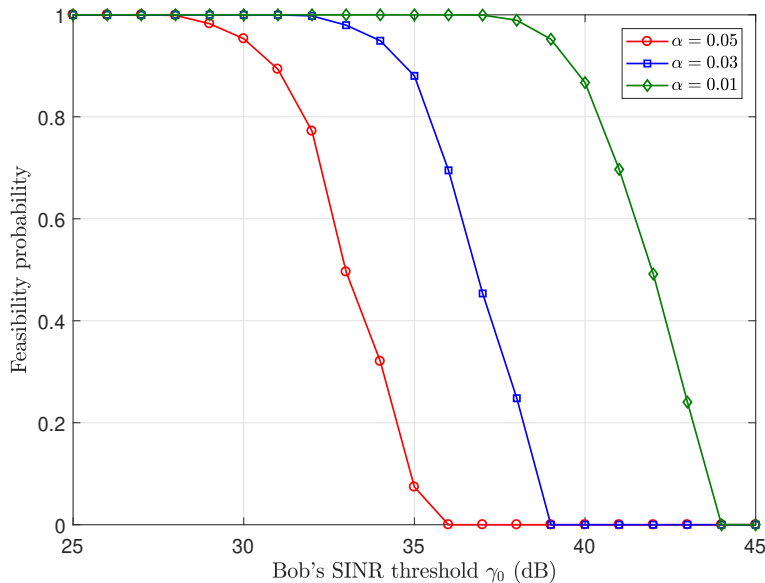
Figure 3: Feasibility probability of $\mathcal{P}9$ versus Bob's SINR threshold $\gamma_0$ for different magnitudes of channel uncertainty. $\rho = 0.5$ and $N_s = 1$.

*B. Known Eves' CSI*

We show in Fig. 5 the feasibility probabilities of $\mathcal{P}8$ with respect to the thresholds of Bob's and Eves' SINRs for two designs: with AN and no-AN. It is observed that the feasibility probabilities of the two designs are not very different. This could be intuitively explained as follows. The constraint (22b) in the case of no-AN scheme is satisfied with a higher probability than that in the case of AN-aided precoding scheme since if $\frac{\mathbf{h}_0^T \mathbf{V} \mathbf{h}_0}{\mathbf{h}_0^T \widetilde{\mathbf{W}} \mathbf{h}_0 + \tilde{\sigma}_0^2} \geq \gamma_0$ holds then $\frac{\mathbf{h}_0^T \mathbf{V} \mathbf{h}_0}{\tilde{\sigma}_0^2} \geq \gamma_0$ also holds. Similarly, one can see that (22c) in the case of AN-aided precoding has a higher probability to be satisfied than that in the case of no-AN. Roughly speaking, compared to the no-AN scheme, using AN decreases the feasibility probability of (22b) yet increases that of (22c). As a result, the feasibility probabilities between AN-aided precoding and no-AN should not very different qualitatively. More specifically, it should be noted that the use of the upper bound (21) is not necessary in the case of no-AN design. This explains why the feasibility probability of no-AN is slightly higher than that of the design with AN. In the design with AN, we notice that the probability decreases following an increase of $\gamma_0$ and a decrease in $\gamma_k$. However, it is interesting that in the case of no-AN at a given $\gamma_0$, the feasibility probability stays almost unchanged with respect to $\gamma_k$. In this known Eves' CSI design, choosing $\gamma_0$ below 30 dB is recommended to
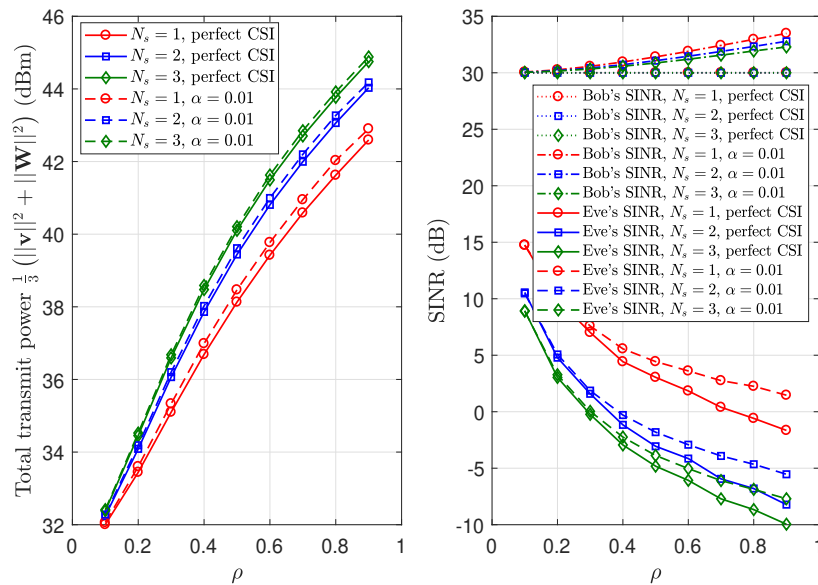
Figure 4: Total transmit power and Eve's SINR versus AN power factor $\rho$ for different AN sizes $N_s$. Bob's SINR threshold $\gamma_0 = 30$ dB.

ensure a higher than $0.9$ feasibility probability.

