# Spoofing Detection for Indoor Visible Light Systems with Redundant Orthogonal Encoding

Jian Chen and Tao Shu
*Department of Computer Science and Software Engineering*
*Auburn University*
Auburn AL, USA
{jzc0111,tshu}@auburn.edu

ICC 2021 - IEEE International Conference on Communications | 978-1-7281-7122-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICC42927.2021.9500335

***Abstract*—As more and more visible light communication (VLC) and visible light sensing (VLS) systems are mounted on today's light fixtures, how to guarantee the authenticity of the visible light (VL) signal in these systems becomes an urgent problem. This is because almost all of today's light fixtures are unprotected and can be openly accessed by almost anyone, and hence are subject to tampering and substitution attacks. In this paper, by exploiting the intrinsic linear superposition characteristics of visible light, we propose VL-Watchdog, a scalable and always-on signal-level spoofing detection framework that is applicable to both VLC and VLS systems. VL-Watchdog is based on redundant orthogonal encoding of the transmitted visible light, and can be implemented as a small hardware add-on to an existing VL system. The effectiveness of the proposed framework was validated through extensive numerical evaluations against a comprehensive set of factors.**

***Index Terms*—spoofing detection, indoor VLC and VLS, multi-link VLC, orthogonal encoding, denial of service**

## I. Introduction

Since the release of IEEE 802.15.7 standard in 2011 [1], visible light (VL) technology has received a lot of interest for both communication and sensing applications. Compared with regular radio frequency (RF) based communication and sensing, visible light communication (VLC) and visible light sensing (VLS) enjoy several unique benefits, such as higher spectrum bandwidth, higher transmission rate, higher energy efficiency, license free and so on. Because of these nice features, VLC/VLS has been considered to be a promising and urgently-needed small-cell solution for offloading the crowded RF bands in 5G systems and beyond. As more and more VLC/VLS systems are mounted on today's light fixtures, how to guarantee the authenticity of the VL signal in these systems becomes an urgent issue. This is due to the fact that almost all of today's light fixtures are unprotected and can be openly accessed by almost anyone, and hence are subject to tampering and substitution attacks. As will be clear shortly in Section II-B, an attacker can easily replace an authentic LED by a rogue LED under his control to inject spoofed VL signal into user's receiver. Unfortunately, most of today's VLS applications do not have a reliable built-in signal authentication mechanism to detect these spoofed signals and hence will mistakenly accept them as authentic sensing inputs, leading to compromised sensing outcome. Similar situation also arises in VLC. For example, the attacker may first block the line of sight (LOS) of the authentic VLC link, and then subsequently point a rogue LED transmitter to the user's receiver (typically a photo-diode) to inject spoofed data to the user [2].

Ensuring the received signals are coming from the legitimate transmitters (LEDs) is the key to address the above problem. Conventionally, this is achieved either at the physical layer – by authenticating the LED hardware, or at the link layer – by authenticating the received data from the LEDs based on cryptographic algorithms. Both methods have their own limitations. In particular, a physical layer authentication method is able to tell from which LEDs a received VL signal is coming by identifying certain physical features pertinent to the LED hardware, such as the light temperature color [3], or the polarization angle [4]. For example, due to the subtle differences in the material and manufacturing conditions, LEDs of the same nominal color temperature actually illuminate light of slightly different wavelengths (i.e., different colors), which could be used as a fingerprint to identify different LEDs. The physical layer methods provide always-on authentication at the signal level, but require each LED to present sufficient and measurable differences in its hardware, which is not scalable in practice [5]. On the other hand, a link-layer data authentication typically relies on cryptography and involves extensive computation (e.g., encryption/decryption) over the transmitted data [6]–[8]. While these cryptographic methods are applicable to VLC applications, as will be clear shortly in Section II-B, they are often irrelevant to VLS, because typically sensing happens at the signal level, and no data (i.e., sequence of 0's and 1's) is transmitted in a VLS application.

