Codebook Design for SCMA-Assisted Visible Light Communication with Shot Noise

Yiming Gui, Zilong Liu, Qu Luo, Lisu Yu, Ziming He, and Pei Xiao

***Abstract*—This paper studies the codebook design for sparse code multiple access (SCMA) based visible light communication (VLC) impaired by shot noise. By focusing on the typical Rician fading channels, we first derive a lower bound of the mutual information of VLC with shot noise and present a design metric called minimum normalized Euclidean distance (MNED). We then propose a novel codebook design approach for VLC including novel non-linear compensation (NLC) constellation and power scheduling matrix under the non-negative and real constraint. Simulation results demonstrate that our proposed codebook design leads to a higher MNED and thus significantly improved bit error performance over the existing SCMA codebooks for VLC systems.**

***Index Terms*—Sparse code multiple access (SCMA), visible light communication (VLC), shot noise, minimum normalized Euclidean distance (MNED).**

I. Introduction

In recent years, visible light communication (VLC) based on light-emitting diode (LED) has received increasing attention from both the academia and industry, thanks to its unique features of safety, unlicensed spectrum, high data rates, low power consumption, and no electromagnetic interference [1], [2]. LED is mainly classified as phosphorous LED and red-green-blue (RGB) LED. The phosphorous LED is most widely used due to its low cost, yet it suffers from a limited bandwidth. To tackle this problem, non-orthogonal multiple access (NOMA) is proposed for offering higher throughput and improved spectral efficiency for given frequency-time resources [3].

NOMA is mainly categorized as power-domain NOMA (PD-NOMA) and code-domain NOMA (CD-NOMA). PD-NOMA aims to multiplex two or more users sharing the same time and frequency resources by allocating them with different levels of power [4]. By contrast, in code-domain NOMA, multiple users are separated through non-orthogonal codebooks/sequences. In this work, we consider a disruptive CD-NOMA scheme, called sparse code multiple access (SCMA), for improving the spectrum efficiency [5]–[8]. In SCMA, the incoming message bits of every user are mapped to a sparse codeword at the transmitter [9]. After passing through a wireless channel, these message bits are efficiently detected by the message passing algorithm (MPA) at the receiver [10], [11]. Legacy codebook design may not be applicable to SCMA-assisted VLC transmissions primarily due to two distinct characteristics of VLC: 1) the modulated signals in VLC systems must be real and positive due to their reliance on intensity modulation/direct detection (IM/DD); 2) the detection of varying photon numbers results in fluctuating currents, which introduce shot noise—a type of noise that is inherently dependent on the incident optical signal [12], [13]. Shot noise fundamentally differs from signal-independent noise (SIDN), which does not vary with the signal strength. Although asymmetrically clipped optical orthogonal frequency-division multiple access OFDMA (ACO-OFDMA) and direct current-biased optical OFDMA (DCO-OFDMA) are proposed to transform complex signals to non-negative and real, such techniques suffer from limited spectral efficiency [12].

The primary objective of this paper is to design enhanced multi-dimensional constellations and sparse codebooks for SCMA-VLC impaired by shot noise. To the best of our knowledge, the effective codebook design criteria for SCMA assisted VLC are largely open. With the aid of the Jensen inequality, our first major contribution is the derivation of a lower bound of mutual information in the presence of shot noise. It is shown that such a lower bound is not only related to the Euclidean distance of the codebooks but also to the codebook magnitudes. In order to approach such a lower bound, a novel minimum normalized Euclidean distance (MNED) metric is introduced.

Our second major contribution is the proposed codebook design scheme based on the derived MNED. By taking advantage of the proposed non-linear compensation (NLC) constellation, piecewise linear (PwL) function and power scheduling matrix (PSM), we show that the MNED can be efficiently maximized by a bare bones particle swarm optimization (BBPSO) algorithm. Compared to traditional unipolar pulse amplitude modulation (UPAM) and square root (SQR) constellations, we show by simulations that the proposed PwL-NLC scheme can significantly improve the system flexibility and robustness against shot noise and channel fading. We further show that the resulting codebooks lead to improved bit error rate (BER) performance by increasing dimensional differences while diversifying the multiplexing points as much as possible.

Y. Gui is with Xihua University, Chengdu 610039, China (e-mail: ymgui_18@163.com). Z. Liu is with the School of Computer Science and Electronics Engineering, University of Essex, UK (e-mail: zilong.liu@essex.ac.uk). Qu Luo and Pei Xiao are with the Institute for Communications, University of Surrey, UK (email: {q.u.luo, p.xiao}@surrey.ac.uk). L. Yu is with the School of Information Engineering, Nanchang University, Nanchang 330031, China (e-mails: lisuyu@ncu.edu.cn). Z. He is with the Samsung Cambridge Solution Centre, System LSI, Samsung Electronics, Cambridge, UK (e-mails: ziming.he@samsung.com).

II. System Model

A. System Model

We consider an SCMA system with K resource elements and J users, thus the overloading factor is calculated as

$\lambda = J/K > 1$ [14]. At the transmitter, every SCMA encoder maps $\log_2 M$ bits into a K-dimensional codeword $\boldsymbol{w}_j$ with N non-zero elements selected from a dedicated codebook of size M. Codewords from different users are multiplexed to drive the LED light. Thermal, amplifier and shot noises are the three prominent sources of noise in VLC. Thermal noise and amplifier noise belong to SIDN, which can be modeled by Gaussian distribution.

However, shot noise is dependent on the signal itself whose variance can be written as

$$\sigma_{sh}^2 = 2qx\Re B + 2qI_B I_2 B, \tag{1}$$

where x denotes the intensity of incident optical signal, q the electronic charge, B the effective bandwidth of the optical receiver, I_B the background noise current, $\Re$ responsivity and I_2 the noise-bandwidth factor.

