# Adaptive User Association for Dense Visible Light Communication Networks in the Presence of Nonlinear Impairments

Pu Miao, *Member, IEEE*, Gaojie Chen, *Senior Member, IEEE*, Yu Yao, Kai-Kit Wong, *Fellow, IEEE* and Jonathon A. Chambers, *Fellow, IEEE*

*Abstract*—User-centric (UC) philosophy is a promising network formation method in light emitting diode enabled visible light communication (VLC) systems. Nevertheless, the nonlinear channel impairments restrict the overall system performance and have not been fully considered in the association structure designing. In this paper, an adaptive user association approach within the UC-cells formation of dense VLC networks is investigated under the consideration of practical nonlinear impairments and adjacent interference. It is mathematically formulated to be an achievable data rate maximization problem by jointly determining the optimal candidates of access point, clipping ratio and information-carrying power. We divide this mixed combinatorial and non-convex optimization problem into two subproblems and delicately transform them to be binary nonlinear programming and constrained linear programming problems, respectively. In addition, we develop an efficient approach to obtain the local optimal solution with low-computational complexity in an alternating iterative way. Simulation results demonstrate that the proposed scheme has relatively fast convergence and shows robustness to the variation of complex interference patterns and nonlinear impairments. Moreover, it can achieve significant throughput gain as compared with the conventional schemes, demonstrating the prospect and validity of this methodology for dense VLC networks with actual nonlinear devices.

*Index Terms*—Visible light communication, user-centric, nonlinear distortion, clipping, power allocation

## I. INTRODUCTION

Light-emitting diode (LED) based visible light communication (VLC) has attracted considerable interest over the last decade due to its advantages of license-free spectrum, low cost and immunity to electromagnetic interference [1]. The VLC is being cast in the spotlight for short-range wireless

communications and regarded as an important supplement for current radio frequency techniques [2]–[4]. Based on intensity modulation and direct detection (IM/DD), extensive studies including spectral-efficient modulation [5], robust coding [6], dimming control [7], channel estimation and equalization [8], [9] have been dedicated to point-to-point VLC systems to improve the achievable data rate. With the growing demands of wireless traffic for emerging indoor applications in 6G communications, how to design an efficient VLC system to simultaneously support multiple users with dynamic-and-flexible connectivity and high-speed transmission is still a critical challenge task in this field [10].

The cellular fixed-shape principle is commonly used for constructing the VLC network to support multiple user equipments (UEs), where the illumination area of an LED-based access point (AP) is considered as an individual attocell and the UEs access their nearest attocell for data transmission [11]. As for an indoor environment, the AP is always deployed in a dense manner to provide comfort illumination and the corresponding number may be much greater than that of UEs, resulting in a highly dense VLC network [12]. Thus, the coverage of multiple attocells would be overlapped and the induced inter-cell interference (ICI) will severely degrade the performance of cell-edge UEs. To alleviate the ICI, several approaches such as subchannel allocation [13] and frequency reuse [14] were proposed accordingly. However, the coordination assignments can be complex and the reuse factor cannot adapt to the variation of UEs' requirements. Moreover, the frequently inter-cell handover problems and outage events are still prominent when the UEs are moving between APs [15]. Therefore, the cellular structure may become inefficient for an indoor VLC network under dense AP deployment.

In contrast, the user-centric (UC) design philosophy relying on an amorphous and irregular-shaped cell structure was proposed for constructing the VLC network [16], where the UE could flexibly connect with the favorable APs based on their specific requirements without any handover and interruption [17]. Within the receiver's field of view (FOV), determining the specific association structure for different UEs and the candidate APs plays a crucial role in UC-cells formation. In [18], the UC-cells formation along with the multiuser scheduling was formulated as a maximum weighted matching problem relying on distance-based weight, and a greedy algorithm was proposed to find the suboptimal yet compelling matching relation and scheduling solution. In [19], a distance-based

amorphous cell formation was investigated and then the power allocation strategy was employed to maximize the energy efficiency of the system. The well-designed power control could indeed enhance the signal strength and alleviate the amount of interference. However, to facilitate the calculation, the power allocation was implemented under the assumption that all other APs transmitted at their maximum permissible optical power. By evaluating the exact interference information, the UC-cells formation and power allocation were jointly treated in [20], and two algorithms in terms of AP-UE clustering and power allocation were proposed for maximization the energy efficiency. In [21], the associated time fraction of UEs that were served by different APs and the corresponding allocated power were jointly optimized to maximize the achievable instantaneous rate of the VLC system, where the UC-cells were formed based on the well-exploited time-domain resource and the associations of different APs and UEs just were executed at different time slots according to time division multiplexing (TDM). Similarly, the joint optimization of AP-grouping based UE association and power allocation algorithm were proposed in [22] to manage the down-link interference by employing the optimal time fraction of association, and finally maximize the overall spectral efficiency. However, the optimal grouping will cost high computational complexity under a dense APs deployment. Apart from the TDM, the frequency division multiplexing (FDM) can be also involved in UC-cell formation. In [23], a joint AP grouping and time-frequency resource allocation was proposed to reduce the intra-group interference and maximize the system throughput for VLC networks. Besides, the subcarrier group, power and modulation formats can be also jointly optimized in [24] and an iterative algorithm was proposed to solve the resource management optimization problem. As for UC-cells formation based on sophisticated resource multiplexing strategies, both the interference mitigation and system performance gain were achieved at the cost of resource utilization. In [25], the UE association based on ingenious on-off collaboration between adjacent APs was investigated without using any time/frequency division multiplexing. By considering the UEs' future locations and dynamic data traffic, an anticipatory association algorithm was proposed to achieve an attractive trade-off between the average delay and average per-user throughput. However, only the uniform power loading was employed in [25] thus the effect of power control on the system performance was not considered.

