On the capacity and optimum signal constellations for VLC system

Linqiong Jia, Ming Chen, Nuo Huang, Feng Shu, and Jiangzhou Wang, *Fellow, IEEE*

***Abstract*—In this paper, the capacity of the point-to-point VLC system is investigated by means of functional analysis subject to amplitude constraint (and average intensity constraint). It is proved that the capacity can be reached by a unique probability density function (PDF). Two sets of necessary and sufficient conditions for the optimum PDF are derived. Moreover, the capacity-achieving PDF are proved to be discrete and finite. Given that the capacity can be achieved by a set of discrete and finite constellations, the capacity-achieving constellation optimization problems under amplitude constraint (and average intensity constraint) are formulated and algorithms are proposed to solve the corresponding problem. Since digital implementation is applied in most practical VLC systems, constellation optimization problem maximizing the mutual information subject to an additional equal probability constraint are put forward and analyzed.**

***Index Terms*—Visible light communication, capacity-achieving PDF, constellation optimization, amplitude constraint, average intensity constraint.**

I. Introduction

Visible light communication (VLC) using light-emitting diodes (LEDs) has draw great attention to VLC from its first birth due to the unauthorized frequency spectrum, high transmission rate, electromagnetic radiation-free property and reusability of LED lighting infrastructures [1]. VLC has been widely recognized as one of the important solutions for the future indoor wireless communication in the fifth-generation (5G) and beyond 5th-generation (B5G) wireless mobile communications [2], [3].

In indoor VLC system with intensity modulation and direct detection (IM/DD), additive white Gaussian noise (AWGN) channel model is applied [4]. Unlike conventional radio frequency (RF) channels, the optical intensity (optical power) in VLC system is represented by the amplitude of the transmit signal. Moreover, the interior brightness is determined by the average intensity which can be adjusted in accordance with users' illumination demand. Thus, the transmit signal has to meet the non-negative constraint, peak intensity constraint and adjustable average intensity constraint.

It is well known that the capacity-achieving distribution is Gaussian for a scalar additive Gaussian noise channel subject to an average power constraint [5]. However, it is infeasible to derive an analytic expression for the capacity of VLC system because of the amplitude and average intensity constraint. Thus, the capacity bounds of VLC and other similar optical wireless communication (OWC) system has been investigated in many literatures. For instance, [6] deduced the capacity bounds subject to both non-negativity constraint and an average optical power constraint in band-limited optical wireless channels using sphere-packing method, which ignored the peak intensity constraint, though. A more tight lower bound and analytical upper bound of PAM wireless optical intensity channels was developed in [7]. The lower bound in [7] was derived by maximizing the input source entropy of a family of discrete non-uniform distributions with equally spaced mass points and an analytical upper bound was derived based on a sphere packing argument. In [8], closed-form upper and lower bounds for the channel capacity of dimmable VLC systems were investigated considering the non-negativity constraint and the average transmitted optical intensity constrained. The capacity bounds were obtained by optimizing the input pulse amplitude modulated input waveform ensemble for band-limited optical intensity channel [9], [10]. Besides, the capacity-approaching equally-spaced non-uniform signalling distribution was proposed by maximizing the entropy of input distribution subject to peak and average optical power constraints in optical Gamma-Gamma channel in [11].

While the most aforementioned capacity analysis focused on the capacity bounds, J. Smith analyzed the channel capacity of a scalar Gaussian channel subject to a peak power constraint and an average electric power constraint directly and came to the conclusion that the capacity-achieving distribution was discrete, with a finite number of probability mass point [12]. By means of [12], it was shown that the capacity-achieving PDFs were discrete for many other channels such as Possion channels, quadrature Gaussian channels and Rayleigh-fading channels [13]–[16]. [17] concluded the previous works with specified channel input constraint and put forward a serious of universal propositions.

To the best of our knowledge, there is so far no comprehensive investigation focused on the capacity analysis and the capacity-achieving input distribution of the point-to-point VLC system. In this paper, we lay emphasis on the capacity and the capacity-achieving distribution of the point-to-point VLC system subject to the amplitude constraint (including non-negative constraint and peak intensity constraint) and adjustable average intensity constraint. As a preliminary research, the channel capacity problem only considering the amplitude constraint

This work is supported in part by National Nature Science Foundation of China under Grant 61771244, 61702258, 61472190, 61501238, 61602245 and 61871128.

Linqiong Jia and Shu Feng are with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, 210094, China (e-mail: jialinqiong, shufeng@njust.edu.cn).

Ming Chen and Nuo Huang are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: chenming, huangnuo@seu.edu.cn).

Jiangzhou Wang is with the School of Engineering and Digital Arts, University of Kent, Canterbury CT2 7NT, U.K. (e-mail: j.z.wang@kent.ac.uk).

is analyzed. Thereafter, the average intensity constraint is included. The main contributions are summarized as follows:

1) The channel capacity of the VLC system only considering the amplitude constraint is analyzed based on the functional analysis and optimization theory by means of [12]. The channel capacity is written as a concave functional optimization problem of the input probability density function (PDF) in the convex feasible region, which leads to the conclusion that a unique optimal solution exists and a set of necessary and sufficient conditions is obtained in terms of the optimization theorem. Moreover, another usable set of necessary and sufficient conditions is provided to verify the optimality of the distribution of the input signal. It is further proved that the capacity-achieving PDF of the input signal is composed of a finite number of discrete points. Based on the properties above, the capacity-achieving constellation optimization problem is formulated and an algorithm to solve this problem is put forward. Unlike the other constellation optimization works [18]–[25] which emphasize on improving symbol error rate (SER) performance, the pragmatic mutual information or the power efficiency, our constellation optimization algorithm aims to maximize the mutual information to access capacity-achieving constellations.
Besides, to accommodate digital implementation easily, constellation optimization problem maximizing the mutual information subject to an additional equal probability constraint is formulated and a corresponding algorithm is provided to solve it.
2) The channel capacity considering both the amplitude constraint and average intensity constraint is investigated. Since it is modeled as a functional optimization problem with an equality constraint, it can be transformed into a new unconstrained concave functional optimization problem with convex feasible domain by Lagrangian multiplier theorem. Thus, it also has a unique optimal solution. The necessary and sufficient conditions for the optimal PDF can be obtained accordingly. Likewise, it can be proved that the capacity-achieving PDF is also combined of a finite number of discrete points. So the parallel capacity-achieving constellation optimization problem and the constellation optimization problem maximizing the mutual information with equal probability constraint are formulated and solved.
In addition, constellation optimization problem maximizing the mutual information subject to an additional equal probability constraint are put forward and solved as well.

The rest of the paper is organized as follows. In Section II, the system model of point to point VLC system is introduced and the channel capacity is modeled as a functional optimization problem. In Section III, the channel capacity and capacity-achieving distritution only considering the amplitude constraint are analyzed based on functional analysis and optimization theory. In Section IV, the parallel channel capacity and capacity-achieving distritution considering both the amplitude constraint and average intensity constraint are investigated based on Section III and Lagrangian multiplier theorem. The constellation optimization problems subject to amplitude constraint and average intensity constraint are put forward and worked out. The numerical results of constellation optimization problems in Section III and Section IV are displayed and discussed. Finally, the paper is concluded in Section VI.

Notations: Boldface lower and upper case letters represent vectors and matrices, respectively. Lowercase non-boldface letters stand for scalars. Uppercase non-boldface letters stand for random variables or constants. The field of real numbers and complex numbers are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. The transpose operator is denoted by $(\cdot)^{\mathrm{T}}$.

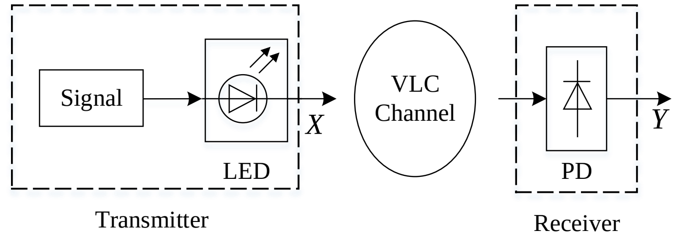

Fig. 1. System model of a point-to-point VLC system.

II. System Description

A. System Model

Consider a typical point-to-point VLC system employing IM/DD modulation as depicted in Fig.1. The received signal Y, measured by the photo-detector (PD) at the receiver, can be expressed as:

$$Y = rgX + Z \tag{1}$$

where X denotes the transmit signal; $Z \in \mathbb{R}$ is the independent zero-mean additive Gaussian noise with variance σ^2, i.e., $Z \sim \mathcal{N}(0, \sigma^2)$ [4]; r represents optoelectronic conversion factor of the PD, which is normally a constant; g represents the channel gain between the LED and the PD, which is assumed to be deterministic and can be calculated according to a specific formula in [4] if a Lambertian emission pattern of LED is adopted and only light-of-sight link between the LED and PD is considered. Without loss of generality, r and g can be normalized to 1 since they scale the SNR only.

The transmit signal X is emitted by directly modulating the light intensity of the LED. Generally speaking, the following constraints for X are taken into consideration due to illumination requirement:

1) Amplitude Constraint: X is non-negative since it represents the instantaneous emitting intensity of the LED [4] and peak-intensity bounded because of the physical limitation of the LED [26], i.e.,

$$0 \leq X \leq A \tag{2}$$

where A denotes the peak intensity of the LED.
2) Adjustable average intensity constraint: The brightness of LED light, determined by the average amplitude of

the transmit signal, should be adjustable in terms of users' demand, i.e.,

$$m_X \triangleq E[X] = \xi A \tag{3}$$

where $\xi \in [0, 1]$ is also called dimming coefficient, representing the illuminating brightness.

