

Low-Complexity Codebook Design for SCMA-Based Visible Light Communication

SAUMYA CHATURVEDI¹, DIL NASHIN ANWAR¹ (Graduate Student Member, IEEE),
 VIVEK ASHOK BOHARA¹ (Senior Member, IEEE), ANAND SRIVASTAVA¹,
 AND ZILONG LIU² (Senior Member, IEEE)

¹Department of Electronics & Communications Engineering, Indraprastha Institute of Information Technology Delhi, New Delhi 110020, India

²School of Computer Science and Electrical Engineering, University of Essex, Colchester CO4 3SQ, U.K.

CORRESPONDING AUTHOR: S. CHATURVEDI (e-mail: saumyac@iitd.ac.in)

This work was supported in part by SPARC from the Ministry of Education, Government of India, under Grant SPARC/2018-2019/P950/SL, and in part by the TCS Research Scholarship Program.

ABSTRACT Sparse code multiple access (SCMA), as a code-domain non-orthogonal multiple access (NOMA) scheme, has received considerable research attention for enabling massive connectivity in future wireless communication systems. In this paper, we present a novel codebook (CB) design for SCMA based visible light communication (VLC) system, which suffers from shot noise. In particular, we introduce an iterative algorithm for designing and optimizing CB by considering the impact of shot noise at the VLC receiver. Based on the proposed CB, we derive and analyze the theoretical bit error rate (BER) expression for the resultant SCMA-VLC system. The simulation results show that our proposed CBs outperform CBs in the existing literature for different loading factors with much less complexity. Further, the derived analytical BER expression well aligns with simulated results, especially in high signal power regions.

INDEX TERMS CB design, input-dependent Gaussian noise (IDGN), sparse code multiple access (SCMA), visible light communication (VLC).

I. INTRODUCTION

IN RECENT years, visible light communication (VLC) has emerged as a reliable alternative to legacy radio-frequency (RF) based communication, especially for indoor scenarios. VLC offers various advantages such as unlicensed spectrum, free of electromagnetic interference, and ease of deployment using existing illumination infrastructure along with high-level security [1]–[3]. Besides, non-orthogonal multiple access (NOMA) has received increasing research attention in the past few years as a key technology for the enabling of massive machine type communication (mMTC) systems, in which a massive number of devices with very high density needs to be served [4], [5]. The primary concept behind NOMA is to serve multiple users over the same resource elements (frequency or time slots) using different codebooks (CBs) or power levels. This paper is concerned with the integration of VLC with a disruptive

NOMA scheme, called sparse code multiple access (SCMA), where multiple users are served simultaneously with different sparsity structures of the CBs [6]. Message passing algorithm (MPA) is used for multi-user detection by taking advantage of the sparse CBs [7]. In SCMA, CB design plays a key role in the enhanced error rate performances [8].

To support multiple access in VLC, conventional orthogonal frequency division multiple access (OFDMA) may not be employed as VLC has the requirement of real and positive signal transmission. In order to overcome the above drawback, direct current-biased optical OFDMA (DCO-OFDMA) and asymmetrically clipped optical OFDMA (ACO-OFDMA) have been studied. However, DCO-OFDMA suffers from deteriorated bit error rate (BER), and ACO-OFDMA suffers from reduced spectral efficiency [9]. In comparison, SCMA can be combined with VLC to increase the spectral efficiency because of the following reasons [10].

- The dominant line-of-sight (LOS) component and the short distance between light-emitting diode (LED) and photodetector (PD) give rise to a high signal-to-noise ratio (SNR) at VLC receiver, which helps improve the performance of SCMA.
- The limited modulation bandwidth of the LEDs can be efficiently utilized by SCMA to achieve high spectral efficiency.

Several research works incorporating SCMA in VLC (SCMA-VLC) systems have been presented in the literature. In [11], [12], experimental demonstration of SCMA-VLC system with overloading factor of 1.5 was reported.¹ In [13], a hybrid power-domain SCMA was investigated for VLC, showing improved data rate at the cost of reduced transmission distance. In [14], color-domain SCMA was proposed for VLC for a large throughput gain over conventional white light SCMA system. All the above works used the predefined RF complex SCMA CBs and made suitable changes (such as Hermitian symmetry and DC-level shifting) so that they can be applied to the VLC system. The authors in [15] focused on SCMA constellation design for VLC system in the presence of thermal noise with six users and eight resource elements (REs). In [16] and [17], SCMA CB has been designed for three users with four REs considering the impact of varying shot noise in VLC systems. Although [15]–[17] have proposed SCMA CBs for different VLC systems, however, there is still a lack of literature on designing SCMA CBs with overloading factor greater than one while considering the presence of shot noise in VLC systems.

Previously, the analytical expression of BER has been discussed [18], [19] either for power-domain NOMA techniques or for single-carrier scheme (such as on-off keying) for shot noise incorporated VLC systems. However, SCMA is a multi-carrier scheme, and its BER analysis will be different from single-carrier schemes. Further, when the SCMA-VLC system incorporates varying shot noise, the conventional closed-form BER expression does not follow the simulation results. Therefore, an exact theoretical BER expression for shot noise incorporated SCMA-VLC system has been proposed in this work. To the best of our knowledge, this paper is the first known work that proposes the SCMA CB design with 1.5 overloading factor for VLC systems considering shot noise. The contributions of this paper are summarized as follows.

- A novel multidimensional, low-complex, and power-efficient SCMA CB designing technique has been proposed for shot noise incorporated VLC system utilizing the log-sum-exponential operation in the optimization objective function of CB design. The proposed CBs outperform CBs in the existing literature [16], [17] for different load factors with much less complexity.

1. The overloading factor of 1.5 means that the number of users is 1.5 times the number of resource elements.

- The proposed optimization problem and its solution lead to flexible SCMA CBs which can meet the need of SCMA-VLC systems with varying loading factors.
- The analytical expression of BER for SCMA-VLC system considering the shot noise has been derived. The simulation results show good agreement with the derived analytical BER expression, especially for higher signal power region.

The remainder of the paper is organized as follows. The SCMA-VLC system model is discussed in Section II. The proposed CB design and the theoretical BER expression for SCMA over shot noise VLC system are discussed in Section III. The results are presented and analyzed in Section IV. Section V concludes the paper.

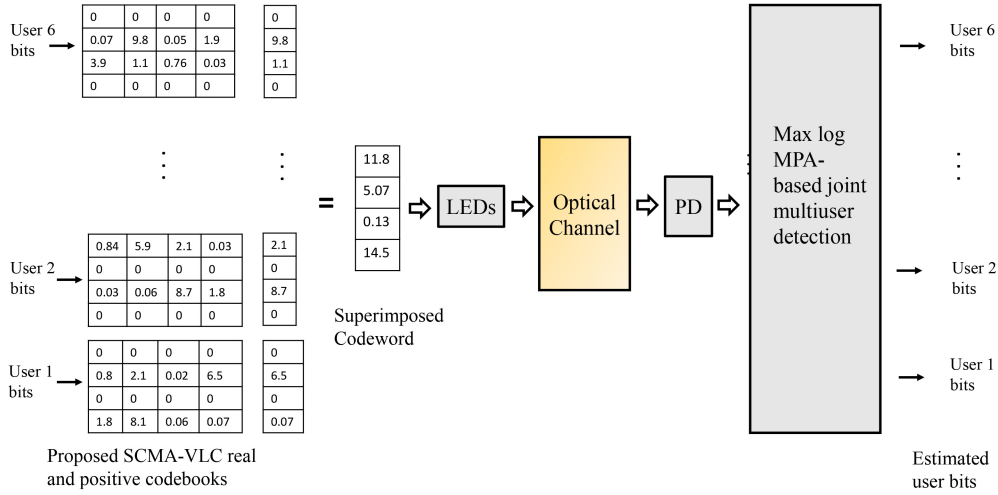
Notations: Throughout this paper, x , \mathbf{x} , \mathbf{X} denote a scalar, vector and a matrix, respectively. Symbols \mathbf{x}^T and \mathbf{X}^T represent transpose of \mathbf{x} and \mathbf{X} , respectively. The i th element of vector \mathbf{x} is denoted by x_i and $(\mathbf{X})_{ij}$ denotes the i th row and j th column element of matrix \mathbf{X} . Symbols \mathbb{B} and $\mathbb{R}_{>0}$ represents the set of binary numbers and, real and positive numbers, respectively. The diagonal matrix is denoted by $\text{diag}(\mathbf{x})$ where the i th diagonal element is x_i and ‘ $\log(\mathbf{x})$ ’ denotes the natural logarithm of each element of \mathbf{x} , respectively. The maximum value of $f(x)$ as x is varied over all its possible values is denoted by ‘ $\max_x f(x)$ ’. The probability density function (PDF) of a Gaussian random variable (RV) with mean μ and variance σ^2 is denoted by $\mathcal{N}[\mu, \sigma^2]$. The trace of a square matrix \mathbf{A} is denoted by $\text{Tr}(\mathbf{A})$ which is calculated by the sum of main diagonal elements of \mathbf{A} . The selection of r elements from a set of n elements is denoted by $\binom{n}{r}$ and $\|\mathbf{x}\|$ denotes the L2 norm of \mathbf{x} , respectively.

II. SYSTEM MODEL

A. INPUT-DEPENDENT GAUSSIAN NOISE

In the VLC system, the optical receiver (i.e., PD) measures the intensity of the incident optical signal and converts it into electrical signal. The interaction of the incoming photons with the matter of PD is a statistical process [20]. The fluctuations in the number of photons detected, results in fluctuating current which cause shot noise at the optical receiver. The generated shot noise depends on the incident optical signal itself [21]–[23]. Further, the thermal motion of the electronic carriers result in fluctuating voltage known as thermal noise.

In case of strong ambient light, the impact of thermal noise is more and the shot noise becomes very low. Thus, the noises at the VLC system are assumed to be independent of the received signal [24]–[26]. However, in case of indoor scenario, the distance between transmitter and receiver is less and the dominant signal is LOS [27], thus the received power is large. However, in practical VLC systems, typical illumination and communication scenarios offer very high SNR [28], [29]. In case of indoor scenario (as considered in our work), since the received power is large, the shot noise is more dominant and cannot be neglected.