Most of these previous studies were conducted under distortion-free transmission and had not fully considered the adverse effect of hardware impairments on the system performance. In fact, the LED has an inherent non-ideal electro-optical conversion characteristic [26] and the available dynamic linear range is always confined [27]. As the signal peaks exceed the dynamic linear range [28], the nonlinear distortions will be introduced which will drastically deteriorate the bit error rate (BER) performance of the whole system [29]. In addition, the commercially available digital-to-analog converters (DACs) and analog-to-digital converters (ADCs) always have limited bit resolution, thus the nonlinear quantization distortion is exacerbated as the signal with high peak-to-average power ratio (PAPR) is imposed. Besides the inherent non-ideal char-

acteristics of the hardware itself, the amount of nonlinearity is also related with the payload information of the modulated signal [30], whose amplitude is driven by information carrying (IC) power. Enhancing the IC power is favorable for achieving higher signal-to-noise ratio (SNR), however, it will result in larger peak-to-peak voltage of a waveform. Deep clipping should be always deployed to accommodate both the available transmission region of the LED and the limited bit resolution of DACs and ADCs, thus the severe clipping distortions would be introduced consequently [31]. Therefore, there indeed exists a trade-off between the clipping and IC power, which should be jointly treated to balance the system performance and the nonlinear impairments. We note that there were relatively insufficient literatures on how to design the optimal UC-cells when facing the inevitable nonlinear impairments derived from the LEDs, DACs and ADCs of all APs and UEs. Although the nonlinearity introduced by both LED saturation and turn-on threshold was considered in [25], the distortions cannot be dynamically adjusted with the variation of UEs distribution and the quantization noises of DACs and ADCs were not involved accordingly.

Motivated by the above analysis, an adaptive user association approach is investigated in this paper by considering three interlinked design aspects in terms of AP selection, clipping and IC power allocation in the presence of practical hardware impairments of multi-user VLC systems. To elaborate, the main contributions are summarized as follows:

- A novel adaptive user association framework involving the nonlinear impairments of LEDs, DACs and ADCs is firstly proposed for UC-cells formation in VLC systems, which is formulated as dynamic achievable data rate maximization relying on the optimization of AP candidates, adaptive clipping and IC power allocation. It is essentially different as compared with the schemes in [19]–[25].
- We conceive an efficient approach by dividing the original mixed combinatorial and non-convex optimization into two sub-problems, and delicately transform them into binary nonlinear programming and constrained linear programming problems, which can be solved by the time-efficient algorithms with low complexity in an alternating iterative way.
- With the joint optimization of APs selection, clipping ratio and IC power, the proposed scheme can alleviate both the interference and nonlinear impairments of the whole system effectively, and then improve the achievable data rate without sacrificing the resource utilization. In addition, it is more flexible to the variation of both UEs' spatial distribution and the unique hardware devices.
- Numerical simulations are performed to verify the associated performance in terms of convergence, optimized stationary points and the achieved sum rate, which shows the applicability and superiority of the proposed scheme to some traditional approaches.

The reminder of this paper is organized as follows. In Section II, the channel model and signal model are introduced, and the statistical analysis of the system nonlinearity in terms of clipping and quantization distortions is also presented.

The problem formulation is described in Section III and the proposed methodology is elaborated in detail in Section IV. Section V demonstrates the simulation results and discussions, and Section VI finally concludes this paper.

*Notations*: The set of real numbers is denoted by $\mathbb{R}$. The terms $(\cdot)^T$ and $\mathbb{E}(\cdot)$ are employed to represent the transpose and the mathematical expectation operators, respectively. In addition, we use $|a|$ to denote the absolute value of $a$ and $|\mathcal{N}|$ to represent the cardinality of set $\mathcal{N}$. Let $\|\cdot\|_p$ denote the $\ell_p$-norm, and $\mathbb{N}(a, \sigma^2)$ is the Gaussian distribution with mean $a$ and variance $\sigma^2$.

## II. SYSTEM MODEL

The conceptual diagram of a typical indoor UC-VLC network consisting of $N$ APs and $M$ UEs is shown in Fig. 1. The APs are uniformly installed on the ceiling for stable illumination and the UEs are randomly distributed in the room. The central controller connecting to the APs via wired lines is responsible for the information collection, analyzing and processing[1]. After decision making, the central controller sends the control and information signal to each AP. Then, the associated APs emit the information signal to the targeted UEs by using the IM/DD-based VLC down-link channel. We assume that each AP is equipped with one high illumination-power LED and each UE has one photodetector (PD), which are oriented vertically downwards and upwards, respectively. In addition, the number of APs is much higher than that of UEs since the dense AP deployment is considered.

### A. Channel Model

The light-of-sight (LOS) path and the first reflection path dominate in the indoor optical propagation. The LOS channel gain between the UE $u_m$ ($m = 1, 2, \cdots, M$) and the AP $a_n$ ($n = 1, 2, \cdots, N$) can be expressed as [13]

$$h_{m,n}^0 = \frac{(\bar{m}_0 + 1) R_{\mathrm{PD}} A_{\mathrm{PD}}}{2\pi D_{m,n}^2} \cos^{\bar{m}_0}(\phi_{m,n}) T_s(\psi_{m,n}) \\ \times g_s(\psi_{m,n}) \cos(\psi_{m,n}), \quad 0 \le \psi_{m,n} \le \psi_C, \tag{1}$$

where $\bar{m}_0 = -1/\log_2(\cos\phi_{1/2})$ is the Lambertian emission order, $\phi_{1/2}$ and $\phi_{m,n}$ are the half power angle and the emission angle of the LED, respectively. In addition, $D_{m,n}$ denotes the transmission distance between $u_m$ and $a_n$, $\psi_{m,n}$ is the incidence angle and $\psi_C$ is the half of the FOV. Let $A_{\mathrm{PD}}$ and $R_{\mathrm{PD}}$ denote the active area and optoelectronic responsivity of PD, respectively. Furthermore, $T_s(\psi_{m,n})$ is the optical filter gain, and $g_s(\psi_{m,n})$ is the concentrator gain, which is calculated by

$$g_s(\psi_{m,n}) = \frac{n_r^2}{\sin^2(\psi_C)}, \quad 0 \le \psi_{m,n} \le \psi_C, \tag{2}$$

where $n_r$ is the internal refractive index.