The relationship between the total transmit power and Bob's SINR threshold for different numbers of Eves is highlighted in Fig. 6. Thresholds for Eves' SINRs are set the same $\gamma_k$'s $= 0$ dB. We observe the benefit of AN-aided precoding over the no-AN design. For example, at $\gamma_0 = 30$ dB and in the case of perfect CSI, the power savings thanks to the use of AN are $3.16$ and $4$ dB when $K = 1, 2$, respectively. According to **Theorem 1**, under some specific setting of $\mathbf{h}_0$, $\mathbf{h}_k$'s, $\gamma_0$, $\gamma_k$'s, the optimal solution to $\mathbf{W}$ is $\mathbf{W} = \mathbf{0}$, which is the same as the case of no-AN design. The superiority of AN-aided precoding shown in this figure implies that of $10,000$ channel realizations taken for the simulation, there is a larger portion that the optimal solution gives rank-1 to $\widetilde{\mathbf{W}}$ (i.e. $\mathbf{W}^* \neq \mathbf{0}$). In other words, the AN-aided precoding is statistically better than the no-AN design. It is seen that the power penalties caused by channel uncertainties increase with respect to $\gamma_0$. We also notice that while the CCP and SDR approaches offer the same solution in the case $K = 1$, the CCP results in a little bit higher transmit power when $K = 2$. This implies that the quality of the CCP solution is getting worse when $K$ increases as we will elaborate further in Fig. 8.

Figure 7 compares the total transmit powers of the known and unknown Eves' CSI design for
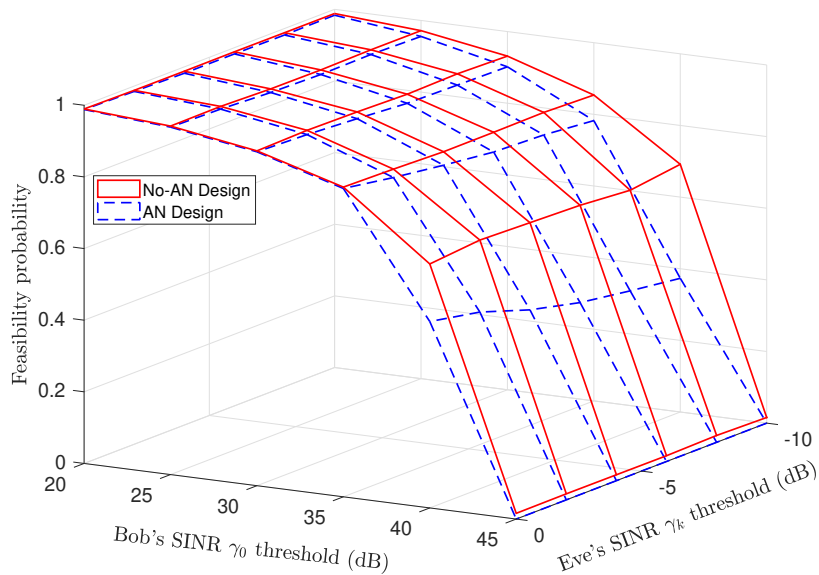
Figure 5: Feasibility probability of $\mathcal{P}8$ with respect to Bob's SINR and Eve's SINR thresholds.

both scenarios of perfect and uncertain channel estimation. For simplicity, we again assume that $\gamma_k$'s are set the same as $\gamma_k$'s $= 0$ dB. For comparison, in the case of unknown CSI design, $\rho$'s are chosen to be $0.2$, $0.21$ for the perfect and uncertain CSI, respectively, when $\gamma_0 = 20$ dB so that the resulting Eves's SINRs are $0$ dB. Similarly, $\rho$'s are $0.38$ and $0.46$ when $\gamma_0 = 25$ dB. We see considerable reductions in the transmit power in the case of known CSI design. Moreover, the total transmitted powers and power penalties due to channel uncertainties increase in accordance with $K$ in the case of known Eves' CSI. This result is intuitive since a larger number of Eves means more constraints are involved in the design. However, in the case of unknown CSI, as the number of Eves is not relevant to the design, the total transmit powers are constant with respect to $K$.