The received signal $\boldsymbol{y}$ can be written as

$$\boldsymbol{y} = \mathrm{diag}(\boldsymbol{h})\sum_{j=1}^{J}\boldsymbol{w}_j + \sqrt{\mathrm{diag}(\boldsymbol{h})\sum_{j=1}^{J}\boldsymbol{w}_j}\boldsymbol{z}^d + \boldsymbol{z}^i, \tag{2}$$

where $\boldsymbol{h} = [h_1, h_2, \cdots, h_K] \in \mathbb{R}_+^{K\times 1}$ denotes Rician fading channel vector between the base station (BS) and users. The probability density function of the normalized Rician random variable is given by [8]

$$\begin{aligned} f_{h_k}(x) &= 2x(1+\kappa)\exp\left(-\kappa - x^2(1+\kappa)\right) \\ &\times I_0\left(2x\sqrt{\kappa(1+\kappa)}\right), \end{aligned} \tag{3}$$

where κ denotes the ratio of average power in the line-of-sight (LOS) path over that in the Non-Line-of-Sight (NLOS) path. $I_0(\cdot)$ is the first order modified Bessel function of the first kind. $\boldsymbol{w}_j = [w_{1,j}, w_{2,j}, \cdots, w_{K,j}] \in \mathbb{R}_+^{K\times 1}$ denotes the transmitting codeword for the j-th user. $\boldsymbol{z}^d = [z_1^d, z_2^d, \cdots, z_K^d]^T$ denotes input-dependent real Gaussian noise vector, where $z_k^d \sim N(0, \varsigma^2\sigma^2)$. $\boldsymbol{z}^i = [z_1^i, z_2^i, \cdots, z_K^i]^T$ denotes input-independent real Gaussian noise vector, where $z_k^i \sim N(0, \sigma^2)$. $\boldsymbol{z}^d$ and $\boldsymbol{z}^i$ are independent with each other and ς^2 denotes the ratio of variance of $\boldsymbol{z}^d$ to that of $\boldsymbol{z}^i$, where ς^2 varies from 1 to 10 [13].

B. Detection Algorithm

Benefiting from the sparsity of SCMA, the transmit codewords and thus the incoming message bits can be decoded by MPA with low complexity. MPA is an iterative algorithm by recursively passing the extrinsic information between the resource nodes (RNs) and variable nodes (VNs) until the maximum number of iterations is reached. The prior probability of each codeword is $\eta_{k\to j}^t = 1/M$. The update process for RN can be expressed as

$$\begin{aligned} \eta_{k\to j}^t(\boldsymbol{w}_j) = &\sum_{\sim\boldsymbol{w}_j}\frac{1}{\sqrt{2\pi}\rho}\exp\left\{-\frac{\left(y_k - \sum_{j\in\varepsilon_k} w_{k,j}\right)^2}{2\rho^2}\right\} \\ &\times \prod_{r\in\varepsilon_k\setminus\{j\}}\eta_{r\to k}^{t-1}(\boldsymbol{w}_r), \end{aligned} \tag{4}$$

where $\rho^2 = \sigma^2(1 + \varsigma^2\sum_{j\in\varepsilon_k}\boldsymbol{w}_j)$. $\eta_{k\to j}^t$ denotes the belief message passing from RN k to VN j in iteration t. $\varepsilon_k \setminus \{j\}$ denotes all the VNs in RN k except VN j.

The update process for VN can be written as

$$\eta_{j\to k}^t(\boldsymbol{w}_j) = \prod_{r\in\zeta_j\setminus\{k\}}\eta_{r\to j}^{t-1}, \tag{5}$$

where $\eta_{j\to k}^t$ denotes the belief message passing from VN j to RN k in iteration t. $\zeta_j \setminus \{k\}$ denotes all the RNs in VN j except RN k.

The complexity of MPA is dominated by the message updating at the RN, which can be approximated as $\mathcal{O}\left(N_{iter}KM^{d_f}\right)$. N_{iter} denotes the number of iterations and d_f denotes the number of users that collide over the RN.

III. Proposed SCMA Codebook Design

A. Proposed Design Criteria

Since there are J users where every user is assigned with a codebook with size of M, one needs to deal with a superimposed constellation with size of M^J, denoted by Φ_{M^J}. Let $\boldsymbol{x} = \sum_{j=1}^{J}\boldsymbol{w}_j$ be a superimposed codeword of Φ_{M^J}. By assuming an equal probability of multiplexing codewords, the entropy can be written as

$$H(\boldsymbol{X}) = -\sum_{s=1}^{M^J}\mathrm{P}(\boldsymbol{x}_s)\log_2\mathrm{P}(\boldsymbol{x}_s) = J\log_2 M, \tag{6}$$

where $\boldsymbol{x}_s$ denotes the s-th superimposed codeword in Φ_{M^J}.