As for the non-LOS (NLOS) paths, only the first reflection is considered because the power of higher order scattering
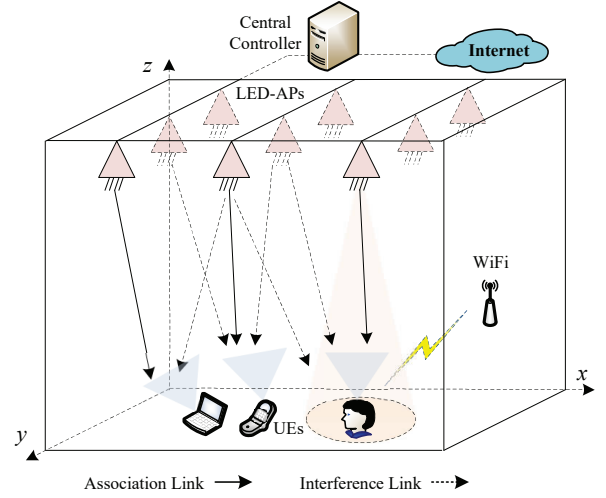
Fig. 1. Illustration of an indoor user-centric VLC system.

components is relatively weak. Explicitly, the gain of the first reflected component is expressed by [21]

$$h_{m,n}^1 = \int_{walls} h_{m,n}^{(1)} \mathrm{d}A_w, \tag{3}$$

where $h_{m,n}^{(1)}$ is the reflection term derived from a small reflection point. Specifically, $h_{m,n}^{(1)}$ is calculated by

$$h_{m,n}^{(1)} = \frac{\rho(\bar{m}_0 + 1) R_{\mathrm{PD}} A_{\mathrm{PD}}}{2\pi^2 D_1^2 D_2^2} A_w \cos^{\bar{m}_0}(\phi_w) \cos(\phi_a) \cos(\phi_b) \\ \times T_s(\psi_w) g_s(\psi_w) \cos(\psi_w), \quad 0 \le \psi_w \le \psi_C, \tag{4}$$

where $D_1$ and $D_2$ represent the distance of $a_n$ to the reflection point and the reflection point to $u_m$, respectively. The reflective area is $A_w$, while $\rho$ denotes the corresponding reflection factor, $\phi_w$ and $\phi_a$ are the angles of irradiance and incidence with respect to $D_1$, respectively. Similarly, $\phi_b$ and $\psi_w$ are the angles of irradiance and incidence with respect to $D_2$, respectively. Finally, the aggregated channel gain between $u_m$ and $a_n$ is given by $h_{m,n} = h_{m,n}^0 + h_{m,n}^1$. Therefore, both the distance and incident angle are the dominant factors which trigger the variation of channel gain. The potential APs for the UEs are not simply determined by their relative spatial distribution, but also dependent on their FOVs.

### B. Signal Model

Let $x_n$ denotes the digital time-domain signal of direct current biased optical orthogonal frequency division multiplexing (DCO-OFDM) that is prepared for $a_n$. According to the central limit theorem, $x_n$ can be modeled as Gaussian distribution following $\mathbb{N}(0, \sigma_{s,n}^2)$, where $\sigma_{s,n}^2$ is the variance of $x_n$ and also represents the IC power. Then, $x_n$ is used to drive $a_n$ by using intensity modulation. Due to the limited bit resolution of the DAC and the limited linear dynamic range of LED, the amplitude clipping is implemented for $x_n$ with high PAPR, shown as

$$x_{\mathrm{clip},n} = \begin{cases} x_n, & \vartheta_l \le x_n \le \vartheta_u \\ \vartheta_u, & x_n > \vartheta_u \\ \vartheta_l, & x_n < \vartheta_l \end{cases}, \tag{5}$$

where $\vartheta_l$ and $\vartheta_u$ are the lower and upper thresholds, respectively. For simplicity, the symmetric clipping, i.e., $\vartheta_u = -\vartheta_l = \vartheta_{th}$, is considered in this paper. Therefore, the clipping ratio (CR) associated with $\vartheta_{th}$ can be defined as

$$\zeta_n = \frac{\vartheta_{th}}{\sqrt{\mathbb{E}\left\{|x_n|^2\right\}}} = \frac{\vartheta_{th}}{\sigma_{s,n}}. \tag{6}$$

The clipped signal $x_{\text{clip},n}$ follows a truncated Gaussian distribution [29], which has a probability density function as

$$P_r(x) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma_x^2}}\exp\left(-\frac{x^2}{2\sigma_x^2}\right), & -\vartheta_{th} \leq x \leq \vartheta_{th} \\ \frac{1}{2}\text{erfc}\left(\frac{\vartheta_{th}}{\sqrt{2\sigma_x^2}}\right)\delta(x - \vartheta_{th}), & x > \vartheta_{th} \\ \frac{1}{2}\text{erfc}\left(\frac{\vartheta_{th}}{\sqrt{2\sigma_x^2}}\right)\delta(x + \vartheta_{th}), & x < -\vartheta_{th} \end{cases} \tag{7}$$

where $\text{erfc}(x) = 1 - \text{erf}(x)$, $\text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$ is the error function and $\delta(\cdot)$ is the Dirac function. For a given $\zeta_n$, therefore, the clipping probability of $x_n$ can be calculated by $\mathcal{P}_{\text{clip}} = 1 - 2\text{erf}\left(\frac{\zeta_n}{\sqrt{2}}\right)$. After clipping, $x_{\text{clip},n}$ is injected into the DAC with a bit resolution of $Q_{\text{DAC},n}$ to produce an analog waveform $x_{\text{clip},n}(t)$. In addition, a DC bias $b_{dc,n}(t)$ is always added to $x_{\text{clip},n}(t)$ to obtain the nonnegative waveform $x_{dc,n}(t)$, which drives the LED to generate the optical signal $x_{o,n}(t) = R_{\text{LED}}x_{dc,n}(t)$. In this paper, we assume the linearized electro-optic conversion gain $R_{\text{LED}} = 1$ for facilitating analysis. In addition, $b_{dc,n}(t)$ can be fixed to $\zeta_n\sigma_{s,n}$ to avoid further clipping from the turn-on threshold. At $u_m$, the received optical signal is detected by a PD and then converted into an electrical signal. After that, the analog electrical signal is sent into an ADC with $Q_{\text{ADC},m}$ bits resolution to generate a digital signal. Then, the demodulation procedure is performed accordingly. Note that the DC component conveys no information thus only the signals related to useful information are analyzed in the following.