 FIGURE 1. The system model diagram of 4×6 SCMA with the proposed CBs at the transmitter in VLC channel.

The variance of shot noise is given as

$$\sigma_{sh}^2 = 2qx_p\Re B + 2qI_B I_2 B, \quad (1)$$

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

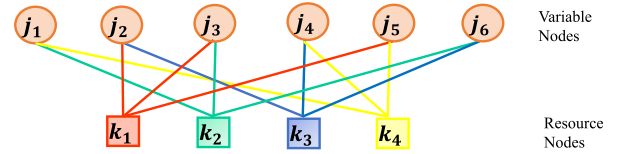
The thermal noise is assumed to be Additive white Gaussian noise (AWGN), which is independent of the intensity of the incident signal [21]. The shot noise affected optical intensity channel can be distinguished as an input-dependent Gaussian noise (IDGN), where variance of IDGN can be used to characterise the physical properties of the VLC system [21], [30]. Consequently, the received signal y can be expressed as

$$y = x_p + \sqrt{x_p}Z_1 + Z_0, \quad y \in \mathbb{R}, \quad x_p \geq 0, \quad (2)$$

where $Z_1 \sim \mathcal{N}_{\mathbb{R}}[0, \zeta^2 \sigma^2]$ denotes a real Gaussian RV with zero mean and $\zeta^2 \sigma^2$ variance describing IDGN (shot noise), and $Z_0 \sim \mathcal{N}_{\mathbb{R}}[0, \sigma^2]$ describes the thermal noise. Here, Z_1 and Z_0 are assumed to be independent, and $\zeta^2 \in [1, 10]$ denotes the strength of shot noise with respect to thermal noise. In an indoor scenario, as the transmitter and the receiver parameters change, the received power changes, and thus the shot noise varies. In this paper, the variation in shot noise has been depicted by considering different shot noise factor ζ^2 values.

B. THE SCMA-VLC MODEL

Consider a downlink indoor LOS VLC system where multiple users are being served, each equipped with a single receiving device. Every J users are grouped together for SCMA, and each group is supported by K REs (e.g., time or frequency slots), providing the load factor $\lambda = J/K$. In Fig. 1, the data transmission and reception of six users ($J = 6$) on four REs ($K = 4$) with $\lambda = 1.5$ is shown.


 FIGURE 2. Factor graph representation of 4×6 SCMA system with $d_f = 3$ and $N = 2$.

At the transmitter, $b = \log_2 M$ bits are mapped to K -dimensional real and positive codewords, with M number of codewords in a CB. These codewords are sparse vectors with N non-zero elements, $K > N$. Each user has a dedicated CB allotted to it. The CB of j th user can be expressed as

$$\mathbb{A}_j = \mathbf{V}_j \mathbf{C}_j, \quad \text{for } j = 1, \dots, J, \quad (3)$$

where $\mathbf{V}_j \in \mathbb{B}^{K \times N}$ and $\mathbf{C}_j \in \mathbb{R}_{>0}^{N \times M}$ denote the mapping matrix and constellation matrix of the j th user, respectively. The constellation matrix $\mathbf{C}_j = [\mathbf{c}_j^1, \dots, \mathbf{c}_j^M]$, where $\mathbf{c}_j^m \in \mathbb{R}_{>0}^{N \times 1}$ denotes the m th constellation point of \mathbf{C}_j . The whole CB structure of SCMA can be represented by the factor graph matrix, $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_J)$, where $\mathbf{f}_j = \text{diag}(\mathbf{V}_j \mathbf{V}_j^T)$. The user j is being served by resource node (RN) k , if $\mathbf{F}_{kj} = 1$ ($1 \leq k \leq K, 1 \leq j \leq J$). For a 4×6 SCMA block having $K = 4$ REs and $J = 6$ users, one example of the factor graph matrix of size 4×6 is given as

$$\mathbf{F}_{4 \times 6} = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}. \quad (4)$$

Fig. 2 shows an example of factor graph representation of six users (or Variable Nodes (VNs)) with four RNs, $N = 2$ and data of d_f users is being superimposed on each RN ($d_f = 3$ for $\mathbf{F}_{4 \times 6}$ shown in (4)). The binary mapping matrices

corresponding to the six users are given below

$$\begin{aligned} \mathbf{V}_1 &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{V}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{V}_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \\ \mathbf{V}_4 &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{V}_5 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{V}_6 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \end{aligned} \quad (5)$$

The set of CBs for J users is dependent on the constellation matrices \mathbf{C}_j , $j = 1, \dots, J$ as the mapping matrices remain fixed as shown in (5). Assuming synchronous multiplexing between users, the received signal can be expressed as

$$\begin{aligned} \mathbf{y} &= \sum_{j=1}^J \text{diag}(\mathbf{h}_j) \mathbf{x}_j + \text{diag} \left(\sum_{j=1}^J \text{diag}(\mathbf{h}_j) \mathbf{x}_j \right)^{\frac{1}{2}} \mathbf{n}_{sh} + \mathbf{n}_{th}, \\ &= \sum_{j=1}^J \text{diag}(\mathbf{h}_j) \mathbf{V}_j \mathbf{c}_j + \text{diag} \left(\sum_{j=1}^J \text{diag}(\mathbf{h}_j) \mathbf{V}_j \mathbf{c}_j \right)^{\frac{1}{2}} \mathbf{n}_{sh} + \mathbf{n}_{th}, \end{aligned} \quad (6)$$

where $\mathbf{x}_j \in \mathbb{R}_{>0}^{K \times 1}$ is a codeword of j th user, $\text{diag}(\mathbf{h}_j)$ denotes the channel matrix of j th user, the second term represents the IDGN with $\mathbf{n}_{sh} \sim \mathcal{N}_{\mathbb{R}}[\mathbf{0}, \zeta^2 \sigma^2 \mathbf{I}]$, and $\mathbf{n}_{th} \sim \mathcal{N}_{\mathbb{R}}[\mathbf{0}, \sigma^2 \mathbf{I}]$. The channels are assumed to be unit vectors for simplicity, and the extension to the general VLC values is straight-forward.

In case of the movement of the users, the channel characteristics of the users (\mathbf{h}_j) of j th user in (6) will be affected. The random waypoint (RWP) mobility model is a simple and straightforward stochastic model that describes a human movement behavior in indoor scenarios [31]. In case of user mobility, the received signal expression remains same as (6), however the channel vector (\mathbf{h}) varies as per the RWP model in (6).

C. SCMA-VLC DECODING

Assuming perfect channel estimation and the set of CBs known to receiver, the SCMA multi-user codeword $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_J]$ can be detected by solving the joint maximum a-posteriori (MAP) probability mass function (PMF) of the transmitted codewords [7], [32]. The estimated \mathbf{X} can be written as

$$\hat{\mathbf{X}} = \underset{\mathbf{x}_j \in \mathbb{A}_j, \forall j}{\text{argmax}} p(\mathbf{X}|\mathbf{y}), \quad (7)$$

where \mathbb{A}_j represents the set of codewords allotted to the j th user. The marginal distribution of (7) with respect to \mathbf{x}_j (i.e., summing the joint PMF of multi-user codeword over all values of \mathbf{X} except \mathbf{x}_j) leads to j th user detected symbol as

$$\hat{\mathbf{x}}_j = \underset{\mathbf{x}_j \in \mathbb{A}_j}{\text{argmax}} \sum_{\mathbf{x}_{-j}} p(\mathbf{X}|\mathbf{y}). \quad (8)$$

Using Bayes' rule

$$\hat{\mathbf{x}}_j = \underset{\mathbf{x}_j \in \mathbb{A}_j}{\text{argmax}} \sum_{\mathbf{x}_{-j}} f(\mathbf{y}|\mathbf{X}) P(\mathbf{X}), \quad (9)$$

where $P(\mathbf{X})$ denotes the joint a-priori PMF of all transmitted codewords and $f(\mathbf{y}|\mathbf{X})$ denotes the conditional pdf of the received vector. Assuming noise components are independent and identically distributed for all REs. Then (9) becomes

$$\hat{\mathbf{x}}_j = \underset{\mathbf{x}_j \in \mathbb{A}_j}{\text{argmax}} \sum_{\mathbf{x}_{-j}} \left(P(\mathbf{X}) \prod_{k=1}^K f(y_k|\mathbf{X}) \right), \quad \text{for } j = 1, \dots, J, \quad (10)$$

where y_k denotes the received signal at the k th RE. Solving the above marginal product of function (MPF) problem with brute force will lead to exponential complexity. Taking advantage of sparse CBs, this problem can be approximately solved by iterative MPA with moderate complexity [7], [33]. MPA works as a decoding algorithm based on passing the extrinsic information between RNs and VNs. The first message is passed from the leaf node (node which has only one neighboring node), and a node of degree d will remain idle until it has received messages from the $d - 1$ number of neighboring nodes. MPA involves a large number of exponential operations which are of high complexity. A mathematical simplification is to move MPA to logarithm domain and use Jacobian logarithm, which gives a simpler version called Max-Log-MPA [34]. With Jacobian logarithm, the log-sum of exponentials can be approximated as a maximum operation as shown below:

$$\log(\exp(a_1) + \dots + \exp(a_J)) \approx \max(a_1, \dots, a_J). \quad (11)$$

The Max-Log-MPA for detection of SCMA codewords in VLC is based on three steps as shown in **Algorithm 1**. Let ζ_j and ξ_k , be the set of nodes connected to VN j and RN k , respectively. Let $\eta_{k \rightarrow j}$ and $\eta_{j \rightarrow k}$ be the message passed from RN k to VN j and vice-versa. Also, $r \in \xi_k \setminus \{j\}$ denotes all the VNs of ξ_k except VN j and $d \in \zeta_j \setminus \{k\}$ denotes the RNs in ζ_j except the k th RN. In **Algorithm 1**, it is assumed that the received signal \mathbf{y} and all the CBs \mathbf{C}_j , $\forall j = 1, \dots, J$ are known at the receiver. The total number of iterations considered in **Step 2** are N_t and, VNs and RNs are updated for each iteration t . Also, the channel coefficient between k th RN and m th VN is denoted by h_{km} and the codeword element transmitted by m th VN on k th RN is denoted by x_{km} , respectively.