The conditional entropy can be written as

$$\begin{aligned} H(\boldsymbol{X}|\boldsymbol{Y},\boldsymbol{h}) = &-\sum_{s=1}^{M^J}\int_{-\infty}^{+\infty} f(\boldsymbol{x}_s,\boldsymbol{y})\log_2\frac{f(\boldsymbol{x}_s,\boldsymbol{y})}{\sum_{t=1}^{M^J} f(\boldsymbol{x}_t,\boldsymbol{y})}\mathrm{d}y \\ = &\frac{1}{M^J}\sum_{s=1}^{M^J}\underbrace{-\int_{-\infty}^{+\infty} f(\boldsymbol{y}|\boldsymbol{x}_s)\log_2 f(\boldsymbol{y}|\boldsymbol{x}_s)\mathrm{d}y}_{\triangleq I_1} \\ &+ \underbrace{\int_{-\infty}^{+\infty} f(\boldsymbol{y}|\boldsymbol{x}_s)\log_2\sum_{t=1}^{M^J} f(\boldsymbol{y}|\boldsymbol{x}_t)\mathrm{d}y}_{\triangleq I_2}. \end{aligned} \tag{7}$$

For I_1 in (7), we have

$$\begin{aligned} I_1 = &-\sum_{k=1}^{K}\int_{-\infty}^{+\infty}\frac{\exp\left\{-\frac{(y_k - h_k x_{k,s})^2}{2(1+h_k x_{k,s}\varsigma^2)\sigma^2}\right\}}{\sqrt{2\pi(1+h_k x_{k,s}\varsigma^2)\sigma^2}} \\ &\times\log_2\frac{\exp\left\{-\frac{(y_k - h_k x_{k,s})^2}{2(1+h_k x_{k,s}\varsigma^2)\sigma^2}\right\}}{\sqrt{2\pi(1+h_k x_{k,s}\varsigma^2)\sigma^2}}\mathrm{d}y \\ = &\frac{K}{2}\log_2 e + \sum_{k=1}^{K}\frac{1}{2}\log_2 2\pi(1+h_k x_{k,s}\varsigma^2)\sigma^2, \end{aligned} \tag{8}$$

where $x_{k,s}$ and $x_{k,t}$ denote the k-th element of K-dimensional superimposed codeword $\boldsymbol{x}_s$ and $\boldsymbol{x}_t$, respectively. y_k denotes the k-th element of receive signal $\boldsymbol{y}$.

For I_2 in (7), we have

$$I_2 \leq \log_2 \sum_{t=1}^{M^J} \int_{-\infty}^{+\infty} f(\boldsymbol{y}|\boldsymbol{x}_s) f(\boldsymbol{y}|\boldsymbol{x}_t) \mathrm{d}y = \log_2 \sum_{t=1}^{M^J} \prod_{k=1}^{K}$$
$$\int_{-\infty}^{+\infty} \frac{\exp\left\{\left(-\frac{(y_k - h_k x_{k,s})^2}{2(1+h_k x_{k,s}\varsigma^2)\sigma^2} - \frac{(y_k - h_k x_{k,t})^2}{2(1+h_k x_{k,t}\varsigma^2)\sigma^2}\right)\right\}}{2\pi\sigma^2 \sqrt{(1+h_k x_{k,s}\varsigma^2)(1+h_k x_{k,t}\varsigma^2)}} \mathrm{d}y \quad (9)$$
$$= \log_2 \sum_{t=1}^{M^J} \frac{\exp\left\{-\sum_{k=1}^{K} \frac{h_k^2 (x_{k,t} - x_{k,s})^2}{4\sigma^2(1+h_k(x_{k,t}+x_{k,s})\varsigma^2/2)}\right\}}{\prod_{k=1}^{K} \sqrt{2\pi\sigma^2(2+h_k(x_t^k + x_s^k)\varsigma^2)}},$$

where the first inequality can be obtained by the Jensen inequality.

Then, the lower bound for mutual information can be obtained as

$$\begin{aligned} I(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{h}) &= H(\boldsymbol{X}) - H(\boldsymbol{X}|\boldsymbol{Y},\boldsymbol{h}) \\ &\geq J\log_2(M) - \frac{K}{2}\log_2 e + \frac{K}{2} \\ &- \frac{1}{M^J}\sum_{s=1}^{M^J} \log_2 \sum_{t=1}^{M^J} \prod_{k=1}^{K} \left(\frac{\sqrt{1+h_k x_{k,s}\varsigma^2}}{\sqrt{1+h_k(x_{k,t}+x_{k,s})\varsigma^2/2}} \right. \\ &\quad \left. \times \exp\left\{ -\frac{h_k^2 (x_{k,t}-x_{k,s})^2}{4\sigma^2(1+h_k(x_{k,t}+x_{k,s})\varsigma^2/2)} \right\} \right) \\ &= I_L(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{h}). \end{aligned} \quad (10)$$

To maximize the $I_L(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{h})$, one can approximate it by minimizing the subsequent formula

$$I_L^p(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{h}) = \frac{1}{M^J}\sum_{s=1}^{M^J} \log_2 \sum_{t=1}^{M^J} \prod_{k=1}^{K} \exp\left\{-h_k d_{k,t,s}\right\}, \quad (11)$$

where $d_{k,t,s} = \frac{(x_{k,t}-x_{k,s})^2}{4\sigma^2(1+(x_{k,t}+x_{k,s})\varsigma^2/2)}$. Then, by taking expectation with respect to $\boldsymbol{h}$ on both sides of (11), we have

$$I_L^p(\boldsymbol{X};\boldsymbol{Y}) = \frac{1}{M^J}\sum_{s=1}^{M^J} \log_2 \sum_{t=1}^{M^J} \Delta_{\boldsymbol{x}_s \to \boldsymbol{x}_t}, \quad (12)$$

with

$$\Delta_{\boldsymbol{x}_s \to \boldsymbol{x}_t} = \prod_{k=1}^{K} \sum_{i=1}^{3} \frac{a_i(1+\kappa)}{1+\kappa+b_i d_{k,t,s}^2} \exp\left\{ \frac{-\kappa b_i d_{k,t,s}^2}{1+\kappa+b_i d_{k,t,s}^2} \right\}, \quad (13)$$

where a_i and b_i are fitting parameters of the following exponential approximation

$$e^{-dh} = \sum_{i=1}^{3} a_i e^{-b_i d^2 h^2}, \quad (14)$$

where $a_1 = 0.3595$, $a_2 = 0.4305$, $a_3 = 0.2048$, $b_1 = 8.8260$, $b_2 = 0.7691$ and $b_3 = 0.1518$.