### C. Statistical Analysis of Nonlinear Impairments

Based on (5) and (7), the average power of $x_{\text{clip},n}$ can be calculated by

$$\begin{aligned} \sigma_{x_{\text{clip},n}}^2 &= \mathbb{E}\left\{|x_{\text{clip},n}|^2\right\} \\ &= 2\int_0^{\vartheta_{th}} \frac{x_n^2}{\sqrt{2\pi\sigma_{s,n}^2}}\exp\left(-\frac{x_n^2}{2\sigma_{s,n}^2}\right)dx + \vartheta_{th}^2\text{erfc}\left(\frac{\zeta_n}{\sqrt{2}}\right), \\ &= C_n\sigma_{s,n}^2 \end{aligned} \tag{8}$$

where $C_n$ is shown as

$$C_n = 1 - \sqrt{\frac{2}{\pi}}\zeta_n\exp\left(-\frac{\zeta_n^2}{2}\right) - \left(1 - \zeta_n^2\right)\text{erfc}\left(\frac{\zeta_n}{\sqrt{2}}\right). \tag{9}$$

According to the Bussgang theorem [32], the clipped $x_{\text{clip},n}$ can be modeled as a sum of two uncorrelated parts, shown as

$$x_{\text{clip},n} = x_{\text{eff},n} + d_n = \alpha_n x_n + d_n, \tag{10}$$

where $x_{\text{eff},n}$ is the effective signal, $\alpha_n$ is the linear attenuation factor, and $d_n$ is the clipping distortion which is statistically

uncorrelated to $x_n$, i.e., $\mathbb{E}\left\{x_n d_n\right\} = 0$. Therefore, $\alpha_n$ can be calculated by

$$\alpha_n = \frac{\mathbb{E}\left\{x_n x_{\text{clip},n}\right\}}{\mathbb{E}\left\{|x_n|^2\right\}} = \text{erf}\left(\frac{\zeta_n}{\sqrt{2}}\right). \tag{11}$$

Moreover, the power of $x_{\text{eff},n}$ is derived as

$$\sigma_{x_{\text{eff},n}}^2 = \text{erf}^2\left(\frac{\zeta_n}{\sqrt{2}}\right)\sigma_{s,n}^2, \tag{12}$$

and the power of $d_n$ can be calculated by

$$\begin{aligned} \sigma_{d,n}^2 &= \mathbb{E}\left\{|x_{\text{clip},n}|^2\right\} - \alpha^2\mathbb{E}\left\{|x_n|^2\right\}, \\ &= G_{d,n}\sigma_{s,n}^2 \end{aligned} \tag{13}$$

where $G_{d,n}$ is related with $\zeta_n$, shown as

$$\begin{aligned} G_{d,n} &= 1 - \sqrt{\frac{2}{\pi}}\zeta_n\exp\left(-\frac{\zeta_n^2}{2}\right) \\ &\quad - \left(1 - \zeta_n^2\right)\text{erfc}\left(\frac{\zeta_n}{\sqrt{2}}\right) - \text{erf}^2\left(\frac{\zeta_n}{\sqrt{2}}\right). \end{aligned} \tag{14}$$

As impinging $x_{\text{clip},n}$ on the DAC with $Q_{\text{DAC},n}$ bits, the signal requires $2^{Q_{\text{DAC},n}}$ levels for quantization, then quantization noise is consequently generated, which can be modeled as additive uniformly distributed white noise with a variance of

$$\sigma_{\text{DAC},n}^2 = \frac{\zeta_n^2\sigma_{s,n}^2\left(1 - \mathcal{P}_{\text{clip}}\right)}{3\cdot 4^{Q_{\text{DAC},n}}} = G_{\text{DAC},n}\sigma_{s,n}^2, \tag{15}$$

where $G_{\text{DAC},n}$ is denoted by

$$G_{\text{DAC},n} = \frac{2\zeta_n^2}{3\cdot 4^{Q_{\text{DAC},n}}}\text{erf}\left(\frac{\zeta_n}{\sqrt{2}}\right). \tag{16}$$

As for $u_m$, the variance of the quantization noise in the ADC can be calculated by

$$\sigma_{\text{ADC},m}^2 = \frac{\zeta_n^2 h_{m,n}^2\sigma_{s,n}^2}{3\cdot 4^{Q_{\text{ADC},m}}} = G_{\text{ADC},n}h_{m,n}^2\sigma_{s,n}^2, \tag{17}$$

where $G_{\text{ADC},n} = \frac{\zeta_n^2}{3\cdot 4^{Q_{\text{ADC},m}}}$. According to (13), (15) and (17), we can foresee that the quantization noise can be reduced by decreasing $\zeta_n$ while at an expense of increasing clipping distortion, and vice versa. Therefore, both the linear attenuation factor and the nonlinear attenuation factor in terms of $\alpha_n$, $C_n$, $G_{d,n}$, $G_{\text{DAC},n}$ and $G_{\text{ADC},m}$ should be well configured to make a good compromise among all APs and UEs for achieving the maximum SNR.

### III. PROBLEM FORMULATION

Let $\mathcal{N}_m$ denotes the pre-defined AP set that could be severed for $u_m$ at the current time slot and $\mathcal{N}_m^{\text{LOS}}$ be the set of APs that has the LOS links with $u_m$. The remaining APs in $\mathcal{N}_m^{\text{LOS}}$ except for $\mathcal{N}_m$ belong to the subset $\bar{\mathcal{N}}_m$. Thus, we have $\mathcal{N}_m \subseteq \mathcal{N}_m^{\text{LOS}}$ and $\mathcal{N}_m \cup \bar{\mathcal{N}}_m = \mathcal{N}_m^{\text{LOS}}$. In this paper, only those links are considered in the UC-cell formation. Although a single AP may provide LOS links for multiple UEs simultaneously, we assume that different UEs have the exclusive associated APs at the current time slot, which should satisfy $\mathcal{N}_m \cap \mathcal{N}_l = \emptyset$, $\forall m \neq l$. Therefore, the received signal of $u_m$ can be formulated as

