Constrained Partial Group Decoding With Max–Min Fairness for Multi-Color Multi-User Visible Light Communication

Guangtao Zheng, Chen Gong, and Zhengyuan Xu

Abstract—Consider a fixed orientation multi-user visible light communication (VLC) system with multi-color light emitting diodes (LEDs) where each transmitter serves a user with limited coordination impeding transmitter-side beamforming. In this paper, a multi-layer coding and constrained partial group decoding (CPGD) method is proposed to tackle strong color interference and increase the system throughput. After channel model formulation, user information rates are allocated and decoding order for all the received data layers is obtained by solving a max-min fairness problem using a greedy algorithm. An achievable rate is derived under the truncated Gaussian input distribution. To reduce the decoding complexity, a map on the decoding order and rate allocation is constructed for all positions of interest on the receiver plane parallel to the transmitter plane, and its size is reduced by a clustering algorithm. Meanwhile, the symmetrical geometry of LED arrays is exploited. Finally, the transmitter-user association problem is formulated and solved by a genetic algorithm. It is observed that the system throughput increases as the receivers are slightly misaligned with corresponding LED arrays due to the reduced interference level, but decreases afterwards due to the weakened link gain.

Index Terms—Visible light communication, multi-layer transmission, constrained partial group decoder, interference cancellation, transmitter-user association.

I. INTRODUCTION

With the increasing amount of data transmitted via wireless links, the spectrum shortage has become a critical problem for the next generation communication systems. Visible light communication (VLC) has emerged as a competent supplement [2], [3] to radio-frequency communications due to its large bandwidth, zero electromagnetic radiation, and simple transceiver structures etc., and has attracted extensive research attentions in recent years [4]–[7].

In a typical indoor environment, a multi-color light emitting diode (LED), which consists of individual color chips, can offer white lighting and support high speed communication [8]. If each chip is used independently to serve one user, multiple users can be supported simultaneously within the coverage of a multi-color LED with little interference. In this paper, an indoor environment with multi-user VLC is considered to support the increasing connectivity requirement of many VLC-enabled user devices. Assume that each color chip is treated as an independent transmitter. Although different user devices within the same LED coverage have the freedom to choose different colors, strong interference still exists from adjacent chips with the same color, and from the leakage through the optical filter equipped on the user device. If not treated properly, e.g. treating interference as noise, it will penalize the system throughput or even cause a communication outage.

To tackle this problem, various existing interference management schemes can be employed, e.g. suppressing the interference by resource allocation [9] or aligning it via transmitter-side signal processing [10], [11]. Among these schemes, multi-layer coding and group decoder [12] provides an advanced interference cancellation framework that can efficiently exploit interference rather than suppressing it. The idea behind the framework is based on the observation that while a receiver is not interested in decoding the messages from interferers, decoding and canceling them are often beneficial for recovering the desired message [13]. To exploit and formalize this idea, at the transmitter, the data stream is split into multiple sub-streams and then encoded into layers individually using independent codebooks. The superposition of those layers is fed into a multi-user channel. At the receiver, multiple user layers are associated with different decoding groups, where those within each group are decoded jointly. Such decoder is called group decoder [12]. The desired signal will be recovered after all its desirable layers are properly decoded. The computational complexity may be intractable when the group size is large. Thus constrained partial group decoder (CPGD) [12], [14], which constrains the group size to be a reasonable number, is proposed.

The optimal group partition and decoding order will be solved under fairness constraints [15], by maximizing the minimum user rate. There are two common fairness con-
straints, namely the symmetric fairness where all users are required to have the same rate guaranteeing the global rate fairness [16], and the max-min fairness which maximizes minimum user throughput. Due to the Han-Kabayashi coding, the users signals need to be divided into multiple layers to increase the user achievable rate. However, since for multiple layer scenario maximizing the minimum user rate is not a submodular optimization problem (convex optimization problem in the discrete variable perspective), in this work, max-min fairness with minimum layer rate is adopted as an approximation to the max-min user rate fairness because of its tractable solutions.

A well-designed CPGD can work very well in traditional wireless settings. In VLC, it is expected that the interference can be properly managed and exploited to ensure a reliable and efficient communication. More importantly, with the assumption of static channel characteristics [17] and the capability of localization in a typical VLC system [18], [19], we make two fundamental changes to the existing interference cancellation framework. First, the rate allocation feedback from receivers to transmitters is relaxed since this information can be precomputed by solving the CPGD problem, thus significant communication overhead during run-time can be avoided, especially for a large number of transmitter and receiver (user) pairs. Second, online solving the CPGD problem, which is usually computationally intensive, is replaced by a less demanding version of look-up table searching and computation. The look-up table, which is named as decoding map, is achieved by solving the CPGD problem of optimal group partition with decoding order and rate allocation at each position in the decoding map.

The major contributions of this paper are as follows:

- For the multi-color multi-user multiple-input multiple-output VLC, we adopt the multi-layer coding and group decoder to perform successive interference cancellation.
- We obtain an achievable rate of a VLC-multiple access channel (VLC-MAC) at each receiver.
- With the assumption of the static nature of VLC channel, and fixing the orientation of the receiver plane parallel to the transmitter plane, we build a decoding map that stores the group partitions with decoding orders and the rate allocations for the transmitters at various typical receiver positions under the max-min fairness. A novel clustering algorithm, which exploits the spatial correlation in the CPGD configurations, is proposed to reduce the map size. A typical receiver location is the center of a cluster. Our key finding is that, under spatial symmetry of the transmitter and receiver arrays, the decoding map is also symmetric, and only 1/8 of the map size is sufficient to represent the decoding order under different locations.
- Finally, with the knowledge of the decoding map and users’ positions, we formulate the transmitter-user association problem. It is observed that receivers at asymmetrical positions will achieve higher sum rate.

The remainder of this paper is organized as follows. The system descriptions are given in Section II, on the channel model, multi-layer coding and CPGD. Section III provides the detailed procedures on the CPGD, and the decoding map construction with size reduction based on the symmetric properties. The problem of transmitter-user association and rate allocation is formulated and solved in Section IV. Numerical results are provided in Section V. We conclude this paper in Section VI.

Notations: We use a bold uppercase letter to represent a matrix, a bold lowercase letter to represent a vector. \(|\cdot|\) represents the size of a set. A backslash between two sets generates a new set which contains the elements of the first set excluding the elements from the second set. The square brackets \([\cdot]\) for a set or vector are used to index their elements.

II. SYSTEM DESCRIPTIONS

A. Channel Model

Consider a multi-color VLC system with \(N_t\) transmitters and \(N_r\) receivers. Each transmitter is an LED chip generating one color, while each receiver is a VLC-enabled user device equipped with a specific optical filter allowing only one color to pass with most power. Each transmitter-receiver pair operates in the same color. The channel gain from the \(i^{th}\) transmitter to the \(j^{th}\) receiver is denoted as \(h_{ji}\) for \(i \in \{1, \ldots, N_t\}\) and \(j \in \{1, \ldots, N_r\}\). Due to the static channel assumption, \(h_{ji}\) remains constant for several symbol blocks and may change to other states independently afterwards. The signal at the \(j^{th}\) receiver at time index \(t \in \{1, \ldots, N\}\) is given by

\[
y_j[t] = \gamma \sum_{i=1}^{N_t} h_{ji} x_i[t] + z_j[t],
\]

where \(\gamma\) is the photodetector responsivity and is assumed to be 1 without loss of generality, \(h_{ji}\) is Lambert channel gain computed by [20]

\[
h_{ji} = \frac{A_r(m_1 + 1)}{2\pi d_{ji}^2} \cos^m(\phi_{ji}) T_s(\psi_{ji}) g(\psi_{ji}) \cos(\psi_{ji}),
\]

where \(0 \leq \psi_{ji} \leq \psi_c\) and \(h_{ji} = 0\) otherwise; \(m_1\) is the Lambert emission order; \(A_r\) is the area of PD; \(\psi_c\) is the field of view (FOV) at the receiver. Given location \(u^i = [u_{i1}, u_{i2}, u_{i3}]\) of transmitter \(i\) and the location \(v^j = [v_{j1}, v_{j2}, v_{j3}]\) of receiver \(j\), the distance \(d_{ji}\) between transmitter \(i\) and receiver \(j\), the radiance angle \(\phi_{ji}\) at transmitter \(i\), and the incidence angle \(\psi_{ji}\) at receiver \(j\) can be computed accordingly.

Moreover, the concentrator gain \(g(\psi_{ji})\) is computed as

\[
g(\psi_{ji}) = \begin{cases} \frac{n^2}{\sin^2 \psi_c}, & 0 \leq \psi_{ji} \leq \psi_c; \\
0, & \psi_{ji} > \psi_c; \end{cases}
\]

where \(n\) is the refractive index. The additive noise term \(z_j[t] \sim \mathcal{N}(0, \sigma_z^2)\) in Eq. (1) satisfies the Gaussian distribution with zero mean and variance \(\sigma_z^2\).