B. Channel Capacity of VLC system

Let $f_X(x)$, $f_Y(y)$ and $f_Z(z)$ represent the PDFs of X, Y and Z, respectively. $\mathcal{F}_X$ denotes the set of all the possibilities for $f_X(x)$ with constraint (2), which means any $f_X(x)$ in $\mathcal{F}_X$ satisfies the following conditions:

$$\begin{cases} f_X(x) = 0, & x \notin [0, A] \\ f_X(x) \geq 0, & x \in [0, A] \\ \int_0^A f_X(x)\mathrm{d}x = 1 \end{cases} \tag{4}$$

As $Z \sim \mathcal{N}(0, \sigma^2)$, $f_Z(z)$ is given by:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2\sigma^2}\right) \tag{5}$$

Therefore $f_Y(y)$ can be calculated by:

$$f_Y(y) = f_X(y) * f_Z(y) = \int_{-\infty}^{\infty} f_Z(y - x) f_X(x)\mathrm{d}x \tag{6}$$

According to (6), $f_Y(y)$ can be regarded as a functional of $f_X(x)$. For notational convenience, let $\psi[y; f_X(\cdot)] = f_Y(y)$ represent the functional from the space $\mathbb{R} \times \mathcal{F}_X$ into the real line $\mathbb{R}$:

$$\psi : \mathbb{R} \times \mathcal{F}_X \ni [y, f_X(x)] \mapsto f_Y(y) \in \mathbb{R} \tag{7}$$

Since x in $f_X(x)$ is an integral dummy variable in (6), it is represented by $\cdot$ to avoid confusion.

The mutual information $I(X;Y)$ of the VLC system can be formulated as [27]:

$$\begin{aligned} I(X;Y) &= \int_{-\infty}^{\infty}\int_0^A f_X(x) f_Z(y-x) \log_2 \frac{f_Z(y-x)}{f_Y(y)} \mathrm{d}x\mathrm{d}y \\ &= H(Y) - H(Y|X) \\ &= H(Y) - H(Z) \end{aligned} \tag{8}$$

where $H(Y)$ is the entropy of Y, which is defined as

$$H(Y) = -\int_{-\infty}^{\infty} f_Y(y) \log_2 f_Y(y)\mathrm{d}y \tag{9}$$

The entropy of Z is a constant related to σ^2, which can be calculated by:

$$H(Z) = -\int_{-\infty}^{\infty} f_Z(z) \log_2 f_Z(z)\mathrm{d}z = \frac{1}{2}\log_2(2\pi e \sigma^2) \tag{10}$$

It can be seen from (8) that $I(X;Y)$ is a functional of $f_X(x)$, too. Let $\phi[f_X(\cdot)] = I(X;Y)$ denote the functional from the space $\mathcal{F}_X$ into the real line $\mathbb{R}$:

$$\phi : \mathcal{F}_X \ni f_X(x) \mapsto I(X;Y) \in \mathbb{R} \tag{11}$$

where the variable x of $f_X(x)$ is denoted by $\cdot$ since it is an integral dummy variable. Furthermore, define a new functional $\zeta[f_X(\cdot)]$ on $\mathcal{F}_X$ as:

$$\zeta[f_X(\cdot)] = \int_0^A x f_X(x)\mathrm{d}x - \xi A \tag{12}$$

It turns out that constraint (3) is written as $\zeta[f_X(\cdot)] = 0$.

As a consequence, the channel capacity of the point-to-point VLC system can be defined as a functional optimization problem:

$$C(A, \sigma, \xi) \triangleq \max_{\substack{f_X(x) \in \mathcal{F}_X \\ \zeta[f_X(\cdot)]=0}} \phi[f_X(\cdot)] \tag{13}$$

Define a new system model of $Y' = X' + Z'$ and assume that Z' is an additive Gaussian noise variable with unite variance and X' is a non-negative input variable with peak intensity constraint $A' = A/\sigma$. It can be verified that $I(X;Y) = I(X';Y')$, where $I(X';Y')$ is the mutual information of $Y' = X' + Z'$. Thus, σ is normalized to one and A represents A/σ for notational simplicity. Besides, it is obvious that ξ has nothing to do with the normalization.

So later in this literature, the channel capacity of VLC system is expressed as

$$C(A, \xi) \triangleq \max_{\substack{f_X(x) \in \mathcal{F}_X \\ \zeta[f_X(\cdot)]=0}} \phi[f_X(\cdot)] \tag{14}$$

where A denotes A/σ.

C. Supplementary Knowledge

As the capacity of VLC system is modeled as a functional optimization problem, some supplementary knowledge of functional analysis and optimization theory that will be used later is introduced in Appendix A [12]. The concepts of weak differentiability and concavity for a functional are given to prove whether a functional is concave or not. The basic Optimization Theorem (Theorem 1), described in Appendix A, is crucial important when the concave functional optimization problem subject to the amplitude constraint only is considered in Section III. As for the functional optimization problem subject to an additional equality constraint in Section IV, the Lagrangian Multiplier Theorem (Theorem 2) can be applied [28].

III. Capacity s.t. Amplitude Constraint

In this section, the channel capacity subject to the amplitude constraint (2) only is analyzed:

$$C(A) \triangleq \max_{f_X(x) \in \mathcal{F}_X} \phi[f_X(\cdot)] \tag{15}$$

where the average intensity constraint (3) is ignored temporarily.

The properties of optimal solution for optimization problem (15) is firstly discussed. Then the capacity-achieving distribution can be obtained by solving the constellation optimization problem using the proposed algorithm. Due to the fact that the capacity-achieving distribution is neither equally-spaced nor equiprobable, it is inconvenient for digital implementation. Therefore, constellation optimization problem maximizing the mutual information subject to an additional equiprobable constraint is formulated and solved.

A. Capacity Analysis s.t. Amplitude Constraint

Since problem (15) is almost the same with the first capacity problem in [12], except for the interval of the amplitude, the propositions for the optimal solution are also. Thus, we are going to make descriptions of the propositions and the rough proof sketch. For the detailed proof, please refer to [12] and [17].

Proposition 1 makes the point that an optimal input PDF exists and provides the corresponding necessary and sufficient condition based on Theorem 1. Proposition 2 puts forward a more usable set of necessary and sufficient conditions. Moreover, it can be concluded from Proposition 2 and Proposition 3 that the optimal input is discrete, taking on a finite number of values in $[0, A]$.

Proposition 1: C is achieved by a unique probability distribution function $f_X^0(x) \in \mathcal{F}_X$, i.e.

$$C = \max_{f_X(x) \in \mathcal{F}_X} \phi[f_X^0(\cdot)] \tag{16}$$

for some unique $f_X^0(x)$ in $\mathcal{F}_X$.

In addition, a necessary and sufficient condition for $f_X^0(x)$ to be optimal is for all $f_X(x)$ in $\mathcal{F}_X$

$$\int_0^{\mathrm{A}} f_X(x) i(x; f_X^0(\cdot)) \mathrm{d}x \leq \phi[f_X^0(\cdot)] \tag{17}$$

where $i(x; f_X^0(\cdot))$ is the marginal information density function defined as:

$$i(x; f_X^0(\cdot)) = \int_{-\infty}^{\infty} f_Z(y - x) \log_2 \frac{f_Z(y - x)}{f_Y(y)} \mathrm{d}y \tag{18}$$

where the variable x of $f_X^0(x)$ is also an integral dummy variable denoted by $\cdot$.

Proof: Obviously, $\mathcal{F}_X$ is a convex and compact set. The functional $\phi[f_X(\cdot)]$ is proved to be strictly concave, continuous and weakly differentiable in $\mathcal{F}_X$ by definition 1, 2 and Lemma 1 in Appendix B. In addition, the necessary and sufficient conditions is derived in terms of Theorem 1. ■

Proposition 2: Let $f_X^0(x)$ be an arbitrary PDF in $\mathcal{F}_X$. Assume that E_0 is the set of all points with nonzero values of $f_X^0(x)$, i.e., $E_0 = \{x \in [0, A] | f_X^0(x) \neq 0\}$. Then $f_X^0(x)$ is optimal if and only if

1) For all $x \in [0, A]$,

$$i(x; f_X^0(\cdot)) \leq \phi[f_X^0(\cdot)] \tag{19}$$

2) For all $x \in E_0$,

$$i(x; f_X^0(\cdot)) = \phi[f_X^0(\cdot)] \tag{20}$$

Proof: (**Proof of Necessity:**) Clearly, if (19) and (20) are both true, the necessary and sufficient conditions in Proposition 1 are satisfied, which makes $f_X^0(x)$ the optimal.

(**Proof of Sufficiency:**) The sufficiency can be proved by contradiction. Assuming that $f_X^0(x)$ is optimal but (19) is not true will lead to a contradiction of Proposition 1. If $f_X^0(x)$ is optimal but (20) is not true, an impossible inequations of $\phi[f_X^0(\cdot)] > \phi[f_X^0(\cdot)]$ comes out. Hence, if $f_X^0(x)$ is optimal, both (19) and (20) are valid. ■

Proposition 3: E_0 is a finite set of points. In other words, the optimal PDF of X can be expressed as the summation of a finite number of scaled impulse functions.

Proof: The proof lies on Proposition 2 and the identity theorem of complex functions [29]. Assuming E_0 is not finite leads to the conclusion that $f_Y(y)$ is a real constant, which is obviously impossible. Therefore, E_0 is finite. ■

B. Constellation Optimization s.t. Amplitude Constraint

Now that it has been proved that the optimal input random variable X for optimization problem (15) takes on a finite number of values, the PDF of X can be written as:

$$f_X(x) = \sum_{i=1}^{N} p_i \delta(x - x_i) \tag{21}$$

where N denotes the constellation number; x_i and p_i represent the position and probability of the i-th $(i = 1, \cdots, N)$ constellation, respectively; $\delta(\cdot)$ is the unit impulse function.

Define a new probability vector $\boldsymbol{p} = (p_1, p_2, \cdots, p_N)^T$ and a new position vector $\boldsymbol{x} = (x_1, x_2, \cdots, x_N)^T$ for the N constellations. The mutual information $I(X;Y)$ can be treated as a function of vector $\boldsymbol{p}$, $\boldsymbol{x}$ and the constellation number N as shown in (23).