In this paper, we present VL-Watchdog, a novel signal-level always-on spoofing detection framework for VLC and VLS systems. VL-Watchdog can be implemented as a small hardware (receiver) add-on to an existing VL system. Once deployed, the watchdog will persistently monitor the light signals in the field to ensure they are sent only from authentic (legitimate) sources. VL-Watchdog supports large-scale VL systems, i.e., one with many smart LEDs, and does not assume any physical or optical difference in the LED hardware. Instead, VL-Watchdog is based on coding. It uses orthogonal codes to encode the illumination of each legitimate LED, so that the transmitted light of a legitimate LED is identifiable by detecting the unique signal structure possessed by the

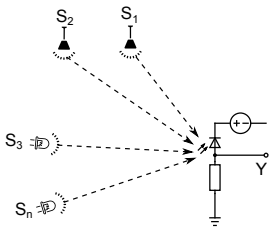


Fig. 1. Indoor Multi-link VL channel

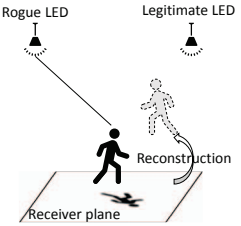


Fig. 2. Spoofing attack scenario

received light. To the best of our knowledge, this is the first signal-level always-on counter-spoofing mechanism applicable to both VLC and VLS systems. Our main contributions are summarized as follows:

- An orthogonal coding based signal-level always-on VL spoofing detection framework VL-Watchdog is proposed. Its optimal detection threshold is also derived by analysis.
- The performance of VL-Watchdog is evaluated based on extensive numerical simulations by taking into account a comprehensive set of parameters, including the number of orthogonal coding basis, the spoofing power to noise ratio, spoofing detection window size, the spoofer's strategies in fabricating its spoofing signals, and random perturbations from the application environment.

The reminder of this paper is organized as follows. The multi-link VL system model and spoofing attack model are introduced in Section II. Spoofing detection framework is presented in Section III. Numerical evaluation is analyzed in Section IV, followed by conclusions in Section V.

## II. Indoor VL System Model and Spoofing Attack Model

### A. Indoor Multi-link VL System Model

We consider a typical multi-link VL channel as shown in Figure 1, which is a general multi-link VL conceptual system that could be used to represent many different applications, such as visible light communication, visible light localization, and visible light sensing. A photo-diode is used to pick up the light and convert it into proportional current, which will be demodulated to a received data $Y$. So, the Multi-link VL channel in Figure 1 can be modelled as

$$Y = \sum_{i=1}^{n} h_i S_i + \omega, \tag{1}$$

where $h$ is VL LOS channel gain that is calculated from geometric attenuation when the light source is assumed to follow a Lambertian radiation pattern [9], $S$ is transmitted signal, $\omega \sim N(0, \sigma^2)$ is the noise processes that are well-modelled as signal-independent, zero-mean, additive, white Gaussian noise (AWGN). The channel gain in this multi-link VL channel model considers only LOS component and ignores reflected components from nearby reflectors. It is valid for most indoor scenarios, because regular building materials (e.g., plaster, wood, and plastic) of walls are diffusive reflectors for light, a unique characteristic of VL channel presents that the LOS component is much stronger than the non-LOS components, leading to a neglectable multipath effect [10], [11].

### B. VL Spoofing Attack Model

Illuminating through an open space, the open nature of VL makes its channel inherently susceptible to spoofing attacks. To make our presentation more concrete, we describe the VL spoofing attack model based on an important example VLS application: Visible Light Human Posture Sensing (VLHPS) [12], which is realized by analyzing complex shadow pattern generated by human body from different light fixtures in the environment. considering the VLHPS spoofing attack scenario illustrated in Figure 2, where for simplicity only one legitimate LED is shown. Suppose that the true posture of the user is "moving to the right", and hence the shadow generated by the rogue LED becomes longer and longer with time. However, because this shadow is mis-identified by the reverse engineering algorithm as one generated by the legitimate LED, it will be mis-mapped to a false posture sensing outcome of "moving to the left", because, based on the relative position of the user and the legitimate LED, moving to the left is the only possible posture under which a shadow generated by the legitimate LED can become longer and longer with time. Note that in this case cryptography-based counter measures are irrelevant, because the attack happens at the signal level (i.e., the shadow) and no logical data is involved in the process.

From the above case study, it is clear that the major reason that such spoofing attack can happen is because there lacks an effective method to authenticate, at the very basic signal level and on an always-on basis, that the received light signals are indeed sent from legitimate devices. Cryptographic authentication methods are either not relevant (because no logical data is involved in the application) or not effective. Considering the wide applications envisioned for VL in the near future and the fact that most existing light fixtures are un-protected, spoofing attack is a highly practical and urgent issue to be addressed for VL systems.

## III. Proposed Spoofing Detection Framework: VL-Watchdog

In this section, we present VL-Watchdog, a novel signal-level always-on spoofing detection framework for VLS and VLC systems. In the following, we first introduce the intuition behind VL-Watchdog and then present its orthogonal-coding based design. Then we formulate the spoofing detection problem as a classical statistical hypothesis test, and determine the test statistic and its optimal threshold by analyzing optimal spoofing detection strategy of the watchdog under ambient light noise.