In a practical indoor environment, the average power budget constraint \(\varepsilon_i\) needs to be selected based on different illumination constraints. It is seen that for the optimal distribution \(\mathbb{E}[x_i] = \varepsilon_i\) for \(\varepsilon_i \leq A_i/2\) and \(\mathbb{E}[x_i] = A_i/2\) for \(\varepsilon_i \geq A_i/2\).
In a multi-color setting, the mutual color interference is reflected by the channel gain $h_{ij}$, which is further determined by the optical filter gain denoting as $T_s(\psi_{ji})$. Note that a certain color LED chip can serve as a transmitter. Its value depends on the color used by transmitter $i$ and the optical filter used by receiver $j$. If transmitter $i$ and receiver $j$ forms a pair, meaning the color matches with the optical filter, $T_s(\psi_{ji})$ may be close to 1. Otherwise, $T_s(\psi_{ji})$ will be small, i.e., small portion of power acting as inter-color interference will be received by receiver $j$. A proper modeling of the color spectrum, e.g., the method in [21], will be helpful in computing $T_s(\psi_{ji})$ and $h_{ij}$, as detailed in Section V.

**B. Input Signals and Layered Encoding**

Multi-layer coding and CPGD are adopted in the system to decode and cancel the interference successively. At each transmitter, the data stream is split into multiple substreams and then encoded by independent codebooks forming layers. The number of layers allocated to transmitter $i$ is denoted as $L_i$, and thus the total number of layers is $L = \sum_{i=1}^{N_t} L_i$. The set of codebooks at transmitter $i$ is denoted as $C_i \triangleq \{C_{i1}, \ldots, C_{iL_i}\}$ for $i \in {1, \ldots, N_t}$. Denoting the coded symbol drawn from layer $k$ at time index $t$ by $x_{ik}[t]$, the received signal in Eq. (1) can be represented by the superposition of multiple layers as follows,

$$y_j[t] = \sum_{i=1}^{N_t} h_{ji} \sum_{k=1}^{L_i} x_{ik}[t] + z_j[t].$$

The rate $R_i$ at transmitter $i$ can be expressed as $R_i = \sum_{k=1}^{L_i} R_{ik}$, where $R_{ik}$ is the rate for layer $k$ encoded by $C_{ik}$. In the following analysis, we omit time index $t$ in Eq. (5) for simplicity. Moreover, we use $(i, k)$ to denote layer $k$ at transmitter $i$, for $1 \leq i \leq N_t$ and $1 \leq k \leq L_i$, and define $\mathcal{K}$ as the ensemble of all layers, i.e., $\mathcal{K} \triangleq \{(i, k)\} | 1 \leq i \leq N_t, 1 \leq k \leq L_i$.

Assuming that the power is equally allocated among the layers at each transmitter and that the layers are independent of each other, we have the following constraints:

$$x_{ik} \geq 0, x_{ik} \leq A_i/L_i, \mathbb{E}[x_{ik}] = \varepsilon_{ik}, \forall i \in {1, \ldots, N_t},$$

where for the optimal distribution $\mathbb{E}[x_{ik}] = \varepsilon_{ik}$ for $\varepsilon_{ik} \leq A_i/(2L_i)$ and $\mathbb{E}[x_{ik}] = A_i/(2L_i)$ for $\varepsilon_{ik} \geq A_i/(2L_i)$. As a truncated Gaussian (TG) distribution [22] can be very tight to the capacity at high a signal-to-noise-ratio (SNR) which is typical for indoor VLC, we adopt this distribution to approximate the capacity. The TG distribution is determined by a three-tuple parameter $(\mu, \nu, A)$ as follows,

$$\theta_{\mu, \nu, A}(x) = \begin{cases} \rho g_{\mu, \nu}(x), & x \in [0, A]; \\ 0, & \text{otherwise}; \end{cases}$$

where $\rho = (G_{\mu, \nu}(A) - G_{\mu, \nu}(0))^{-1}$; $g_{\mu, \nu}(x)$ is Gaussian distribution with mean $\mu$ and variance $\nu^2$; $G_{\mu, \nu}(x)$ is the corresponding cumulative Gaussian distribution. Based on the constraint between the mean and peak power in Eq. (6), for a $(\mu, \nu, A)$-TG distribution, the mean $\mu$ and variance $\nu^2$ are given by

$$\hat{\mu} = \nu^2[\theta_{\mu, \nu}(0) - \theta_{\mu, \nu}(A)] + \mu,$$

$$\nu^2 = \nu^2[1 - A\theta_{\mu, \nu}(A) - \hat{\mu}(\theta_{\mu, \nu}(0) - \theta_{\mu, \nu}(A))].$$

For each layer, a $(\mu_{ik}, \nu_{ik}, A_{ik})$-TG distribution is adopted for the input signal $x_{ik}$, where the TG parameters are delicately selected such that $\hat{\mu}_{ik}$ satisfies Eq. (8).

In real communications, transmitter-side cooperation can increase the system achievable rate if all LEDs are connected to a central controller. However, if such central coordination is not available, for example in a workshop where the LEDs belong to different types of machines, the interference channel with independent data streams is assumed. We assume weak coordination where the transmitter-receiver association can be established before the data transmission. However, strong coordination during the transmission, for example beamforming, is not supported.

**C. Constrained Partial Group Decoder (CPGD)**

Due to the multi-user nature of the system, each receiver decodes not only the layers sent from its associated transmitter, but also other layers if the decoding can improve the achievable rate of layers from its associated transmitter. We use ordered sequence $q_l^i = [Q_{l1}^i, \ldots, Q_{lp}^i]$ to represent the group partition of layers and the corresponding decoding order at receiver $j$, where $Q_{l1}^i \subseteq \mathcal{K}$ denotes the set of layers decoded in the order $l = 1, 2, \ldots, p$. There are $p_j$ groups and $m_j$ decoding stages with $m_j \leq p_j$. For any $b$, $1 \leq b \leq L_i$, we have $(i, b) \in Q_{m_j}^i$ and $(i, b) \notin \bigcup_{k=m_j+1}^{p_j} Q_k^i$, i.e., all layers from transmitter $i$ have been decoded before stage $m_j + 1$. At the $i^{th}$ decoding stage, the CPGD employed at receiver $j$ jointly decodes layer $Q_{l1}^i$ via maximum likelihood (ML), while treating the remaining layers $\bigcup_{\ell=1}^{l-1} Q_{\ell}^i$ as additive noise, until all desired layers from the associated transmitter are decoded. To control the decoding complexity, the size of each group should follow $|Q_{\ell}^i| \leq \tau$ for $\ell = 1, \ldots, p_j$, hence the name constrained partial group decoder. Formally, the CPGD decodes the desired messages from the received signal $y_j$ as follows [12]

1. Initialize $m = 1$.
2. Compute $\Sigma_m^i = \sigma_j^2 + \sum_{b=m+1}^{p_j} \sum_{(i,k) \in Q_m^i} h_{ji}^2 \rho_{ik}^2$ as the noise variance and jointly decode the layers in $Q_m^i$ via ML decoding.
3. Update $y_j \leftarrow y_j - \sum_{(i,k) \in Q_m^i} h_{ji} \hat{x}_{ik}$, where $\hat{x}_{ik}$ is the decoded symbol of $x_{ik}$.
4. If $m = m_j$, stop and output the information of desired layers. Otherwise, update $m \leftarrow m + 1$ and go to step 2.

The optimal group partitioning and decoding order for each user depends on the objective function to be optimized. In this paper, the max-min objective function among the rates of all user-associated transmitters is adopted. Usually this optimization problem is solved in a distributed manner, where each receiver $j$ obtains the optimal ordered group partition $q_j^i$, also called decoding order in this paper, and the locally maximal rates $r_j^i$ achievable for all transmission layers. Then,
the transmitters need feedbacks from the receivers to obtain the global rate allocation via taking the minimum of all locally obtained rates for each transmission layer. A decoding outage may be declared if some layer’s rate exceeds its corresponding channel capacity. More details on CPGD can be found in [12] and references therein.

### III. Decoding Map and Size Reduction

Assuming accurate knowledge on the receiver’s position, we can pre-compute $q^j$ and $r^j$ at each receiver position of interest. Such results can be obtained by optimizing the maximum rate objective function to achieve a tradeoff between throughput and fairness, as compared to the symmetric fairness. In this section, we first calculate the achievable rate of a VLC-MAC, which originates the scenario where multiple layers are sent from the same transmitter to a receiver. Then, we provide an algorithm to solve the objective function, based on which the concept of decoding map is introduced. To effectively represent the decoding map and show the symmetric properties, a size reduction algorithm is proposed.