As a result, the capacity of the point-to-point VLC system only subject to the amplitude constraint can be achieved by the optimal solution of the following optimization problem:

$$\begin{aligned} &\max_{\boldsymbol{p}, \boldsymbol{x}, N} \bar{\phi}(\boldsymbol{p}, \boldsymbol{x}, N) \\ \text{s.t.} \quad & 0 \leq x_1 < x_2 < \cdots < x_N \leq A \\ & 0 \leq p_i \leq 1, i = 1, \cdots, N \\ & \sum_{i=1}^{N} p_i = 1 \end{aligned} \tag{24}$$

where constraint (2) becomes $0 \leq x_i \leq A$; $0 < p_i < 1$ and $\sum_{i=1}^{N} p_i = 1$ are imposed because p_i is the probability for the i-th constellation.

The optimization problem is a mixed non-convex optimization problem with $3N+1$ inequality constraints and 1 equality constraint, where N is also an optimized variable confined to be a positive integer. Therefore, it is impracticable to get the analytical solution. However, it is known that the unique optimal solution of (24) definitely exists according to Proposition 1 and Proposition 2 provides an optimality verification method. Thus, the Constellation Optimization Algorithm (**Algorithm 1**) is developed to solve (24) and get the optimal PDF of X. In **Algorithm 1**, the constellation number is initialized to be

$$\phi[f_X(\cdot)] = -\int_{-\infty}^{\infty} \sum_{i=1}^{N} p_i \frac{1}{\sqrt{2\pi}} \exp\left[\frac{-(y - x_i)^2}{2}\right] \log\left[\sum_{j=1}^{N} p_j \frac{1}{\sqrt{2\pi}} \exp\left[\frac{-(y - x_j)^2}{2}\right]\right] \mathrm{d}y - \frac{1}{2}\log_2(2\pi \mathrm{e}) \triangleq \bar{\phi}(\boldsymbol{p}, \boldsymbol{x}, N) \tag{23}$$

Algorithm 1 Constellation Optimization Algorithm s.t. Constraint (2)

Input: A: peak intensity of LED

Initialize: $N = 2$, $x_1 = 0$, $x_2 = A$, $p_1 = p_2 = 1/2$.

While: 1

if: $i(x; f_X^0(\cdot)) \leq i(x_i^*; f_X^0(\cdot)), \forall x \in [0, A]$ & $i(x_1^*; f_X^0(\cdot)) = \cdots = i(x_N^*; f_X^0(\cdot)), \forall x_i^* \in E_0$

  Break;

else:

  $N = N + 1$

  Calculate the new optimal $\boldsymbol{p}^*$ and $\boldsymbol{x}^*$ by interior point algorithm

endif

End While

Return: $N^* = N$, $\boldsymbol{p}^*$, $\boldsymbol{x}^*$

---

2 with $p_1 = p_2 = 1/2$, $x_1 = 0$ and $x_2 = A$ according to Proposition 4.

Proposition 4: Provided that the constellation number N is fixed to be 2, the optimal constellation probability vector and position vector are $\boldsymbol{p}^* = (1/2, 1/2)^T$ and $\boldsymbol{x}^* = (0, A)^T$, respectively, which means, the two optimal constellation points are equiprobable and located at the double ends of the signal interval $[0, A]$.

Proof: See Appendix C ■

C. Constellation Optimization s.t. Amplitude Constraint and Equal Probability Constraint

In order to make digital implementation easier, the constellation points are set to be equally distributed. Therefore, the constellation optimization problem subject to equal probability constraint is formulated and solved by the proposed algorithm. Numerical results show that the optimal distributions make the mutual information very close to the channel capacity over a wide range of A and the constellations are equally spaced, which is preferred in practical implementation.

As the equal probability constraint is imposed, the probability for each constellation is $1/N$, i.e., $p_1 = \cdots = p_N = 1/N$. Hence, it leaves the constellation number N and the position vector $\boldsymbol{x} = (x_1, \cdots, x_N)^T$ the optimization variables. The PDF of X is written as:

$$f_X(x) = \frac{1}{N}\sum_{i=1}^{N}\delta(x - x_i) \tag{25}$$

Define the mutual information as a function of $\boldsymbol{x}$ and N denoted by $\bar{\bar{\phi}}(\boldsymbol{x}, N)$:

$$\bar{\bar{\phi}}(\boldsymbol{x}, N) = \bar{\phi}(\boldsymbol{p}, \boldsymbol{x}, N)|_{\boldsymbol{p}=(\frac{1}{N}, \cdots, \frac{1}{N})} \tag{26}$$

The optimization problem subject to equal probability constraint is written as:

$$\max_{\boldsymbol{x}, N} \bar{\bar{\phi}}(\boldsymbol{x}, N) \tag{27}$$
$$\text{s.t.} \quad 0 \leq x_1 < x_2 < \cdots < x_N \leq A$$

The optimization problem (27) can be solved by the interior point method when N is given. However, N is also an optimization variable, thus, an algorithm (**Algorithm 2**) is proposed to find the optimal constellation number N by comparing the mutual information of the optimal constellation for every N started from $N = 2$ till the terminal condition is meet.

Algorithm 2 Constellation Optimization Algorithm s.t. Constraint (2) & Equal Probability Constraint

---

Input: A: peak intensity of LED

Initialize: $N = 2$, $x_1 = 0$, $x_2 = A$.

  Calculate $C_N = \bar{\bar{\phi}}(\boldsymbol{x}, N)$

While: 1

  $N = N + 1$

  Calculate the optimal $\boldsymbol{x}_N^*$ by interior point method

  Calculate $C_N = \bar{\bar{\phi}}(\boldsymbol{x}_N^*, N)$

if: $\underbrace{C_{N-1} \leq C_N}_{\boldsymbol{a}}\ \underbrace{\&0.1 < \frac{x_i^* - x_{i-1}^*}{x_j^* - x_{j-1}^*} < 10, \forall i, j \in [2, N]}_{\boldsymbol{b}}$

  Continue;

else:

  Break;

endif

End While

Return: $N^* = N - 1$, $\boldsymbol{x}_{N-1}^*$

In **Algorithm 2**, the initial value of $\boldsymbol{x} = (0, A)^T$ for $N = 2$ is optimal according to Proposition 4. Then the constellation number is increased by 1. The new optimal solution $\boldsymbol{x}_N^*$ is worked out by interior point method (or other algorithms) and the new mutual information $C_N = \bar{\bar{\phi}}(\boldsymbol{x}_N^*, N)$ is computed accordingly. If $\boldsymbol{a}$ and $\boldsymbol{b}$ are both true, the **while** loop is continued, or else, the **while** loop is breaked and $\boldsymbol{x}_{N-1}^*$ is the optimal solution. The judgement statement $\boldsymbol{b}$ is applied to guarantee that the mutual information is not increased because of some overlapped constellations.

Since the constellation probability is fixed to be $1/N$, the constellation distribution obtained by solving (27) must be a lower capacity bound for (24). According to the numerical results in Section IV, the constellations subject to the additional equal probability constraint happen to be equally spaced in this scenario. Besides, we realize that the optimal constellations distribution for (27) can be termed as capacity-approaching by numerical analysis.

IV. Capacity s.t. Amplitude & Average Intensity Constraint

In this section, the original channel capacity (14) considering both amplitude constraint and average intensity constraint is investigated. Parallel to Section III, the optimal solution for optimization problem (14) is firstly analyzed. Then, capacity-achieving constellation optimization problem is put forward and worked out using the proposed algorithm. Moreover, constellation optimization problem maximizing the mutual information subject to the additional equiprobable constraint is formulated and solved to make practical implementation easier.

A. Capacity Analysis s.t. Amplitude & Average Intensity Constraint

It can be proved that the optimization problem (14) is a concave function with an equality constraint imposed by the illumination requirement of VLC systems. Therefore, the following propositions are raised on the basis of propositions 1-3 and Theorem 2 in the Appendix.

Proposition 5: The value $C(A,\xi)$ is achieved by a unique PDF $f_X^0(x) \in \mathcal{F}_X$ satisfying constraint (3), i.e.,

$$\begin{aligned} C(A,\xi) &= \max_{f_X \in \mathcal{F}_X} [\phi[f_X(\cdot)] - \nu\zeta[f_X(\cdot)]] \\ &= \phi[f_X^0(\cdot)] - \nu\zeta[f_X^0(\cdot)] \end{aligned} \tag{28}$$

A necessary and sufficient condition for $C(A,\xi) = \phi[f_X^0(\cdot)]$ being optimal is that for some constant ν,

$$\int_0^{\mathrm{A}} [f_X(x) i(x; f_X^0(\cdot)) - \nu x f_X(x)] \mathrm{d}x \le \phi[f_X^0(\cdot)] - \nu\xi A \tag{29}$$

where $f_X(x)$ is any PDF in $\mathcal{F}_X$.

Proof: It is already known by Proposition 1 that $\mathcal{F}_X$ is a convex and compact set and that functional $\phi[f_X(\cdot)]$ is strictly concave, continuous and weakly differentiable in $\mathcal{F}_X$.

In terms of the definition, the weak derivative of functional $\zeta[f_X(\cdot)]$ is given by:

$$\begin{aligned} \zeta'_{f_X^0(x)}[g_X(\cdot)] &= \lim_{\theta \to 0^+} \frac{\zeta[(1-\theta) f_X^0(\cdot) + \theta g_X(\cdot)] - \zeta[f_X^0(\cdot)]}{\theta} \\ &= \zeta[g_X(\cdot)] - \zeta[f_X^0(\cdot)] \end{aligned} \tag{30}$$

In addition, letting $f_X^1(x) = (1-\xi)\delta(x) + \xi\delta(x-A)$ implies that $\zeta[f_X^1(x)] = 0$. Thus, functional $\zeta[f_X(\cdot)]$ is linear, bounded and weakly differentiable in $\mathcal{F}_X$ and there exists at least one $f_X^1(x) \in \mathcal{F}_X$ such that $\zeta[f_X^1(x)] = 0$. As a result, function $\phi[f_X(\cdot)] - \nu\zeta[f_X(\cdot)]$ is strictly concave, continuous and weakly differentiable. By Theorem 1 and Theorem 2, (14) is equivalent to optimization problem (28) for some constant ν and there exists a unique PDF $f_X^0(x) \in \mathcal{F}_X$ satisfying constraint (3) such that

$$C(A,\xi) = \phi[f_X^0(\cdot)] - \nu\zeta[f_X^0(\cdot)] = \phi[f_X^0(\cdot)] \tag{31}$$

for some constant ν, which means, (28) holds.