In the case of shot noise, the variance at each RN is different and dependent on the superimposed codeword element at the corresponding RN, while in the presence of only thermal noise, $\zeta^2 = 0$, so $\rho^2 = \sigma^2$ and variance will be the same for all RNs. In case of three users scenario, the factor graph would not be cyclic and number of elements being superimposed on each RN would be different ($d_f = 2$ for $k = 1, 2$ and $d_f = 1$ for $k = 3, 4$). Since d_f is different for different REs, therefore, **Algorithm 1** will change accordingly [35].

Algorithm 1: Max-Log-MPA**Input:** \mathbf{y} , \mathbf{C}_j , $\forall j = 1, \dots, J$.**Output:** Estimated value of bits for j th user, $\forall j = 1, \dots, J$.**1: Step 1: Initialization**The prior probability of each codeword: $\eta_{j \rightarrow k}^0 = \log \frac{1}{M}$.**2: Step 2: Message passing along the nodes****for** $t = 1$ **to** N_t **do**

a) Update RN:

$$\eta_{k \rightarrow j}^t(\mathbf{x}_j) = \max_{\mathbf{x}_j} \left(\frac{-1}{2\rho^2} \left\| \mathbf{y}_k - \sum_{m \in \xi_k} h_{km} \mathbf{x}_{km} \right\|^2 + \sum_{r \in \xi_k \setminus \{j\}} \eta_{r \rightarrow k}^{t-1}(\mathbf{x}_r) \right),$$

where, $\rho^2 = \sigma^2 + \varsigma^2 \sigma^2 \sum_{m \in \xi_k} h_{km} \mathbf{x}_{km}$.

b) Update VN:

$$\eta_{j \rightarrow k}(\mathbf{x}_j) = \log \left(\frac{1}{M} \right) + \sum_{d \in \xi_j \setminus \{k\}} \eta_{d \rightarrow j}^{t-1}(\mathbf{x}_j).$$

end**3: Step 3: Bit log-likelihood ratio (LLR)**

The final belief computed at each VN is:

$$\log \mathbf{I}_j(\mathbf{x}_j) = \log \left(\frac{1}{M} \right) + \sum_{k \in \xi_j} \eta_{k \rightarrow j}(\mathbf{x}_j)$$

The κ th bit LLR at j th VN is

$$LLR_{\kappa}^j = \max_{\{\mathbf{x}_j \in \mathbb{A}_j | \kappa=0\}} (\log(\mathbf{I}_j(\mathbf{x}_j))) - \max_{\{\mathbf{x}_j \in \mathbb{A}_j | \kappa=1\}} (\log(\mathbf{I}_j(\mathbf{x}_j))).$$

where, $\hat{\kappa}_j = 0$ if LLR_{κ}^j is positive otherwise $\hat{\kappa}_j = 1$.

The computational complexity of standard MPA and Max-Log-MPA considering the effect of shot noise in decoding is shown in Table 1 [36]. The number of operations is shown for the update step of RN as shown in **Step 2** (a) of **Algorithm 1**. Max-Log-MPA requires more addition operations and less multiplication operations than MPA. This gap of increase in addition and decrease in multiplication operations between MPA and Max-Log-MPA increases with the overloading factor.

III. CB DESIGN AND OPTIMIZATION

For SCMA CB design and optimization, most researchers aim for maximization of the minimum Euclidean distance (MED) between the superimposed codewords [37]–[39]. In this section, a method of designing constellation matrices $\mathbf{C} = [\mathbf{C}_1, \dots, \mathbf{C}_J] \in \mathbb{R}_{>0}^{N \times MJ}$ is introduced and discussed as an optimization problem.

A. THE PROBLEM FORMULATION

In the AWGN channel, MED is taken as one of the key performance indicators for determining the performance of the designed SCMA CB. Let \mathbf{s}_i be a superimposed codeword from the combined constellation matrix $\mathbb{M}_v \in \mathbb{R}_{>0}^{K \times MJ}$. Then, the MED between $\mathbf{s}_i \in \mathbb{M}_v$ and $\mathbf{s}_j \in \mathbb{M}_v$ is given as

$$d'_{ij} = (\mathbf{s}_i - \mathbf{s}_j)^T (\mathbf{s}_i - \mathbf{s}_j), \quad \forall i < j \in [1, M^J]. \quad (12)$$

TABLE 1. Complexity analysis of MPA and Max-Log-MPA considering shot noise.

Operations	MPA	Max-Log-MPA
Exponential	$M^{d_f} K d_f N_{it}$	0
Multiplication	$(d_f + 3) M^{d_f} K d_f N_{it}$	$4 M^{d_f} K d_f N_{it}$
Addition	$(2d_f + 2) M^{d_f} K d_f N_{it}$	$(3d_f + 1) M^{d_f} K d_f^2 N_{it}$
Comparison	0	$M^{d_f} K d_f N_{it}$

The MED emerge as a critical parameter because of the pair-wise error probability (PEP) for AWGN channel which is given as

$$P'(\mathbf{s}_i \rightarrow \mathbf{s}_j) = Q \left(\sqrt{\frac{\|\mathbf{s}_i - \mathbf{s}_j\|^2}{2\sigma^2}} \right). \quad (13)$$

However, in case of IDGN, the PDF of the received signal \mathbf{y} is given as (assuming unit channel gain)

$$f(\mathbf{y}|\mathbf{s}_i) = \frac{1}{(2\pi)^{K/2}} \frac{1}{|\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{s}_i)^T \Sigma_i^{-1} (\mathbf{y} - \mathbf{s}_i) \right\}, \quad (14)$$

where $\Sigma_i = \text{diag}\{\varsigma^2 \sigma^2 \mathbf{s}_i + \sigma^2 \mathbf{1}\}$ denotes the covariance matrix for the \mathbf{s}_i th superimposed codeword. It can be seen from (14) that the variance is now dependent on the superimposed codeword.

In the case of only thermal noise, MED is taken as a critical parameter for designing CB where variance is independent of the superimposed codeword. However, in the case of shot noise, when variance is no longer independent, the conventional MED cannot serve as an optimal performance indicator. With the effect of shot noise, the shape of constellation points gets distorted. To mitigate the effect of IDGN, the authors in [16] introduced an improved optimization objective based on Rotated-MED for designing constellation matrices. The rotated Euclidean distance (RED) between superimposed codewords \mathbf{s}_i and \mathbf{s}_j can be expressed as

$$d_{ij} = (\mathbf{s}_i - \mathbf{s}_j)^T \mathbf{G}_i^{-\frac{1}{2}} \mathbf{G}_j^{-\frac{1}{2}} (\mathbf{s}_i - \mathbf{s}_j), \quad \forall i < j \in [1, M^J], \quad (15)$$

where $\mathbf{G}_i = \varsigma^2 \text{diag}\{\mathbf{s}_i\} + \mathbf{I}$, $\mathbf{G}_j = \varsigma^2 \text{diag}\{\mathbf{s}_j\} + \mathbf{I}$ and $\mathbf{I} \in \mathbb{R}^{K \times K}$ represents the Identity matrix. It is to be noted from (15) that the new parameter RED incorporates the shot noise effect while computing the Euclidean distance between superimposed codewords. The authors in [16] aimed at maximizing the rotated minimum Euclidean distance (R-MED) subject to the power constraints. The objective function proposed in [16] for SCMA CB design is given as

$$\max_{\mathbf{L} \in \mathbb{R}_{>0}^{NMJ \times 1}} \left\{ \min_{1 \leq i < j \leq M^J} d_{ij}(\mathbf{L}) \right\}, \quad (16)$$

where $\mathbf{L} = \text{vec}(\mathbf{C})$. This was further simplified as

$$\max_{\mathbf{L} \in \mathbb{R}_{>0}^{NMJ \times 1}, d_{\min} > 0} d_{\min}, \quad (17)$$

where d_{\min} denotes the minimum value of RED. It has been verified in [16] that R-MED based CB outperforms the MED based CB for VLC systems affected by shot noise. The results in [16] have shown that for $J = 3, K = 4, M = 4, \varsigma^2 = 5$ and $\sigma^2 = 0.01$, BER of 10^{-3} is achieved at $P_e = 10$ and $P_e = 7$ for MED and R-MED based CBs, where P_e denotes the maximum electrical power allotted to a user. Since from the literature we can understand that RED works as an important parameter for designing CB than MED, we have included RED in our optimization problem. Further, the authors in [17] introduced the concept of distance range (DR) based on the maximum and the minimum values of the RED. They have mentioned that for the low to medium SNR region, the BER is no longer dependent on just R-MED. The optimization problem proposed in [17] for CB designing is

$$\max_{d_{\min} > 0, d_{\max} > 0} \{d_{\min} - \alpha d_{\max}\} \quad (18a)$$

$$\text{s.t. } L_i > 0, \quad \forall i \in [1, NMJ], \quad (18b)$$

$$\frac{1}{M} \text{Tr}(\mathbf{C}_j^T \mathbf{C}_j) \leq P_e, \quad \forall j \in [1, J], \quad (18c)$$

$$d_{ij} \geq d_{\min}, \quad \forall i < j, \quad (18d)$$

$$d_{ij} \leq d_{\max}, \quad \forall i < j, \quad (18e)$$

where d_{\max} denotes the maximum value of RED, and $\alpha > 0$ denotes the weighting factor. The objective function shown in (18a) is dependent on d_{\min} and d_{\max} . The real and positive optical signal constraint is considered in (18b) and maximum electrical power constraint is considered in (18c). All the REDs will lie between d_{\min} and d_{\max} (18d), (18e). Also, it is to be noted that, lesser the overlapping between the constellation points of a user, better the performance of the designed CBs.