#### A. Achievable Rate Calculation

The link from transmitter $i$ to the associated receiver $j$ contains multiple layers which is equivalent to a MAC. Given $\mathcal{V} \subseteq \mathcal{K}_j$ and $\mathcal{G} \subseteq \mathcal{K}_j$ with $\mathcal{V} \cap \mathcal{G} = \emptyset$ and $\mathcal{V} \cup \mathcal{G} = \mathcal{K}_j$, where the former represents the signal set and the latter is treated as interference, Eq. (5) can be rewritten as follows

$$y_j = \sum_{(i,k) \in \mathcal{V}} h_{ji}x_{ik} + \sum_{(i,k) \in \mathcal{G}} h_{ji}x_{ik} + z_j.$$

Due to the signal’s non-negativity with peak and average power constraints, the closed form of a VLC channel capacity is still open [22]. However, a closed-form achievable rate approximation that approaches the capacity can be obtained using the truncated Gaussian distribution given in Section II-B. Note that the achievable rate of MAC for VLC is proposed in [23]. In this work, we investigate the achievable rate of interference channels, where the interferences are decoded only if the decoding is helpful to decoding the desired signal. Define the sum achievable rate for signals $X_V = \{X_{ik}|(i,k) \in \mathcal{V}\}$ treating those in $\mathcal{G}$ as noise at the receiver $j$ as $\mathcal{R}_j(\mathcal{V}, \mathcal{G}) = I(X_V; Y_j)$. Then

$$\mathcal{R}_j(\mathcal{V}, \mathcal{G}) = h(X_V) - h(X_V|Y_j) = \sum_{(i,k) \in \mathcal{V}} h(X_{ik}) - h(X_V|Y_j).$$

Moreover, letting $\mathcal{V}[m]$ denote the $m^{th}$ elements in set $\mathcal{V}$, we expand $h(X_V|Y_j)$ with chain rule [24] to get

$$h(X_V|Y_j) = \sum_{m=1}^p h(X_{\mathcal{V}[m]}|Y_j, X_{\mathcal{V}[1]}, \ldots, X_{\mathcal{V}[m-1]}),$$

where $p = |\mathcal{V}|$. We then invoke the following lemma [25].

**Lemma 1:** Let $X_1, X_2, \ldots, X_k$ be an arbitrary set of random variables with mean $\mu$ and covariance matrix $\mathbf{K}$. Let $\mathcal{S}$ be any subset of $\{1, 2, \ldots, k\}$ and $\mathcal{S}'$ be its complement. Then

$$h(X_S|X_{\mathcal{S}'} \leq h(X_S^2|X_{\mathcal{S}'})$$

where $(X_1^*, X_2^*, \ldots, X_k^*) \sim \mathcal{N}((\mu, \mathbf{K}))$. $\square$

According to Lemma 1, $h(X_V|Y_j)$ can be upper bounded by $h(X_V^*|Y_j^*)$, where $(X_1^*, X_2^*, \ldots, X_k^*)$ is jointly Gaussian with the same covariance matrix as that of $(X_V, Y_j)$. Then, we have

$$h(X_V|Y_j) \leq \sum_{m=1}^p h(X_{\mathcal{V}[m]}|Y_j^*, X_{\mathcal{V}[1]}^*, \ldots, X_{\mathcal{V}[m-1]}^*),$$

where $(X_{\mathcal{V}[m]}^*, X_{\mathcal{V}[1]}^*, \ldots, X_{\mathcal{V}[m-1]}^*)$ are jointly Gaussian with covariance matrix $\mathbf{K}_m$.

We calculate covariance matrix $\mathbf{K}_m$ under the group size constraint $p \leq \tau$, according to CPGD. We write $X_V$ in a sequence $X_{\mathcal{V}[1]} \ldots, X_{\mathcal{V}[p]}$ with the corresponding variance $\nu^2_{\mathcal{V}[1]}, \ldots, \nu^2_{\mathcal{V}[p]}$, for $p \leq \tau$. Similarly, we write $X_G$ in a sequence $X_{\mathcal{G}[1]} \ldots, X_{\mathcal{G}[q]}$, where $\mathcal{G}[m], \forall m = 1, \ldots, q$ is an element in $\mathcal{G}$, with the corresponding variance $\nu^2_{\mathcal{G}[1]}, \ldots, \nu^2_{\mathcal{G}[p]}$. Since $X_{\mathcal{V}[m]}^*, X_{\mathcal{V}[1]}^*, \ldots, X_{\mathcal{V}[m-1]}^*$ are assumed mutually independent, the $m$-by-$m$ submatrix in the upper left corner of $\mathbf{K}_m$ is a diagonal matrix with diagonal elements being the variances of the corresponding variables. Knowing that $X_{\mathcal{V}[m]}^*, X_{\mathcal{V}[1]}^*, \ldots, X_{\mathcal{V}[m-1]}^*$ each follows a truncated Gaussian distribution, $m$ by $m$ diagonal matrix $\mathbf{K}_m$ is located in the upper left corner of $\mathbf{K}_m$. Next, we compute the $(m+1)^{th}$ column (the $(m+1)^{th}$ row can be directly obtained from the $(m+1)^{th}$ column by symmetry). For the element in the $k^{th}$ row $(m+1)^{th}$ column, since $Y_j = \sum_{m=1}^{|\mathcal{V}|} h_{jV_{[m]}}X_{\mathcal{V}[m]} + \sum_{m=1}^{|\mathcal{G}|} h_{jG_{[m]}}X_{\mathcal{G}[m]} + Z_j$, the covariance can be computed by $\mathbb{E}[(X_{\mathcal{V}[k]} - \mathbb{E}[X_{\mathcal{V}[k]])](Y_j - \mathbb{E}[Y_j])] = \mathbb{E}[h_{jV_{[k]}}^2\nu^2_{\mathcal{V}[k]}]$. Similarly, for the element in the $(m+1)^{th}$ row and $(m+1)^{th}$ column, we have $\mathbb{E}[(Y_j - \mathbb{E}[Y_j])(Y_j - \mathbb{E}[Y_j])] = \sum_{\ell=1}^p h_{jV_{[\ell]}}^2\nu^2_{\mathcal{V}[\ell]} + \sum_{\ell=1}^q h_{jG_{[\ell]}}^2\nu^2_{\mathcal{G}[\ell]} + \sigma^2$. Hence $\mathbf{K}_m$ can be written in Eq. (15), as shown at the bottom of this page. According to Eq. (40) in [25], $h(X_{\mathcal{V}[m]}^*|Y_j^*, X_{\mathcal{V}[1]}^*, \ldots, X_{\mathcal{V}[m-1]}^*)$ can be obtained by

$$h(X_{\mathcal{V}[m]}^*|Y_j^*, X_{\mathcal{V}[1]}^*, \ldots, X_{\mathcal{V}[m-1]}^*) = \frac{1}{2} \log 2\pi e (\nu^2_{\mathcal{V}[m]} - \mathbf{b}\mathbf{s}^{-1}\mathbf{b}^T),$$

where

$$\mathbf{K}_m = \begin{bmatrix}
\hat{\nu}^2_{V[1]} & 0 & \ldots & 0 \\
0 & \hat{\nu}^2_{V[2]} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \hat{\nu}^2_{V[m-1]}
\end{bmatrix}.$$
where \( \mathbf{b} = [0, \ldots, 0, h_{jV[m]} \hat{v}_{i}^{2}] \), \( h_{jV[n]} \) equals \( h_{ji} \) with \( i \) being determined by \( V[n] \); \( \mathbf{S} \) is the matrix \( \mathbf{K}_{m} \) with its first row and column being deleted. By exploiting the sparsity of the matrix \( \mathbf{K}_{m} \), we can explicitly express each conditional entropy at the right hand side of Eq. (14) as

\[
\begin{align*}
    h(X_{V[m]}^{*} | Y_{V}^{*}, X_{V[1]}^{*}, & \ldots, X_{V[m-1]}^{*}) = \\
    \frac{1}{2} \log 2\pi e \left( \hat{v}_{i}^{2} \right) - \\
    - \sum_{m=1}^{p} \frac{1}{2} \log \left( \hat{v}_{i}^{2} \right) \left( h_{jV[m]} \hat{v}_{i}^{2} \right) + \\
    + \sum_{\ell=1}^{m} h_{jV[m]}^{2} (\hat{v}_{i}^{2} + \sigma_{j}^{2}) \right),
\end{align*}
\]

and thus \( h(X_{V}|Y_{j}) \) can be upper bounded by

\[
\begin{align*}
    h(X_{V}|Y_{j}) & \leq \frac{p}{2} \log 2\pi e + \sum_{m=1}^{p} \frac{1}{2} \log \left( \hat{v}_{i}^{2} \right) \left( h_{jV[m]} \hat{v}_{i}^{2} \right) + \\
    & \quad - \sum_{\ell=1}^{m} h_{jV[m]}^{2} (\hat{v}_{i}^{2} + \sigma_{j}^{2}) \right),
\end{align*}
\]

Therefore, \( \mathcal{R}_{j}(V, G) \) can be computed as follows,

\[
\mathcal{R}_{j}(V, G) = - \sum_{m=1}^{p} \phi_{V[m]} + \frac{1}{2} \sum_{m=1}^{p} \left[ \log \hat{v}_{i}^{2} \left( h_{jV[m]} \hat{v}_{i}^{2} \right) + \\
    - \sum_{\ell=1}^{m} h_{jV[m]}^{2} (\hat{v}_{i}^{2} + \sigma_{j}^{2}) \right),
\]

where \( \phi_{V[m]} = \phi_{ik} \) if \( V[m] = (i, k) \); and following similar steps in [23], we have \( \phi_{ik} = \log(\rho_{ik} + \frac{1}{2} ((A_{ik} - \mu_{ik})^{2} + \mu_{ik}^{2} + \nu_{ik} \nu_{ik}(0))) \).