$$y_m = \sum_{n\in\mathcal{N}_m} h_{m,n}x_{clip,n} + \sum_{j\in\bar{\mathcal{N}}_m} h_{m,j}x_{clip,j} + z_m, \tag{18}$$

where $z_m$ is the additive background white noise with the variance of $\sigma_{z_m}^2$ and has the power spectral density (PSD) of $10^{-22}$ A$^2$/Hz [25]. It can be seen that the first term in (18) is the useful signal from the associated APs and the second term is adjacent interference from the undesired APs. Therefore, the signal to interference plus distortion and noise ratio (SIDNR) of $u_m$ can be expressed by

$$
\mathrm{I}_m = \frac{\sum_{n=1}^{|\mathcal{N}_m|} G_B h_{m,n}^2 \sigma_{x_{\mathrm{eff}},n}^2}{\sum_{j=1}^{|\bar{\mathcal{N}}_m|} h_{m,j}^2 \sigma_{x_{\mathrm{clip}},j}^2 + \sum_{n=1}^{|\mathcal{N}_m|} h_{m,n}^2 \left(\sigma_{d,n}^2 + \sigma_{\mathrm{DAC},n}^2\right) + \sum_{n=1}^{|\mathcal{N}_m|} \sigma_{\mathrm{ADC},m}^2 + \sigma_{z_m}^2},
\tag{19}
$$

where $G_B \approx 1$ is the bandwidth utilization factor for DCO-OFDM. By substituting (8), (12), (13), (15) and (17) into (19), we can rewrite (19) as

$$
\mathrm{I}_m = \frac{\sum_{n=1}^{|\mathcal{N}_m|} h_{m,n}^2 \alpha_n^2 \sigma_{s,n}^2}{\sum_{j=1}^{|\bar{\mathcal{N}}_m|} h_{m,j}^2 C_j \sigma_{s,j}^2 + \sum_{n=1}^{|\mathcal{N}_m|} h_{m,n}^2 G_n \sigma_{s,n}^2 + \sigma_{z_m}^2},
\tag{20}
$$

where $G_n = G_{d,n} + G_{\mathrm{DAC},n} + G_{\mathrm{ADC},n}$.

Therefore, we can infer that the performance of $\mathrm{I}_m$ would be improved by recruiting more participating APs that offer high-quality links to $u_m$, while switching off unexpected APs that belongs to $\bar{\mathcal{N}}_m$. However, these recruited APs for $u_m$ would degrade the performance of $\mathrm{I}_l$ by emitting undesired interference to $u_l$. While these APs disabled for $u_m$ could be the potential dominant links for $u_l$. Besides, the performance of $\mathrm{I}_m$ is also affected by the nonlinear distortion derived from the APs that belongs to both $\mathcal{N}_m$ and $\bar{\mathcal{N}}_m$, which should be well balanced to guarantee the overall system performance. Therefore, it seems more complex to construct the optimal association structure among all APs and UEs because the potential candidates are mutually constrained, and the interference patterns and nonlinear impairments of the whole system are also interlinked. Intuitively, we speculate that if the clipping ratio, IC power and the participating APs can be adaptively configured with the variation of all UEs' spatial distribution, the adjacent interference and nonlinear impairments could be effectively alleviated to some extent.

To formulate the aforementioned problem, the binary variable $\beta_{m,n} \in \{0,1\}$ is introduced to denote the association relationship between the $u_m$ and $a_n$, where $\beta_{m,n} = 1$ if the $a_n$ is assign to $u_m$ and $\beta_{m,n} = 0$ for otherwise. Let $\tilde{\boldsymbol{\beta}}_n = \left[\beta_{1,n}, \beta_{2,n}, \cdots, \beta_{M,n}\right]^T \in \mathbb{R}^M$ be the association vector of $a_n$ and $\boldsymbol{\beta} = \left[\tilde{\boldsymbol{\beta}}_1, \tilde{\boldsymbol{\beta}}_2, \cdots, \tilde{\boldsymbol{\beta}}_N\right] \in \mathbb{R}^{M \times N}$ be the association matrix. Besides, we use $\boldsymbol{\zeta} = [\zeta_1, \zeta_2, \cdots, \zeta_N]^T \in \mathbb{R}^N$ to denote the clipping ratio vector and $\boldsymbol{\sigma}_s = \left[\sigma_{s,1}^2, \sigma_{s,2}^2, \cdots, \sigma_{s,N}^2\right]^T \in \mathbb{R}^N$ the IC power vector. Then the achievable instantaneous rate of $u_m$ can be expressed by

$$
R_m = B \log_2 \left(1 + \tilde{\mathrm{I}}_m\right),
\tag{21}
$$

where $B$ denotes the modulation bandwidth of DCO-OFDM, and $\tilde{\mathrm{I}}_m$ the reformulated SIDNR, shown as

$$
\tilde{\mathrm{I}}_m = \frac{\Omega_{m,1}\left(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)}{\Omega_{m,2}\left(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)},
\tag{22}
$$

where

$$
\Omega_{m,1}\left(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right) = \sum_{n=1}^{N} \beta_{m,n} h_{m,n}^2 \alpha_n^2 \sigma_{s,n}^2,
\tag{23}
$$

and

$$
\begin{aligned}
\Omega_{m,2}\left(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right) = &\sum_{l \neq m, l=1}^{M} \sum_{j=1}^{N} \beta_{l,j} h_{m,j}^2 C_j \sigma_{s,j}^2 \\
&+ \sum_{n=1}^{N} \beta_{m,n} h_{m,n}^2 G_n \sigma_{s,n}^2 + \sigma_{z_m}^2.
\end{aligned}
\tag{24}
$$