B. THE OBJECTIVE FUNCTION

The objective function mentioned in (16) considers all the REDs while designing and optimizing the CBs to get optimum results. However, considering all the REDs makes the objective function non-convex and non-differentiable [40]. To solve (16), authors in [17] have proposed the objective function as shown in (18). It should be noted that the numerical methods used for solving the nonlinear maxi-min problem aim at approximating the $F(x) = \min_i f_i(x)$ by close enough smooth functions, and the closeness of approximation decide the nearness to optimal solutions [41]. If we consider $\sum_i \phi(f_i(x))$, the approximation to $F(x)$ becomes better when $\phi(\cdot)$ is a strong convex function (i.e., large positive second derivative). One approach to solve this non-convex problem is to use well known ‘‘log sum-exponentials’’ continuous approximation [42], [43]. With this approximation, the objective function in (16) transforms to

$$\min_{\mathbf{L} \in \mathbb{R}_{>0}^{NMJ \times 1}} \frac{1}{\beta} \ln \left\{ \sum_{i \neq j} e^{-\beta d_{ij}(\mathbf{L})} \right\}, \quad (19)$$

where $\beta > 0$ is a tuning parameter which results in the smoothness of (19) and affects the accuracy of the model.

Algorithm 2: SCMA CB Optimization Algorithm

Step 1: Initialization:

Randomly initialize $\mathbf{L} = \mathbf{L}^0$ satisfying the intensity and power constraints, $\beta = 1$.

Step 2: Optimization

1: **for** $\beta \leq 30$ **do**

2: (a) Solve the optimization problem

$$[\mathbf{L}'^t, \mathbf{L}'] \leftarrow \min \frac{1}{\beta} \ln \left\{ \sum_{i \neq j} e^{-\beta d_{ij}(\mathbf{L})} \right\}$$

subject to Constraints.

where f_v^t denotes the value of the objective function at the obtained solution \mathbf{L} at t' th iteration.

3: (b)

4: **if** $|f_v^t - f_v^{t'-1}| \geq 10^{-3}$; **then**

5: $t' = t' + 1$, go to **Step 2** (a);

6: **else**

7: **exit**

8: **endif**

9: **end for**

Output: \mathbf{L}' as the SCMA output constellation stacked vector.

- Example 1: Let $\beta = 1$, and there are two distances $d_1 \gg d_2$, then

$$-\ln(e^{-d_1} + e^{-d_2}) \approx -\ln(e^{-d_2}) = d_2 = \min\{d_1, d_2\}.$$

The function in (19) is differentiable but still non-convex; however, optimization algorithms based on the derivation of the Hessian of the objective function can find the local optimal solution [44]. Also, if the objective function is optimized number of times with initialization from small set of random numbers and choosing the β parameter wisely, the likelihood of obtaining the global optimal solution increases [45]. Thus, the optimization problem is formulated as

$$\min_{\mathbf{L} \in \mathbb{R}_{>0}^{NMJ \times 1}} \frac{1}{\beta} \ln \left\{ \sum_{i \neq j} e^{-\beta d_{ij}(\mathbf{L})} \right\} \quad (20a)$$

$$\text{s.t. } L_i > 0, \quad \forall i \in [1, NMJ], \quad (20b)$$

$$\frac{1}{M} \text{Tr}(\mathbf{C}_j^T \mathbf{C}_j) \leq P_e, \quad \forall j \in [1, J], \quad (20c)$$

where (20b) denotes the real positive intensity constraint of the transmitted codeword elements and (20c) denotes that the average electrical power of the generated codewords should be less than the maximum electrical power (P_e) allotted per user. The optimization problem in (20) can be solved using commercially available solvers like MATLAB's *fmincon*. The procedure for solving the above optimization problem is summarized in **Algorithm 2**.

C. COMPLEXITY ANALYSIS

For the case of the proposed optimization problem in (20), initially, the set of all multiplexed vectors (or superimposed codewords) \mathbb{M}_v is computed. With J users and M number of codewords in the CB, the number of superimposed codewords obtained are M^J , i.e., $|\mathbb{M}_v| = M^J$. Since the complexity of operations such as addition and subtraction is $O(1)$ therefore, the complexity of computing multiplexed

vectors is $O(JM^J)$. Next, for generating the covariance matrix (Σ_i for \mathbf{s}_i superimposed codeword), each multiplexed vector is multiplied with the scalar ς^2 which results in the complexity $O(KM^J)$. The complexity of computing RED (as shown in (15)) between two multiplexed vectors is $O(K)$ as the length of each vector is K and the total number of multiplexed vectors are M^J . Thus, the total number of REDs between M^J multiplexed vectors is $\binom{M^J}{2}$ and the total complexity for calculating REDs will be $O(K \binom{M^J}{2})$. Next, the complexity of computing the objective function (20a) is $\binom{M^J}{2}$. The complexity of constraints in (20b) and (20c) is $O(1)$ since it includes addition and square operations. The maximum value of β at which the algorithm converges is denoted by β_{max} and number of iterations for the MATLAB's *fmincon* solver is N_{it} , then total complexity to generate the constellation matrices for J users using **Algorithm 2** is $O_{Logsum} = [\beta_{max}N_{it}(JM^J + KM^J + K \binom{M^J}{2} + \binom{M^J}{2} + 1)]$. If we take the most dominant part resulting for complexity, it is equal to $O(\beta_{max}N_{it}K \binom{M^J}{2})$.

In case of the DR based CB [17] optimization problem (as shown in (18)), the total complexity comes from the computation of the objective function and four constraints. The third and fourth distance based constraints (18d), (18e) are non-convex and are approximated using first order Taylor series [16]. From the calculation of multiplexed vectors, covariance matrices and REDs, the complexity is $O(JM^J + KM^J + K \binom{M^J}{2})$. The complexity of intensity and power constraints (18b), (18c) is $O(1)$. For the approximation of two non-convex constraints (18d), (18e), firstly the coefficient matrix Π_i is calculated to obtain multiplexed vector \mathbf{s}_i from \mathbf{L} [16], and its complexity is M^JKNMJ . Next, we found that the derivative of RED has the highest complexity because of the kronecker product operation involved. The derivative term has the complexity $O((KNMJ)^2 \binom{M^J}{2})$. So, with N_{out} as total number of outer loop iterations, the total complexity is $O_{DR} = [O(N_{out}N_{it}(JM^J + KM^J + K \binom{M^J}{2} + \binom{M^J}{2} + 1 + M^JKNMJ + (KNMJ)^2 \binom{M^J}{2}))]$. Taking only the dominant part, the complexity becomes $O(N_{out}N_{it}(KNMJ)^2 \binom{M^J}{2})$. It can be observed that the proposed Log-sum-exp scheme (20) has reduced the computation complexity of SCMA CB design by a significant amount from DR based CB optimization problem shown in (18).

D. THEORETICAL BER

SCMA is a multi-carrier scheme, i.e., the superimposed transmitted codeword $\mathbf{s}_i = [s_i^1, \dots, s_i^K]$ is a K dimensional real codeword as discussed in Section II-B. In this section, the theoretical BER expression has been derived for the IDGN incorporated SCMA-VLC system. For only the thermal noise case, the PEP is given as (13). However, in the case of shot noise, the overall variance changes and is dependent on ς^2 and the intensity of the incident optical

signal. The error rate for the superimposed codeword \mathbf{s}_i is

$$P'(\mathbf{s}_i) = \sum_{\substack{i'=1 \\ i' \neq i}}^{M^J-1} P'(\mathbf{s}_i \rightarrow \mathbf{s}_{i'}) \quad (21)$$

$$= \sum_{\substack{i'=1 \\ i' \neq i}}^{M^J-1} Q\left(\sqrt{\sum_{k=1}^K \frac{\|s_i^k - s_{i'}^k\|^2}{2v_{i,k}}}\right), \quad (22)$$

where s_i^k and $s_{i'}^k$, denotes the k th element of the superimposed codeword \mathbf{s}_i and $\mathbf{s}_{i'}$, respectively. The k th diagonal element of the covariance matrix Σ_i of the superimposed codeword \mathbf{s}_i is denoted by $v_{i,k}$. The diagonal element $v_{i,k}$ denotes the total variance at k th RE when \mathbf{s}_i was transmitted. The total probability of error is given as [46]

$$P_{err} = \frac{1}{J \log_2 M} \sum_{i=1}^{M^J} P_a(\mathbf{s}_i) \sum_{\substack{i'=1 \\ i' \neq i}}^{M^J-1} h_d(\mathbf{s}_{i'}, \mathbf{s}_i) P'(\mathbf{s}_i \rightarrow \mathbf{s}_{i'}), \quad (23)$$

where $P_a(\mathbf{s}_i)$ denotes the *a priori* probability of \mathbf{s}_i , $h_d(\mathbf{s}_{i'}, \mathbf{s}_i)$ denotes the bit error when $\mathbf{s}_{i'}$ is chosen over \mathbf{s}_i and $P'(\mathbf{s}_i \rightarrow \mathbf{s}_{i'})$ denotes the PEP between \mathbf{s}_i and $\mathbf{s}_{i'}$.

IV. RESULTS AND DISCUSSION

In this section, the simulation and theoretical results of the BER performance of the proposed SCMA CB are presented. Firstly, using **Algorithm 2**, we can generate the proposed CBs for different users. Further, Max-Log-MPA (as shown in **Algorithm 1**) has been used for multi-user detection.

A. RF BASED CBS PERFORMANCE IN VLC SYSTEM

We consider the case where RF-based CBs are used for transmitting data over a VLC system. We consider the 4×6 SCMA CBs from [37] and [47]. RF CBs cannot be employed as it is, as the VLC system has the requirement of real and positive signal transmission. So, the complex codeword elements are divided into real and imaginary parts, which are then DC biased to get real and positive codeword elements [37], [47]. These real and positive codewords are transmitted over different REs. This way, the data of six users is transmitted over eight REs. Further, we have generated the Log-sum-exp based CB for six users and eight REs scenario using **Algorithm 2**. Fig. 3 shows the bit error rate (BER) performance of the Star QAM based CB [37], Huawei CB [47] and Log-sum-exp CBs for $K = 8$ and $\sigma^2 = 0.01$. It can be understood from Fig. 3 that CBs designed solely for the VLC system (proposed Log-sum-exp based CB) outperforms the RF-based CBs for shot noise incorporated VLC system. The performance gap observed in Fig. 3 motivated the authors to design CB especially for the SCMA-VLC system.