### A. Overview

In the proposed VL-Watchdog framework, signals illuminated from legitimate LEDs are made orthogonal between each other. VL-Watchdog determines the authenticity of received signals by checking whether the expected orthogonality still holds in the received signals. More specifically, consider

a VL system that has $n$ legitimate LEDs $T_1, \ldots, T_n$ and a $m$-dimensional signal space spanned by $m$ base vectors $A_1, \ldots, A_m$, where $m > n$, and $A_i$ is orthogonal with $A_j$ for any $1 \leq i, j \leq m$ and $i \neq j$. From a geometric perspective, a $m$-axis Cartesian coordinate system is used to represent the space, where an axis $i$ corresponds to the base vector $A_i$, for $1 \leq i \leq m$. Let the whole set of all axes be denoted by $\mathbf{A} \stackrel{\text{def}}{=} (A_1, \ldots, A_m)$.

At time $t$, a subset of $n$ axes are selected from the whole set $\mathbf{A}$ and are used to modulate the data bits sent by the $n$ legitimate LEDs, one for each LED. In the case of VLC, these bits are the data to be communicated. Denote the $n$ axes that are selected at time $t$ by the set $\mathbf{R}(t) \stackrel{\text{def}}{=} (r_1(t), \ldots, r_n(t))$, and $\mathbf{R}(t) \subset \mathbf{A}$. Without loss of generality, let us suppose that axis $r_i(t)$ is used to modulate the data of $T_i$, so a bit $S_i(t) \in (-1, +1)$ to be sent by $T_i$ at time $t$ will be modulated as $S_i(t)A_{r_i(t)}$, for $1 \leq i \leq n$. The signal $S_i(t)A_{r_i(t)}$ will be transmitted by $T_i$ over VL channel using intensity modulation. $\mathbf{R}(t)$ is referred to as the transmission mode of the VL system at time $t$.

Given the absence of any illegitimate transmissions, the received signal at the VL-Watchdog at time $t$, say $Y(t)$, is simply a linear combination of $S_i(t)A_{r_i(t)}$'s, for all $1 \leq i \leq n$. Such a received signal resides in the sub-space spanned by vectors $A_{r_i(t)}$'s, where $1 \leq i \leq n$, and therefore should be orthogonal to any axis $j$ that is not in $\mathbf{R}(t)$, i.e., $\forall j \in \mathbf{A} - \mathbf{R}(t)$, which is defined as spare basis. Such an orthogonality condition can be efficiently checked by VL-Watchdog by projecting $Y(t)$ to each of the $m$ axes and verifying that

$$\begin{cases} Y(t) \bullet A_i \neq 0 & \text{if } i \in \mathbf{R}(t) \\ Y(t) \bullet A_i = 0 & \text{if } i \in \mathbf{A} - \mathbf{R}(t) \end{cases} \tag{2}$$

where the operator $\bullet$ denotes inner product between two vectors.

Clearly, when an illegitimate LED presents, $Y(t)$ will include a component contributed by the spoofing signal. The only way for the orthogonality condition in (2) to continue to hold (so the attack can elude from being detected), is for the spoofer to generate its signal at time $t$ only in the sub-space spanned by vectors $A_{r_i(t)}$'s. This requires the spoofer to follow every orthogonal axis that is selected for modulation at every moment of time. But this is difficult to achieve, as $\mathbf{R}(t)$ appears to be a random process from the spoofer's viewpoint, especially when $m$ is sufficiently greater than $n$.

### *B. Orthogonal Coding Based VL-Watchdog Design*

VL-Watchdog implements the aforementioned Cartesian coordinate system by using orthogonal coding. In particular, Walsh-Hadamard codes are used due to their simplicity and great popularity in real-world applications [13]. Walsh-Hadamard codes can be efficiently generated because they correspond to rows of the Hadamard matrix. In particular, given a Hadamard matrix $\mathbf{H}$ with size of $m$ ($2^k, k = 1, 2, 3...$), up to $m$ orthogonal codes, say $\mathbf{C}_1, \ldots, \mathbf{C}_m$ can be generated as each row of $\mathbf{H}$. From a geometric perspective, if the Hadamard matrix expands to be a $m$ dimensional space, each pair of orthogonal codes represents two perpendicular vectors in it, so the $m$ orthogonal codes constitute the $m$ orthogonal basis in such a space.

In VL-Watchdog, the aforementioned base vector set $\mathbf{A}$ is implemented as the set of orthogonal codes $(\mathbf{C}_1, \ldots, \mathbf{C}_m)$, so an axis $i$ in the Cartesian coordinate system is represented by code $\mathbf{C}_i$, for $1 \leq i \leq m$. The modulation process is simply implemented by convolving each transmitted signal with the assigned orthogonal code, which will expand the transmitted signal into a much higher frequency band (e.g., at 100 KHz level). The same orthogonal code is used by VL-Watchdog to perform the signal projection defined in (2).