In terms of Theorem 1, the necessary and sufficient condition for $f_X^0(x)$ being optimal is $\phi'_{f_X^0(\cdot)}[f_X(\cdot)] - \nu\zeta'_{f_X^0(\cdot)}[f_X(\cdot)] \le 0$ for all $f_X(x) \in \mathcal{F}_X$, which is expanded as

$$\begin{aligned} &\phi'_{f_X^0(\cdot)}[f_X(\cdot)] - \nu\zeta'_{f_X^0(\cdot)}[f_X(\cdot)] \\ &= \int_0^A [-f_X^0(x) + f_X(x)] i(x, f_X^0(\cdot)) \mathrm{d}x \\ &\quad - \nu \int_0^{\mathrm{A}} x f_X(x) \mathrm{d}x + \nu \int_0^{\mathrm{A}} x f_0(x) \mathrm{d}x \le 0 \end{aligned} \tag{32}$$

Since $\int_0^{\mathrm{A}} x f_X^0(x) \mathrm{d}x = \xi A$, (29) is proved. ■

Proposition 6: Let $f_X^0(x)$ be an arbitrary PDF in $\mathcal{F}_X$ satisfying constraint (3). Assume that E_0 is the set of all points with nonzero values of $f_X^0(x)$, i.e., $E_0 = \{x \in [0,A] | f_X^0(x) \neq 0\}$. Then, $f_X^0(x) \in \mathcal{F}_X$ is optimal if and only if, for some ν,

1) For all $x \in [0,A]$,

$$i(x; f_X^0(\cdot)) \le \phi[f_X^0(\cdot)] - \nu(\xi A - x) \tag{33}$$

2) For all $x \in E_0$

$$i(x; f_X^0(\cdot)) = \phi[f_X^0(\cdot)] - \nu(\xi A - x) \tag{34}$$

Proof: The proof parallels that of Proposition 2.

(**Proof of Necessity:**) If conditions (33) and (34) both hold, $f_X^0(x)$ is optimal because the necessary and sufficient condition of Proposition 5 is satisfied.

(**Proof of Sufficiency:**) Sufficiency is proved by contradiction.

Assume that $f_X^0(x)$ is optimal but (33) is not true. Then, there exists at least one $x_1 \in [0,A]$ such that $i(x_1; f_X^0(\cdot)) > \phi[f_X^0(\cdot)] - \nu(\xi A - x_1)$. Let $f_X(x) = \delta(x - x_1)$ (a unit impulse function at x_1), then

$$\begin{aligned} \int_0^A f_X(x) i(x; f_X^0(\cdot)) \mathrm{d}x &= i(x_1; f_X^0(\cdot)) \\ &> \phi[f_X^0(\cdot)] - \nu(\xi A - x_1) \end{aligned} \tag{35}$$

which contradicts Proposition 5. Hence, if $f_X^0(x)$ is optimal, (33) is valid.

Similarly, We assume that $f_X^0(x)$ is optimal but (34) is not true. Suppose there exists a nonempty subset E_1 of E_0 in which (34) is not true, i.e. , $\int_{E_1} f_X^0(x)\mathrm{d}x = \rho > 0$ and $i(x; f_X^0(\cdot)) > \phi[f_X^0(\cdot)] - \nu(\xi A - x)$ for $x \in E_1$. Since (34) is true for all $x \in E_0 - E_1$, we have $i(x; f_X^0(\cdot)) = \phi[f_X^0(\cdot)] - \nu(\xi A - x)$ and $\int_{E_0 - E_1} f_X^0(x)\mathrm{d}x = 1 - \rho$.

Rewrite $\phi[f_X^0(\cdot)]$ as (36). Clearly, when $\rho \neq 0$, $\phi[f_X^0(\cdot)] > \phi[f_X^0(\cdot)]$ is a contradiction. Thus, $\rho = 0$ and (34) is true. ■

Proposition 7: The value of E_0 is a finite set of points.

Proof: The proof parallels the proof of Proposition 3 with some difference.

Define the marginal entropy density function

$$h(x; f_X(\cdot)) \triangleq -\int_{-\infty}^{\infty} f_Z(y-x) \log_2 f_Y(y) \mathrm{d}y \tag{37}$$

Then we have

$$i(x; f_X(\cdot)) = h(x; f_X(\cdot)) - \frac{1}{2}\log_2(2\pi \mathrm{e}) \tag{38}$$

Assuming that $f_X^0(\cdot)$ is optimal, (34) and (38) are true for all $x \in E_0$. Thus, for all $x \in E_0$, $h(x; f_X^0(\cdot))$ satisfies

$$h(x; f_X^0(\cdot)) - \nu x = \phi[f_X^0(\cdot)] + \frac{1}{2}\log_2(2\pi \mathrm{e}) - \nu\xi A = c \tag{39}$$

Obviously, the right side is a constant which can be denoted by c. Consequently, $h(x; f_X^0(\cdot)) - \nu x = c$ holds for all $x \in E_0$.

Let $h(x; f_X^0(\cdot)) - \nu x$ extend to the entire complex plane. Then it can be proved that $h(x; f_X^0(\cdot)) - \nu x$ is analytic. E_0 can be treated as a bounded set on the real line $\mathbb{R}$ in the complex plane. If E_0 is infinite, then $h(x; f_X^0(\cdot)) - \nu x = c$ holds on the entire complex plane in accordance with the Identity Theorem of complex functions. In particular, $h(x; f_X^0(\cdot)) - \nu x = c$ is true for all $x \in \mathbb{R}$, i.e.,

$$-\int_{-\infty}^{\infty} f_Z(y-x) \log_2 f_Y(y) \mathrm{d}y = c + \nu x \tag{40}$$

This is possible if and only if $f_Y(y) = 2^{-c-\nu x}$. Since $f_Z(y-x)$ is Gaussian and the input amplitude is constrained, $f_Y(y)$ being exponential on the real line as $f_Y(y) = 2^{-c-\nu x}$ is obviously impossible. Therefore, E_0 must be finite. ■

B. Constellation Optimization s.t. Amplitude & Average Intensity Constraint

It has been proved that the optimal input random variable X for optimization problem (14) also takes on a finite number of values in Proposition 5-7. Thus the PDF of X in the form of (21) is a function of the constellation number N, the position of i-th constellation x_i and the probability of the i-th constellation p_i with $i = 1, \cdots, N$.

Since constraint (3) is considered, x_i and p_i have to satisfy

$$\zeta[f_X(\cdot)] = \int_0^A x f_X(x) \mathrm{d}x - \xi A = \sum_{i=1}^{N} x_i p_i - \xi A = 0 \tag{41}$$

As before the mutual information $\phi[f_X(\cdot)] = I(X;Y)$ is written as a function denoted by $\bar{\phi}(\boldsymbol{p}, \boldsymbol{x}, N)$ with variables $\boldsymbol{p}$, $\boldsymbol{x}$ and N as shown in (23).

The constellation optimization problem subject to (2) and (3) achieving the channel capacity is formulated as:

$$\begin{aligned} &\max_{\boldsymbol{p},\boldsymbol{x},N} \bar{\phi}(\boldsymbol{p}, \boldsymbol{x}, N) \\ \text{s.t.} \quad & 0 \le x_1 < x_2 < \cdots < x_N \le A \\ & 0 \le p_i \le 1, i = 1, \cdots, N \\ & \sum_{i=1}^{N} p_i = 1 \\ & \sum_{i=1}^{N} p_i x_i = \xi A \end{aligned} \tag{42}$$

An analytic expression for optimization problem (42) is difficult to acquire as problem (24). However, **Algorithm 3** is put forward to achieve the optimal solution of (42) for any A and ξ based on Proposition 5-7. The algorithm starts from $N = 2$ with $\boldsymbol{x}$ and $\boldsymbol{p}$ initialized to be $\boldsymbol{x} = (0, A)^T$ and $\boldsymbol{p} = (1-\xi, \xi)^T$, respectively. The initialization is imposed because they are optimal when A is not large based on the numerical results. In cases of a large A, the optimal $\boldsymbol{p}$ and $\boldsymbol{x}$ for $N = 2$ can be computed by the interior point method at first in the **while** loop. Then, Proposition 6 provides optimality verification of $(\boldsymbol{p}, \boldsymbol{x}, N)$. If the test is valid, the program is terminated. Otherwise, let $N = N + 1$ and the **while** loop continues. This procedure is repeated until the optimal solution is found.

Algorithm 3 Constellation Optimization Algorithm s.t. Constraint (2) & (3)

Input: A: peak intensity of LED
ξ: the dimming coefficient

Initialize: $N = 2$, $x_1 = 0$, $x_2 = A$, $p_1 = 1 - \xi$, $p_2 = \xi$.

While: 1

Calculate the new optimal $\boldsymbol{p}^*$ and $\boldsymbol{x}^*$ by interior point algorithm

if: $\exists \nu$, $i(x; f_X^0(\cdot)) \le \phi[f_X^0(\cdot)] - \nu(\xi A - x), \forall x \in [0, A]$
& $i(x_i^*; f_X^0(\cdot)) = \phi[f_X^0(\cdot)] - \nu(\xi A - x_i^*), \forall x_i^* \in E_0$

Break;

else:

$N = N + 1$

endif

End While

Return: $N^* = N$, $\boldsymbol{p}^*$, $\boldsymbol{x}^*$

In practical VLC system subject to constraints (2) and (3), channel capacity is related to not only A but also the dimming coefficient ξ. The next proposition shows the relationship of the optimal constellation distributions for ξ and $1-\xi$ when A is fixed.