B. THE CASE OF THREE USERS

For the initial simulation, we consider the case where three users share four REs with $M = 4$ and the number of non-zero

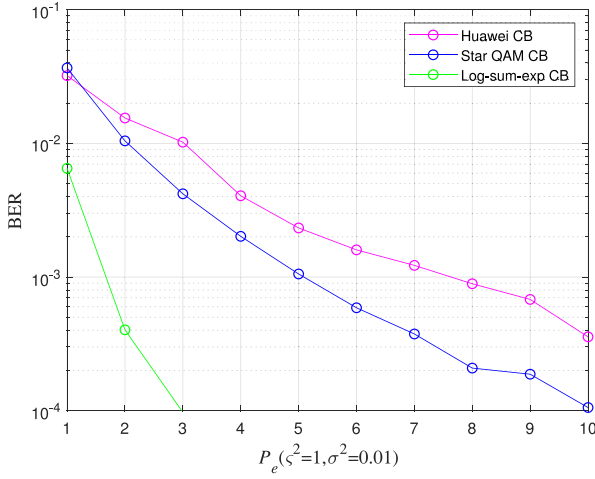


FIGURE 3. BER performance comparison for $J = 6$ between Star QAM based CB [37], Huawei CB [47] and Log-sum-exp CBs for $K = 8$, $\sigma^2 = 0.01$, $M = 4$.

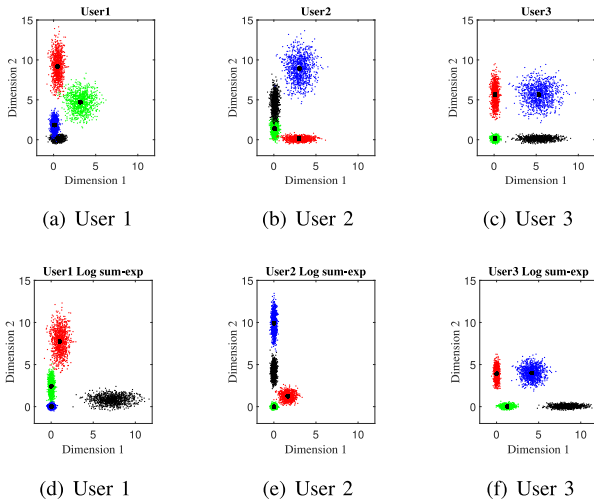


FIGURE 4. SCMA constellations scatter plots with $J = 3$, $K = 4$, $P_e = 30$, $\zeta^2 = 10$ (a)-(c): DR based constellation sets, (d)-(f): Log-sum-exp based constellation sets.

elements in the codeword is $N = 2$. The mapping matrices $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3$ have the same values as shown in (5) (i.e., $d_f = 2$ for $k = 1, 2$ and $d_f = 1$ for $k = 3, 4$). For three users, the CBs (DR based and Log-sum-exp based) are generated by solving the optimization problems (18), (20) (CBs are shown in Appendix). For $J = 3$, **Algorithm 2** is found to be converging for low values of β (i.e., $\beta = 10$). Fig. 4. shows the 2-dimensional (2D) scatter plots of the generated DR based CBs (18) and the proposed Log-sum-exp based CBs (20) for $\zeta^2 = 10$ and $P_e = 30$. It can be noted that the constellation points under DR criterion (18) are substantially overlapping as shown in Fig. 4(a)-4(c). However, the overlapping between proposed Log-sum-exp based constellation points for all three users is negligible as shown in Fig. 4(d)-4(f). As the overlapping increases between different constellation points, it becomes difficult for the receiver to detect the symbols correctly, and thus, the overall performance deteriorates.

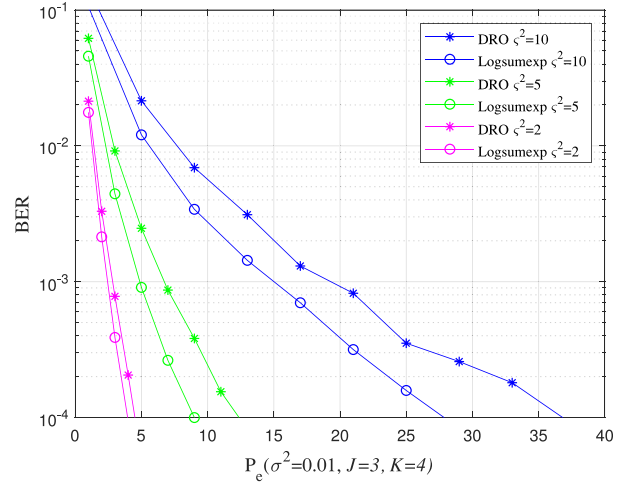


FIGURE 5. BER performance comparison for $J = 3$ between DR based CBs and Log-sum-exp CBs for $\sigma^2 = 0.01$, $M = 4$.

Fig. 5 shows the bit error rate (BER) performance for three users of the proposed Log-sum-exp based (20) SCMA CBs and the DR based CBs (18), respectively. The negligible overlapping in the case of log-sum-exp CBs as compared to DR CBs (in Fig. 4), reasons the performance gain obtained in Fig. 5. For both CBs (DR based and Log-sum-exp based), Max-Log-MPA (**Algorithm 1**) has been used for multi-user detection. It can be noticed that for the low value of ζ^2 (e.g., $\zeta^2 = 2$), BER performance of both CBs is close, and for high values of ζ^2 (e.g., $\zeta^2 = 5, 10$), the proposed Log-sum-exp based CBs yield better performance gain as compared to DR based CBs. This suggests that the objective function with all the REDs results in better CB (i.e., Log-sum-exp based CB) than the CB with only minimum and maximum values of REDs (i.e., DR based CB). It is shown that our proposed CBs achieve the BER of 10^{-3} at $P_e = 15$ while DR based CB achieve it at $P_e \approx 20$ for $\zeta^2 = 10$. This suggests that for large values of ζ^2 , the advantage of the proposed Log-sum-exp based CB is more significant than DR based CB. Further, from the results it can inferred that, we can reduce the dimming level to $\frac{30-4}{30} = 86.6\%$ of designed power ($P_e = 30$) for $\zeta^2 = 2$.

The bivariate normal density is specified by the 2D constellation points and the variance of these points in different dimensions. Fig. 6 shows the lines of the equal probability density (EPD) functions of the bivariate Gaussian. The center of the ellipse is the 2D constellation point and its shape is determined by the covariance matrix of the 2D constellation points. The principal axes of these ellipses are given by the eigenvectors of the covariance matrix and, semi-major and semi-minor axis are dependent on the eigenvalues, respectively [48]. For $N = 2$, the m th constellation point of the j th user is $\mathbf{c}_j^m \in \mathbb{R}_{>0}^{2 \times 1}$ and the covariance matrix of this 2D constellation point is denoted by $\mathbf{S}_{c_j^m}$. So, the center of the ellipse is \mathbf{c}_j^m and its shape is determined by $\mathbf{S}_{c_j^m}$. Fig. 6(a)-6(c) shows the lines of EPD for DR based constellation points and Fig. 6(d)-6(f) shows lines of EPD for Log-sum-exp based constellation points, respectively.

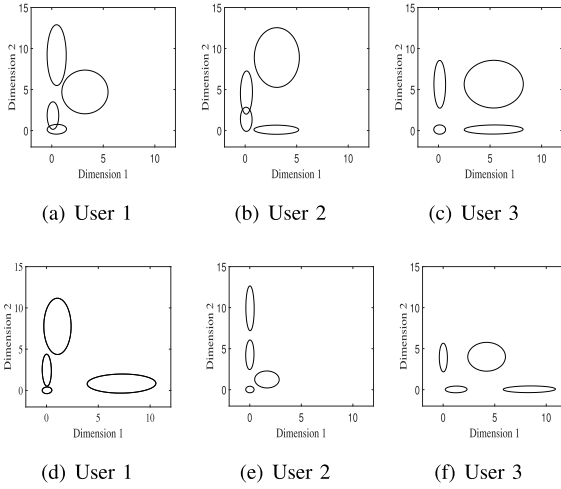


FIGURE 6. Lines of equal probability density functions of 2D SCMA constellations points with 95% confidence interval for $J = 3$, $K = 4$, $P_e = 30$, $\zeta^2 = 10$ (a)-(c): DR based EPDLs, (d)-(f): Log-sum-exp based EPDLs.

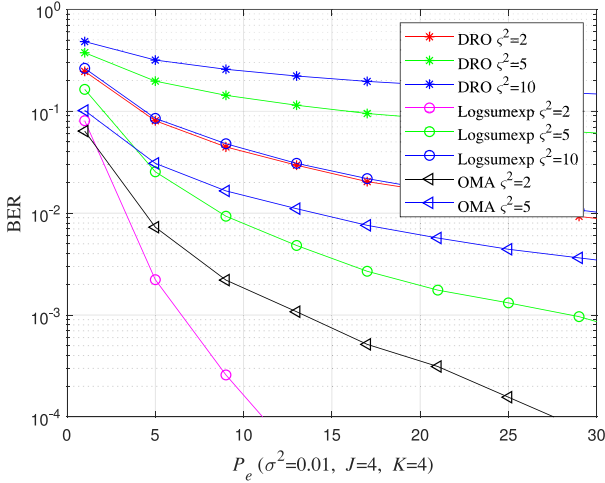


FIGURE 7. BER performance comparison for $J = 4$ between SCMA DR based CBs, Log-sum-exp CBs and OMA scheme for $\sigma^2 = 0.01$, $M = 4$.