Additionally, we employ the weight $w_m$ to represent the priority of $u_m$. Thus, the UC-cells formation can be mathematically formulated by maximizing the weighted sum rate $\mathcal{F}\left(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)$ with respect to the variables $\boldsymbol{\beta}$, $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$, which can be expressed by

$$
\mathcal{P}_1 : \max_{\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}} \mathcal{F}\left(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right) = \max_{\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}} \sum_{m=1}^{M} w_m R_m
\tag{25a}
$$

$$
\text{s.t.} \quad \sum_{m=1}^{M} \beta_{m,n} \leq 1, \quad \forall n
\tag{25b}
$$

$$
\sum_{n=1}^{N} \beta_{m,n} \leq |\mathcal{N}_m|, \quad \forall m
\tag{25c}
$$

$$
\beta_{m,n} = \{0,1\}, \quad \forall m, \ n \in \mathcal{N}_m^{LOS}
\tag{25d}
$$

$$
\beta_{m,n} = 0, \quad \forall m, \ n \notin \mathcal{N}_m^{LOS}
\tag{25e}
$$

$$
0 \leq \sigma_{s,n}^2 \leq \sigma_{s,\max}^2, \quad \forall n
\tag{25f}
$$

$$
\sum_{n=1}^{N} \sigma_{s,n}^2 \leq P_{\mathrm{IC}}
\tag{25g}
$$

$$
1 \leq \zeta_n \leq \zeta_{\max}, \quad \forall n
\tag{25h}
$$

$$
\frac{\vartheta_{\min}}{2} \leq \zeta_n \sigma_{s,n} \leq \frac{\vartheta_{\max}}{2}, \quad \forall n
\tag{25i}
$$

To elaborate, constraint (25b) meets the condition that an AP can only serve one UE at the current time slot. As for a particular UE, constraint (25c) corresponds to the fact that the amount of participating APs cannot exceed the pre-defined cardinality of $\mathcal{N}_m$, and constraints (25d) and (25e) guarantee that only the APs providing the LOS components will have the chance to be recruited, otherwise, they will be abandoned by $u_m$. Besides, (25f) constraint the IC power of the modulated signal falling within its feasible region and (25g) corresponds to the total power budget $P_{\mathrm{IC}}$ with the consideration of energy consumption. Since severe clipping will result in invalid communication with contaminated BER, constraint (25h) limits the clipping ratio from the perspective of effective demodulation. Moreover, constraint (25i) guarantees that the DC bias should be operated within the available range $\left[\frac{\vartheta_{\min}}{2}, \frac{\vartheta_{\max}}{2}\right]$ for the sake of linear modulation and uniform illumination.

According to the objective function and the constraints of (25), the problem $\mathcal{P}_1$ is therefore a mixed combinational and non-convex optimization problem with respect to the discrete variable $\boldsymbol{\beta}$, and continuous variables $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$, which is non-trivial and challenging to solve. Concretely, the feasible region of $\boldsymbol{\beta}$ is a kind of finite set of points and non-convex. Then the optimal solution can be obtained by enumerating all feasible $\boldsymbol{\beta}$ and then calculates the corresponding achievable rate under

various clipping and power loading strategies. Intuitively, the $\{\boldsymbol{\beta}^\star, \boldsymbol{\sigma}_s^\star, \boldsymbol{\zeta}^\star\}$ which provides the largest data rate will be selected as the optimal solution for the problem $\mathcal{P}_1$. However, the amount of feasible solutions increases exponentially with the number of APs and UEs, and the computational complexity becomes huge. Moreover, since the adjacent interference and nonlinear impairment are mutually constrained and interlinked, the trade-off between $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$ is dynamically varied with the spatial distribution of all UEs and the active APs. Furthermore, the objective function with respect to $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$ is also non-concave, and the joint optimization is still a highly nonlinear and non-convex problem. Therefore, finding the exact optimal solution of the mixed combinational non-convex optimization problem will be intractable and nontrivial in a reasonable time.

## IV. PROPOSED METHODOLOGY

In this section, an efficient approach is proposed to find the local optimal solution of problem $\mathcal{P}_1$ from a practical point of interference management and nonlinearity mitigation. Firstly, the problem $\mathcal{P}_1$ is separated into two subproblems with respect to the discrete and continuous variables, respectively. Then, we apply a metric to measure the corresponding participation qualification of each AP and transform the complex multi-AP combinatorial problem into a binary linear programming problem. In addition, an improved fractional programming is also employed to transform the non-convex clipping and power allocation subproblem into a series of constrained linear programming problems. Finally, the optimization of $\boldsymbol{\beta}$, $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$ are alternately solved by the time-efficient algorithms and updated within sequential iterations.

Specifically, the AP selection subproblem is shown as

$$\max_{\boldsymbol{\beta}} \sum_{m=1}^{M} w_m R_m^{\bar{\boldsymbol{\sigma}}_s, \bar{\boldsymbol{\zeta}}} \tag{26a}$$

$$\text{s.t.} \quad (25b), (25c), (25d), (25e) \tag{26b}$$

where $R_m^{\bar{\boldsymbol{\sigma}}_s, \bar{\boldsymbol{\zeta}}}$ is the achievable rate of $u_m$ for the fixed $\bar{\boldsymbol{\sigma}}_s$ and $\bar{\boldsymbol{\zeta}}$. Subsequently, the clipping and power allocation subproblem is expressed by

$$\max_{\boldsymbol{\sigma}_s, \boldsymbol{\zeta}} \mathcal{F}(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}) = \max_{\boldsymbol{\sigma}_s, \boldsymbol{\zeta}} \sum_{m=1}^{M} w_m R_m^{\bar{\boldsymbol{\beta}}} \tag{27a}$$

$$\text{s.t.} \quad (25f), (25g), (25h), (25i) \tag{27b}$$

where $R_m^{\bar{\boldsymbol{\beta}}}$ denotes the corresponding achievable rate of $u_m$ on the condition of a fixed $\bar{\boldsymbol{\beta}}$.