### B. Problem Statement

Assuming that at position \( V^{j} = [\mathbf{v}^{j}_{1}, \mathbf{v}^{j}_{2}, \mathbf{v}^{j}_{3}] \), the decoding order is denoted as \( \mathbf{q}^{j} = [\mathbf{Q}^{j}_{1}, \ldots, \mathbf{Q}^{j}_{m}] \). The optimal \( \mathbf{q}^{j} \) is obtained via solving the following optimization problem,

\[
\hat{\mathbf{q}}^{j} = \arg\max_{\mathbf{q}^{j} \in \mathcal{Q}} \psi(\mathbf{q}^{j}),
\]

where \( \mathcal{Q} \) is the set of all legitimate decoding orders and \( \psi(\mathbf{q}^{j}) \) is the max-min rate of all layers given by

\[
\psi(\mathbf{q}^{j}) = \max_{\ell \in [K]} \min_{\ell \leq m} \mathbf{r}^{j}[\ell],
\]

where the function \( \mathbb{E}(\mathbf{Q}^{j}_{m}) \) transforms each \( (i, b) \in \mathbf{Q}^{j}_{m} \) to a linear index \( \ell \) in the rate vector \( \mathbf{r}^{j} \) with relation \( \ell = \sum_{b=1}^{m} L_{a} + b \); \( \mathbf{r}^{j} \mathbb{E}(\mathbf{Q}^{j}_{m}) \) can be conceived as a point within the capacity region specified by \( \mathbb{C}(\mathbf{Q}^{j}_{m}, \mathbb{U}^{j}_{m+1} \mathbf{Q}^{j}_{m}) \) with \( \mathbf{Q}^{j}_{m} \) being the signal set, \( \mathbb{U}^{j}_{m+1} \mathbf{Q}^{j}_{m} \) the noise; \( p_{j} \) is the length of vector \( \mathbf{q}^{j} \). Note that the max-min fairness on the user rate is of more interest than that on the layer rate. However, in case of multiple layers for one user, it is intractable to solve the max-min fairness problem on user rate, since it is not a submodular optimization problem (convex optimization problem in the discrete variable perspective). Thus, we resort to a suboptimal solution on the max-min fairness on the layer rate as a tractable solution. We will show in Section V that for layered signals adopted at the transmitters, the minimum rate will also increase over single layer optimization.

By solving problems (20) and (21), we can obtain the optimal decoding order \( \mathbf{q}^{j} \) and the corresponding local rate allocation \( \mathbf{r}^{j} \) for all the transmission layers. For those layers not in \( K^{j} \), their corresponding rates are usually set to infinity such that layers are not needed to decode.

It would be helpful to invoke the rate increment margin defined as follows [15],

\[
\Delta_{j}(V, G, r) \triangleq \min_{\mathbf{r} \in \mathbb{R}^{r}} \frac{\mathcal{R}_{j}(U, G) - \|\mathbf{r}\|_{1}}{|U|},
\]

where \( \mathbb{R}^{r} \) contains the rates selected by \( U \) from any decodable rate vector \( r \) satisfying the constraint in Eq. (21); \( V \) acts as the signal layer set and \( G \) is treated as the noise layer set with \( V, G \subseteq K^{j} \) and \( V \cap G = \emptyset \); \( \|\cdot\|_{1} \) denotes the norm-1 operation. The definition means at position \( \mathbf{q}^{j} \), if the rate of layers in \( V \) increases beyond \( \Delta_{j}(V, G, r) \), then decoding \( V \) while treating \( G \) as noise may cause an outage. Thus given \( \mathbf{q}^{j} \), for \( 1 \leq m \leq p_{j} \), the maximum achievable rate allocated to \( \mathbf{Q}^{j}_{m} \) is \( \Delta_{j}(\mathbf{Q}^{j}_{m}, \mathbb{U}^{j}_{m+1} \mathbf{Q}^{j}_{m}, \mathbf{0}) \).

Problem (20) can be optimally solved by a greedy method under Gaussian signaling [26], where the achievable rate function \( \mathcal{R}_{j}(V, G) \) exhibits submodularity and transitivity which play the key role in guaranteeing the optimality, shown as follows [26],

\[
\begin{align*}
\mathcal{R}_{j}(D_{1}, G) + \mathcal{R}_{j}(D_{2}, G) & \geq \mathcal{R}_{j}(D_{1} \cup D_{2}, G) \quad + \mathcal{R}_{j}(D_{1} \cup D_{2}, G),
\end{align*}
\]

for disjoint sets \( U, V, G \). Although due to the non-Gaussian input distributions adopted in this paper, the optimality of original decoding order may not be guaranteed, we still adopt it as an approximate solution with low computational complexity. We will perform successive decoding and rate allocation as those in the Gaussian case. The procedure is listed in Algorithm 1 based on [12], where the user positions are explicitly incorporated into the algorithm.

**Algorithm 1** Compute the Decoding Order at \( \mathbf{v}^{j} = (\mathbf{v}_{1}^{j}, \mathbf{v}_{2}^{j}, \mathbf{v}_{3}^{j}) \).

1. Given \( D = K^{j} \), \( G = \emptyset \), \( \ell = 1 \), \( p_{j} = 0 \), set \( \mathbf{r}^{j} \) to be an all-zero vector.
2. repeat
3. \( \delta_{\ell} = \min_{\mathbf{r} \neq 0, \mathbf{r} \subseteq D, |V| \leq \tau} \mathcal{R}_{j}(V, G) / |V| \);
4. \( \mathbf{G}_{\ell}^{j} = \arg\max_{\mathbf{r} \neq 0, \mathbf{r} \subseteq D, |V| \leq \tau} \mathcal{R}_{j}(V, G) / |V| \);
5. \( D \leftarrow D \setminus \mathbf{G}_{\ell}^{j} \) and \( G \leftarrow G \cup \mathbf{G}_{\ell}^{j} \);
6. \( \mathbf{r} \mathbb{E}(\mathbf{Q}_{m}) = \delta_{\ell}, p_{j} = p_{j} + 1 \);
7. \( \ell = \ell + 1 \);
8. until \( D = \emptyset \);
9. \( \mathbf{q}^{j}[m] \leftarrow \mathbf{G}_{\ell-m}^{j}, 1 \leq m \leq p_{j} \);
10. Output \( \hat{\mathbf{q}}^{j}, \hat{\mathbf{r}}^{j} = \mathbf{r}^{j} \).
The input is the set of layers $K_j$ received at $j^{th}$ user position $v^j$. Algorithm 1 runs in a greedy manner such that in each step, it finds the largest rate increment margin $\delta_l$ (set $r = 0$ in Eq. (22)) and the corresponding layer set $G_l^j$. The algorithm keeps executing this step until all the layers are grouped and the corresponding rates are assigned. By doing this, the algorithm will obtain a solution to problem (20) with low computational complexity.

C. Decoding Map

As the decoding order is discrete, we can quantize the user location in the feasible space into various regions, and presume that the decoding order in the center of each region may well represent that for users in that region if the region size is sufficiently small. Based on this, we introduce the concept of decoding map, which is a look-up table saving $q^j$ and $\hat{r}^j$ at each position $v^j$ obtained using Algorithm 1. Given space length $d_l$, width $d_w$, and sampling gap $d_s$, the computational complexity is on the order of $O(L^2 + \frac{d_s}{d_l})$ for $\tau$ not large.

Once knowing the users’ positions, the decoding map provides prior knowledge about the decoding orders and rate allocations to configure the transmitter and receiver, such that the online computation complexity and communication overhead can be significantly reduced.

Note that for two receivers $j_1$ and $j_2$ operating with different colors, the resulting $q^{j_1}$ and $q^{j_2}$ are different, the same as $r^{j_1}$ and $r^{j_2}$. Thus for each color, we maintain a decoding map by sampling the positions uniformly among the possible positions of users. At each position, we run Algorithm 1 and store the resultant $\hat{q}$ and $\hat{r}$ with superscript $j$ being omitted since at the same position using the same color, the results are the same for different receivers.

D. Size Reduction for the Decoding Map

The size of a decoding map is typically large especially for a large number of transmitters and small distance between the sampling points in the decoding map. This will pose challenges to storage and searching. In this work, we aim to reduce the decoding map size by clustering the decoding orders at different positions into clusters, and use one decoding order $q^*$ in each cluster to represent those at all points in the cluster. We concede that optimal orders are not necessarily the same as $q^*$. However, if using the decoding order at one position to represent those at other positions will not significantly degrade the rate from the optimal decoding order, the two positions can be clustered into the same decoding order. This is the main motivation and idea of using the decoding order in one position to represent those in other positions. Hence, the decoding map size reduction falls into finding a clustering algorithm on the decoding orders at different positions within an acceptable rate loss.