In the high SNR regime, $I_L^p(\boldsymbol{X};\boldsymbol{Y})$ can be improved by minimizing the maximum distance of $\Delta_{\boldsymbol{x}_s \to \boldsymbol{x}_t}$ among all pairs. Hence, the MNED can be given as

$$\Im_{\min} = \min_{\boldsymbol{x}_s \neq \boldsymbol{x}_t} \{ \frac{1}{\sqrt[K]{\Delta_{\boldsymbol{x}_s \to \boldsymbol{x}_t}}} | \forall \boldsymbol{x}_s, \boldsymbol{x}_s \in \Phi_{M^J}, s \neq t \}. \quad (15)$$

1) When $\kappa \to \infty$ for AWGN channel, (15) reduces to

$$\Im_{\min} = \min_{\boldsymbol{x}_s \neq \boldsymbol{x}_t} \left\{ \sum_{k=1}^{K} d_{k,t,s} | \forall \boldsymbol{x}_s, \boldsymbol{x}_s \in \Phi_{M^J}, s \neq t \right\}. \quad (16)$$

2) When $\kappa = 0$ for Rayleigh channel, (15) reduces to

$$\Im_{\min} = \min_{\boldsymbol{x}_s \neq \boldsymbol{x}_t} \left\{ \sqrt[K]{\prod_{k=1}^{K} d_{k,t,s}} | \forall \boldsymbol{x}_s, \boldsymbol{x}_s \in \Phi_{M^J}, s \neq t \right\}. \quad (17)$$

B. Proposed Codebook Design

When there is only thermal noise, one-dimensional UPAM constellations given below can achieve the best BER performance, i.e.,

$$\boldsymbol{\mathcal{C}}_{1\times M}^{\mathrm{UPAM}} = \begin{bmatrix} 0, & 1, & \cdots, & M-1 \end{bmatrix}. \quad (18)$$

Different from traditional wireless communications, the shot noise makes the equivalent noise variance larger as the signal amplitude increases. Instead of using the uniform design of UPAM, intuitively the distance between neighboring constellation points should increase with the amplitude to meet the same BER performance requirements. To compensate the detrimental impact of the shot noise, the SQR constellation is proposed as

$$\boldsymbol{\mathcal{C}}_{1\times M}^{\mathrm{SQR}} = \begin{bmatrix} s_0, & s_1, & \cdots & s_{M-1} \end{bmatrix}, \quad (19)$$

where $\sqrt{s_i} - \sqrt{s_{i-1}} = 1$ and $s_0 = 0$.

To fully exploit the advantages of the UPAM constellation and SQR constellation, we introduce a novel NLC constellation as follows:

$$\boldsymbol{\mathcal{C}}_{M\times 1}^{\mathrm{NLC}} = \begin{bmatrix} c_0, & c_1, & \cdots & c_{M-1} \end{bmatrix}^T, \quad (20)$$

where $c_m = m + \alpha m^{\beta+1}$ with m and $\alpha m^{\beta+1}$ denote the linear and nonlinear items, respectively. $\alpha > 0$ serves as a multiplicative factor and $\beta > 0$ acts as an exponential factor, thus enabling fine control over the nonlinear compensation. α and β are determined by maximizing the MNED of the NLC. Note that the distance between two adjacent constellation point is given by $\delta_m = c_m - c_{m-1} = 1 + \alpha\left(m^{\beta+1} - (m-1)^{\beta+1}\right)$, which increases as the increase of the m.

Based on the above analysis, we present below a class of multi-dimensional mother constellation (MC) which can be obtained by interleaving the one-dimensional constellation $\boldsymbol{\mathcal{C}}_{1\times M}$. The steps to design the proposed MC are briefly summarized as follows:

Step 1: Design a one-dimensional vector as:

$$\boldsymbol{t}_{i,k}^{-} = [t_{M-i}, t_{M-i-k}, \cdots, t_{\max(0, M-i-k\lceil \frac{M}{k} \rceil)}], \quad (21)$$

where $t_i = c_{\tau m + \gamma\lfloor \frac{m}{q} \rfloor}$, $q = \left\lceil \sqrt[N]{M} \right\rceil$. τ and γ denote the slope and intercept parameters of the PwL function $\tau m + \gamma\left\lfloor \frac{m}{q} \right\rfloor$, respectively. $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ represent the floor and ceiling operations, respectively.

Step 2: Based on the one-dimensional vector obtained in Step 1, the nth dimension constellation is generated as:

$$\boldsymbol{g}_n = \begin{cases} [t_0, t_1, \cdots, t_{M-1}] & , n \text{ is odd} \\ [\boldsymbol{t}_{1,q}^{-}, \boldsymbol{t}_{2,q}^{-}, \cdots, \boldsymbol{t}_{q,q}^{-}] & , n \text{ is even} \end{cases}. \quad (22)$$

Step 3: Then, the N–dimensional MC with size M can be given by

$$\boldsymbol{\mathcal{C}}_{N\times M} = [\boldsymbol{g}_1^T, \boldsymbol{g}_2^T, \cdots, \boldsymbol{g}_N^T]^T. \quad (23)$$

For example, for $M = 4$, $N = 2$, the MC is given by

$$\boldsymbol{\mathcal{C}}_{2\times4} = \begin{bmatrix} c_0, c_\tau, c_{2\tau+\gamma}, c_{3\tau+\gamma} \\ c_{3\tau+\gamma}, c_\tau, c_{2\tau+\gamma}, c_0 \end{bmatrix}. \quad (24)$$

When $\tau = 1$ and $\gamma = 1$, (24) reduces to

$$\boldsymbol{\mathcal{C}}_{2\times4} = \begin{bmatrix} 0, 0, 1+\alpha, 1+\alpha \\ 0, 1+\alpha, 0, 1+\alpha \end{bmatrix}. \quad (25)$$

In addition, MCs with $M = 8$ and $M = 16$ are given in (26) and (27) at the bottom of the page, respectively.