The convergence behavior of **Algorithm 1** is illustrated in Fig. 8 for different number of Eves. We observe a significant difference in the convergence error between $K = 1$ and $K > 1$. Specifically, to target the convergence error of $\epsilon = 10^{-2}$, the algorithm needs, on average, only 3 iterations in the case $K = 1$ while it requires 8 and 9 iterations when $K = 2$ and 3, respectively. This explains the high computational expense of the CCP approach when the number of Eves is large as we mentioned in the complexity analysis part.
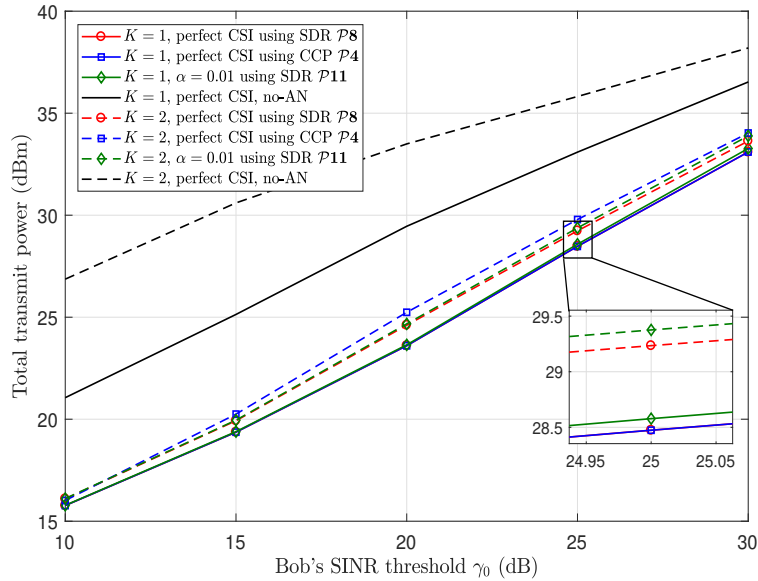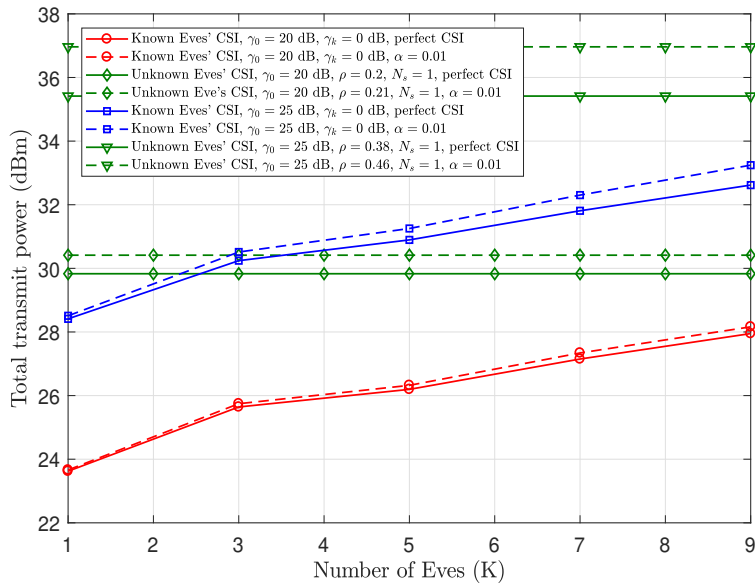
Figure 6: Total transmit power versus Bob's SINR threshold $\gamma_0$. Eves' SINR thresholds $\gamma_k$'s = 0 dB.

## VI. CONCLUSIONS

In this paper, we studied AN-aided precoding designs for enhancing PLS in multi-user VLC systems from the perspective of energy efficiency. Depending on the availability of unauthorized users' CSI at the transmitter, two different design approaches were investigated. In the case of unknown unauthorized users' CSI, the design problem was convex and it was revealed that increasing the AN size is more beneficial in saving power. In the case of known unauthorized users' CSI, however, the design problem was shown to be non-convex. Thus, two different sub-optimal complementary approaches, namely: CCP and SDR, were presented to solve the design. Though the SDR is not always applicable, it does not involve an iterative procedure as the CCP does. Numerical results showed that while the two approaches offered the same solution in the case $K = 1$, the SDR exhibited a better solution quality when $K > 1$.

## ACKNOWLEDGEMENTS

Figure 7: Total transmit power versus the number of Eves $K$.

## APPENDIX A

### PROOF OF THEOREM 1

We rewrite $\mathcal{P}8$ in the following form

$$\mathcal{P}8.1 \quad \underset{\mathbf{V}, \widetilde{\mathbf{W}}}{\text{minimize}} \quad \text{Tr}\left(\mathbf{V}\right) + \text{Tr}\left(\widetilde{\mathbf{W}}\right) \tag{43a}$$

$$\text{subject to} \quad \text{Tr}\left(\mathbf{V}\mathbf{H}_0\right) \geq \gamma_0 \left(\text{Tr}\left(\widetilde{\mathbf{W}}\mathbf{H}_0\right) + \tilde{\sigma}_0^2\right), \tag{43b}$$
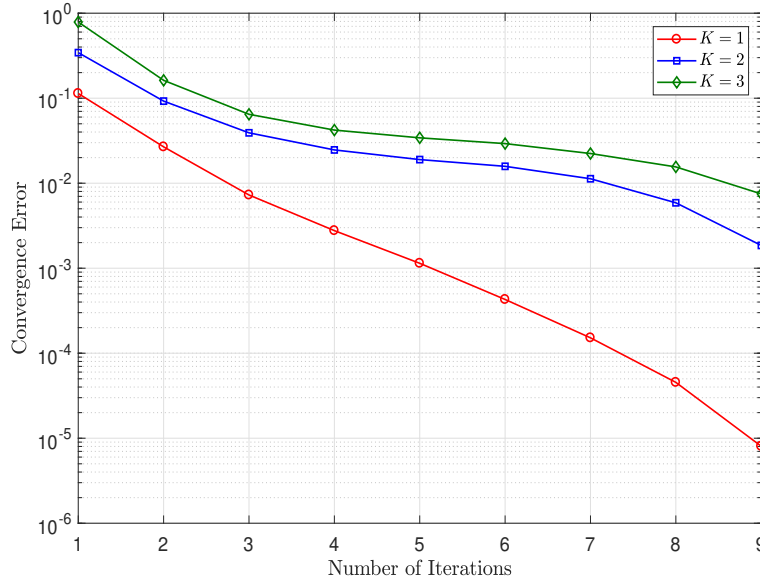
$$\text{Tr}\left(\mathbf{V}\mathbf{H}_k\right) \leq \gamma_k \left(\text{Tr}\left(\widetilde{\mathbf{W}}\mathbf{H}_k\right) + \tilde{\sigma}_k^2\right), \quad k = 1, 2, ..., K, \tag{43c}$$

$$\text{Tr}\left(\mathbf{V}\mathbf{E}_n\right) + \text{Tr}\left(\widetilde{\mathbf{W}}\mathbf{E}_n\right) \leq \Delta_n^2, \quad n = 1, 2, ..., N_t, \tag{43d}$$