To achieve this, we first define a normalized distance between two user positions $\mathbf{v}^{j_1}$ and $\mathbf{v}^{j_2}$,

$$d(\mathbf{v}^{j_1}, \mathbf{v}^{j_2}) \triangleq \| \mathbf{v}^{j_2} - \mathbf{v}^{j_1}(q^{j_1}) \|_2 / \| \mathbf{v}^{j_2} \|_2,$$

which characterizes the normalized variation between the rate allocation vector at position $\mathbf{v}^{j_2}$ and rate allocation vector $\mathbf{r}^{j_2}(q^{j_1})$ at position $\mathbf{v}^{j_2}$ using decoding order $q^{j_1}$ at $\mathbf{v}^{j_1}$. Rate vector $r^{j_2}(q^{j_1})$ can be obtained by successively setting $\mathbf{v} = \mathbf{Q}^j m$, $\mathcal{G} = \cup_{m=1}^{p_{j_2}} \mathcal{Q}^j m$, for $m = 1, \ldots, p_{j_2}$, and computing the rate increment using Eq. (22) with $r = 0$ under decoding order $q^{j_1} = [\mathcal{Q}_1^j, \ldots, \mathcal{Q}_{p_{j_2}}^j]$. The size reduction algorithm is given in Algorithm 2, which reduces the decoding map size of each color.

In Algorithm 2, $\tau_{\text{diff}}$ is the threshold that determines whether two positions are in the same category, while $\tau_{\text{loss}}$ is the maximal tolerable loss and greater than $\tau_{\text{diff}}$ in general. We label each position sequentially using integers as the indices. Let $\mathcal{N}_k$ denote the index set for the $k^{th}$ category, and $\mathcal{N}_K$ denote the vector of all index sets. The algorithm starts with category number $K = 1$, which implies that $\mathcal{N}_K$ contains all points. In each category, we compute the average distance of each point to all the other points. If the maximum average distance is greater than the preset threshold $\tau_{\text{loss}}$, we need to further divide this category into smaller ones. The partition proceeds in such a way that the points with the difference smaller than the given threshold $\tau_{\text{diff}}$ are put into a new category. On the other hand, if the maximal average distance is lower than $\tau_{\text{loss}}$, no further partition is needed and the algorithm proceeds to evaluate the next category. The algorithm stops when in all categories no further partition is needed or the category number $K$ does not change.

The algorithm will terminate in finite steps and is effective in reducing the size of a decoding map and producing clear boundaries for all categories. However, as the number of LEDs increases in a fixed area, it becomes more difficult to reduce the size as adjacent transmitter positions bear more variations.

Algorithm 2 Size reduction of a decoding map

1: Set $K=1$, $\ell = 1$, initialize $\tau_{\text{diff}}$, $\tau_{\text{loss}}$, $\mathcal{N}_K = [\mathcal{N}_k], k = 1, \ldots, K$;
2: repeat
3: for $k = 1 : K$ do
4: $\mathcal{N}_k = \mathcal{N}_K[k]$;
5: $d_k[i] = \frac{1}{|\mathcal{N}_k|} \sum_{j \in \mathcal{N}_k} d(u^i, v^j), \forall i \in \mathcal{N}_k$;
6: if $\max(d_k) > \tau_{\text{loss}}$ then
7: repeat
8: $\xi = d_k[1]$;
9: $\mathcal{N}_l = \{ i : |\xi - d_k[i]| < \tau_{\text{diff}}, i \in \mathcal{N}_k \}$;
10: $\mathcal{N}_k = \mathcal{N}_k \setminus \mathcal{N}_l$;
11: $d_k = d_k[\mathcal{N}_k]$;
12: $\ell = \ell + 1$;
13: until $\mathcal{N}_k = \emptyset$;
14: else
15: $\mathcal{N}_l = \mathcal{N}_k$;
16: $\ell = \ell + 1$;
17: end if
18: end for
19: $K = \ell - 1$, $\mathcal{N}_K = [\mathcal{N}_k], k = 1, \ldots, K$;
20: until $K$ does not change
21: Output $K$ and $\mathcal{N}_K$. 
E. Symmetry Constraint and Size Reduction

As we observe from Eq. (19), the decoding order is determined by the link gain $h_{ji}$, if the noise variances at all positions are the same. Moreover, if the link gain vector $h_{ji}$ is the permutation of another, the decoding order can be obtained via certain permutation and we don’t need to run Algorithm 1 again. Thus, the computational complexity to obtain the decoding map can be further reduced.

If the transmitters are uniformly spaced and arranged in a regular shape, e.g., a square which is typical for indoor lighting, the channel gains exhibit certain symmetry properties with respect to the transmitters. By exploiting these symmetry properties, we can compute the decoding order directly at other rotated or reflected positions based on the results of the current position.

More specifically, given the square geometry and $N_h \times N_v$ index matrix $A$, we use $A[r, c] = \pi_A(r, c)$ to denote the transmitter index at the $r^{th}$ row and $c^{th}$ column of the LED array, where $\pi_A(r, c)$ is the function that maps position $(r, c)$ into an integer index. Conversely, $(r, c) = \pi_A^{-1}(A[r, c])$ computes the position given the transmitter index. If we know the layer $(i, k) \in Q_{th}$ at the $m^{th}$ decoding stage, the corresponding layer at the symmetrical position is $(A_p,i,k)$, where $A_p$ is the matrix generated by $A$ in the following three ways: (1) $A_p(r, c) = A(N_h + 1 - c, N_v + 1 - r)$ $(N_h = N_v$ for a square geometry), if two points are symmetrical with respect to (w.r.t) the diagonal; (2) $A_p(N_h - r + 1, c) = A(r, c)$ if two points are symmetrical w.r.t. the horizontal center line; (3) $A_p(r, N_h - c + 1) = A(r, c)$ if two points are symmetrical w.r.t. the vertical center line.

Figure 1 gives an example. For simplicity, assuming single color LED with single layer at each position, then the index matrix is given by $A = [2, 4, 1, 3]$. At point $v^1$, we have $q^1 = [(4,1), (2,1), (3,1), (1,1)]$. Then at point $v^2$, which is symmetrical to $v^1$ w.r.t. $y = x$, we first obtain $A_p = [3, 4, 1, 2, 2, 3, 1]$, apply the transformation $(A_p, \pi_A^{-1}(i), k)$, and get the decoding order $q^2 = [(4,1), (3,1), (2,1), (1,1)]$. Similarly, at point $v^4$ which is symmetrical w.r.t. $y = 0$, we obtain $q^4 = [(2,1), (4,1), (1,1), (3,1)]$ with $A_p = [4, 2, 3, 1]$. Given $q^x$, we obtain at point $v^x$, which is symmetrical w.r.t. $x = 0$, the decoding order $q^3 = [(3,1), (4,1), (1,1), (2,1)]$ with $A_p = [1, 3, 2, 4]$. By doing the above computation iteratively, we only need to compute $\frac{1}{8}$ of the original space to get the entire decoding map.

We concede that the map needs to be recomputed if the channel link gains change, which happens in case of people/object moving close to the transceiver. Since the first-order reflection component of the channel gain is significantly weaker than the line-of-sight component, it can be expected that the link gains do not change significantly.

IV. Transmitter-User Association

With Rate Allocation

Note that in Algorithm 1, no explicit transmitter-user associations are assumed, but it is important to find the good associations to give the final system configurations, since different user associations may give different global rate allocation schemes, some of which will have higher sum rates than others. To further enhance the transmission rate, the transmitter-user association should be optimized.

We aim to maximize the sum rate of the system. Given the decoding order $q_j$ of the $j^{th}$ user, which can be obtained from the decoding map, and the corresponding local rate allocation $v_j$, we can formulate the optimal transmitter-user association and rate allocation problem as follows,

$$\begin{align*}
\text{max} & \quad \sum_{i=1}^{N_t} \sum_{j=1}^{N_r} L_t \left[ \eta_{ij} \left( \sum_{b=1}^{i-1} L_h + k \right) \right], \\
\text{s.t.} & \quad \eta_{ij} \in \{0, 1\}, \quad \forall i \in \{1, \ldots, N_t\}; \\
& \quad \sum_{i=1}^{N_r} \eta_{ij} = 1, \quad \forall i \in \{1, \ldots, N_t\}; \\
& \quad \sum_{j=1}^{N_r} \eta_{ij} = 1, \quad \forall j \in \{1, \ldots, N_r\}; \\
& \quad m_j = \varphi(\eta_j, q^j), \quad \forall j \in \{1, \ldots, N_r\}; \\
& \quad iv \left( \bigcup_{m=m_j+1} Q^j \right) = +\infty, \\
& \quad \forall j \in \{1, \ldots, N_r\}; \\
& \quad \bar{r}[\ell] = \min_j r^j[\ell], \quad \ell = 1, \ldots, \sum_{i=1}^{N_r} L_t; \\
& \quad \bar{r}[0] = 0.
\end{align*}$$