C. THE CASE WITH FOUR, FIVE AND SIX USERS ($J = 4, 5, 6$)

Consider the case where four users share four REs. The Log-sum-exp based CB is designed for four users using **Algorithm 2**, and its performance is compared with DR based CB. Fig. 7 shows the BER performance of four users where Log-sum-exp based CB outperforms DR based CBs for different values of ζ^2 . The reason behind the performance difference can be understood from the level of overlapping between the constellation points for both CBs in Fig. 8. It can be observed in Fig. 8 that the DR based 2D constellation points suffer from significant overlapping, especially for user 1 (Fig. 8(a)) and user 4 (Fig. 8(d)), while Log-sum-exp based 2D constellation points are barely overlapping. This results in significant performance improvement for Log-sum-exp based CBs. Also in Fig. 7, the comparison with pulse amplitude modulation orthogonal multiple

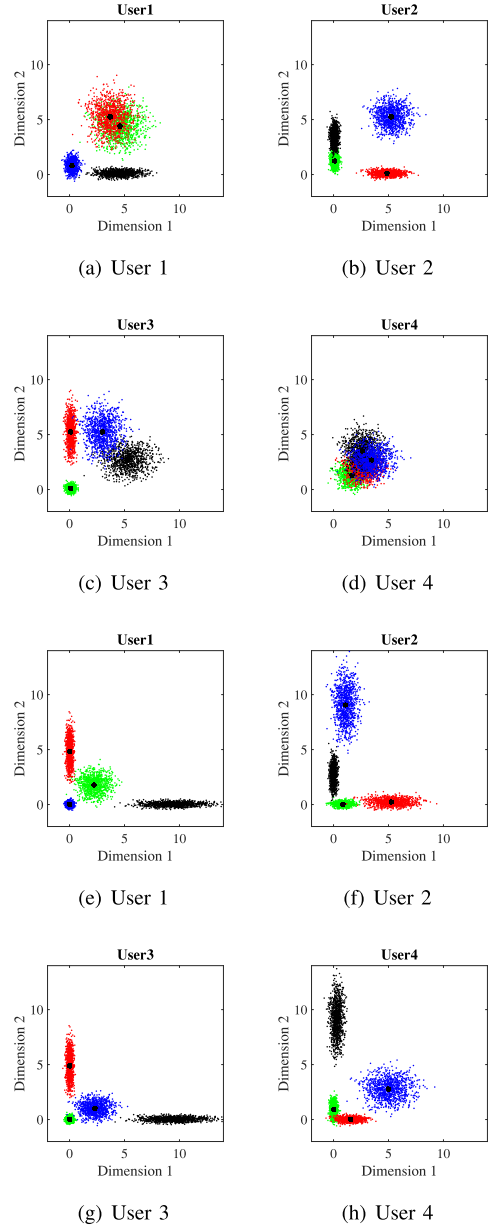


FIGURE 8. SCMA constellations with $J = 4$, $K = 4$, $P_e = 30$, $\zeta^2 = 10$ (a)-(d): DR based constellation sets, (e)-(f): Log-sum-exp based constellation sets.

access (4 PAM-OMA) scheme is shown. It is noted that the proposed algorithm outperforms the OMA scheme for different values of ζ^2 . This is expected as the OMA scheme offers no frequency diversity in contrast to the SCMA scheme.

For five users, the complexity of generating DR based CB increases significantly. Further, the performance degrades drastically and requires a very high P_e at BER of 10^{-3} . Therefore, the effect of varying shot noise factor ζ^2 with only Log-sum-exp based CBs, when five users data is being transmitted over four REs ($\lambda = 5/4$) has been shown in Fig. 9. The BER performance with respect to increasing P_e is shown in Fig. 10 for six users. Here, the effect of varying shot noise factor ζ^2 is shown when the proposed

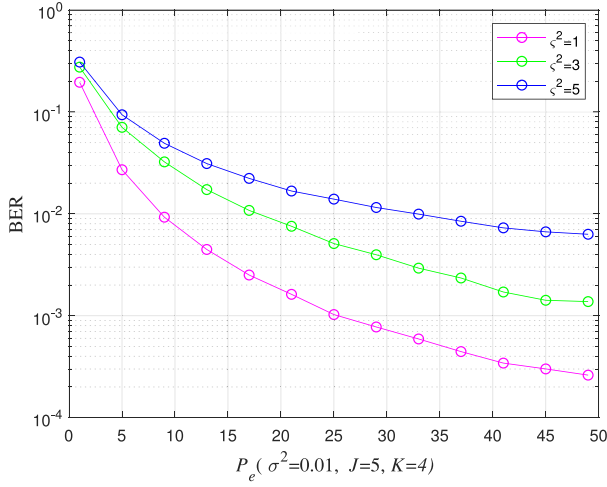


FIGURE 9. BER performance comparison for $J = 5$ based on Log-sum-exp CBs for $\sigma^2 = 0.01$, $M = 4$.

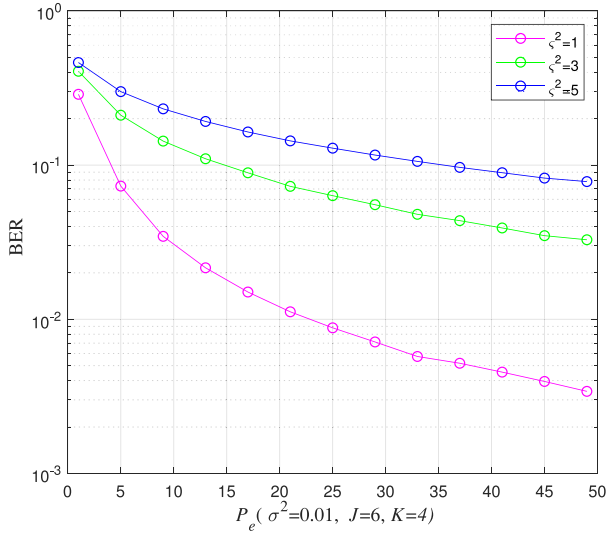


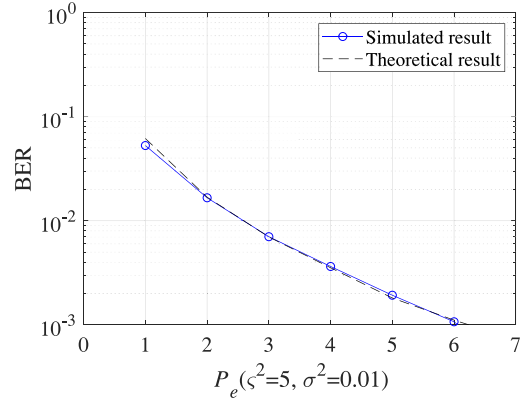
FIGURE 10. BER performance comparison for $J = 6$ Log-sum-exp CBs for $\sigma^2 = 0.01$, $M = 4$.

Log-sum-exp based CBs have been used for transmitting data over four REs.

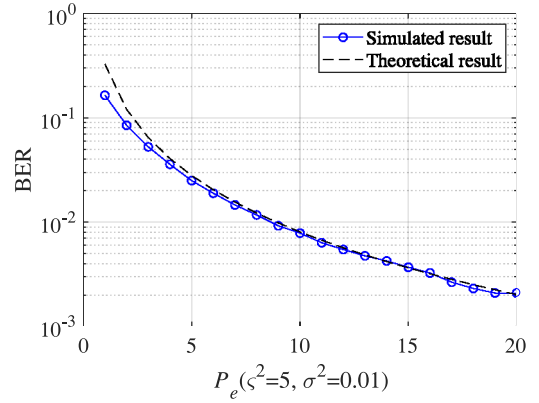
In Fig. 11, the theoretical and simulated BER performances are compared. We find that the simulated curves match with the theoretical curves, especially for the high signal power region. These results verify the correctness of the simulated results obtained in Fig. 5 and Fig. 7. The results from Fig. 5, 7, 9 and 10 have been summarized in Table 2. From Table 2, we have the electrical power required to achieve BER of 10^{-3} for different users and VLC systems with different shot noise factors.

V. CONCLUSION

In this paper, a novel low-complex CB design technique has been proposed for SCMA-VLC system impaired by shot noise. This work emphasizes that the designing of specific CBs is required for the SCMA-VLC system. The proposed



(a) Three Users



(b) Four Users

FIGURE 11. Theoretical and Simulated BER performance comparison for (a) $J = 3$ (b) $J = 4$ for $\zeta^2 = 5$, $\sigma^2 = 0.01$ and $M = 4$.

TABLE 2. Electrical power required to get BER of 10^{-3} , for different users and ζ^2 at $K = 4$.

Number of Users	ζ^2	P_e
3	1	3
3	5	5
3	10	15
4	1	6.5
4	5	29
5	1	25
5	3	>50
6	1	>50

CBs are generated by optimizing the Log-sum of exponentials of all REDs between the superimposed codewords. With the inclusion of shot noise in the SCMA system, the overall noise variance is dependent on the strength of the incident signal, as shown in **Algorithm 1**. The proposed CBs provide improved BER performance as compared to CBs proposed in the existing literature along with reduced complexity. Further, the theoretical analysis of BER is also performed for the shot noise incorporated SCMA-VLC system. The simulated results were validated, and both the theoretical and simulated curves are in good agreement for high signal power regions. In the future, the performance of SCMA codebook can further be enhanced for lower P_e by

designing CBs considering lower dimming requirements, and improve the performance of higher load factor scenarios in the SCMA-VLC system.