$$\mathbf{V} \succeq \mathbf{0}, \ \widetilde{\mathbf{W}} \succeq \mathbf{0}, \tag{43e}$$

where $\mathbf{H}_k = \mathbf{h}_k \mathbf{h}_k^T \succeq 0 \ \forall k = 0, 1, .., K$ and $\mathbf{E}_n = \mathbf{e}_n \mathbf{e}_n^T \succeq 0 \ \forall n = 1, 2, ..., N_t$. The Lagrangian function of $\mathcal{P}8.1$ can then be defined as

$$\mathcal{L}\left(\mathbf{V}, \widetilde{\mathbf{W}}, \{\mu_k\}, \{\nu_n\}, \mathbf{X}_1, \mathbf{X}_2\right) = \text{Tr}\left(\mathbf{V}\right) + \text{Tr}\left(\widetilde{\mathbf{W}}\right)$$

$$- \text{Tr}\left(\mathbf{X}_1\mathbf{V}\right) - \text{Tr}\left(\mathbf{X}_2\widetilde{\mathbf{W}}\right) - \mu_0 \left(\text{Tr}\left(\mathbf{V}\mathbf{H}_0\right) - \gamma_0\text{Tr}\left(\widetilde{\mathbf{W}}\mathbf{H}_0\right)\right)$$

$$+ \sum_{k=1}^{K} \mu_k \left(\text{Tr}\left(\mathbf{V}\mathbf{H}_k\right) - \gamma_k\text{Tr}\left(\widetilde{\mathbf{W}}\mathbf{H}_k\right)\right)$$

Figure 8: Convergence behavior of **Algorithm 1**.

$$+ \sum_{n=1}^{N_t} \nu_n \left( \mathrm{Tr}\left(\mathbf{V}\mathbf{E}_n\right) + \mathrm{Tr}\left(\widetilde{\mathbf{W}}\mathbf{E}_n\right) \right) + \mu_0 \gamma_0 \tilde{\sigma}_0^2 - \sum_{k=1}^{K} \mu_k \sigma_k \tilde{\sigma}_k^2 - \sum_{n=1}^{N_t} \nu_n \Delta_n^2, \tag{44}$$

where $\{\mu_k\} \geq 0$, $\{\nu_n\} \geq 0$, $\mathbf{X}_1 \succeq \mathbf{0}$, and $\mathbf{X}_2 \succeq \mathbf{0}$ are the dual variables associated with the constraints in (43b) - (43e), respectively. The Karush-Kuhn-Tucker (KKT) equations relevant to the optimal $\mathbf{V}^*$ and $\widetilde{\mathbf{W}}^*$ are given by

$$\mathbf{X}_1^* = \mathbf{I}_{N_t} - \mu_0^* \mathbf{H}_0 + \sum_{k=1}^{K} \mu_k^* \mathbf{H}_k + \sum_{n=1}^{N_t} \nu_n^* \mathbf{E}_n, \tag{45}$$

$$\mathbf{X}_2^* = \mathbf{I}_{N_t} + \mu_0^* \gamma_0 \mathbf{H}_0 - \sum_{k=1}^{K} \mu_k \gamma_k \mathbf{H}_k + \sum_{n=1}^{N_t} \nu_n^* \mathbf{E}_n, \tag{46}$$

$$\mathbf{X}_1^* \mathbf{V}^* = \mathbf{0}, \quad \mathbf{X}_2^* \widetilde{\mathbf{W}}^* = \mathbf{0}, \tag{47}$$

Firstly, we prove that $\mathrm{rank}(\mathbf{V}^*) = 1$. Since $\{\mu_k\} \geq 0$, $\{\nu_n\} \geq 0$, $\{\mathbf{H}_k\} \succeq \mathbf{0}$, and $\{\mathbf{E}_n\} \succeq \mathbf{0}$, it holds that $\mathbf{I}_{N_t} + \sum_{k=1}^{K} \mu_k^* \mathbf{H}_k + \sum_{n=1}^{N_t} \nu_n^* \mathbf{E}_n \succeq \mathbf{0}$. As a result, $\mathrm{rank}\left( \mathbf{I}_{N_t} + \sum_{k=1}^{K} \mu_k^* \mathbf{H}_k + \sum_{n=1}^{N_t} \nu_n^* \mathbf{E}_n \right) = N_t$. Moreover $\mathrm{rank}(\mu_0^* \mathbf{H}_0) \leq 1$. Thus, from (45), $\mathrm{rank}(\mathbf{X}_1^*) \geq N_t - 1$. From $\mathbf{X}_1^* \mathbf{V}^* = \mathbf{0}$ in (47), it is seen that the columns of $\mathbf{V}^*$ must lie in the null-space of $\mathbf{X}^*$. Then, according to the rank-nullity theorem

$$\mathrm{rank}\left(\mathbf{V}^*\right) \leq \mathrm{Nullity}\left(\mathbf{X}_1^*\right) = N_t - \mathrm{rank}\left(\mathbf{X}_1^*\right). \tag{48}$$

As rank$(\mathbf{X}_1^*) \geq N_t - 1$, we get rank$(\mathbf{V}^*) \leq 1$. If rank$(\mathbf{V}^*) = 0$ then $\mathbf{V}^* = \mathbf{0}$. Obviously, this does not satisfy the constraint in (43b). Therefore, rank$(\mathbf{V}^*) = 1$.

Using the same arguments, one can also show that rank$\left(\widetilde{\mathbf{W}}^*\right) \leq 1$. This completes the proof.