For codebook generation, we also need certain properly selected PSMs. Denote $K \times J$ by the SCMA setting. In this paper, 4×6 and 5×10 SCMA settings with overloading given by 150% and 200% are considered, where the PSMs are respectively given as follows:

$$\boldsymbol{\mathcal{PS}}_{4\times6} = \begin{bmatrix} p_1 & 0 & p_2 & 0 & p_3 & 0 \\ 0 & p_1 & p_2 & 0 & 0 & p_3 \\ p_3 & 0 & 0 & p_2 & 0 & p_1 \\ 0 & p_3 & 0 & p_2 & p_1 & 0 \end{bmatrix}. \quad (28)$$

$$\boldsymbol{\mathcal{PS}}_{5\times10} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_4 & 0 & 0 & 0 & p_1 & p_2 & p_3 & 0 & 0 & 0 \\ 0 & p_3 & 0 & 0 & p_4 & 0 & 0 & p_1 & p_2 & 0 \\ 0 & 0 & p_2 & 0 & 0 & p_3 & 0 & p_4 & 0 & p_1 \\ 0 & 0 & 0 & p_1 & 0 & 0 & p_2 & 0 & p_3 & p_4 \end{bmatrix}. \quad (29)$$

Due to the real and non-negative constraints in the VLC system, phase rotation is not allowed for codebook generation. Therefore, we can only distinguish different users by their magnitude. In addition, the power imbalance in different dimensions can facilitate MPA decoding, thus leading to improved BER performance. The corresponding mapping matrix for $\boldsymbol{\mathcal{PS}}_{4\times6}$ can be written as

$$\boldsymbol{\mathcal{V}}_1 = \begin{bmatrix} p_1 & 0 \\ 0 & 0 \\ 0 & p_3 \\ 0 & 0 \end{bmatrix}, \boldsymbol{\mathcal{V}}_2 = \begin{bmatrix} 0 & 0 \\ p_1 & 0 \\ 0 & 0 \\ 0 & p_3 \end{bmatrix}, \boldsymbol{\mathcal{V}}_3 = \begin{bmatrix} p_2 & 0 \\ 0 & p_2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$
$$\boldsymbol{\mathcal{V}}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ p_2 & 0 \\ 0 & p_2 \end{bmatrix}, \boldsymbol{\mathcal{V}}_5 = \begin{bmatrix} p_3 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & p_1 \end{bmatrix}, \boldsymbol{\mathcal{V}}_6 = \begin{bmatrix} 0 & 0 \\ p_3 & 0 \\ 0 & p_1 \\ 0 & 0 \end{bmatrix}, \quad (30)$$

where $\boldsymbol{\mathcal{V}}_j = ezc\left(\text{diag}\left(\boldsymbol{\mathcal{PS}}_j\right)\right)$, $\boldsymbol{\mathcal{PS}}_j$ denotes the j-th column of $\boldsymbol{\mathcal{PS}}$. diag$(\cdot)$ stands for the vector diagonalization and $ezc(\cdot)$ denotes the resultant matrix after removing the all-zero columns. The mapping matrix $\boldsymbol{\mathcal{V}}_j$ maps the N-dimensional vector to a K-dimensional sparse SCMA codeword.

Therefore, the proposed PwL-NLC codebook of user j can be obtained by

$$\text{CB}_j = \boldsymbol{\mathcal{V}}_j \boldsymbol{\mathcal{C}}. \quad (31)$$

Based on the derived MNED, an SCMA codebook optimization problem can be formulated as

$$\begin{aligned} \max_{\boldsymbol{p},\alpha,\beta} \quad & \Im_{\min}(\text{CB}) \\ s.t. \quad & \|\text{CB}\|_2^2 = \frac{NMJ}{d_f}, \\ & \boldsymbol{p} \geq \boldsymbol{0}, \\ & \alpha \geq 0,\ \beta \geq 0, \end{aligned} \quad (32)$$

where $\boldsymbol{p}$ denotes the parameter vector in PSM and $\boldsymbol{0}$ denotes the zero vector with the same length as $\boldsymbol{p}$. For example, the parameter vectors for $\boldsymbol{\mathcal{PS}}_{4\times6}$ and $\boldsymbol{\mathcal{PS}}_{5\times10}$ can be expressed as $\boldsymbol{p} = [p_1, p_2, p_3]$ and $\boldsymbol{p} = [p_1, p_2, p_3, p_4]$, respectively.

Algorithm 1 The BBPSO Algorithm for Maximizing the MNED

Input: N_s: the size of the swarm. D: the number of parameters. I_t: the number of iterations.