APPENDIX

The DR criteria based SCMA CB for VLC system obtained for three users for $\varsigma^2 = 5$ and $P_e = 30$ is shown below:

$$\begin{aligned}\mathbf{CB}_1^{DR} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 3.1908 & 0.4595 & 0.1090 & 0.4907 \\ 0 & 0 & 0 & 0 \\ 4.6727 & 9.1620 & 1.8083 & 0.1476 \end{bmatrix}, \\ \mathbf{CB}_2^{DR} &= \begin{bmatrix} 0.1074 & 2.9890 & 3.0405 & 0.1300 \\ 0 & 0 & 0 & 0 \\ 1.3517 & 0.1077 & 8.8601 & 4.5986 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \mathbf{CB}_3^{DR} &= \begin{bmatrix} 0.1119 & 0.1105 & 5.3128 & 5.3146 \\ 0.1123 & 5.6158 & 5.6235 & 0.1178 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.\end{aligned}$$

The Log-sum-exp based SCMA CB for VLC system obtained for three users for $\varsigma^2 = 5$ and $P_e = 30$ is shown below:

$$\begin{aligned}\mathbf{CB}_1^{LS} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2.7712 & 0.0100 & 0.0100 & 2.6317 \\ 0 & 0 & 0 & 0 \\ 4.4089 & 9.1626 & 1.4151 & 0.0100 \end{bmatrix}, \\ \mathbf{CB}_2^{LS} &= \begin{bmatrix} 0.0100 & 2.7785 & 2.5709 & 0.0100 \\ 0 & 0 & 0 & 0 \\ 0.0100 & 1.4036 & 9.1888 & 4.3893 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \mathbf{CB}_3^{LS} &= \begin{bmatrix} 0.0100 & 5.4749 & 0.0100 & 5.4747 \\ 0.0100 & 5.4797 & 5.4796 & 0.0100 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.\end{aligned}$$

Next, the SCMA CB for VLC system generated for four users based on Log-sum-exp for $\varsigma^2 = 5$ and $P_e = 30$ is shown below:

$$\begin{aligned}\mathbf{CB}_1^{LS} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1.5383 & 3.0405 & 0.0167 & 8.9796 \\ 0 & 0 & 0 & 0 \\ 1.3426 & 5.0483 & 0.0173 & 0.0374 \end{bmatrix}, \\ \mathbf{CB}_2^{LS} &= \begin{bmatrix} 0.0191 & 3.5636 & 4.1652 & 0.9455 \\ 0 & 0 & 0 & 0 \\ 0.0156 & 0.5631 & 8.9683 & 2.7765 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \mathbf{CB}_3^{LS} &= \begin{bmatrix} 0.0278 & 6.5059 & 0.0257 & 6.4581 \\ 0.0296 & 4.1903 & 4.2063 & 0.0268 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \mathbf{CB}_4^{LS} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.0207 & 1.5398 & 5.0348 & 0.0264 \\ 0.0154 & 2.3810 & 0.3892 & 9.2690 \end{bmatrix}.\end{aligned}$$

Similarly, the SCMA CB for VLC system based on the proposed scheme for five users for $\varsigma^2 = 5$ and $P_e = 30$ is shown below:

$$\begin{aligned}\mathbf{CB}_1^{LS} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1.8229 & 0.8713 & 0.0707 & 9.3633 \\ 0 & 0 & 0 & 0 \\ 0.2044 & 4.4911 & 1.5872 & 0.5904 \end{bmatrix}, \\ \mathbf{CB}_2^{LS} &= \begin{bmatrix} 1.3717 & 4.8155 & 1.2190 & 0.1399 \\ 0 & 0 & 0 & 0 \\ 0.0725 & 1.1399 & 9.0214 & 2.4409 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \mathbf{CB}_3^{LS} &= \begin{bmatrix} 0.1790 & 0.3468 & 2.3449 & 8.1443 \\ 0.0915 & 2.9830 & 5.7936 & 0.3253 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \mathbf{CB}_4^{LS} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.0936 & 1.8345 & 4.8909 & 0.6543 \\ 1.1709 & 0.1005 & 2.2914 & 9.0194 \end{bmatrix}, \\ \mathbf{CB}_5^{LS} &= \begin{bmatrix} 3.1438 & 0.1500 & 6.2322 & 1.5089 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.1397 & 1.4186 & 3.3720 & 6.4569 \end{bmatrix}.\end{aligned}$$

At last, we show the SCMA CB for VLC system based on the proposed scheme for six users for $\varsigma^2 = 5$ and $P_e = 30$ as follows:

$$\begin{aligned}\mathbf{CB}_1^{LS} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.8677 & 2.1631 & 0.0295 & 6.5262 \\ 0 & 0 & 0 & 0 \\ 1.8603 & 8.1426 & 0.0594 & 0.0741 \end{bmatrix}, \\ \mathbf{CB}_2^{LS} &= \begin{bmatrix} 0.8433 & 5.9356 & 2.0682 & 0.0335 \\ 0 & 0 & 0 & 0 \\ 0.0379 & 0.0580 & 8.6673 & 1.8618 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \mathbf{CB}_3^{LS} &= \begin{bmatrix} 0.1106 & 0.0293 & 1.9845 & 9.2333 \\ 0.0687 & 4.7442 & 2.2192 & 0.0930 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \mathbf{CB}_4^{LS} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.0476 & 1.5733 & 5.8671 & 0.0408 \\ 1.0104 & 0.0347 & 0.4297 & 8.9561 \end{bmatrix}, \\ \mathbf{CB}_5^{LS} &= \begin{bmatrix} 3.2891 & 0.0679 & 3.8873 & 0.1111 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.0304 & 1.3519 & 3.6688 & 5.7811 \end{bmatrix}, \\ \mathbf{CB}_6^{LS} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.0788 & 9.8115 & 0.0528 & 1.9884 \\ 3.9862 & 1.1829 & 0.7601 & 0.0323 \\ 0 & 0 & 0 & 0 \end{bmatrix}.\end{aligned}$$

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor Dr. Nestor Chatzidiamantis and the anonymous reviewers for their constructive suggestions which have greatly helped to improve the quality of this work. They are also thankful to Vyacheslav P. Klimentyev for the publicly available codes and various discussions.

REFERENCES

- [1] S. Rajagopal, R. D. Roberts, and S.-K. Lim, "IEEE 802.15.7 visible light communication: Modulation schemes and dimming support," *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 72–82, Mar. 2012.
- [2] M. Kavehrad, "Sustainable energy-efficient wireless applications using light," *IEEE Commun. Mag.*, vol. 48, no. 12, pp. 66–73, Dec. 2010.
- [3] A. Jovicic, J. Li, and T. Richardson, "Visible light communication: Opportunities, challenges and the path to market," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 26–31, Dec. 2013.
- [4] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct. 2017.
- [5] S. Chen, Y.-C. Liang, S. Sun, S. Kang, W. Cheng, and M. Peng, "Vision, requirements, and technology trend of 6G: How to tackle the challenges of system coverage, capacity, user data-rate and movement speed," *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 218–228, Apr. 2020.
- [6] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE 24th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun.*, London, U.K., Sep. 2013, pp. 332–336.
- [7] R. Hoshyari, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1616–1626, Apr. 2008.
- [8] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "SCMA codebook design," in *Proc. IEEE 80th Veh. Technol. Conf.*, Vancouver, BC, Canada, Sep. 2014, pp. 1–5.
- [9] S. Dimitrov, S. Sinanovic, and H. Haas, "Clipping noise in OFDM-based optical wireless communication systems," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1072–1081, Apr. 2012.
- [10] H. Marshoud, V. M. Kapinas, G. K. Karagiannidis, and S. Muhaidat, "Non-orthogonal multiple access for visible light communications," *IEEE Photon. Technol. Lett.*, vol. 28, no. 1, pp. 51–54, Jan. 1, 2016.
- [11] B. Lin, X. Tang, Z. Zhou, C. Lin, and Z. Ghassemloooy, "Experimental demonstration of SCMA for visible light communications," *Opt. Commun.*, vol. 419, pp. 36–40, Jul. 2018.
- [12] J. An and W.-Y. Chung, "Single-LED multichannel optical transmission with SCMA for long range health information monitoring," *J. Lightw. Technol.*, vol. 36, no. 23, pp. 5470–5480, Dec. 1, 2018.
- [13] B. Lin, X. Tang, and Z. Ghassemloooy, "A power domain sparse code multiple access scheme for visible light communications," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 61–64, Jan. 2020.
- [14] R. Mitra, S. Sharma, G. Kaddoum, and V. Bhatia, "Color-domain SCMA NOMA for visible light communication," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 200–204, Jan. 2021.
- [15] S. Lou, C. Gong, Q. Gao, and Z. Xu, "SCMA with low complexity symmetric CB design for visible light communication," in *Proc. Int. Conf. Commun. (ICC)*, Jul. 2018, pp. 1–6.
- [16] S. Hu, Q. Gao, C. Gong, Z. Xu, R. Boluda-Ruiz, and K. Qaraqe, "Energy-efficient modulation for visible light SCMA system with signal-dependent noise," in *Proc. Int. Symp. Commun. Syst. Netw. Digit. Signal Process. (CSNDSP)*, Budapest, Hungary, Jul. 2018, pp. 1–6.
- [17] Q. Gao, S. Hu, C. Gong, E. Serpedin, K. Qaraqe, and Z. Xu, "Distance-range-oriented constellation design for VLC-SCMA downlink with signal-dependent noise," *IEEE Commun. Lett.*, vol. 23, no. 3, pp. 434–437, Mar. 2019.
- [18] J. Y. Wang, J. B. Wang, and Y. Wang, "Fundamental analysis for visible light communication with input-dependent noise," in *Optical Fiber and Wireless Communications*, R. Róka, Ed. Rijeka, Croatia: InTech, 2017, doi: [10.5772/68019](https://doi.org/10.5772/68019).
- [19] V. Dixit and A. Kumar, "An exact BER analysis of NOMA-VLC system with imperfect SIC and CSI," *AEU Int. J. Electron. Commun.*, vol. 138, Aug. 2021, Art. no. 153864, doi: [10.1016/j.aeu.2021.153864](https://doi.org/10.1016/j.aeu.2021.153864).
- [20] J. M. Senior and M. Y. Jamro, *Optical Fiber Communications: Principles and Practice*. 3rd ed. New York, NY, USA: Prentice-Hall, 2008.
- [21] S. M. Moser, "Capacity results of an optical intensity channel with input-dependent Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 207–223, Jan. 2012.
- [22] Q. Gao, S. Hu, C. Gong, and Z. Xu, "Modulation designs for visible light communications with signal-dependent noise," *J. Lightw. Technol.*, vol. 34, no. 23, pp. 5516–5525, Dec. 1, 2016.
- [23] J.-Y. Wang, H. Ge, J.-X. Zhu, J.-B. Wang, J. Dai, and M. Lin, "Adaptive spatial modulation for visible light communications with an arbitrary number of transmitters," *IEEE Access*, vol. 6, pp. 37108–37123, 2018.
- [24] K.-I. Ahn and J. K. Kwon, "Capacity analysis of M-PAM inverse source coding in visible light communications," *J. Lightw. Technol.*, vol. 30, no. 10, pp. 1399–1404, May 15, 2012.
- [25] J.-B. Wang, Q.-S. Hu, J. Wang, M. Chen, and J.-Y. Wang, "Tight bounds on channel capacity for dimmable visible light communications," *J. Lightw. Technol.*, vol. 31, no. 23, pp. 3771–3779, Dec. 1, 2013.
- [26] A. Chaaban, Z. Rezki, and M.-S. Alouini, "Fundamental limits of parallel optical wireless channels: Capacity results and outage formulation," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 296–311, Jan. 2017.
- [27] D. N. Anwar, R. Ahmad, and A. Srivastava, "Energy-efficient coexistence of LiFi users and light enabled IoT devices," *IEEE Trans. Green Commun. Netw.*, early access, Sep. 29, 2021, doi: [10.1109/TGCN.2021.3116267](https://doi.org/10.1109/TGCN.2021.3116267).
- [28] J. Grubor, S. Randel, K.-D. D. Langer, and J. Walewski, "Broadband information broadcasting using LED-based interior lighting," *J. Lightw. Technol.*, vol. 26, no. 24, pp. 3883–3892, Dec. 15, 2008.
- [29] L. Hanzo, H. Haas, S. Imre, D. O'Brien, M. Rupp, and L. Gyongyosi, "Wireless myths, realities, and futures: From 3G/4G to optical and quantum wireless," *Proc. IEEE*, vol. 100, no. Special Centennial Issue, pp. 1853–1888, May 2012.
- [30] H. Chen and Z. Xu, "A two-dimensional constellation design method for visible light communications with signal-dependent shot noise," *IEEE Commun. Lett.*, vol. 22, no. 9, pp. 1786–1789, Sep. 2018.
- [31] C. Bettstetter, H. Hartenstein, and X. Pérez-Costa, "Stochastic properties of the random waypoint mobility model," *Wireless Netw.*, vol. 10, no. 5, pp. 555–567, 2004.
- [32] Y. Lin, Y. Liu, and Y. Siu, "Low complexity message passing algorithm for SCMA system," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2466–2469, Dec. 2016.
- [33] A. Bayesteh, H. Nikopour, M. Taherzadeh, H. Baligh, and J. Ma, "Low complexity techniques for SCMA detection," in *Proc. IEEE Globecom Workshops*, San Diego, CA, USA, Dec. 2015, pp. 1–6.
- [34] W. B. Ameur, P. Mary, M. Dumay, J.-F. Héland, and J. Schwoerer, "Performance study of MPA, Log-MPA and max-log-MPA for an uplink SCMA scenario," in *Proc. 26th Int. Conf. Telecommun. (ICT)*, Hanoi, Vietnam, 2019, pp. 411–416.
- [35] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [36] J. Liu, G. Wu, S. Li, and O. Tirkkonen, "On fixed-point implementation of Log-MPA for SCMA signals," *IEEE Wireless Commun. Lett.*, vol. 5, no. 3, pp. 324–327, Jun. 2016.
- [37] L. Yu, P. Fan, D. Cai, and Z. Ma, "Design and analysis of SCMA codebook based on star-QAM signaling constellations," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10543–10553, Nov. 2018.
- [38] V. P. Klimentyev and A. B. Sergienko, "SCMA CBs optimization based on genetic algorithm," in *Proc. Eur. Wireless 23th Eur. Wireless Conf.*, Dresden, Germany, 2017, pp. 1–6.
- [39] M. Vameghestahbanati, I. D. Marsland, R. H. Gohary, and H. Yanikomeroglu, "Multidimensional constellations for uplink SCMA systems—A comparative study," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2169–2194, 3rd Quart., 2019.