## APPENDIX B

### PROOF OF THEOREM 2

To facilitate taking differentiation of the Lagrangian function with respect to $\mathbf{V}_{\text{rb}}$, we rewrite $\mathcal{P}10$ as follows

$$\mathcal{P}10.1 : \underset{\substack{\mathbf{V}_{\text{rb}}, \tau_0, \omega_0, \\ \lambda_0, \xi_0}}{\text{minimize}} \quad \text{Tr}\left(\mathbf{V}_{\text{rb}}\right) \tag{49a}$$

subject to

$$\begin{bmatrix} \lambda_0 \mathbf{I}_{N_t} & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t}^T & -\lambda_0 \delta_0^2 - \tau_0 \end{bmatrix} + \widehat{\mathbf{H}}_{0,\mathbf{I}}^T \mathbf{V}_{\text{rb}} \widehat{\mathbf{H}}_{0,\mathbf{I}} \succeq \mathbf{0}, \tag{49b}$$

$$\begin{bmatrix} \xi_0 \mathbf{I}_{N_t} - \overline{\mathbf{W}}\,\overline{\mathbf{W}}^T & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t}^T & -\xi_0 \delta_0^2 + \omega_0 \end{bmatrix} \succeq \mathbf{0}, \tag{49c}$$

$$\mathbf{V}_{\text{rb}} \succeq \mathbf{0} \tag{49d}$$

$$\text{Tr}\left(\mathbf{V}_{\text{rb}} \mathbf{E}_n\right) \leq \Delta_{n,\rho}, \quad n = 1, 2, ..., N_t, \tag{49e}$$

$$\tau_0 - \gamma_0 \left( \left( \rho \frac{\underset{n}{\min}\Delta_n}{\underset{n}{\max}\left\| \left[\overline{\mathbf{W}}\right]_{n,:} \right\|_1} \right)^2 \omega_0 + \tilde{\sigma}_0^2 \right) \geq 0, \tag{49f}$$

$$\lambda_0 \geq 0, \xi_0 \geq 0, \tag{49g}$$

where $\widehat{\mathbf{H}}_{0,\mathbf{I}} = \begin{bmatrix} \mathbf{I}_{N_t} & \hat{\mathbf{h}}_0 \end{bmatrix}$ and $\Delta_{n,\rho} = \left( \Delta_n - \rho \frac{\underset{n}{\min}\Delta_n}{\underset{n}{\max}\left\|\left[\overline{\mathbf{W}}\right]_{n,:}\right\|_1} \left\| \left[\overline{\mathbf{W}}\right]_{n,:} \right\|_1 \right)^2$. The Lagrangian function of $\mathcal{P}10.1$ is then given by

$$\mathcal{L}\left(\mathbf{V}_{\text{rb}}, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3, \{a_n\}, b_0, c_0, d_0\right) = \text{Tr}\left(\mathbf{V}_{\text{rb}}\right)$$

$$- \text{Tr}\left(\mathbf{Y}_1 \begin{bmatrix} \lambda_0 \mathbf{I}_{N_t} & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t}^T & -\lambda_0 \delta_0^2 - \tau_0 \end{bmatrix}\right) - \text{Tr}\left(\mathbf{Y}_1 \widehat{\mathbf{H}}_{0,\mathbf{I}}^T \mathbf{V}_{\text{rb}} \widehat{\mathbf{H}}_{0,\mathbf{I}}\right)$$

$$- \text{Tr}\left(\mathbf{Y}_2 \begin{bmatrix} \xi_0 \mathbf{I}_{N_t} - \overline{\mathbf{W}}\,\overline{\mathbf{W}}^T & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t}^T & -\xi_0 \delta_0^2 + \omega_0 \end{bmatrix}\right)$$

$$-\operatorname{Tr}\left(\mathbf{Y}_3 \mathbf{V}_{\mathrm{rb}}\right)+\sum_{n=1}^{N_t} a_n \operatorname{Tr}\left(\mathbf{V}_{\mathrm{rb}} \mathbf{E}_n\right)-\sum_{n=1}^{N_t} a_n \Delta_{n, \rho}$$

$$-b_0\left(\tau_0-\gamma_0 \omega_0\right)-c_0 \lambda_0-d_0 \xi_0, \tag{50}$$

where $\mathbf{Y}_1 \succeq \mathbf{0}$, $\mathbf{Y}_2 \succeq \mathbf{0}$, $\mathbf{Y}_3 \succeq \mathbf{0}$, $\{a_n\} \geq 0$, $b_0 \geq 0$, $c_0 \geq 0$, $d_0 \geq 0$ are the dual variables associated with the constraints in (49b) - (49g). The KKT equations relevant to the proof are given by

$$\mathbf{Y}_1^*\left(\begin{bmatrix} \lambda_0^* \mathbf{I}_{N_t} & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t}^T & -\lambda_0^* \delta_0^2-\tau_0^* \end{bmatrix}+\widehat{\mathbf{H}}_{0, \mathbf{I}}^T \mathbf{V}_{\mathrm{rb}}^* \widehat{\mathbf{H}}_{0, \mathbf{I}}\right)=\mathbf{0}, \tag{51}$$

$$\mathbf{Y}_3^*=\mathbf{I}_{N_t}-\widehat{\mathbf{H}}_{0, \mathbf{I}} \mathbf{Y}_1^* \widehat{\mathbf{H}}_{0, \mathbf{I}}^T+\sum_{n=1}^{N_t} a_n^* \mathbf{E}_n=\mathbf{0}, \tag{52}$$