In the above problem formulation, Eq. (28) means that a transmitter serves only one user; Eq. (29) means a user decodes the signal from only one transmitter. The function $m_j = \varphi(\eta_j, q^j)$ in Eq. (30) computes the subset’s index given $\eta_j$, which is the $j^{th}$ row of the matrix $\eta = [\eta_{ij}]$, such that $Q^j_{m_j}$ contains at least one signal layer and the subsets after $m_j$ contain no signal layers. Formally, $m_j$ satisfies the following conditions

$$\begin{align*}
i^* & = \arg \max_i \eta_{ij} = 1; \\
\exists b = 1, \ldots, L_{i^*}, (i^*, b) & \in Q^j_{m_j}; \\
\forall b = 1, \ldots, L_{i^*}, (i^*, b) & \notin \bigcup_{m=m_j+1} Q^j_m.
\end{align*}$$

The layers $\bigcup_{m=m_j+1} Q^j_m$ are discarded since the desired layers have been recovered. Furthermore, Eq. (31) sets the rates of
Algorithm 3 Iterative Rate Update for $N_r$ Users

1: Given channel-user association $\eta$, initial rate allocation $\mathbf{r}$ and $N_r$ users at $\mathbf{v}^1, \ldots, \mathbf{v}^{N_r}$, set $\hat{\mathbf{r}} = \mathbf{r}$;
2: repeat
3:   for $j = 1$ to $N_r$ do
4:     $\mathcal{D} = \mathcal{K}_j$, $\mathcal{G} = \emptyset$, $i = \arg\{\eta_{jb} = 1\}$, $p_j = 0$;
5:     $\mathbf{r}^j[k] = +\infty$, $k = 1, \ldots, \sum_{b=1}^{N_r} L_b$;
6:     $\mathbf{r}^j[\Xi((i, b))] = 0$, $b = 1, \ldots, L_i$;
7:     $\ell = 1$, $m_j = 0$;
8:     repeat
9:       $\delta = \min_{\mathbf{G} \neq \emptyset, |\mathbf{G}| \leq \tau} \Delta_j(\mathbf{V}, \mathbf{G}, \hat{\mathbf{r}})$;
10:      $\mathbf{G}_j^l = \arg\min_{\mathbf{G} \neq \emptyset, |\mathbf{G}| \leq \tau} \Delta_j(\mathbf{V}, \mathbf{G}, \hat{\mathbf{r}})$;
11:     $\mathcal{D} \leftarrow \mathcal{D} \setminus \mathbf{G}_j^l$ and $\mathcal{G} \leftarrow \mathcal{G} \cup \mathbf{G}_j^l$;
12:     if $\Xi((i, b)) \neq \mathcal{G}$, $b = 1, \ldots, L_i$ then
13:       $\mathbf{r}^j[\Xi(\mathbf{G}_j^l)] \leftarrow \delta$;
14:       $m_j \leftarrow m_j + 1$;
15:     end if
16:     $\ell = \ell + 1$;
17:   until $\mathcal{D} = \emptyset$
18:   Set $\mathbf{q}^j[m] \leftarrow \mathbf{G}_j^l$, $1 \leq m \leq m_j$;
19:   $\mathbf{q}^j[m_j + 1] \leftarrow \cup_{m > m_j} \mathbf{G}_j^l$;
20: end for
21: $\hat{\mathbf{r}}[k] = \min_{j=1}^{N_r} \mathbf{r}^j[k]$, $k = 1, \ldots, \sum_{b=1}^{N_r} L_b$;
22: $\hat{\mathbf{r}}[k] \leftarrow \hat{\mathbf{r}}[k] + \hat{\mathbf{r}}[k]$, $k = 1, \ldots, \sum_{b=1}^{N_r} L_b$;
23: until $\hat{\mathbf{r}}$ converges
24: Output $\hat{\mathbf{r}}$ and $\mathbf{q}^j$, $j = 1, \ldots, N_r$.

inactive layers $\mathcal{K} \setminus \mathcal{K}_j$ and discarded layers $\cup_{m=m_j+1}^{N_r} \mathcal{Q}_m$ to infinity since they do not affect the global rate allocation computed by Eq. (32). Then function $\Xi(\mathbf{Q}_m)$ defined in Section III-B transforms index set to a linear index.

To solve the above problem, we need to find the optimal association matrix $\eta$ that maximizes the system’s throughput given all users’ decoding orders and the corresponding local rate allocations. Since all elements of $\eta$ are 0 or 1, such problem is a 0-1 combinatorial optimization problem. Moreover, this problem cannot be solved efficiently using auction method [27], since the rate for each layer in the summation is not fixed for each possible association matrix $\eta$ because of Eq. (31). On the other hand, an exhaustive search is often prohibitive due to the large size of matrix $\eta$. A feasible approach is to adopt the genetic algorithm [28] to heuristically find a good solution. The details on the genetic algorithm are standard, and thus omitted in this paper.

Note that the above optimization problem provides a basic rate allocation and the associated decoding order. However, the sum rate gain provided by multi-layer coding and CPGD hasn’t been fully exploited. An iterative rate update will be conducted at the transmitter side as detailed in Algorithm 3. Different from Algorithm 1, at the end of an iteration, the rate allocation $\hat{\mathbf{r}}$ in $\Delta_j(\mathbf{V}, \mathbf{G}, \hat{\mathbf{r}})$ in the new iteration of Algorithm 3 is set to the value of current iteration. So the sum rate of $\hat{\mathbf{r}}$ is guaranteed to increase after each iteration. The algorithm terminates when $\hat{\mathbf{r}}$ converges.

The outputs of Algorithm 3 are the enhanced global rate allocation vector $\hat{\mathbf{r}}$ and the improved decoding order $\mathbf{q}^j$ for $j = 1, \ldots, N_r$. The system can be configured through two phases. In the first phase, the system employs the rate allocation and decoding orders obtained via solving the transmitter-user association problem, to establish the initial communication links. In the second phase, the transmitters distribute the updated decoding orders to the corresponding users, followed by updating the rate allocation for each transmission layer. The above two phases can enhance the system throughput promised by the proposed interference cancellation scheme.

The computational complexity of Algorithm 3 is on the order of $O(L^2+1 N_r N_c)$ for $\tau$ smaller than a constant, where $N_r$ is the maximum number of iterative steps for each user in Algorithm 3. In practice, $N_c$ and $\tau$ are small and thus limit the complexity. Note that although the association needs to be performed as the link gain changes, considering the low speed and low link gain variation speed in the indoor scenario, online computational burden is still implementable.

V. NUMERICAL RESULTS

Consider a practical indoor environment shown in Figure 3(a), with 16 LEDs with $4 \times 4$ squared layout with $d_h = d_v = 0.2$m. At each LED, there are four LED chips acting as transmitters with red, green, blue, and yellow colors whose spectra are shown in Figure 2. The wavelength of four colors ranges [380nm, 480nm], [500nm, 550nm], [560nm, 600nm], [600nm, 680nm], where each spectrum is modeled by Gaussian distribution [21] with out-band leakage ratios [0.1, 0.2, 0.2, 0.1], respectively. The four color spectra are shown in Figure 2.

To compute the interference under multi-color transmission, we may need to obtain the filtering gain matrix $\mathbf{F}$ [29] with its $p^{th}$ row and $q^{th}$ column element being the power ratio from the transmitter with the $p^{th}$ color to the $q^{th}$ optical filter. Letting $N_c$ be the number of colors, we first model the spectrum of the $p^{th}$ ($p = 1, \ldots, N_c$) color mainly within the wavelength $[\lambda_{p1}, \lambda_{p2}]$ with Gaussian distribution $\mathcal{N}(\mu_p, \sigma_p^2)$, where $\mu_p = (\lambda_{p1} + \lambda_{p2})/2$ and $\sigma_p$ is obtained by solving $\int_{\lambda_{p2}}^{\infty} \mathcal{N}(\mu_p, \sigma_p^2) d\lambda = P_p/2$ for leakage probability $P_p$ out of band $[\lambda_{p1}, \lambda_{p2}]$. Assume that the $q^{th}$ optical filter has the passband over wavelength $[\lambda_{q1}, \lambda_{q2}]$. Then $\mathbf{F}_{pq}$ can be
obtained by integrating the $p^{th}$ color spectra over the band $[\lambda_{p1}, \lambda_{p2}]$, i.e., $F_{pq} = \int_{\lambda_{p1}}^{\lambda_{p2}} \mathcal{N}(\mu_p, \sigma_p^2) d\lambda$.

The filtering gain matrix is given in the following

$$
\begin{bmatrix}
0.900 & 0.011 & 0.000 & 0.000 \\
0.011 & 0.800 & 0.036 & 0.000 \\
0.000 & 0.027 & 0.800 & 0.100 \\
0.000 & 0.000 & 0.050 & 0.900
\end{bmatrix},
$$

(37)

where each row denotes the power ratio received at four optical filters for each color band. It is observed that the receiver receives not only the signals from the same color band, but also signals from other colors. If the $i^{th}$ transmitter uses the $p^{th}$ color band and the $j^{th}$ receiver uses the $q^{th}$ optical filter, $T_s(\psi_{ji}) = F_{pq}$.