Proposition 8: Suppose the optimal constellation number, location and probability for the given A and ξ, are N, $\boldsymbol{x} = (x_1, \cdots, x_N)^T$ and $\boldsymbol{p} = (p_1, \cdots, p_N)^T$, respectively. If the dimming coefficient becomes $1 - \xi$ while A remains unchanged, we have

$$C(A, \xi) = C(A, 1-\xi) \tag{43}$$

Besides, the optimal constellation number, location and probability, denoted by N', $\boldsymbol{x}' = (x_1', \cdots, x_N')^T$ and $\boldsymbol{p}' = (p_1', \cdots, p_N')^T$ can be obtained by

$$N' = N \tag{44}$$

$$\boldsymbol{x}' = (A - x_N, A - x_{N-1} \cdots, A - x_1)^T \tag{45}$$

$$\boldsymbol{p}' = (p_N, p_{N-1}, \cdots, p_2, p_1)^T \tag{46}$$

Proof: Under the assumption that the constellation number N, location $\boldsymbol{x} = (x_1, \cdots, x_N)^T$ and probability $\boldsymbol{p} = (p_1, \cdots, p_N)^T$ are optimal, the channel capacity is expressed

---

$$\begin{aligned} \phi[f_X^0(\cdot)] &= \int_{E_1} f_X^0(x) i(x; f_X^0(\cdot)) \mathrm{d}x + \int_{E_0 - E_1} f_X^0(x) i(x; f_X^0(\cdot)) \mathrm{d}x \\ &> \int_{E_1} f_X^0(x)[\phi[f_X^0(\cdot)] - \nu(\xi A - x)] \mathrm{d}x + \int_{E_0 - E_1} f_X^0(x)[\phi[f_X^0(\cdot)] - \nu(\xi A - x)] \mathrm{d}x \\ &= \rho \phi[f_X^0(\cdot)] + (1-\rho)\phi[f_X^0(\cdot)] = \phi[f_X^0(\cdot)] \end{aligned} \tag{36}$$

as:
$$C(A,\xi) = -\int_{-\infty}^{\infty} \Lambda(y)\mathrm{d}y - \frac{1}{2}\log_2(2\pi \mathrm{e}) \tag{47}$$
where, the integrand is given by
$$\Lambda(y) = \sum_{i=1}^{N} p_i f_Z(y-x_i)\log_2\left[\sum_{j=1}^{N} p_j f_Z(y-x_j)\right]. \tag{48}$$
In addition, N, $\boldsymbol{x}$ and $\boldsymbol{p}$ satisfy conditions (33) and (34).

If the dimming coefficient becomes $\xi' = 1-\xi$ while A remains unchanged, $\boldsymbol{x}' = (A-x_N, A-x_{N-1}\cdots, A-x_1)^T$ and $\boldsymbol{p}' = (p_N, p_{N-1}, \cdots, p_2, p_1)^T$ are feasible because they meet all the constraints in (42). The optimality of N', $\boldsymbol{p}'$ and $\boldsymbol{x}'$ can be proved by by verifying that $\boldsymbol{x}'$ and $\boldsymbol{p}'$ satisfy the necessary and sufficient condition of (33) and (34) based on the fact that $\boldsymbol{x}$ and $\boldsymbol{p}$ satisfy (33) and (34).

We assume that the integrand in $C(A, 1-\xi)$ is denoted by
$$\Lambda'(y) = \sum_{i=1}^{N} p_i f_Z(y-(A-x_i))\log_2\left[\sum_{j=1}^{N} p_j f_Z(y-(A-x_j))\right] \tag{49}$$
Obviously, $\Lambda(y)$ and $\Lambda'(y)$ are axial symmetry about $y = A/2$. Therefore, the integrals of $\Lambda(y)$ and $\Lambda'(y)$ in $(-\infty, +\infty)$ are the same. $C(A,\xi) = C(A, 1-\xi)$ is proved. ■

C. Constellation Optimization s.t. Amplitude & Average Intensity Constraint & Equal Probability Constraint

As constellations are preferred to distributed with equal probability to make implementation easier in practice, constellation optimization problem subject to constraints (2) and (3) is investigated when each constellation is confined to be equally distributed.

When the constellation number is N, the PDF of input variable X is in expression of (25) as before. The average intensity constraint (41) subject to equal probability constraint is expressed as:
$$E[x_i] = \frac{1}{N}\sum_{i=1}^{N} x_i = \xi A \tag{50}$$

The mutual information is a function of N and $\boldsymbol{x}$, which remains to be denoted by $\bar{\bar{\phi}}(\boldsymbol{x}) = \bar{\phi}(\boldsymbol{p},\boldsymbol{x})|_{\boldsymbol{p}=(\frac{1}{N},\cdots,\frac{1}{N})}$ as in (26). As a result, the constellation position optimization problem is given by
$$\begin{aligned} &\max_{\boldsymbol{x}} \bar{\bar{\phi}}(\boldsymbol{x}) \\ \text{s.t.}\quad & 0 \le x_1 < x_2 < \cdots < x_N \le A \\ & \frac{1}{N}\sum_{i=1}^{N} x_i = \xi A \end{aligned} \tag{51}$$
Algorithm 2 for (27) can also be used to solve (51) with the linear constraint as stated in (50). Hence, the details are omitted here.

V. Numerical Results and Discussions

In this section, the optimal distributions for optimization problem (24), (27), (42) and (51) are illustrated and discussed, respectively.

A. Optimal Distributions s.t. Amplitude Constraint

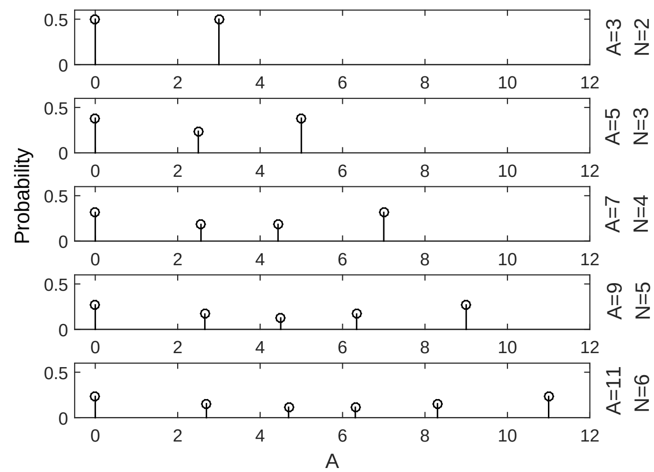

Fig. 2. Optimal constellation distribution at selected values of A

The optimal constellation distributions of optimization problem (24) at values of $A = 3, 5, 7, 9, 11$ are shown in Fig. 2. The optimal number N^* and the optimal probability and location of the constellations at each A are obtained by **Algorithm 1**.

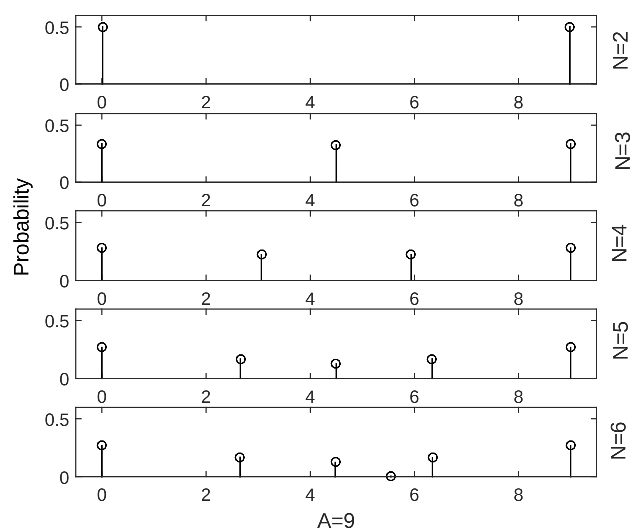

Fig. 3. Optimal constellation distribution for different N when $A = 9$

It is known that the optimal N^* is 5 when $A = 9$ according to Fig. 2. Since N will start from 2 in **Algorithm 1**, thus the optimal constellation distributions are recorded and shown in Fig. 3 from $N = 2$ to $N = 6$. It can be seen that when $N = 6$, 5 of the optimal constellations located at the same place as the optimal case for $N = 5$ with one other constellation distributed elsewhere with probability tending to zero. In this way, the mutual information for $N = 6$ is almost capacity-achieving. According to other simulation results for different 'A's with various 'N's, in the case of $N > N^*$, N^* of the optimal constellations are located on the same point as the optimal case whereas the other $N - N^*$ constellations lie in other point with probabilities being almost zero. Beside, it is worth noting that all the optimal constellation positions and probabilities satisfy (19) when $N \leq N^*$. However, (20) is not qualified when $N < N^*$, which is used to verify the global optimality for the optimal distribution of $N = N^*$.

Besides, Hypothesis 1 is concluded from Fig. 2 and Fig. 3 that all the optimal constellation distributions $\boldsymbol{x}^* = (x_1^*, x_2^*, \cdots, x_N^*)^T$ and $\boldsymbol{p}^* = (p_1^*, p_2^*, \cdots, p_N^*)^T$ meet when $N \leq N^*$.

Hypothesis 1: As long as $N \leq N^*$, $\boldsymbol{x}^*$ and $\boldsymbol{p}^*$ follow the laws below.

1) The elements of vector $\boldsymbol{x}^* = (x_1^*, x_2^*, \cdots, x_N^*)^T$ in $[0, A]$ are symmetric about the center $A/2$, i.e.,

$$\frac{x_n^* + x_{N+1-n}^*}{2} = \frac{A}{2}, n = 1, \cdots, N \tag{52}$$

2) The elements of vector $\boldsymbol{p}^* = (p_1^*, p_2^*, \cdots, p_N^*)^T$ are even symmetric:

$$p_n^* = p_{N+1-n}^*, n = 1, \cdots, N \tag{53}$$

3) On the left of the center $A/2$, the space between constellations is strictly decrease monotonically, whereas it is strictly increase monotonically on the right of $A/2$. Define

$$\begin{aligned} d_1 &= x_2^* - x_1^* = x_N^* - x_{N-1}^* \\ d_2 &= x_3^* - x_2^* = x_{N-1}^* - x_{N-2}^* \\ &\vdots \\ d_{\lfloor \frac{N}{2} \rfloor} &= x_{\lfloor \frac{N}{2} \rfloor + 1}^* - x_{\lfloor \frac{N}{2} \rfloor}^* = x_{\lceil \frac{N}{2} \rceil + 1}^* - x_{\lceil \frac{N}{2} \rceil}^* \end{aligned} \tag{54}$$

We have $d_1 > d_2 > \cdots > d_{\lfloor \frac{N}{2} \rfloor}$.