For a VL system of $n$ legitimate LEDs, the received signal at VL-Watchdog at time $t$ can be modelled as

$$Y_j(t) = \sum_{i=1}^{n} h_i S_i(t) C_{(r_i(t),j)} + \omega_j \text{ for } 1 \leq j \leq m \tag{3}$$

where $h$ is the VL LOS channel gain, $S(t)$ is the signal bit to be sent by each legitimate LED, $C_{ij}$ is the orthogonal code chips from $\mathbf{C}_i$ and $r_i(t)$ is the selected orthogonal code index from $\mathbf{R}(t)$ at time $t$, $\omega$ is the ambient light noise and interference that could be well-modelled as AWGN. The projection process mathematically constitutes a correlation of the received signal with all the orthogonal codes. So, the detected signal at VL-Watchdog at time $t$ can be mathematically calculated as

$$\begin{aligned} S_i(t)' &= \frac{1}{m} h_i^{-1} \sum_{j=1}^{m} Y_j(t) C_{ij} \\ &= \begin{cases} S_i(t) + \frac{1}{m} h_i^{-1} \sum_{j=1}^{m} \omega_j C_{ij} & \text{if } i \in \mathbf{R}(\mathbf{t}) \\ \frac{1}{m} h_i^{-1} \sum_{j=1}^{m} \omega_j C_{ij} & \text{if } i \in \mathbf{A} - \mathbf{R}(\mathbf{t}) \end{cases} \end{aligned} \tag{4}$$

Since there are non-zero projections on the complementary subset of $\mathbf{R}(t)$ caused by AWGN interference in (4), a certain threshold $\tau$ is essential to reduce the probability of false detection.

### *C. Spoofing Detection under Noise*

For a given indoor multi-link VL system, the proposed VL-Watchdog aims to determine whether there is a spoofing attack or not in a reasonable amount of time. Under the proposed VL-Watchdog framework, any non-zero projection detected on the spare basis could be only due to noise or spoofing. In order to differentiate the spoofing attack from noise, we propose a statistical hypothesis test based on the average signal power projected on all spare basis during a given time window $T$ that consists of $s$ time slots, say $t_1, \ldots, t_s$, each with its randomly assigned transmission mode $\mathbf{R}(t)$. More specifically, the null hypothesis is given as

$\mathcal{H}_0$ : no spoofing (i.e., noise induced non-zero projection),

and the alternate hypothesis is given as

$\mathcal{H}_1$ : presence of spoofing (i.e., spoofing induced non-zero projection).

In this significance testing, the test statistic $P$ is defined as the average total signal power projected on all spare basis in each time slot $t_j (j = 1, 2, \ldots, s)$. So, the observed test statistic $P_{obs}$ can be mathematically expressed as

$$P_{obs} = \frac{1}{s} \sum_{j=1}^{s} \sum_{i} |S_i(t)'|^2, \forall i \in \mathbf{A} - \mathbf{R}(t). \tag{5}$$

For a given hypothesis test threshold $\tau$, the presence of spoofing attack is declared under the condition:

$$P_{obs} > \tau. \tag{6}$$

The threshold $\tau$ plays an important role in the proposed spoofing detection framework, and an optimal threshold $\tau$ would maximize the spoofing detection accuracy of the VL-Watchdog. According to the maximum a posteriori (MAP) criteria, the optimal threshold $\tau$ is decided by the test statistic distribution under the null hypothesis $\mathcal{H}_0$ and alternative hypothesis $\mathcal{H}_1$, respectively, which can be analyzed as follows.

To calculate the test statistic distribution under the null hypothesis $\mathcal{H}_0$, we model the noise as AWGN shown in (3), whose amplitude projection on each spare basis $i$ in $\mathbf{A} - \mathbf{R}(t)$ is i.i.d. and follows normal distribution, i.e.,

$$S_i(t)' \sim N(0, \frac{\sigma^2}{m}), \tag{7}$$

where $\sigma^2$ is the average power of the AWGN. So the total detected power $P_1^j$ on all spare basis given the presence of only noise in time slot $t_j$ can be calculated as

$$P_1^j = \sum_{i \in \mathbf{A} - \mathbf{R}(t)} |S_i(t)'|^2. \tag{8}$$

Therefore the random variable $P_1^j \frac{m}{\sigma^2}$ follows chi-square distribution with $m - n$ degrees of freedom, i.e.,

$$P_1^j \frac{m}{\sigma^2} = \sum_{i} \left( S_i(t)' \frac{\sqrt{m}}{\sigma} \right)^2 \sim \chi^2(m - n) \tag{9}$$