$$\mathbf{Y}_3^* \mathbf{V}_{\mathrm{rb}}^*=\mathbf{0}. \tag{53}$$

We first prove that $\operatorname{rank}\left(\mathbf{Y}_1^*\right) \leq 1$. At the optimal solution, let $\mathbf{u}_0^*$ be the error vector that $\left(\hat{\mathbf{h}}_0^T+\mathbf{u}_0^{*T}\right) \mathbf{V}_{\mathrm{rb}}^*\left(\hat{\mathbf{h}}_0+\mathbf{u}_0^*\right)=\tau_0^*=\min_{\|\mathbf{u}_0\|_2 \leq \delta_0}\left(\hat{\mathbf{h}}_0^T+\mathbf{u}_0^T\right) \mathbf{V}_{\mathrm{rb}}^*\left(\hat{\mathbf{h}}_0+\mathbf{u}_0\right)$ and let

$$\mathbf{A}_0=\begin{bmatrix} \lambda_0^* \mathbf{I}_{N_t}+\mathbf{V}_{\mathrm{rb}}^* & \mathbf{V}_{\mathrm{rb}}^* \hat{\mathbf{h}}_0 \\ \hat{\mathbf{h}}_0^T \mathbf{V}_{\mathrm{rb}}^* & -\lambda_0^* \delta_0^2+\hat{\mathbf{h}}_0^T \mathbf{V}_{\mathrm{rb}}^* \hat{\mathbf{h}}_0-\tau_0^* \end{bmatrix}.$$ Due to the fact that $\mathbf{A}_0 \succeq \mathbf{0}$, we have

$$\begin{bmatrix} \mathbf{u}_0^{*T} & 1 \end{bmatrix} \mathbf{A}_0 \begin{bmatrix} \mathbf{u}_0^* \\ 1 \end{bmatrix}$$

$$=\left(\left(\hat{\mathbf{h}}_0^T+\mathbf{u}_0^{*T}\right) \mathbf{V}_{\mathrm{rb}}^*\left(\hat{\mathbf{h}}_0+\mathbf{u}_0^*\right)-\tau_0^*\right)+\lambda_0^*\left(\|\mathbf{u}_0^*\|_2^2-\delta_0^2\right)$$

$$=\lambda_0^*\left(\|\mathbf{u}_0^*\|_2^2-\delta_0^2\right) \geq 0. \tag{54}$$

Since $\|\mathbf{u}_0^*\|_2^2 \leq \delta_0^2$ and $\lambda_0^* \geq 0$, then $\lambda_0^*\left(\|\mathbf{u}_0^*\|_2^2-\delta_0^2\right)=0$. As a result, we get $\begin{bmatrix} \mathbf{u}_0^{*T} & 1 \end{bmatrix} \mathbf{A}_0=\mathbf{0}$. This implies that

$$\operatorname{rank}\left(\mathbf{A}_0\right)=\operatorname{rank}\left(\begin{bmatrix} \lambda_0^* \mathbf{I}_{N_t}+\mathbf{V}_{\mathrm{rb}}^* & \mathbf{V}_{\mathrm{rb}}^* \hat{\mathbf{h}}_0 \end{bmatrix}\right), \tag{55}$$

due to $\begin{bmatrix} \lambda_0^* \mathbf{I}_{N_t}+\mathbf{V}_{\mathrm{rb}}^* & \mathbf{V}_{\mathrm{rb}}^* \hat{\mathbf{h}}_0 \end{bmatrix}$ can be linearly expressed by the row vectors of $\begin{bmatrix} \lambda_0^* \mathbf{I}_{N_t}+\mathbf{V}_{\mathrm{rb}}^* & \mathbf{V}_{\mathrm{rb}}^* \hat{\mathbf{h}}_0 \end{bmatrix}$. Similarly, since $\mathbf{A}_0 \begin{bmatrix} \mathbf{u}_0^* \\ 1 \end{bmatrix}=\mathbf{0}$, $\mathbf{V}_{\mathrm{rb}}^* \hat{\mathbf{h}}_0$ can also be linearly expressed by the column vectors of $[\lambda_0^* \mathbf{I}_{N_t}+\mathbf{V}_{\mathrm{rb}}^*]$. Together with (55), we get

$$\operatorname{rank}\left(\mathbf{A}_0\right)=\operatorname{rank}\left(\begin{bmatrix} \lambda_0^* \mathbf{I}_{N_t}+\mathbf{V}_{\mathrm{rb}}^* \end{bmatrix}\right). \tag{56}$$

30

Now, to determine rank $\left(\left[\lambda_0^* \mathbf{I}_{N_t} + \mathbf{V}_{\mathrm{rb}}^*\right]\right)$, we need to verify whether $\lambda_0^*$ is positive or not. If $\lambda_0^* = 0$ then

$$\begin{bmatrix} \mathbf{u}_0^{*T} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{V}_{\mathrm{rb}}^* & \mathbf{V}_{\mathrm{rb}}^* \hat{\mathbf{h}}_0 \\ \hat{\mathbf{h}}_0^T \mathbf{V}_{\mathrm{rb}}^* & \hat{\mathbf{h}}_0^T \mathbf{V}_{\mathrm{rb}}^* \hat{\mathbf{h}}_0 - \tau_0^* \end{bmatrix} \begin{bmatrix} \mathbf{u}_0^* \\ 1 \end{bmatrix} = 0 \tag{57}$$