4) For any $N \geq 2$, the element of $\boldsymbol{p}^* = (p_1^*, p_2^*, \cdots, p_N^*)^T$ is strictly decrease monotonically when $n \leq \lceil N/2 \rceil$ and strictly increase monotonically if $n \geq \lceil N/2 \rceil$.

The laws are summarized from numerical results and have not been proved yet.

B. Optimal Constellation Distribution s.t. Amplitude Constraint and Equal Probability Constraint

The optimal constellation distributions of (27) at values of $A = 3, 5, 7, 9, 11$ are shown in Fig. 4. It is worth mentioning that the optimal constellation distribution are equally spaced subject to the constraint of equal probability.

In Fig. 4, the optimal N is 4 for $A = 9$. As N begins with 2 in **Algorithm 2**, we investigate the optimal constellation distributions from $N = 2$ to $N = 5$ as given in Fig. 5. It is shown that when $N \leq 4$, the optimal constellations are equally spaced. When $N = 5$, though the constellations are also equally spaced, the mutual information is smaller than $N = 4$. Under the condition of $N = 6$, the constellations are not equiprobable since the first two constellations are overlapped, which means that there are actually 5 non-equiprobable constellations. That's why condition ***b*** is deployed in **Algorithm 2**. As a matter of fact, as $N \to \infty$, the constellation distributions with overlapped constellations are tending to the optimal distributions obtained by **Algorithm 1**.

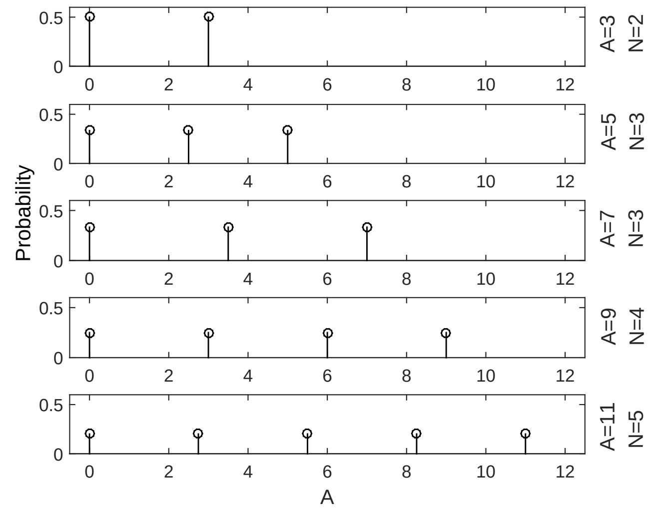

Fig. 4. Optimal constellation distribution at selected values of A with equal probability constraint

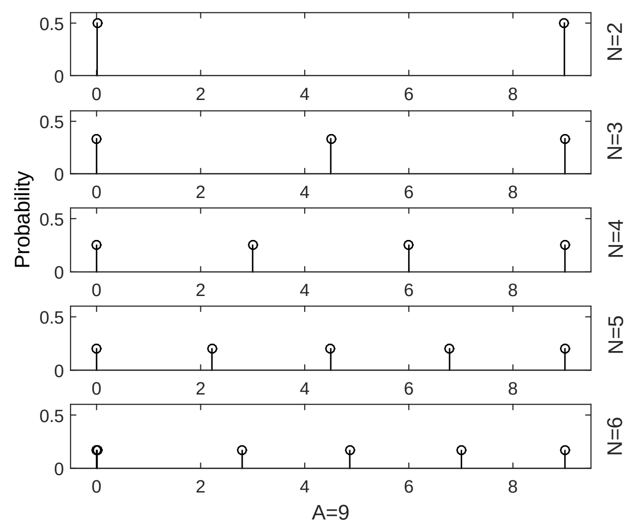

Fig. 5. Optimal constellation distributions at $A = 9$ s.t. equal probability constraint

The capacity lines of optimization problem (24) and (27) versus A are shown in Fig. 6 in bits/symbol. It can be seen that the optimal solution for (27) is a very tight lower bound of the channel capacity subject to constraint (2). Therefore, the optimal distribution of (27) is actually a capacity-approaching constellation distribution, which is very meaningful and useful in practice since digital implementation prefers equiprobable and equal-spaced constellations. Besides, the information rate of uniformly distributed input is a asymptote of $C(A)$ for large A according to the information theory [27], i.e., $H(Y) \cong H(X)$ as A increases:

$$\tilde{C} = \log_2 A - \frac{1}{2} \log_2(2\pi \mathrm{e}) \tag{55}$$

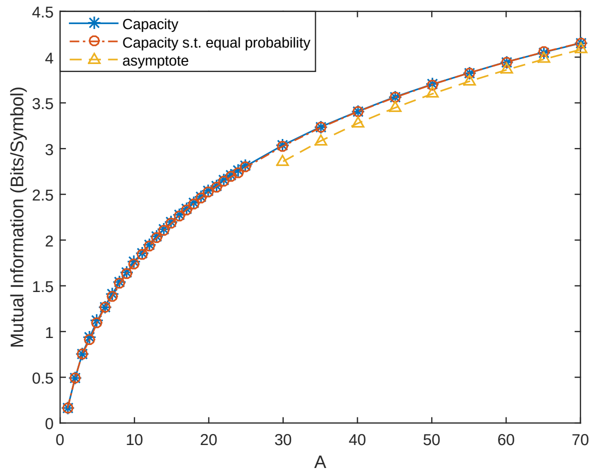

Fig. 6. Capacity (s.t. equal probability constraint) and asymptote

C. Optimal Distribution s.t. Amplitude & Average Intensity Constraint

The optimal constellation distributions of (42) at values of $A = 3, 5, 7, 9, 11$ for $\xi = 0.5$ are the same as the optimal solution for (24) in Fig. 2. This coincides with the condition that the mean of the optimal constellations in Fig. 2 equals $0.5A$. The optimal constellation distribution of (42) at values of $A = 3, 5, 7, 9, 11$ for $\xi = 0.3$ is shown in Fig. 7. The optimal constellation number N and the optimal probabilities and positions at each A are determined by **Algorithm 3**. It can be seen in Fig. 7 that more weight is assigned to the constellation points located in left side of $0.5A$ since $\xi < 0.5$.

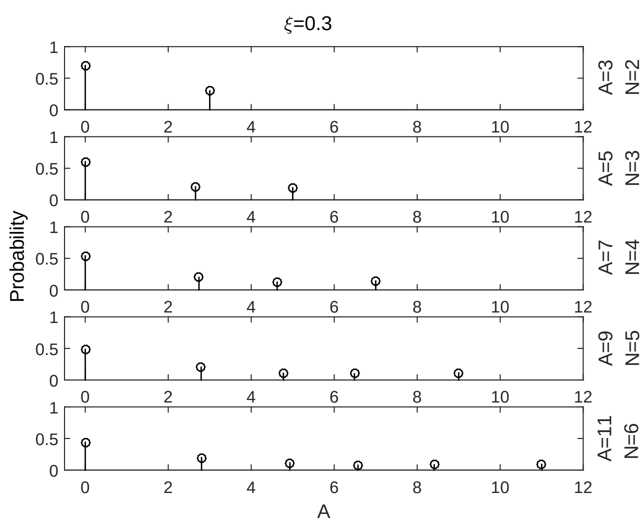

Fig. 7. Optimal constellation distributions at selected values of A for $\xi = 0.3$

Fig. 8 shows the optimal constellation distributions with different values of ξ for $A = 7$. It can be seen that ξ has impact on both constellation number and distribution. Besides, the symmetry of optimal constellation distribution for ξ and $1 - \xi$ in Proposition 8 is verified according to Fig. 8 as well.

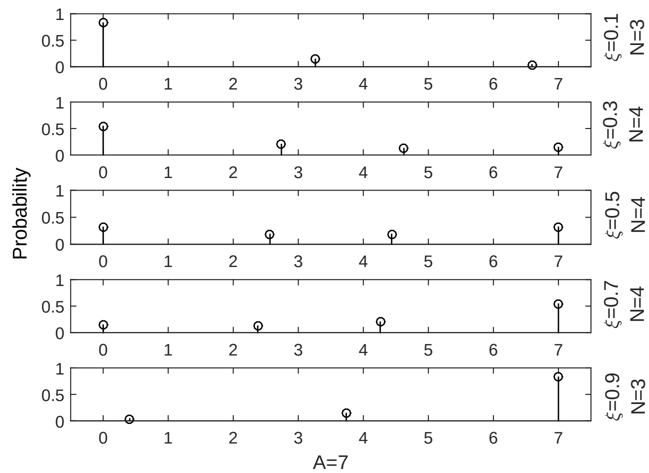

Fig. 8. Optimal constellation distributions at different values of ξ for $A = 7$

D. Optimal Distribution s.t. Amplitude & Average Intensity Constraint & Equal Probability Constraint

Fig. 9 shows the optimal equiprobable constellation distribution for (51) when $\xi = 0.3$. The optimal constellation number is decided by **Algorithm 2**. More constellations are located near zero since $\xi = 0.3 < 0.5$.