Therefore $P_1^j$ follows Gamma distribution with a shape parameter of $\frac{m-n}{2}$ and a scale parameter of $\frac{2\sigma^2}{m}$, i.e., $P_1^j \sim$ Gamma($\frac{m-n}{2}, \frac{2\sigma^2}{m}$). Over the time window $T$, the average total detected power on all spare basis given the presence of only noise, denoted by $P_1$, can be calculated as

$$P_1 = \frac{1}{s} \sum_{j=1}^{s} P_1^j. \tag{10}$$

So the random variable $P_1$ follows Gamma distribution with a shape parameter of $\frac{s(m-n)}{2}$ and a scale parameter of $\frac{2\sigma^2}{sm}$, i.e., $P_1 \sim$ Gamma($\frac{s(m-n)}{2}, \frac{2\sigma^2}{sm}$), and its probability density function is given by

$$f_{P_1}(x) = \frac{\sigma^2}{sm} \frac{x^{\frac{s(m-n)}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{s(m-n)}{2}} \Gamma(\frac{s(m-n)}{2})} \tag{11}$$

where $\Gamma(\bullet)$ denotes the gamma function. So, the detected test statistic distribution for given $\mathcal{H}_0$ will be calculated as

$$f_{P|\mathcal{H}_0}(x|P_1) = f_{P_1}(x). \tag{12}$$

To calculate the test statistic distribution under the alternative hypothesis $\mathcal{H}_1$, we consider a blind-guess spoofing strategy, in which the attacker randomly chooses $k (1 \leq k \leq m)$ orthogonal codes from the whole base vector set $\mathbf{A}$ in each time slot to generate its spoofing signal. Among the $k$ chosen orthogonal codes, let $\xi$ denote the number of those that happen to be in the spare basis set $\mathbf{A} - \mathbf{R}(t)$ and $k - \xi$ denote the rest of the chosen codes that are in the transmission mode set $\mathbf{R}(t)$. Clearly, $\xi$ is a random variable that takes value from the set $0 \leq \xi \leq k$. The attacker then equally allocates its transmission power $P_s$ onto the $k$ chosen orthogonal codes to generate the spoofing signal. In our following analysis, we first consider the basic case that $k$ is a deterministic number known to the hypothesis test. Based on the result of this basic case, we will then extend our analysis subsequently to the more general case that $k$ is a random variable.

*1) The Case of Deterministic $k$:* In this case, the probability mass function of $\xi$ in each time slot can be calculated as

$$\text{Prob}(\xi = k_s) = \begin{cases} \frac{C_{m-n}^{k_s} C_n^{k-k_s}}{C_m^k}, k_s = \begin{cases} 0, \ldots, k; & 1 \leq k \leq n \\ k-n, \ldots, k; & n < k < m-n \\ k-n, \ldots, m-n-1; & m-n \leq k \leq m \end{cases} \\ \sum_{k'=m-n}^{k} \frac{C_n^{k'-m+n}}{C_m^{k'}}, k_s = m-n; & m-n \leq k \leq m \end{cases} \tag{13}$$

where $C_i^j = \frac{i!}{(i-j)!j!}$ is the binomial coefficient of $i$ choose $j$. Its expectation and variance can be calculated as

$$\begin{aligned} E(\xi) &= \sum_{\xi=0}^{m-n} \xi \text{Prob}(\xi) \\ Var(\xi) &= \frac{1}{m-n+1} \sum_{\xi=0}^{m-n} [\xi - E(\xi)]^2 \end{aligned} \tag{14}$$

As the attacker is randomly selecting $k$ orthogonal codes in each time slot, $\xi$'s in time slots $t_j (j = 1, 2, \ldots, s)$ are i.i.d. Given a sufficiently large number of slots in the time window $T$ (e.g., greater than 10 slots in $T$), according to the central limit theorem, the average number of orthogonal codes that are chosen by the attacker in a time slot but are not in the underlying transmission mode set of that slot should approximately follow a normal distribution, i.e.,

$$\overline{\xi} \sim N(E(\xi), \frac{1}{s} Var(\xi)). \tag{15}$$

Thus, the average total detected power on all spare basis given the presence of spoofing in an arbitrary slot is given by $P_2 = \frac{P_s}{k} \overline{\xi}$. Clearly, $P_2$ also follows a normal distribution:

$$P_2 \sim N(\frac{P_s}{k} E(\xi), \frac{P_s^2}{sk^2} Var(\xi)) \tag{16}$$

and its probability density function is

$$f_{P_2}(x) = \frac{1}{\sqrt{\frac{2\pi P_s^2}{sk^2} Var(\xi)}} e^{-\frac{1}{2} \frac{sk^2 (x - \frac{P_s}{k} E(\xi))^2}{P_s^2 Var(\xi)}}. \tag{17}$$