which implies $\mathbf{V}_{\mathrm{rb}}^* \left(\hat{\mathbf{h}}_0 + \mathbf{u}_0^*\right) = \mathbf{0}$, or, in other words $\left(\hat{\mathbf{h}}_0^T + \mathbf{u}_0^{*T}\right) \mathbf{V}_{\mathrm{rb}}^* \left(\hat{\mathbf{h}}_0 + \mathbf{u}_0^*\right) = 0$. As a result, $\begin{bmatrix} \mathbf{u}_0^{*T} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{V}_{\mathrm{rb}}^* & \mathbf{V}_{\mathrm{rb}}^* \hat{\mathbf{h}}_0 \\ \hat{\mathbf{h}}_0^T \mathbf{V}_{\mathrm{rb}}^* & \hat{\mathbf{h}}_0^T \mathbf{V}_{\mathrm{rb}}^* \hat{\mathbf{h}}_0 - \tau_0^* \end{bmatrix} \begin{bmatrix} \mathbf{u}_0^* \\ 1 \end{bmatrix} = \left(\hat{\mathbf{h}}_0^T + \mathbf{u}_0^{*T}\right) \mathbf{V}_{\mathrm{rb}}^* \left(\hat{\mathbf{h}}_0 + \mathbf{u}_0^*\right) - \tau_0^* = -\tau_0^*$. This means $\tau_0^* = 0$, which obviously violates the constraint in (42h). Therefore, $\lambda_0^* > 0$, which results in rank $(\mathbf{A}_0) = N_t$.

From (51), according to the rank-nullity theorem

$$\mathrm{rank}\,(\mathbf{Y}_1^*) \le \mathrm{Nullity}\,(\mathbf{A}_0) = (N_t + 1) - \mathrm{rank}\,(\mathbf{A}_0) = 1. \tag{58}$$

From (52), since $\{a_n^*\} \ge 0$ we have

$$\mathrm{rank}\left(\mathbf{Y}_3^* + \widehat{\mathbf{H}}_{0,\mathbf{I}} \mathbf{Y}_1^* \hat{\mathbf{H}}_{0,\mathbf{I}}^T\right) = \mathrm{rank}\left(\mathbf{I}_{N_t} + \sum_{n=1}^{N_t} a_n^* \mathbf{E}_n\right) = N_t. \tag{59}$$

Also, $\mathrm{rank}\left(\mathbf{Y}_3^* + \widehat{\mathbf{H}}_{0,\mathbf{I}} \mathbf{Y}_1^* \widehat{\mathbf{H}}_{0,\mathbf{I}}^T\right) \le \mathrm{rank}\,(\mathbf{Y}_3^*) + \mathrm{rank}\left(\widehat{\mathbf{H}}_{0,\mathbf{I}} \mathbf{Y}_1^* \widehat{\mathbf{H}}_{0,\mathbf{I}}^T\right) \le \mathrm{rank}\,(\mathbf{Y}_3^*) + 1$. Thus, rank $(\mathbf{Y}_3^*) \ge N_t - 1$. This, together with (53), leads to rank $(\mathbf{V}_{\mathrm{rb}}^*) \le 1$. If rank $(\mathbf{V}_{\mathrm{rb}}^*) = 0$ then $\mathbf{V}_{\mathrm{rb}}^* = \mathbf{0}$, hence $\tau_0^* = 0$, which does not satisfy the constraint in (42h). Therefore rank $(\mathbf{V}_{\mathrm{rb}}^*) = 1$.

## REFERENCES

[1] Cisco Systems Inc., Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2017-2022 White Paper. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-738429.html

[2] A. Jovicic, L. Junyi, T. Richardson, "Visible light communication: opportunities, challenges and the path to market," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 26–32, December 2013.

[3] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: a survey, potential and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2047–2077, September 2015.

[4] L. Zeng, D. C. O'Brien, H. L. Minh, G. E. Faulkner, K. Lee, D. Jung, Y. Oh, E. T. Won, "High data rate multiple input multiple output (MIMO) optical wireless communications using white LED lighting," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 9, pp. 1654–1662, December 2009.

[5] H. L. Minh, D. C. O'Brien, G. E. Faulkner, L. Zeng, K. Lee, D. Jung, Y. Oh, E. T. Won, "100-Mb/s NRZ visible light communications using a postequalized white LED," *IEEE Photon. Technol. Lett.*, vol. 21, no. 15, pp. 1063–1065, August 2009.

[6] A. H. Azhar, T-A. Tran, D. C. O'Brien, "A gigabit/s indoor wireless transmission using MIMO-OFDM visible-light communications," *IEEE Photon. Technol. Lett.*, vol. 25, no. 2, pp. 171–174, January 2013.

31

[7] D. Tsonev, H. Chun, S. Rajbhandari, J. J. D. McKendry, S. Videv, E. Gu, M. Haji, S.Watson, A. E. Kelly, G. Faulkner, M. D. Dawson, H. Haas, D. C. O'Brien, "A 3-gb/s single-LED OFDM-based wireless VLC link using a gallium nitride μLED," *IEEE Photon. Technol. Lett.*, vol. 26, no. 7, pp.637–640, April 2014.

[8] A. Nuwanpriya, S. W. Ho, C. S. Chen, "Indoor MIMO visible light communications: novel angle diversity receivers for mobile users," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 9, pp. 1780–1792, September 2015.

[9] B. Schrenk, C. Pacher, "1 Gb/s all-LED visible light communication system," in *Proc. of the 2018 Optical Fiber Communications Conference and Exposition*, San Diego, USA, March 2018.

[10] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, October 1975.