The transmit array includes totally 64 transmitters, which is located at the center of the transmitter plane with margins $d_{q1} = d_{q2} = 1m$. Assume that the receivers can move, but they are always oriented upwards to ensure that their plane is parallel to that of the transmitters. The receiver plane of interest consists of a plane with width $d_w$ and length $d_l$ both 2.6m, and is colored in yellow with vertical distance to transmitter plane $d_z$. 2m.

At the transmitter, the peak power of an LED chip is set to $A_v = 1$ and the average power is set to $\varepsilon_i = 0.5$, for all $i = 1, \ldots, N_i$. The Lambertian order is $m_1 = 1$ with the transmitter semi-angle at half-power $\Phi_{1/2} = 60^\circ$. At the receiver, a non-imaging detector is used with the active area $A_r = 1 \times 10^{-4} m^2$ and the refractive index of optical concentrator $n = 1.5$. Assume the receiver FOV $\psi_c = 30^\circ$ (semi-angle). Let $h_{11}$ denote the channel gain from (0,0,0) to (0,2,0) through the optical filter with passband [380nm, 480nm]. The noise variance $\sigma^2$ is set such that the receiver-side SNR $A_r h_{11}/\sigma = 15$dB, which is typical for indoor VLC with strong light intensity. The noise variance is assumed to be the same for all positions on the receiver plane.

Multi-layer encoding and CPGD are employed at the transmitter and the receiver, respectively. For each transmitter, two-layer encoding is adopted to achieve the tradeoff between multi-layer coding gain and computational complexity. The layer indices are labeled sequentially in Figure 3(b) for illustration. At each position, the layer index increases from the LED with the smallest wavelength to that with the largest wavelength. For each layer, the input signal obays a TG distribution defined in Eq. (7) with $\mu_{ik}$ and $\nu_{ik}$ to be determined. We choose $\mu_{ik} = 3\nu_{ik}$ as suggested in [22], and combine Eq. (6) and Eq. (8) to compute $\hat{\mu}_{ik}$ and $\hat{\nu}_{ik}$ for the TG distribution. For each receiver, a CPGD with group size one is employed, since it has been reported that the group size of one has a low complexity to achieve the most of the performance gain, especially in the scenario with channel coding [12].

To obtain the decoding map, Algorithm 1 is adopted to compute the decoding order and local rate allocation at each position uniformly sampled with the sampling gap $d_s = 0.1m$ on the receiver plane, which are then stored. Note that any of the four colors can be adopted by a user.

Figure 4 gives the result at position $(0.8,0.6,2)$ of the decoding map constructed from the signal passing an optical filter with passband [380nm, 480nm]. Figure 4(a) shows the signal layer decoded at each decoding stage. From Figure 4(a) we see that layers sent from the same color are decoded successively and layers from color with wavelength [500nm, 550nm] will not be decoded until all the layers from other colors are recovered successfully. The reason is the weak inter-color interference and large difference in signal strength induced by optical filters. The layers from color [560nm, 600nm] and color [600nm, 680nm] are not presented since the signal interference passing through the optical filter within passband [380nm, 480nm] is negligible. Figure 4(b) presents the maximum achievable rates for all layers when the decoding order is given in Figure 4(a). This gives the local rate allocation for the layers at position $(0.8,0.6,2)$.

Furthermore, Figure 5 shows $\min_{i \in \{1,2,3,4\}} d(\mathbf{v}^i, \mathbf{v}^4)$, where $\mathbf{v}^i$ is the receiver position sampled at $d_s = 0.02m$; $\mathbf{v}^4$ are the four closest positions to $\mathbf{v}^i$ on the receiver plane sampled at $d_s = 0.1m$; and $d(\cdot,\cdot)$ is defined in Eq. (25). We can see that the distances are very small, indicating that using the decoding order at one position for a neighbor position leads to negligible loss in terms of the layer rates.

For each decoding map of a receiver plane, we employ Algorithm 2 to reduce its size. In our simulation, $\tau_{\text{diff}}$ was set to $1 \times 10^{-5}$ and $\tau_{\text{loss}}$ was set to 0.1. Figure 6 shows the clustering result of the decoding map of color [380nm, 480nm] under the indoor setting shown in Figure 3(a). The black empty circles represent the positions that will yield decoding outages because of the weak signal reception. The filled circles with the same color are in the same cluster denoting that they share the same decoding order. In this map, a total of $N_s = 681$ points are classified into $N_k = 511$ clusters, and
the compression ratio defined as $(N_s - N_k)/N_s$ is 25.0%. The maximum average loss (or the maximum average normalized distance) is close to zero, meaning that sharing the same decoding order will not penalize the system's sum rate.

We show that the density of the transmitters affects the effectiveness of Algorithm 2. The compression ratio will increase if we reduce the number of transmitters within the same area. Figure 7(a) gives the clustering result of the decoding map with only four transmitters at four corners located at $(0, 0, 0)$, $(0, 0, 6, 0)$, $(0, 6, 0, 0)$, $(0, 6, 0, 6)$. The decoding map has been clustered into 28 categories with the maximum average loss close to zero and the compression ratio being 95.9%, which demonstrates the effectiveness of the size reduction algorithm. Figure 7(b) gives the clustering result when transmitters are located at $(0, 0, 0)$ and $(0, 0, 6, 0)$. The decoding map has been clustered into 4 categories with maximum average loss close to zero and the compression ratio being 99.2%.

Note that Algorithm 2 uses the distance defined in (25) to characterize the similarity between two decoding orders, to perform clustering and reduce the size of a decoding map. But it ignores the possible symmetry of channel gains, which can establish an explicit relation between two decoding orders at two symmetrical positions. As demonstrated in Section III-E, the transmitters arranged in a square shape will have symmetrical positions having the same channel gain set on the receiver plane, which shows an explicit relationship between the corresponding two decoding orders. The compression ratio of Figure 6 with a higher density of transmitters can be further increased to 7/8 or 87.5% if we only compute the upper right triangle region of the decoding map and obtain the rest based on that region. The number of decoding orders needed to be stored in Figure 7(a) can be reduced from 28 to 7 yielding an increased compression ratio of 99.0%.

We further pursue the optimal transmitter-user association and rate allocation to maximize the system throughput. In the following, we consider users using color band [380nm, 480nm] as an example. Our system has two configuration phases as described at the end of Section IV, i.e., determining transmitter-user associations using decoding maps and iterative rate update. The decoding map used in the following simulations is constructed from the symmetric properties defined in Section III-E, i.e., we only compute the upper right 1/8 area of the whole decoding map, and construct the remaining 7/8 area based on the 1/8 results.

Figure 8 gives positions of 16 users uniformly over the receiver plane either parallel or perpendicular to the transmitter plane in a squared grid layout with the distance between adjacent users being 0.2m. On the $x-y$ plane, the position of the bottom left user $(x, y, z)$ with $z = 2$ represents all the 16 users' positions, which are constrained in the region $-1 \leq x \leq 1.6$ and $-1 \leq y \leq 1.6$. On the $y-z$ plane, the position of the bottom left user $(x, y, z)$ with $x = 0.3$ represents all the 16 users' positions. Users are confined in region $-1 \leq x \leq 1.6$ and $1 \leq z \leq 3$. Then on each plane, we move all users via step 0.1m, and solve the maximization problem of Eq. (26) subject to the constraints from Eq. (28) to Eq. (33) by genetic algorithm.

Figure 9 gives the sum and minimum rates on the $x-y$ receiver plane parallel to the transmitter plane when the transmitter-user association is taken into account and iterative rate update is conducted. Similar to the technique used in [12], we adopt a user rate threshold $r_{ud} = 0.05$ to inactivate...
users whose rates are below $r_{ud}$. This threshold value is used throughout the following simulations. In Figure 9(a), we observe lower rates at $x = 0$ and $y = 0$ than those of their adjacent positions. This can be justified as follows. Assuming the associated transmitter for user 1 is $i_1$, then for user 2, which is symmetrical to 1, the associated transmitter is $i_2 = A_p\left(\pi^{-1}_A(i_1)\right)$. Then at user 1, generally we have $R_1 > R_2$, where $R_1 = r_1^z\mathbb{E}(i_1, k)$ and $R_2 = r_2^z\mathbb{E}(i_2, k)$; at user 2, by symmetry, we have $R_1' = R_2$, $R_2' = R_1$ with $R_1' = r_1^z\mathbb{E}(i_1, k)$ and $R_2' = r_2^z\mathbb{E}(i_2, k)$. By taking the element-wise minimum operation to $r_1$ and $r_2$, the rates of layers $(i_1, k)$ and $(i_2, k)$ are both constrained to the lower rate $R_2$. Thus strong symmetric interference exhibited by user locations may constrain the throughput of the system. Moreover, it is observed that based on the allocation results, each user is allocated to the colors matching the corresponding wavelengths. On the other hand, if the majority of users are at asymmetrical positions, this rate constraint will be lessened since some other user may suffer weaker interference for the same transmitter layer. So at the center of the receiver plane, the sum rate is the lower and it increases as the signal strength decays. Furthermore, we observe similar rate variation in Figure 12(b) on the minimum rate of adjacent positions. The number of symmetrical users is $y = 0$, which constrains the throughput. The sum rate increases as the users move to left or right positions, and decreases as the signal strength decays. Furthermore, we observe similar rate variation in Figure 12(b) on the minimum rate distribution over the $y - z$ plane.