So, the test statistic distribution given $\mathcal{H}_1$ can be calculated as

$$f_{P|\mathcal{H}_1}(x|P_1 + P_2) = f_{P_1+P_2}(x) \approx f_{P_2}(x). \tag{18}$$

Here the approximation is due to the fact that the power of spoofing signal is usually much stronger than that of the AWGN (i.e., $\frac{P_s}{\sigma^2} \gg 1$), so noise power can be safely neglected from the test statistic.

Given an equally-probable a priori distribution between $\mathcal{H}_0$ and $\mathcal{H}_1$, i.e., $\text{Prob}(\mathcal{H}_0) = \text{Prob}(\mathcal{H}_1) = 0.5$, the MAP criteria downgrades to the maximum likelihood (ML) criteria. Therefore the optimal detection threshold $\tau^o$ can be determined by solving the following equation

$$\frac{f_{P|\mathcal{H}_1}(\tau^o|P_1+P_2)}{f_{P|\mathcal{H}_0}(\tau^o|P_1)} = 1. \tag{19}$$

In practice, because $P_s \gg \sigma^2$, the solution to the above equation always exists and is unique.

*2) The Case of Random $k$:* In this case, let $p_k$ denote the probability by which the attacker selects $k$ orthogonal codes in a time slot, where $1 \le k \le m$ and $\sum_{k=1}^{m} p_k = 1$. The probability mass function of $\xi$ in a time slot can be calculated as

$$\begin{aligned}\text{Prob}(\xi = k_s) &= \sum_{k=1}^{m} p_k \text{Prob}(\xi = k_s|k) \\ &= \begin{cases} \sum_{k=1}^{n} p_k \frac{C_n^k}{C_m^k}, & k_s = 0 \\ \sum_{k=k_s}^{k_s+n} p_k \frac{C_{m-n}^{k_s} C_n^{k-k_s}}{C_m^k}, & 0 < k_s < m-n \\ \sum_{k=m-n}^{m} p_k \sum_{k'=m-n}^{k} \frac{C_n^{k'-m+n}}{C_m^{k'}}, & k_s = m-n \end{cases}\end{aligned} \tag{20}$$

Its expectation and variance can be calculated by substituting (20) into (14). By following a similar derivation in the previous deterministic case, we can calculate the test statistic distribution given $\mathcal{H}_1$ from (17) with the updated expectation and variance in this random case. Therefore the optimal detection threshold $\tau^o$ can be determined by solving (19) in this case.

## IV. Numerical Evaluation

To evaluate the performance of VL-Watchdog, we resort to simulations, which allow us to measure how the proposed spoofing detector performs against a set of attack parameters.

### A. Performance Metrics

We use the following spoofing detection rate $PD$, miss detection rate $MD$, and false warning rate $FW$ to characterize the accuracy of the proposed VL-Watchdog detector:

$$\begin{aligned} PD &= \int_{\tau^o}^{\infty} f_{P|\mathcal{H}_1}(x|P_1+P_2)dx, \\ MD &= \int_{-\infty}^{\tau^o} f_{P|\mathcal{H}_1}(x|P_1+P_2)dx, \\ FW &= \int_{\tau^o}^{\infty} f_{P|\mathcal{H}_0}(x|P_1)dx. \end{aligned} \tag{21}$$

where $\tau^o$ is the optimal detection threshold as defined in (19). Based on these quantities, the precision and sensitivity measures of the detector are defined as follows:

$$Precision = \frac{PD}{PD+FW}, Sensitivity = \frac{PD}{PD+MD}. \tag{22}$$

The overall performance is measured by the $F_1$ score [14], which is defined as

$$F_1 = 2\frac{Precision \times Sensitivity}{Precision + Sensitivity} = \frac{2PD}{2PD+FW+MD}. \tag{23}$$