### A. AP Selection

As for a specific $u_m$, the potential AP candidates can be roughly divided into two categories. One is that the AP has only a LOS link to $u_m$ but zero channels to other UEs, and the other one is that the AP also delivers the LOS links to multiple UEs in addition to $u_m$. Therefore, the first category must be recruited by $u_m$ because its participation will improve the associated performance of $u_m$ significantly whereas the second category should be carefully treated for recruiting or switching off. As for the multi-AP association case, i.e., $|\mathcal{N}_m| > 1$, the

problem (26) can be considered as a combinatorial optimization for choosing the most appropriate subset for each UE so as to achieve the maximization of the transmission rate at the given $\bar{\boldsymbol{\sigma}}_s$ and $\bar{\boldsymbol{\zeta}}$. However, the corresponding solution is not as straightforward as that of the single-AP case. For one thing, the cardinality of $\mathcal{N}_m$ is essentially different for each UE and it is difficult to manually determine the optimal one just based on the empirical trials. For another, even though the number of APs is not very large, examining all combinations will lead to a considerably large solution space, which will introduce unacceptable computational complexity.

To this end, we introduce a metric to approximately evaluate the participating eligibility for each AP. Let $\Gamma_{m,n}$ denote the contribution ratio of $a_n$ towards the association for $u_m$, which can be calculated by

$$\Gamma_{m,n} = \frac{h_{m,n}^2 \bar{\sigma}_{s,n}^2}{\sum_{l \neq m, l=1}^{M} h_{l,n}^2 \bar{\sigma}_{s,n}^2}. \tag{28}$$

High strength of $\Gamma_{m,n}$ indicates that $a_n$ provides more useful signal to $u_m$ than the interference offered to the other UEs, and vice versa. Then, the average contribution of $a_n$ can be obtained by

$$\tilde{\Gamma}_n = \frac{1}{M} \sum_{m=1}^{M} \Gamma_{m,n}. \tag{29}$$

As for $u_m$, the AP whose contribution ratio is higher than the average value is encouraged to be a reliable candidate. Therefore, a criterion is adopted to select the potential serving APs for each UE, shown as

$$\mathcal{A}_m = \left\{ n : \Gamma_{m,n} > \tilde{\Gamma}_n \right\}. \tag{30}$$

As a result, the searching space can be reduced to a certain extent as compared with the original set. To determine the optimal cardinality for each UE, an auxiliary variable $\eta_{m,q} \in \{0,1\}$ is employed to denote the active flag of association mode $q \in \mathcal{Q}_m$ ( $\mathcal{Q}_m = \{1, 2, \cdots, |\mathcal{A}_m|\}$), where $\eta_{m,q} = 1$ indicates that the number of APs serving $u_m$ is $q$, and $\eta_{m,q} = 0$ otherwise. Furthermore, $\beta_{m,n}$ is evolved into $\beta_{m,n,q}$ to denote the association behavior, where $\beta_{m,n,q} = 1$ indicates that $a_n$ is associated with $u_m$ at the $q$th association mode, and $\beta_{m,n,q} = 0$ otherwise. Hence, the achievable rate of $u_m$ can be considered as the sum rate of all feasible modes, which can be reformulated by

$$\tilde{R}_m^{\bar{\boldsymbol{\sigma}}_s, \bar{\boldsymbol{\zeta}}} = B \sum_{q=1}^{|\mathcal{A}_m|} \log_2 \left(1 + \tilde{\mathrm{I}}_m^q\right), \tag{31}$$

where $\tilde{\mathrm{I}}_m^q$ is the SIDNR when the $q$th association mode is employed for $u_m$, which can be expressed as

$$\tilde{\mathrm{I}}_m^q = \frac{\sum_{n=1}^{N} \beta_{m,n,q} h_{m,n}^2 \alpha_n^2 \bar{\sigma}_{s,n}^2}{\Xi_m^q + \sigma_{z_m}^2}, \tag{32}$$

where $\Xi_m^q$ is the interference and nonlinear distortion seen at $u_m$, shown as

$$\Xi_m^q = \sum_{q=1}^{|\mathcal{A}_m|} \sum_{l \neq m, l=1}^{M} \sum_{j=1}^{N} \beta_{l,j,q} h_{m,j}^2 C_j \bar{\sigma}_{s,j}^2 + \sum_{n=1}^{N} \beta_{m,n,q} h_{m,n}^2 G_n \bar{\sigma}_{s,n}^2. \tag{33}$$

Therefore, the problem in (26) can be reformulated as

$$\mathcal{P}_2 : \max_{\beta_{m,n,q}} \sum_{m=1}^{M} w_m \tilde{R}_m^{\bar{\boldsymbol{\sigma}}_s, \bar{\boldsymbol{\zeta}}} \tag{34a}$$

$$\text{s.t.} \quad \sum_{m=1}^{M} \beta_{m,n,q} \leq 1, \ \forall n \in \mathcal{A}_m, \ q \in \mathcal{Q}_m \tag{34b}$$

$$\sum_{n=1}^{N} \beta_{m,n,q} + \left(1 - \eta_{m,q}\right) q = q, \quad \forall m, q \in \mathcal{Q}_m \tag{34c}$$

$$\sum_{q=1}^{|\mathcal{A}_m|} \eta_{m,q} \leq 1, \quad \forall m \tag{34d}$$

$$\eta_{m,q} = \{0,1\}, \quad \forall m, q \in \mathcal{Q}_m \tag{34e}$$

$$\beta_{m,n,q} = \{0,1\}, \quad \forall m, n \in \mathcal{A}_m, q \in \mathcal{Q}_m \tag{34f}$$

$$\beta_{m,n,q} = 0, \quad \forall m, n \notin \mathcal{A}_m, q \in \mathcal{Q}_m. \tag{34g}$$

For each mode, constraint (34b) ensures that an AP can only participate in serving of one UE at most, and constraint (34c) guarantees that the UE can only be served by a total $q$ APs once the mode $q$ is selected. Besides, constraint (34d) requires that only one multi-AP association mode can be deployed for one UE or none of them are employed. What's more, constraint (34f) and (34g) require that only the AP demonstrating the dominate eligibility can be selected. Therefore, the subproblem (26) becomes a binary nonlinear programming problem $\mathcal{P}_2$. In this paper, branch and bound is employed to search the optimal solution. The details can be found in [33], which are not further elaborated here. Finally, $\beta_{m,n}$ can be obtained from the optimized $\beta_{m,n,q}$ by merging all association modes since only one mode can employed for one UE at most.