Initialization: Randomly initialize particle $q_{i,j}$, where $i \in \{1, \cdots, N_s\}$ and $j \in \{1, \cdots, D\}$. $\boldsymbol{p}_i(0) = \{q_{i,1}, \cdots, q_{i,1}\}$. $\boldsymbol{g}(0) = \arg\max_{\boldsymbol{p}_i(0)} \Im_{\min}(\boldsymbol{p}_i(0))$.
for $t = 0 : I_t - 1$ **do**
  for $i = 1 : N_s$ **do**
    for $j = 1 : D$ **do**
      $u_{i,j} = (p_{i,j}(t) + g_j(t))/2$
      $\sigma_{i,j} = |p_{i,j}(t) - g_j(t))|$
      $q_{i,j}(t+1) = u_{i,j} + \sigma_{i,j} \cdot N(0,1)$
    end for
  end for
  for $i = 1 : N_s$ **do**
    if $\Im_{\min}(\boldsymbol{q}_i(t+1)) > \Im_{\min}(\boldsymbol{p}_i(t))$ **then**
      $\boldsymbol{p}_i(t+1) = \boldsymbol{q}_i(t+1)$
    else
      $\boldsymbol{p}_i(t+1) = \boldsymbol{p}_i(t)$
    end if
  end for
  $\boldsymbol{g}(t+1) = \arg\max_{\boldsymbol{p}_i(t+1)} \Im_{\min}(\boldsymbol{p}_i(t+1))$
end for

Output: Optimal results $\boldsymbol{g}(I_t)$

Unfortunately, the optimization problem in (32) is still hard to solve due to the non-convex objective function. In this paper, we propose to use the BBPSO algorithm to obtain good codebooks with maximum MNED, as shown in **Algorithm 1**. The BBPSO is modeled as a swarm of vectors whose trajectories oscillate around a field defined by the previous best value of each individual and the best values of all particles. However, as M and J increase, the computational complexity for calculating MNED becomes prohibitively high. A novel

$$\boldsymbol{\mathcal{C}}_{2\times8} = \begin{bmatrix} c_0, c_\tau, c_{2\tau}, c_{3\tau+\gamma}, c_{4\tau+\gamma}, c_{5\tau+\gamma}, c_{6\tau+2\gamma}, c_{7\tau+2\gamma} \\ c_{7\tau+2\gamma}, c_{4\tau+\gamma}, c_\tau, c_{6\tau+2\gamma}, c_{3\tau+\gamma}, c_0, c_{5\tau+\gamma}, c_{2\tau} \end{bmatrix}. \quad (26)$$

$$\boldsymbol{\mathcal{C}}_{2\times16} = \begin{bmatrix} c_0, c_\tau, c_{2\tau}, c_{3\tau}, c_{4\tau+\gamma}, c_{5\tau+\gamma}, c_{6\tau+\gamma}, c_{7\tau+\gamma}, c_{8\tau+2\gamma}, c_{9\tau+2\gamma}, c_{10\tau+2\gamma}, c_{11\tau+2\gamma}, c_{12\tau+3\gamma}, c_{13\tau+3\gamma}, c_{14\tau+3\gamma}, c_{15\tau+3\gamma} \\ c_{15\tau+3\gamma}, c_{11\tau+2\gamma}, c_{7\tau+\gamma}, c_{3\tau}, c_{14\tau+3\gamma}, c_{10\tau+2\gamma}, c_{6\tau+\gamma}, c_{2\tau}, c_{13\tau+3\gamma}, c_{9\tau+2\gamma}, c_{5\tau+\gamma}, c_{1\tau}, c_{12\tau+3\gamma}, c_{8\tau+2\gamma}, c_{4\tau+\gamma}, c_0 \end{bmatrix}. \quad (27)$$

spatial partitioning approach is employed to alleviate this problem. Firstly, the superimposed constellation points are sorted according to their spatial locations and partitioned into Z blocks. Secondly, the MNED of the M^J/Z superimposed points in each block is calculated. Finally, the minimum MNED of all the blocks is calculated to approximate the MNED of the whole constellation. This approach has demonstrated excellent performance and significantly reduced the complexity from $O(M^{2J})$ to $O(M^{2J}/Z)$.

IV. Numerical Results

In this section, we perform numerical evaluations of the proposed PwL-NLC codebooks for the (4×6) and (5×10) SCMA systems characterized by (28) and (29), respectively. These two systems give rise to overloading factors of $\lambda = 150\%$ and $\lambda = 200\%$, respectively. Let us set $\kappa = 10$ and $\varsigma^2 = 4$ for the optimization of the proposed codebooks. Given that the SNR in the research of codebook design for VLC systems typically falls within the range of 30 dB to 50 dB [13], we consider SNR=40 dB which is sufficiently large in VLC systems. The proposed codebooks are all available at our GitHub project[1]. The main codebook for comparison with the proposed ones is the logsumexp codebook in [13].

To support multiple access in VLC, conventional complex-valued SCMA codebooks should not be employed, as VLC has the requirement of real and positive signal transmission. Consequently, the positive signals possess amplitude but lack phase, and the variance of the shot noise is solely dependent on the amplitude of the signal. Fig. 1 illustrates the constellations of the UPAM, SQR and proposed NLC with $M \in \{4, 8, 16, 32\}$. We can see that UPAM constellation keeps the same spacing between adjacent constellation points. Conversely, the spacing in the proposed and SQR constellations grows as the signal amplitude increases. As far as the shot noise is concerned, one can see that the UPAM constellation is too conservative, while the SQR constellation is too sensitive. The proposed NLC constellation is optimized for MNED, striking a delicate balance and hence achieving better performance.

In Fig. 2, we evaluate the BER performances of the UPAM, SQR, and NLC constellations with $\varsigma^2 = 4$ and $M \in \{4, 8, 16, 32\}$. $\kappa \to \infty$ is the typical value for VLC system. As shown in Fig. 2, the proposed NLC constellation always achieves the best performance compared to the UPAM and SQR constellations due to the higher MNED. Although the SQR constellation slightly outperforms the UPAM constellation at low signal-to-noise ratios (SNRs), it performs poorly in the medium and high SNR regimes, especially when M is large. Specifically, the proposed constellation achieves about 3 dB and 9 dB gains over the UPAM and SQR constellations at BER=10^{-5} with $M = 8$, respectively.