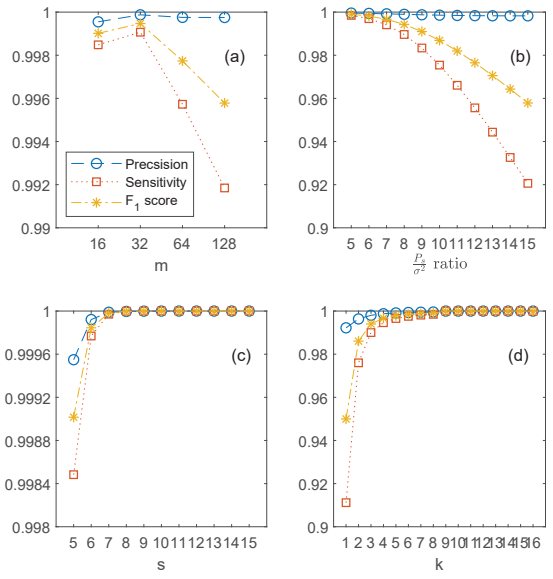


Fig. 3. Performance evaluation for spoofing detection under varying factors.

$F_1$ score is calculated as harmonic mean between precision and sensitivity and it represents the overall accuracy of the detector.

### B. Simulation Results

We simulate a multi-link VLC system with 8 legitimate LED transmitters ($n = 8$). In each time window $T$, we assume that a spoofer will present randomly with a 0.5 probability. We are interested in evaluating how the VL-Watchdog will perform against a set of parameters, including the number of base vectors $m$, the spoofing power to noise ratio $\frac{P_s}{\sigma^2}$, the number of time slots $s$ within the given time window, and the number of orthogonal codes $k$ that the spoofer chooses in fabricating its spoofing signal. In each simulation we vary the value of one of the above parameters while keeping the others constant. To this end, we assume the following default value for the parameters in our simulation: $m = 16$, $\frac{P_s}{\sigma^2} = 5$, $s = 5$, and $k = 8$. The simulation results are shown in Figure 3.

*1) Impact of the Number of Base Vectors:* Figure 3(a) shows the impact of the number of base vectors $m$ on the spoofing detection performance. We can see that there is an optimal number of base vectors ($m = 32$), which is about four times of the number of transmitters and it maximizes the overall spoofing detection performance ($F_1$ score). It could be used to determine the optimal number of base vectors that should be used for a given number of transmitters in a multi-link VLC system. Additionally, there is a slight increase of overall performance before the optimal $m$, which could be explained by the fact that adequate increase of spare basis would benefit the overall performance. After the optimal $m$, we can see that with the increase of $m$, $F_1$ score decreases rapidly and the $Sensitivity$ measurement drops off while the $Precision$ measurement maintains at approximately same level. It turns out that the decline of the $Sensitivity$ measurement is mainly induced by the rapid decrease of $PD$, which leaves $FW$ almost unchanged. It is not surprising because as the increase

of $m$, the average power assigned to each base vectors from spoofing will decrease rapidly, which will make the spoofing signal behaves much more similar with the background noise.

*2) Impact of Spoofing Power:* Figure 3(b) shows the impact of the spoofing power to noise ratio $\frac{P_s}{\sigma^2}$ on the spoofing detection performance. As for the numerical simulation, we fix the noise power $\sigma^2 = 1$, so the spoofing power $P_s$ changes accordingly with the $\frac{P_s}{\sigma^2}$ ratio. We can see from the figure that with the increase of the $\frac{P_s}{\sigma^2}$ ratio, the overall spoofing detection performance degrades gradually, i.e., the $F_1$ score decreases gradually. We can also observe that the $Precision$ measurement remains almost unchanged while the $Sensitivity$ measurement decreases rapidly. It might be a little surprising at first sight, but it would be still in line with our intuition if we take a thorough consideration on (18). Although the mean of the detected average spoofing power on spare basis increases with $P_s$, the variance increases quadratically, so the enlarged variance would eventually induce the decrease of $PD$, which is represented as $Sensitivity$ measurement in the figure.

*3) Impact of the Number of Time Slots:* Figure 3(c) shows the impact of the number of time slots $s$ within a given time window on the spoofing detection performance. We can see the increase of the overall spoofing detection performance from $F_1$ score with the increase of $s$, but it has a very limited impact which is about 0.1%. In practice, as the power of spoofing signal differs significantly from that of the background noise, we can always expect using a large $s$ would differentiate a spoofing attack from noise with less randomness. It is worth noting that once $s$ exceeds a certain number, e.g., $s \geq 8$ in this case, it won't impact the spoofing detection performance anymore. This could be utilized to explore an minimum $s$ as we always prefer to detect a potential spoofer in an efficient way, given the condition that the received power projection process in VL-Watchdog is performed in each time slot.