[11] I. Csiszar, J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[12] S. L. Y. Cheong, M. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, July 1978.

[13] C. Mitrpant, A. J. Vinck,and L. Yuan, "An achievable region for the Gaussian wiretap channel with side information," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2181–2190, May 2006.

[14] P. Gopala, L. Lai, H. El Gamal, "On the secrecy capacity of fading channels", *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, October 2008.

[15] 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, July 2010.

[16] A. Khisti and G. Wornell, "Secure transmission with multiple antennas II: the MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, November 2010.

[17] J-Y. Wang, C. Liu, J-B. Wang, Y. Wu, M. Lin, J. Cheng, "Physical-layer security for indoor visible light communications: secrecy capacity analysis," in *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6423–6436, December 2018.

[18] X. Zhao, H. Chen, J. Sun, "On physical-layer security in multiuser visible light communication systems with non-orthogonal multiple access," *IEEE Access*, vol. 6, pp. 34004–34017, June 2018.

[19] L. Yin, H. Haas, "Physical-layer security in multiuser visible light communication networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 1, pp. 162–174, January 2018.

[20] A. Mostafa, L. Lampe, "Physical-layer security for MISO visible light communication channels," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 9, pp. 1806–1818, September 2015.

[21] ————————, "Optimal and robust beamforming for secure transmission in MISO visible light communication links," *IEEE Trans. Signal Process.*, vol. 64, no. 24, pp. 6501–6516, December 2016.

[22] S. Ma, Z-L. Dong, H. Li, Z. Lu, S. Li "Optimal and robust secure beamformer for indoor MISO visible light communication," *J. Lightwave Technol.*, vol. 34, no. 21, pp. 4988–4998, November 2016.

[23] T. V. Pham, A. T. Pham, "Secrecy sum-rate of multi-user MISO visible light communication systems with confidential messages," *Elsevier Optik,* vol. 151, pp. 65–76, December 2017.

[24] A. Arafa, E. Panayirci, H. V. Poor, "Relay-aided secure broadcasting for visible light communications," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 4227-4239, June 2019.

[25] M. A. Arfaoui, A. Ghrayeb, C. M. Assi, "Secrecy performance of multi-user MISO VLC broadcast channels with confidential messages," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7789–7800, November 2018.

[26] A. Mostafa, L. Lampe, "Securing visible light communications via friendly jamming," in *Proc. of the 2014 IEEE Global Communications Conference, Workshop on Optical Wireless Communications*, December 2014.

[27] H. Zaid, Z. Rezki, A. Chaaban, M. S. Alouini, "Improved achievable secrecy rate of visible light communication with

32

cooperative jamming," in *Proc. of the 2015 IEEE Global Conference on Signal and Information Processing*, Orlando, USA, December 2015.

[28] H. Shen, Y. Deng, W. Xu, C. Zhao, "Secrecy-oriented transmitter optimization for visible light communication systems," *IEEE Photon J.*, vol. 8, no. 5, October 2016, Art. ID 7905914.

[29] T. V. Pham, T. Hayashi, A. T. Pham, " Artificial-noise-aided precoding design for multi-user visible light communication channels," *IEEE Access*, vol. 7, pp. 3767–3777, December 2018.

[30] S. Cho, G. Chen, J. P. Coon, "Enhancement of physical layer security with simultaneous beamforming and jamming for visible light communication systems," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 10, pp. 2633–2638, October 2019.

[31] M. A. Arfaoui, H. Zaid, Z. Rezki , A. Ghrayeb, A. Chaaban, M. S. Alouini, "Artificial noise-based beamforming for the MISO VLC wiretap channel," *IEEE Trans. Commun.*, vol. 67, no. 4, pp. 2866–2879, April 2019.

[32] T. V. Pham, A. T. Pham, "Energy efficient artificial noise-aided precoding design for visible light communication systems," in *Proc. of the International Conference on Computing, Networking and Communications*, Hawaii, USA, February 2020.

[33] M. A. Arfaoui, Z. Rezki, A. Ghrayeb, M. S. Alouini, "On the secrecy capacity of MISO visible light communication channels," in *Proc. of the IEEE Global Communications Conference*, Washington, DC, USA, December 2016.

[34] T. V. Pham, H. L. Minh, A. T. Pham, "Multi-user visible light communication broadcast channels with zero-forcing precoding," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2509–2521, June 2017.

[35] A. L. Yuille and A. Rangarajan, "The concave-convex procedure (CCCP)," *Neural Comput.*, vol. 15, no. 4, pp. 915–936, Apr. 2003.

[36] B. K. Sriperumbudur and G. R. G. Lanckriet, "On the convergence of the concave-convex procedure," *in Proc. Neural Inf. Process. Syst.*, 2009, pp. 1759–1767.

[37] J. Lofberg, "YALMIP: a toolbox for modeling and optimization in MATLAB," in *Proc. IEEE Int. Symp. on Computer Aided Control Systems Design*, pp. 284–289, September 2004.

[38] M. Grant, S. Boyd, "CVX: Matlab software for disciplined convex programming version 2.1," *http://cvxr.com/cvx/*, January 2015.

[39] Z. -Q. Luo, W. -K. Ma, A. M. -C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Sig. Proc. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[40] W. Xu, Y. Cui, H. Zhang, G. Y. Li, X. You, "Robust beamforming with partial channel state information for energy efficient networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2920–2935, December 2015.

[41] S. Boyd, L. Vandenberghe, "Convex Optimization," *Cambridge University Press*, 2004.