Table I compares the MNEDs of the proposed PwL-NLC codebooks and logsumexp codebook [13] with overloading factors $\lambda = 150\%$ and $\lambda = 200\%$, respectively. In addition, since the minimum Euclidean distance (MED) is widely used as a performance metric [7], we also present the MEDs of different codebooks. Clearly, the proposed PwL-NLC codebook achieves larger MED and MNED values compared to the logsumexp codebook.

[1]https://github.com/Guixiaoming/SCMACodebook/tree/main/VLC

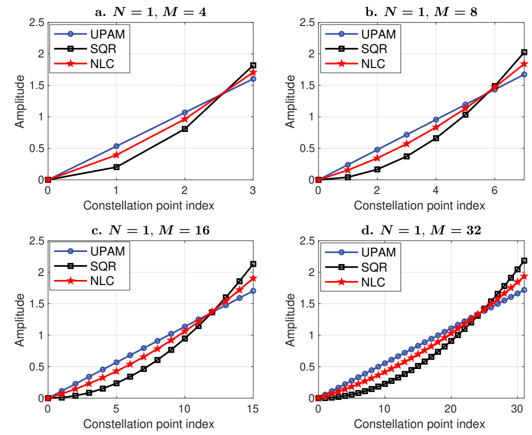

Fig. 1: Plots of UPAM, SQR, and proposed constellation with $M \in \{4, 8, 16, 32\}$.

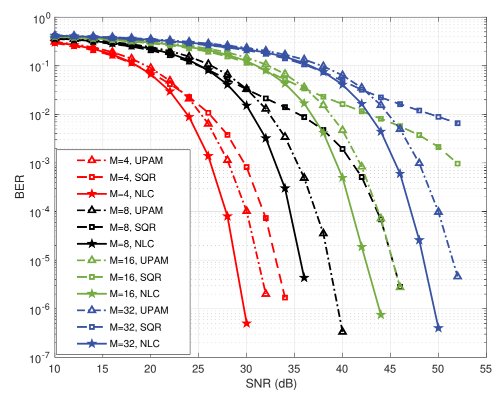

Fig. 2: BER comparison of the UPAM, SQR, and proposed constellation with $\varsigma^2 = 4$ and $\kappa \to \infty$.

TABLE I: Comparison of different codebooks in terms of MED and MNED with $\varsigma^2 = 4$ and $\kappa = 10$.

System setting (M, λ)	Codebooks	MED	MNED
(4,150%)	Logsumexp [13]	0.0540	1.1582
(4,150%)	PwL-NLC	0.1734	10.0733
(8,150%)	PwL-NLC	0.0216	1.2161
(16,150%)	PwL-NLC	0.0051	1.0070
(4,200%)	PwL-NLC	0.0270	1.2061
(8,200%)	PwL-NLC	0.0163	1.1091

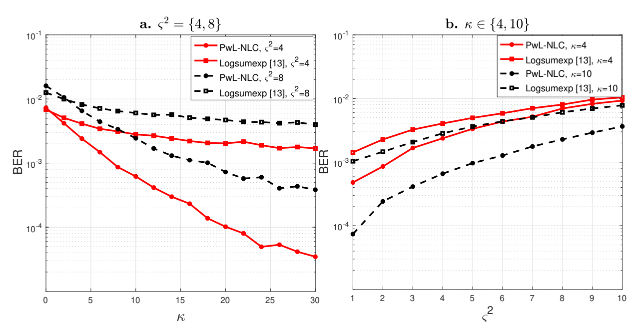

Fig. 3: Uncoded BER performance of different codebooks with $M = 4$ and $\lambda = 150\%$ (SNR = 35 dB).

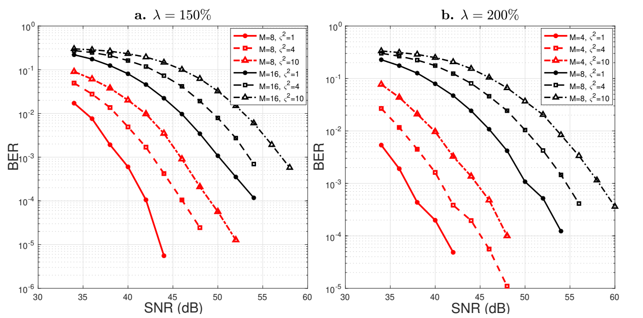

Fig. 4: Uncoded BER performance of different codebooks with $M \in \{4, 8, 16\}$ and $\lambda \in \{150\%, 200\%\}$.

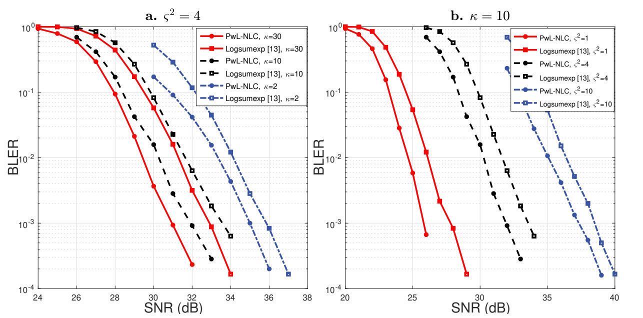

Fig. 5: LDPC coded BLER performance of different codebooks with $M = 4$ and $\lambda = 150\%$.