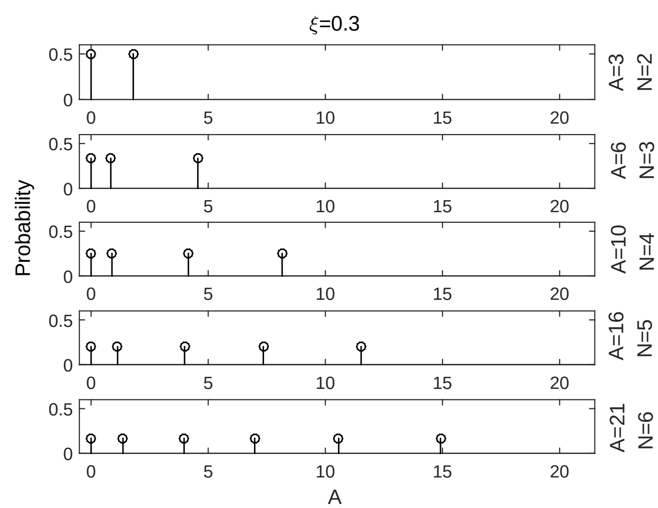

Fig. 9. The optimal constellation distribution s.t. equiprobable constraint for $\xi = 0.3$

The optimal constellation distributions with the equal probability constraint for different ξ when $A = 11$ are shown in Fig. 10. The optimal constellation distribution is the same with the optimal solution for optimization problem (27) when $\xi = 0.5$.When $\xi \neq 0.5$, the constellations are not equally spaced any more. They are concentrated on the zero side if $\xi < 0.5$ and on the A side if $\xi > 0.5$. It can be seen from Fig. 10 that the symmetry of optimal constellation for ξ and $1-\xi$ is still true when the equal probability constraint is considered.

The capacity lines with $\xi = 0.5, 0.4, 0.3, 0.2$ of optimization problem (42) and (51) versus A are plotted in Fig. 11 in

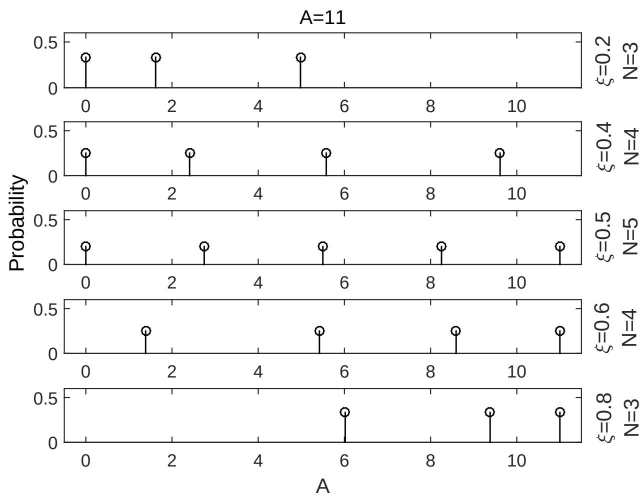

Fig. 10. The optimal constellation distribution for different ξ when $A = 11$

bits/symbol. Since the capacity is symmetric about $\xi = 0.5$, the capacity lines with $\xi = 0.6, 0.7, 0.8$ should be the same as the ones with $\xi = 0.4, 0.3, 0.2$ according to Proposition 8, thus are omitted here. It can be seen that the channel capacity achieves the maximum when $\xi = 0.5$ and declines with the increase of $|\xi - 0.5|$. This is because the deviation of ξ from 0.5 confines the constellation positions to the 0's end or the A's end. Furthermore, the maximized mutual information subject to the equal probability constraint is capacity-approaching in the case of $\xi = 0.5$. It deviate from the capacity when $\xi \neq 0.5$. The gap between the capacity and the maximized mutual information subject to the equal probability constraint goes larger as $|\xi - 0.5|$ increases.

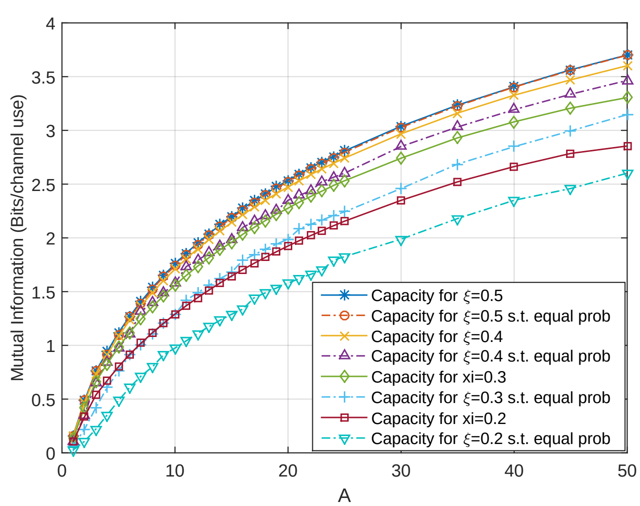

Fig. 11. Capacity (s.t. equal probability constraint) for different ξ

VI. Conclusion

In this paper, the capacity of the point to point VLC system considering the amplitude constraint and an adjustable average intensity constraint has been investigated by means of functional analysis. It was proved that a unique optimal PDF of the input variable for the functional optimization problem exists and the necessary and sufficient conditions were obtained. In addition, it was further proved that the capacity-achieving PDF was composed of a finite number of discrete points. Therefore, the constellation optimization problems to reach the channel capacity were put forward and algorithms were proposed to work out the optimization problems. Besides, the capacity-approaching constellation optimization problems subject to additional equal probability constraint were investigated to make digital implementation easier.

Appendix A
Supplementary Knowledge

Definition 1: $\phi[f_X(\cdot)]$ is weakly differentiable at $f_X^0(x) \in \mathcal{F}_X$ if there exists a map $\phi'_{f_X^0(\cdot)}[g_X(\cdot)] : \mathcal{F}_X \mapsto \mathbb{R}$ such that [12]:

$$\phi'_{f_X^0(\cdot)}[g_X(\cdot)] = \lim_{\theta \to 0^+} \frac{\phi[(1-\theta)f_X^0(\cdot) + \theta g_X(\cdot)] - \phi[f_X^0(\cdot)]}{\theta} \tag{56}$$

where $f_X^0(x)$ is a fixed element in $\mathcal{F}_X$ and θ a number in $[0, 1]$. $\phi[f_X(\cdot)]$ is said to be weakly differentiable in $\mathcal{F}_X$ as long as $\phi[f_X(\cdot)]$ is weakly differentiable at all $f_X^0(\cdot)$ in $\mathcal{F}_X$.

Definition 2: On the basis that $\mathcal{F}_X$ is a compact, convex, topological space, $\phi[f_X(\cdot)]$ is said to be concave (convex-cap in some references) if for all $f_X(x), g_X(x) \in \mathcal{F}_X$ and all $\theta \in [0, 1]$

$$\phi[(1-\theta)f_X(\cdot)+\theta g_X(\cdot)] \geq (1-\theta)\phi[f_X(\cdot)]+\theta\phi[g_X(\cdot)] \tag{57}$$

Theorem 1: Let $\phi[f_X(\cdot)]$ be a continuous, weakly differentiable strictly concave map from a compact, convex, topological space Ω to $\mathbb{R}$. Define

$$C \triangleq \max_{f_X(x)\in\Omega} \phi[f_X(\cdot)] \tag{58}$$

Then

- $C = \phi[f_X^0(x)]$ for some unique $f_X^0(x)$ in Ω;
- A necessary and sufficient condition for $\phi[f_X^0(\cdot)] = C$ is $\phi'_{f_X^0(\cdot)}[g_X(\cdot)] \leq 0$ for all $g_X(x) \in \Omega$.

Theorem 2: Let Ω be a convex set, $\phi[f_X(\cdot)]$ and $\zeta[f_X(\cdot)]$ concave functionals on Ω to $\mathbb{R}$. Assume that there exists at least a $f_X^1(x) \in \Omega$ such that $\zeta[f_X^1(\cdot)] = 0$ and let

$$C' = \max_{\substack{f_X(x)\in\Omega \\ \zeta[f_X(\cdot)]=0}} \phi[f_X(\cdot)] \tag{59}$$

Then there exists a constant ν such that

$$C' = \max_{f_X(x)\in\Omega} \phi[f_X(\cdot)] - \nu\zeta[f_X(\cdot)] \tag{60}$$

Furthermore, if the maximum in (59) is achieved by $f_X^0(x)$ in Ω and $\zeta[f_X^0(x)] = 0$, it is achieved by $f_X^0(x)$ in (60) and $\nu\zeta[f_X^0(x)] = 0$.

APPENDIX B
LEMMA FOR PROPOSITION 1

Lemma 1: Supposing that $f(x)$ is a function in $\mathcal{F}_X$, let $\beta_{f(x)}[g(x)]$ be a functional on $\mathcal{F}_X$ defined as:

$$\beta_{f(x)}[g(x)] = \int_0^A f(x)\log g(x)\mathrm{d}x \tag{61}$$

Then, if and only if $g(x) = f(x)$, $\beta_{f(x)}[g(x)]$ is maximized, i.e.,

$$\int_0^A f(x)\log g(x)\mathrm{d}x \le \int_0^A f(x)\log f(x)\mathrm{d}x \tag{62}$$

which means the equality is true if and only if $g(x) = f(x)$.

Proof: Assume that $\beta_{f(x)}[g(x)]$ is maximized by $g(x) = g_0(x)$. Define a new function of θ:

$$\bar{\beta}(\theta) = \int_0^A f(x)\log[(1-\theta)g_0(x) + \theta g(x)]\mathrm{d}x \tag{63}$$

where $g(x)$ is any function in $\mathcal{F}_X$.

Obviously, $\bar{\beta}(\theta)$ is maximized when $\theta = 0$. Therefore, the derivative of $\bar{\beta}(\theta)$ at $\theta = 0$ is zero, i.e.,

$$\bar{\beta}'(\theta)|_{\theta=0} = \int_0^A f(x)\frac{-g_0(x) + g(x)}{g_0(x)}\mathrm{d}x = 0 \tag{64}$$

(64) can be simplified as:

$$\int_0^A g(x)\frac{f(x)}{g_0(x)}\mathrm{d}x = 1 \tag{65}$$

where $g(x)$ represent any PDF with $\int_0^A g(x)\mathrm{d}x = 1$. As a consequence, only $f(x)/g_0(x) \equiv 1$ guarantees that (65) is always true, which means $g_0(x) = f(x)$.