*4) Impact of the Number of Random Selections:* In order to simplify the calculation, we only simulate the deterministic selection case, in which $k$ is a random number but it's deterministic to be the same for all the $s$ time slots. Figure 3(d) shows the impact of the number of orthogonal codes $k$ that the spoofer chooses in fabricating its spoofing signal on the spoofing detection performance. We can see an improvement of the overall spoofing detection performance from $F_1$ score as the increase of $k$. It can be also observed that there is a significant improvement of overall performance when $k$ is relative small and then the overall performance saturates once $k$ exceeds the number of transmitters ($k > 8$), which is in line with the intuition that for a fixed $m$ and $n$, with the increase of $k$, there could be much more proportions of the average spoofing power projected onto the spare basis to be detected since it's assumed that the spoofing power is equally assigned to $k$ orthogonal basis.

## V. Conclusions

In conclusion, to secure the indoor multi-link VL system from spoofing attack, we proposed a signal-level always-on spoofing detection framework VL-Watchdog in this paper, which piggybacks on the redundant orthogonal encoding. By exploiting the intrinsic linear superposition properties of VL, the transmission mode consisting of periodically changed orthogonal codes was used to identify encoded data transmitted by multiple LED transmitters in case of rogue LED transmitters. The proposed VL-Watchdog was numerically evaluated under different factors and it was proved to be effective. In terms of implementation, the proposed VL-Watchdog can be easily integrated into the current VL system with a small hardware add-on of minimum overhead under existing infrastructure.

## Acknowledgment

This work is supported in part by NSF under grants CNS-2006998, CNS-1837034, CNS-1745254. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of NSF.

## References

[1] S. Rajagopal, R. D. Roberts, and S. Lim, "IEEE 802.15.7 visible light communication: modulation schemes and dimming support," *IEEE Communications Magazine*, vol. 50, pp. 72–82, Mar. 2012.

[2] G. J. Blinowski, "The Feasibility of Launching Rogue Transmitter Attacks in Indoor Visible Light Communication Networks," *Wireless Personal Communications*, pp. 1–19, Aug. 2017.

[3] A. T. L. Lee, H. Chen, S.-C. Tan, and S. Y. Hui, "Precise dimming and color control of LED systems based on color mixing," *IEEE Transactions on Power Electronics*, vol. 31, pp. 65–80, Jan. 2016.

[4] Y. Wang, C. Yang, Y. Wang, and N. Chi, "Gigabit polarization division multiplexing in visible light communication," *Optics Letters*, vol. 39, no. 7, p. 1823, 2014.

[5] C. Zhang and X. Zhang, "LiTell: Robust Indoor Localization Using Unmodified Light Fixtures," in *Proceedings of the 22Nd Annual International Conference on Mobile Computing and Networking*, MobiCom '16, (New York, NY, USA), pp. 230–242, ACM, 2016.

[6] A. Bittau, M. Handley, and J. Lackey, "The final nail in WEP's coffin," in *2006 IEEE Symposium on Security and Privacy (S P'06)*, pp. 15 pp.–400, May 2006.

[7] K. Bhargavan, A. D. Lavaud, C. Fournet, A. Pironti, and P. Y. Strub, "Triple handshakes and cookie cutters: Breaking and fixing authentication over tls," in *2014 IEEE Symposium on Security and Privacy*, pp. 98–113, IEEE, 2014.

[8] C. Meyer, J. Somorovsky, E. Weiss, J. Schwenk, S. Schinzel, and E. Tews, "Revisiting ssl/tls implementations: New bleichenbacher side channels and attacks," in *23rd USENIX Security Symposium (USENIX Security 14)*, pp. 733–748, 2014.

[9] Z. Ghassemlooy, W. Popoola, and S. Rajbhandari, *Optical Wireless Communications: System and Channel Modelling with MATLAB*. CRC Press, Aug. 2012.

[10] T. Komine and M. Nakagawa, "Fundamental analysis for visible-light communication system using LED lights," *IEEE Transactions on Consumer Electronics*, vol. 50, pp. 100–107, Feb. 2004.

[11] J. M. Kahn, W. J. Krause, and J. B. Carruthers, "Experimental characterization of non-directed indoor infrared channels," *IEEE Transactions on Communications*, vol. 43, pp. 1613–1623, Feb. 1995.

[12] T. Li, C. An, Z. Tian, A. T. Campbell, and X. Zhou, "Human sensing using visible light communication," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, MobiCom '15, (New York, NY, USA), p. 331–344, Association for Computing Machinery, 2015.

[13] M. Noshad and M. Brandt-Pearce, "Hadamard-Coded Modulation for Visible Light Communications," *IEEE Transactions on Communications*, vol. 64, pp. 1167–1175, Mar. 2016.

[14] D. M. Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation," 2011.