In Fig. 3, we evaluate the uncoded BER of the proposed PwL-NLC codebook and logsumexp codebook with $M = 4$ and $\lambda = 150\%$ at SNR = 35 dB. Fig. 3(a) compares the BER performance across different channel models with $\varsigma^2 \in \{4, 8\}$. With the increasing of κ, the communication quality of Rician channel progressively improves, leading to a significant enhancement in the BER performance. The proposed PwL-NLC codebook demonstrates a marked superiority over the logsumexp codebook. In addition, we are interested in BER performance comparison with shot noise, which is shown in Fig. 3(b). One can see that the proposed PwL-NLC codebook outperforms the logsumexp codebooks significantly.

Moreover, we extend our codebook design to higher modulation orders with $M \in \{8, 16\}$ and a larger overloading factor with $\lambda = 200\%$. In Fig. 4, we compare the uncoded BER performance with $\lambda = 150\%$ at different ς^2. On can see that the proposed PwL-NLC codebook with $\varsigma^2 = 1$ achieves about 7 dB gain over that with $\varsigma^2 = 10$ at BER=10^{-4}.

Finally, we evaluate the coded BER of the proposed PwL-NLC codebook and logsumexp codebook with $M = 4$ and $\lambda = 150\%$. We use the 5G NR LDPC code with rate of 0.8462 and block length of 260 [15]. In Fig. 5(a), the proposed PwL-NLC codebook exhibits improved error performance than that of the logsumexp codebook for $\varsigma^2 = 4$ with different channel fading. In particular, the proposed PwL-NLC codebook outperforms the logsumexp codebook by 2 dB gains at BER=10^{-3}. When there exists shot noise, as expected, the error performance of the proposed PwL-NLC codebook deteriorates. In particular, the PwL-NLC achieves about 2 dB gains over the logsumexp codebook at BER=10^{-3}.

V. Conclusions

In this paper, we have introduced a novel class of SCMA codebooks which can achieve excellent BER performance for downlink VLC system with shot noise. We have derived a lower bound of mutual information under a generalized Rician fading channel model and proposed a novel codebook design criteria. In addition, we have introduced an efficient codebook design method through the proposed PwL-NCL constellation, PSM and parameter optimization. Numerical results demonstrated the superiority of the proposed codebook in terms of uncoded BER performance and LDPC coded BLER performance improvements in different VLC system. In particular, the proposed codebooks with $\lambda = 150\%$ and $\lambda = 200\%$ achieve significant performance improvements for typical κ and ς^2 in VLC systems.

References

[1] "IEEE Approved Draft Standard for Information Technology–Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 7: Light Communications," *IEEE P802.11bb/D7.0*, pp. 1-29, Mar. 2023.

[2] A. Jovicic, J. Li and T. Richardson, "Visible light communication: opportunities, challenges and the path to market," *IEEE Communications Magazine*, vol. 51, no. 12, pp. 26-32, Dec. 2013.

[3] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347-2381, Dec. 2017.

[4] H. Marshoud, V. M. Kapinas, G. K. Karagiannidis, and S. Muhaidat,"Non-Orthogonal Multiple Access for Visible Light Communications," *IEEE Photonics Technology Letters*, vol. 28, no. 1, pp. 51-54, Jan. 2016.

[5] Z. Liu and L. -L. Yang, "Sparse or dense: a comparative study of code-domain NOMA systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 4768-4780, Aug. 2021.

[6] Q. Luo, P. Gao, Z. Liu, L. Xia, Z. Mheich, P. Xiao, and A. Maaref, "An Error Rate Comparison of Power Domain Non-Orthogonal Multiple Access and Sparse Code Multiple Access," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 500-511, Mar. 2021.

[7] X. Li, Z. Gao, Y. Gui, Z. Liu, P. Xiao, and L. Yu, "Design of Power-Imbalanced SCMA Codebook," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 2, pp. 2140-2145, Feb. 2022.

[8] Q. Luo, Z. Liu, G. Chen, P. Xiao, Y. Ma, and A. Maaref, "A Design of Low-Projection SCMA Codebooks for Ultra-Low Decoding Complexity in Downlink IoT Networks," *IEEE Transactions on Wireless Communications*, vol. 22, no. 10, pp. 6608-6623, Oct. 2023.

[9] L. Yu, P. Fan, D. Cai, and Z. Ma, "Design and analysis of SCMA codebook based on star-QAM signaling constellations," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10543-10553, Nov. 2018.

[10] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Transactions on Signal Process*, vol. 56, no. 4, pp. 1616-1626, Apr. 2008.

[11] S. Chaturvedi, Z. Liu, V. A. Bohara, A. Srivastava, and P. Xiao, "A Tutorial on Decoding Techniques of Sparse Code Multiple Access," *IEEE Access*, vol. 10, pp. 58503-58524, May. 2022.

[12] S. D. Dissanayake and J. Armstrong, "Comparison of ACO-OFDM, DCO-OFDM and ADO-OFDM in IM/DD systems," *Journal of Lightwave Technology*, vol. 31, no. 7, pp. 1063-1072, Apr. 2013.

[13] S. Chaturvedi, D. N. Anwar, V. A. Bohara, A. Srivastava, and Z. Liu, "Low-complexity codebook design for SCMA-based visible light communication," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 106-118, Jan. 2022.

[14] M. Vameghestahbanati, I. D. Marsland, R. H. Gohary, and H. Yanikomeroglu, "Multidimensional Constellations for Uplink SCMA Systems—A Comparative Study," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2169-2194, 2019.

[15] GPP TS 38.212, Rel. 16, "5G NR, multiplexing and channel coding," Oct. 2020. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38series/38.212