Besides, it can be verified that the second derivative of $\bar{\beta}(\theta)$ with respect to θ is negative. Thus, $g_0(x) = f(x)$ maximizes $\beta_{f(x)}[g(x)]$. ∎

APPENDIX C
PROOF FOR PROPOSITION 4

Assuming that the PDF of X is $f_X(x) = p_1\delta(x - x_1) + p_2\delta(x - x_2)$ when $N = 2$, the PDF of Y can be written as:

$$f_Y(y) = p_1 f_Z(y - x_1) + p_2 f_Z(y - x_2) \tag{66}$$

The entropy of Y is expressed as:

$$\begin{aligned} H(Y) = -\int_{-\infty}^{\infty} &[p_1 f_Z(y - x_1) + p_2 f_Z(y - x_2)] \\ &\times \log[p_1 f_Z(y - x_1) + p_2 f_Z(y - x_2)]\mathrm{d}y \end{aligned} \tag{67}$$

Due to $p_1 + p_2 = 1$, substitute $p_2 = 1 - p_1$ into $H(Y)$ and then take the partial derivative of $H(Y)$ respect to p_1,

$$\begin{aligned} \frac{\mathrm{d}H(Y)}{\mathrm{d}p_1} &= -\int_{-\infty}^{\infty} [f_Z(y - x_1) - f_Z(y - x_2)] \\ &\quad \{\log[p_1 f_Z(y - x_1) + (1 - p_1) f_Z(y - x_2)] + 1\}\mathrm{d}y \\ &\triangleq -\int_{-\infty}^{\infty} \Lambda(y; p_1, x_1, x_2)\mathrm{d}y \end{aligned} \tag{68}$$

It can be verified that if $p_1 = 1/2$, the independent variable of the integrand $\Lambda(y; p_1, x_1, x_2)$ is odd-symmetric with respect to the point $y = \frac{x_1 + x_2}{2}$, i.e.,

$$\Lambda(y; p_1, x_1, x_2) = -\Lambda(x_1 + x_2 - y; p_1, x_1, x_2) \tag{69}$$

Thus $\frac{\mathrm{d}H(Y)}{\mathrm{d}p_1} = 0$ when $p_1 = 1/2$. $p_1 = 1/2$ is an extreme point of $H(Y)$.

Then, taking the second partial derivative of $H(Y)$ with respect to p_1, we have:

$$\frac{\mathrm{d}^2 H(Y)}{\mathrm{d}^2 p_1} = -\int_{-\infty}^{\infty} \frac{[f_Z(y - x_1) - f_Z(y - x_2)]^2}{p_1 f_Z(y - x_1) + (1 - p_1) f_Z(y - x_2)}\mathrm{d}y \tag{70}$$

It can be seen that $\frac{\mathrm{d}^2 H(Y)}{\mathrm{d}^2 p_1}$ is negative for all p_1. Thus $H(Y)$ is concave and $p_1 = \frac{1}{2}$ the maximum.

Now that we have $p_1 = p_2 = \frac{1}{2}$, $H(Y)$ can be expressed as:

$$\begin{aligned} H(Y) = -\frac{1}{2}\int_{-\infty}^{\infty} &[f_Z(y - x_1) + f_Z(y - x_2)] \\ &\times \log\left[\frac{1}{2}f_Z(y - x_1) + \frac{1}{2}f_Z(y - x_2)\right]\mathrm{d}y \end{aligned} \tag{71}$$

Let $z = y - x_1$ and $t = x_2 - x_1$, $H(Y)$ turns into:

$$\begin{aligned} H(Y) = -\frac{1}{2}\int_{-\infty}^{\infty} &[f_Z(z) + f_Z(z - t)] \\ &\times \log\left[\frac{1}{2}f_Z(z) + \frac{1}{2}f_Z(z - t)\right]\mathrm{d}z \end{aligned} \tag{72}$$

Then take the partial derivative of $H(Y)$ with respect to t, we have (73). It can be verified that $\frac{\mathrm{d}H(Y)}{\mathrm{d}t} > 0$ for all $t > 0$, so $H(Y)$ is a monotone increasing function of t. Therefore, $t = A$ is the maximum for $H(Y)$, which means $x_1 = 0$, $x_2 = A$ is the optimal constellation position.

Proposition 4 is proved.

REFERENCES

[1] "IEEE standard for local and metropolitan area networks–part 15.7: Short-range wireless optical communication using visible light," *IEEE Std 802.15.7-2011*, pp. 1–309, Sept. 2011.

[2] M. Ayyash, H. Elgala, A. Khreishah, V. Jungnickel, T. Little, S. Shao, M. Rahaim, D. Schulz, J. Hilt, and R. Freund, "Coexistence of wifi and lifi toward 5g: concepts, opportunities, and challenges," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 64–71, February 2016.

[3] C. Yan, Y. Xu, J. Shen, and J. Chen, "A combination of vlc and wifi based indoor wireless access network and its handover strategy," in *2016 IEEE International Conference on Ubiquitous Wireless Broadband (ICUWB)*, Oct 2016, pp. 1–4.

[4] T. Komine and M. Nakagawa, "Fundamental analysis for visible-light communication system using LED lights," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 100–107, Feb. 2004.

[5] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[6] S. Hranilovic and F. R. Kschischang, "Capacity bounds for power- and band-limited optical intensity channels corrupted by gaussian noise," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 784–795, May 2004.

$$\frac{\mathrm{d}H(Y)}{\mathrm{d}t} = -\int_{-\infty}^{\infty}\left\{\frac{y}{2\sqrt{2\pi}\sigma^3}\exp\left[-\frac{y^2}{2\sigma^2}\right]\right\}\left\{\log\left\{\frac{1}{2\sqrt{2\pi}\sigma}\exp\left[-\frac{(y+t)^2}{2\sigma^2}\right] + \frac{1}{2\sqrt{2\pi}\sigma}\exp\left[-\frac{y^2}{2\sigma^2}\right]\right\} + 1\right\}\mathrm{d}y \tag{73}$$

[7] A. A. Farid and S. Hranilovic, "Capacity bounds for wireless optical intensity channels with gaussian noise," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6066–6077, Dec 2010.

[8] J. B. Wang, Q. S. Hu, J. Wang, M. Chen, and J. Y. Wang, "Tight bounds on channel capacity for dimmable visible light communications," *Journal of Lightwave Technology*, vol. 31, no. 23, pp. 3771–3779, Dec 2013.

[9] J. Zhou and W. Zhang, "On the capacity of bandlimited optical intensity channels with gaussian noise," *IEEE Transactions on Communications*, vol. 65, no. 6, pp. 2481–2493, June 2017.

[10] D. Zhang and S. Hranilovic, "Bandlimited optical intensity modulation under average and peak power constraints," *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 3820–3830, Sep. 2016.

[11] A. A. Farid and S. Hranilovic, "Channel capacity and non-uniform signalling for free-space optical intensity channels," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 9, pp. 1553–1563, December 2009.

[12] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar gaussian channels," *Information and Control*, vol. 18, no. 3, pp. 203 – 219, 1971.

[13] S. Shamai, "Capacity of a pulse amplitude modulated direct detection photon channel," *IEE Proceedings I - Communications, Speech and Vision*, vol. 137, no. 6, pp. 424–430, Dec 1990.

[14] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature gaussian channels," *IEEE Transactions on Information Theory*, vol. 41, no. 4, pp. 1060–1071, July 1995.

[15] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[16] M. Katz and S. Shamai, "On the capacity-achieving distribution of the discrete-time non-coherent additive white gaussian noise channel," in *Proceedings IEEE International Symposium on Information Theory,*, June 2002, pp. 165–.

[17] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-achieving probability measure for conditionally gaussian channels with bounded inputs," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2073–2088, June 2005.

[18] J. Wang, J. Wang, Y. Wu, M. Lin, and M. Chen, "Constellation optimization for spatial modulation based indoor optical wireless communications," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, Dec 2017, pp. 1–6.

[19] B. Zhang, C. Gong, S. Li, and Z. Xu, "Constellation design for led-based full-duplex vlc," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Nov 2018, pp. 186–190.

[20] H. Chen and Z. Xu, "A two-dimensional constellation design method for visible light communications with signal-dependent shot noise," *IEEE Communications Letters*, vol. 22, no. 9, pp. 1786–1789, Sep. 2018.

[21] N. Huang, J.-B. Wang, B. Zheng, R. Guan, and M. Chen, "Design of two-dimensional signal constellations for visible light communication," *Optics Communications*, vol. 385, pp. 167–171, 2017.

[22] E. Monteiro and S. Hranilovic, "Design and implementation of color-shift keying for visible light communications," *J. Lightwave Technol*, vol. 32, no. 10, pp. 2053–2060, May 2014.

[23] R. J. Drost and B. M. Sadler, "Constellation design for channel precompensation in multi-wavelength visible light communications," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1995–2005, jun 2014.

[24] L. Jia, F. Shu, M. Chen, W. Zhang, J. Li, and J. Wang, "Joint constellation-labeling optimization for vlc-csk systems," *IEEE Wireless Communications Letters*, pp. 1–1, 2019.

[25] M. Beko and R. Dinis, "Systematic method for designing constellations for intensity-modulated optical systems," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 6, no. 5, pp. 449–458, May 2014.

[26] J. Wang, J. Dai, R. Guan, L. Jia, Y. Wang, and M. Chen, "Channel capacity and receiver deployment optimization for multi-input multi-output visible light communications," *Optics Express*, vol. 24, no. 12, pp. 13 060–13 074, 2016.

[27] T.M.Cover and J.A.Thomas, *Elements of Information Theory*. Wiley, 1991.

[28] W. Rudin, *Functional Analysis*, 2nd ed. McGraw-Hill, 1991.

[29] K. Knopp, *Theory of Functions*. New York: Dover, 1945.