- [40] J. M. Luna-Rivera, V. Guerra, J. Rufo-Torres, R. Perez-Jimenez, C. Suarez-Rodriguez, and J. Rabadan-Borges, "Low-complexity colour-shift keying-based visible light communications system," *IET Optoelectron.*, vol. 9, no. 5, pp. 191–198, Oct. 2015.
- [41] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York, NY, USA: Springer-Verlag, 1999.
- [42] R. Chen, "Solution of minimax problems using equivalent differentiable functions," *Comput. Math. Appl.*, vol. 11, no. 12, pp. 1165–1169, Dec. 1985.
- [43] D. U. Campos-Delgado, J. M. Luna-Rivera, R. Perez-Jimenez, C. A. Gutierrez, V. Guerra, and J. Rabadan, "Constellation design for color space-based modulation in visible light communications," *Phys. Commun.*, vol. 31, pp. 154–159, Dec. 2018.
- [44] E. Monteiro and S. Hranilovic, "Constellation design for color-shift keying using interior point methods," in *Proc. IEEE Globecom Workshops*, Anaheim, CA, USA, Dec. 2012, pp. 1224–1228.
- [45] D. N. Anwar and A. Srivastava, "Constellation design for single photodetector based CSK with probabilistic shaping and white color balance," *IEEE Access*, vol. 8, pp. 159609–159621, 2020.
- [46] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels*, 1st ed. Hoboken, NJ, USA: Wiley, 2000.
- [47] "5G Algorithm Innovation Competition." Altera Innovate Asia FPGA Design Contest. 2015. [Online]. Available: <http://www.innovateasia.com/5g/images/pdf/1%20%20Algorit%20Innovati%20Competition-ENV1%2%20SCMA.pdf>
- [48] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York, NY, USA: Wiley, 2001.

SAUMYA CHATURVEDI received the B.E. degree in electronics and communication engineering from the University of Technology of Madhya Pradesh in 2014, and the M.Tech. degree in electronics and communication engineering from the National Institute of Technology Delhi in 2018. She is currently pursuing the Ph.D. degree with the Wireless and Radio Communication Research Lab, Indraprastha Institute of Information Technology Delhi. Her research interests include next-generation communication system, nonorthogonal multiple access, and visible light communication.

DIL NASHIN ANWAR (Graduate Student Member, IEEE) received the B.Tech. and M.Tech. degrees from the Department of Electronics and Communication Engineering, Aligarh Muslim University, Aligarh, in 2013 and 2015, respectively. She is currently pursuing the Ph.D. degree with the Department of Electronics and Communication Engineering, Indraprastha Institute of Information Technology (IIIT) Delhi. As a Research Assistant, she has worked in the classification of EEG signals with IIIT Delhi, till 2017. She is also a Visvesvaraya Ph.D. Fellow, supported by DeitY, Government of India. Her research interests include enabling green communication in resource, power and hardware constrained systems, such as IoT and UAVs with light as a communication medium. She is further invested in LiFi-based coexistence and resource allocation of heterogeneous users.

VIVEK ASHOK BOHARA (Senior Member, IEEE) received the Ph.D. degree from Nanyang Technological University, Singapore, in 2011. From 2011 to 2013, he was a Postdoctoral Researcher (Marie Curie Fellowship) with ESIEE Paris, University Paris-East. In 2013, he joined the Indraprastha Institute of Information Technology Delhi, India, where he is currently an Associate Professor. He has authored and coauthored over 50 publications in major IEEE/IET journals and refereed international conferences, two book chapters, and one patent. His research interests are toward next-generation communication technologies, such as device-to-device communication, carrier aggregation, and visible light communications. He received the First Prize in National Instruments ASEAN Virtual Instrumentation Applications Contest in 2007 and 2010. He was also the recipient of the Best Poster Award at the IEEE ANTS 2014 and the IEEE Comsnet 2015 and 2016 conferences.

ANAND SRIVASTAVA received the M.Tech. and Ph.D. degrees from IIT Delhi. Before joining IIIT Delhi, he was the Dean and the Professor with the School of Computing and Electrical Engineering, Indian Institute of Technology Mandi, India, and also an Adjunct Faculty with IIT Delhi. Prior to this, he was with Alcatel-Lucent-Bell Labs, India, as a Solution Architect for access and core networks. Before joining Alcatel Lucent, he had a long stint (20 years) with the Center for Development of Telematics (CDOT), a telecom research center of Govt. of India, where he was the Director and a member of CDOT Board. During his stay in CDOT, he provided technical leadership and motivation to extremely qualified team of engineers engaged in the development of national-level projects in the areas of telecom security systems, network management system, intelligent networks, operations support systems, access networks (GPON), and optical technology-based products. Majority of these projects were completed successfully and commercially deployed in the public network. He was also closely involved with ITU-T, Geneva, in Study Group 15 and represented India for various optical networking standards meetings. His research work is in the area of optical core and access networks, vehicle-to-vehicle communications, fiber-wireless architectures, optical signal processing, and visible light communications.

ZILONG LIU (Senior Member, IEEE) received the bachelor's degree from the School of Electronics and Information Engineering, Huazhong University of Science and Technology, China, in 2004, the master's degree from the Department of Electronic Engineering, Tsinghua University, China, in 2007, and the Ph.D. degree from the School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore, in 2014. He is a Lecturer (Assistant Professor) with the School of Computer Science and Electronics Engineering, University of Essex. From January 2018 to November 2019, he was a Senior Research Fellow with the Institute for Communication Systems, Home of the 5G Innovation Centre, University of Surrey. Prior to his career in the U.K., he spent nine and half years with NTU, first as a Research Associate from July 2008 to October 2014 and then a Research Fellow from November 2014 to December 2017. His Ph.D. thesis "Perfect- and Quasi-Complementary Sequences," focusing on fundamental limits, algebraic constructions, and applications of complementary sequences in wireless communications, has settled a few long-standing open problems in the field. His research lies in the interplay of coding, signal processing, and communications, with a major objective of bridging theory and practice as much as possible. Recently, he has developed an interest in applying machine learning for wireless communications. He is an Associate Editor of IEEE WIRELESS COMMUNICATIONS LETTERS, IEEE ACCESS, and *Frontiers in Communications and Networks*. He is a General Co-Chair of the 10th International Workshop on Signal Design and its Applications in Communications 2022 and a TPC Co-Chair of the 2020 IEEE International Conference on Advanced Networks and Telecommunications Systems. Besides, he was/is a TPC Member of a number of IEEE Conferences/Workshops, such as ICC, WCSP, GLOBECOM, ICCS, SETA, and ISIT. Details of his research can be found at: <https://sites.google.com/site/zilongliu2357>.