Figure 10 gives the rate gain in achievable sum rates and minimum user rates at different positions over a time division multiplexing (TDM) baseline system. It is observed that our system can achieve higher sum user rate and minimum user rate than the TDM system except at some center positions on the receiver plane, where the interference is stronger than other positions.

As mentioned at the end of Section III-B, we consider the objective of max-min rate of layers as a suboptimal and tractable solution. Because of the intractable optimal solution to max-min rate of users under multi-layer setting, comparison between maximizing the minimum layer rate and maximizing the minimum user rate under multi-layer setting is difficult. However, we can see how our scheme improves the minimum user rate over an unlayered one. Figure 11 shows the comparison of sum rate and minimum user rate of the proposed system against the unlayered system, where minimum rate of layers is equivalent to minimum rate of users. Both systems adopt the same configuration used in Figure 9. It is observed that the minimum user rate in the proposed system is larger than that of the unlayered one, at the cost of lower achievable sum rate. This is because we adopt the max-min rate of layers as the objective on obtaining decoding orders.

Figure 12(a) gives the sum rate distribution on the $y - z$ receiver plane perpendicular to the transmitter plane. Similarly, we observe lower rates at $y = 0$ compared with those of adjacent positions. The number of symmetrical users is highest at $y = 0$, which constrains the throughput. The sum rate increases as the users move to left or right positions, and decreases as the signal strength decays. Furthermore, we observe similar rate variation in Figure 12(b) on the minimum rate distribution over the $y - z$ plane.

VI. Conclusion

In this paper, we have employed multi-layer coding and CPGD for interference cancellation in VLC, where signals are transmitted via multiple LEDs with different colors. We model the multi-color multi-user interference channel with Gaussian-like spectra and ideal optical filters. A max-min fairness optimization problem is formulated, based on the achievable rate of a VLC-MAC obtained from truncated Gaussian distribution. The problem is solved by greedy algorithm with no preset transmitter-user associations. A decoding map is constructed to reduce the online computational complexity. A clustering algorithm is proposed to reduce the map size and
facilitate storage and searching, where the compression ratio can be further increased based on the symmetrical geometry of the transmitters. Finally, the transmitter-user association problem is formulated as a 0-1 optimization problem solved by genetic algorithm, followed by an iterative update of the rate allocations and decoding orders. The sum and minimum achievable rates are evaluated by numerical results.

REFERENCES

[1] G. Zheng, C. Gong, and Z. Xu, “Multi-layer coding and map-assisted partial group decoding for multi-color multi-user visible light communication,” in Proc. IEEE Int. Conf. Commun. (ICC), May 2019, pp. 1-6.

[2] H. Elgala, R. Mesleh, and H. Haas, “Indoor optical wireless communication: Potential and state-of-the-art,” IEEE Commun. Mag., vol. 49, no. 9, pp. 56-62, Sep. 2011.

[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. 26-32, Dec. 2013.

[4] Q. Gao, R. Wang, Z. Xu, and Y. Hua, “DC-informative joint color-frequency modulation for visible light communications,” J. Lightw. Technol., vol. 33, no. 11, pp. 2181-2188, Jun. 1, 2015.

[5] M. A. Khalighi and M. Uysal, “Survey on free space optical communication: A communication theory perspective,” IEEE Commun. Surveys Tuts., vol. 16, no. 4, pp. 2231-2258, Nov. 2014.

[6] A. Lapidoth, S. M. Moser, and M. A. Wigger, “On the capacity of free-space optical intensity channels,” IEEE Trans. Inf. Theory, vol. 55, no. 10, pp. 4449-4461, Oct. 2009.

[7] A. A. Farid and S. Hranilovic, “Channel capacity and non-uniform signalling for free-space optical intensity channels,” IEEE J. Sel. Areas Commun., vol. 27, no. 9, pp. 1553-1563, Dec. 2009.

[8] L. Grobe et al., “High-speed visible light communication systems,” IEEE Commun. Mag., vol. 51, no. 11, pp. 60-66, Dec. 2013.

[9] R. Bai, H. Tian, B. Fan, and S. Liang, “Coordinated transmission based interference mitigation in VLC network in Proc. IEEE 82nd Veh. Technol. Conf. (VTC-Fall), Sep. 2015, pp. 1–5.

[10] X. Zhang, Q. Gao, and Z. Xu, “Optical interference alignment for an indoor visible light communication X-channel in Proc. Global Conf. Signal Inf. Process. (GlobalSIP), Dec. 2015, pp. 1175–1179.

[11] A. S. Motahari, S. Oveis-Gharan, M.-A. Maddah-Ali, and A. K. Khandani, “Real interference alignment: Exploiting the potential of single antenna systems,” IEEE Trans. Inf. Theory, vol. 60, no. 8, pp. 4799–4810, Aug. 2014.

[12] C. Gong, A. Tajer, and X. Wang, “Interference channel with constrained partial group decoding,” IEEE Trans. Commun., vol. 59, no. 11, pp. 3059–3071, Nov. 2011.

[13] T. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” IEEE Trans. Inf. Theory, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.

[14] C. Gong, O. Abu-Ella, X. Wang, and A. Tajer, “Constrained group decoder for interference channels,” J. Commun., vol. 7, no. 5, pp. 382–390, May 2012.

[15] M. Ashraphijuo, A. Tajer, C. Gong, and X. Wang, “A receiver-centric approach to interference management: Fairness and outage optimization,” IEEE Trans. Inf. Theory, vol. 62, no. 10, pp. 5619–5642, Oct. 2016.

[16] N. Prasad and X. Wang, “Outage minimization and rate allocation for the multiuser Gaussian interference channels with successive group decoding,” IEEE Trans. Inf. Theory, vol. 55, no. 12, pp. 5540–5557, Dec. 2009.

[17] P. Chvojka, S. Zvanovec, P. A. Haigh, and Z. Ghassemlooy, “Channel characteristics of visible light communications within dynamic indoor environment,” J. Lightw. Technol., vol. 33, no. 9, pp. 1719–1725, May 1, 2015.

[18] S. H. Yang, E. M. Jeong, D. R. Kim, H. S. Kim, Y. H. Son, and S. K. Han, “Indoor three-dimensional location estimation based on LED visible light communication,” Electron. Lett., vol. 49, no. 1, pp. 54–56, Jan. 2013.

[19] K. Qiu, F. Zhang, and M. Liu, “Let the light guide us: VLC-based localization,” IEEE Robot. Autom. Mag., vol. 23, no. 4, pp. 174–183, Dec. 2016.

[20] Z. Ghassemlooy, W. Popoola, and S. Rajbhandari, Optical Wireless Communications: System and Channel Modelling With MATLAB, Boca Raton, FL, USA: CRC Press, 2012.

[21] L. Cui, Y. Tang, H. Jia, J. Luo, and B. Gnade, “Analysis of the multichannel WDM-VLC communication system,” J. Lightw. Technol., vol. 34, no. 24, pp. 5627–5634, Dec. 1, 2016.

[22] A. Chaaban, J.-M. Morvan, and M.-S. Alouini, “Free-space optical communications: Capacity bounds, approximations, and a new sphere-packing perspective,” IEEE Trans. Commun., vol. 64, no. 3, pp. 1176–1191, Mar. 2016.

[23] O. M. S. Al-Ebraheemy, A. Chaaban, T. Y. Al-Naffouri, and M. S. Alouini, “Capacity bounds for the 2-user Gaussian IM-DD optical multiple-access channel,” in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), May 2016, pp. 2823–2826.

[24] T. M. Cover and J. A. Thomas, Elements of Information Theory. Hoboken, NJ, USA: Wiley, 2012.

[25] J. Thomas, “Feedback can at most double Gaussian multiple access channel capacity,” IEEE Trans. Inf. Theory, vol. IT-33, no. 5, pp. 711–716, Sep. 1987.

[26] C. Gong, A. Tajer, and X. Wang, “Group decoding for multi-relay assisted interference channels,” IEEE J. Sel. Areas Commun., vol. 30, no. 8, pp. 1489–1499, Sep. 2012.

[27] K. Yang, N. Prasad, and A. Farid, “An auction approach to resource allocation in uplink OFDMA systems,” IEEE Trans. Signal Process., vol. 57, no. 11, pp. 4482–4496, Nov. 2009.

[28] D. Whitley, “A genetic algorithm tutorial,” Statist. Comput., vol. 4, no. 2, pp. 65–85, Jun. 1994.

[29] K. X. Zhou, C. Gong, and Z. Y. Xu, “Color planning and inter-cell interference coordination for multicolor visible light communication networks,” J. Lightw. Technol., vol. 35, no. 22, pp. 4980–4993, Nov. 15, 2017.