With the aid of participating eligibility, the sparsity structure of the solution space is well exploited and the search range can be narrowed to a certain extent. In a sense, the employed $\mathcal{A}_m$ and $\mathcal{Q}_m$ can also provide priors for branch search, which would speed up the convergence of the entire branch boundary. Moreover, the multi-AP association can be jointly optimized for all UEs and the amount of recruited APs in serving of a specific UE can be adaptively configured as considering the relative interference pattern of the whole system, rather than the preset integer shown in the original problem.

### B. Adaptive Clipping and Power Allocation

For a given $\bar{\boldsymbol{\beta}}$, $\mathcal{F}(\boldsymbol{\sigma}_s, \boldsymbol{\zeta})$ is generally considered as a non-convex function. Solving these non-convex optimization problems in a direct manner is computationally intractable in a reasonable time. However, as shown in subproblem (27), each logarithm term contains a ratio of two functions, which are nondecreasing and concave for $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$. Therefore, we employ an improved fractional programming based on a Lagrangian dual transform and quadratic transform to recast the original (27) as a sequence of simple constrained linear programming problems, and then to find its local optimal solution with low computational complexity.

Based on a Lagrangian dual transform, the subproblem (27) can be rewritten as

$$\max_{\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}} f_1\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}\right) \tag{35a}$$

$$\text{s.t.} \quad (25\text{f}), (25\text{g}), (25\text{h}), (25\text{i}), \tag{35b}$$

where $\boldsymbol{\gamma} = [\gamma_1, \cdots, \gamma_M]^T$ refers to a collection of auxiliary variables. In addition, $f_1\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}\right)$ is given by

$$f_1\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}\right) = \sum_{m=1}^{M} w_m \left(\log_2\left(1 + \gamma_m\right) - \frac{\gamma_m}{\ln 2}\right) + \sum_{m=1}^{M} \frac{w_m}{\ln 2} \left(\gamma_m + 1\right) \Theta_m, \tag{36}$$

where

$$\Theta_m = \frac{\Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)}{\Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)}, \tag{37}$$

and $\Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right) = \Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right) + \Omega_{m,2}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)$.

*Lemma 1:* The maximum objective values of (27) and (35) are the same.

*Proof 1:* The proof is given in Appendix I. ∎

As shown in (36), $\Theta_m$ involved in the last item of $f_1\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}\right)$ contains a sum-of-ratio form, which involves the variables of $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$. To this end, we invoke the quadratic transform to decouple the denominator and the numerator from the multiple fractional terms. Specifically, problem (35) can be further transformed as

$$\mathcal{P}_3 : \max_{\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu}} f_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu}\right) \tag{38a}$$

$$\text{s.t.} \quad (25\text{f}), (25\text{g}), (25\text{h}), (25\text{i}), \tag{38b}$$

where $\boldsymbol{\mu} = [\mu_1, \cdots, \mu_M]^T$ is the auxiliary variables and $f_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu}\right)$ is given as

$$f_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu}\right) = \sum_{m=1}^{M} w_m \log_2\left(1 + \gamma_m\right) - \frac{1}{\ln 2} \sum_{m=1}^{M} w_m \gamma_m + \frac{1}{\ln 2} \sum_{m=1}^{M} 2\mu_m \sqrt{\varpi_m \Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)} - \frac{1}{\ln 2} \sum_{m=1}^{M} \mu_m^2 \Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right), \tag{39}$$

where $\varpi_m = w_m\left(1 + \gamma_m\right)$. As $\boldsymbol{\sigma}_s$ is fixed, $\zeta_n$ is mainly determined by $\vartheta_{\max}$ and $\sigma_{s,n}^2$ of $a_n$. In addition, the problem $\mathcal{P}_3$ is an objective maximization over variables $\boldsymbol{\sigma}_s$, $\boldsymbol{\zeta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\mu}$. Fortunately, it can be verified that $f_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu}\right)$ is concave with respect to those variables. Therefore, the optimization of $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$ can be done alternately at a given $\boldsymbol{\gamma}$ and $\boldsymbol{\mu}$. To solve this problem, we propose an iterative algorithm to sequentially optimize $\boldsymbol{\gamma}$, $\boldsymbol{\mu}$, $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$ in an alternative manner while fixing the others.

Firstly, according to the Karush Kuhn Tucker (KKT) conditions, $\gamma_m$ can be updated by solving the following equation as all the other variables are fixed, shown as

$$\frac{\partial f_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu}\right)}{\partial \gamma_m} = 0, \quad \forall m = 1, 2, \cdots, M. \tag{40}$$

After simplifying, $\gamma_m^*$ can be calculated by

$$\gamma_m^* = \frac{\mu_m^2 \Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)}{2w_m}$$
$$+ \frac{\sqrt{\left(\mu_m^2 \Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)\right)^2 + 4w_m \mu_m^2 \Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)}}{2w_m}. \tag{41}$$

By substituting (41) into (39), the optimal $\mu_m$ can be updated by setting

$$\frac{\partial f_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}\right)}{\partial \mu_m} = 0, \quad \forall m = 1, 2, \cdots, M. \tag{42}$$

Thus, $\mu_m^*$ can be calculated by

$$\mu_m^* = \frac{\sqrt{\varpi_m^* \Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)}}{\Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)}, \tag{43}$$

where $\varpi_m^* = w_m\left(1 + \gamma_m^*\right)$.

After substituting $\boldsymbol{\gamma}^*$ and $\boldsymbol{\mu}^*$ into (39), the first two terms of $f_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right)$ can be regarded as constants that are only related with $\boldsymbol{\gamma}^*$. Then, we use $\hat{f}_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right)$ to denote the components of the last two terms in $f_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right)$, which can be expressed by

$$\hat{f}_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right) = \frac{1}{\ln 2} \sum_{m=1}^{M} 2\mu_m^* \sqrt{\varpi_m^* \Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right)}$$
$$- \frac{1}{\ln 2} \sum_{m=1}^{M} \left(\mu_m^*\right)^2 \Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta}\right). \tag{44}$$

Then, the optimal $\boldsymbol{\sigma}_s$ can be derived by solving

$$\boldsymbol{\sigma}_s^* = \arg\max_{\boldsymbol{\sigma}_s} \hat{f}_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right). \tag{45}$$

Let $\frac{\partial}{\partial \boldsymbol{\sigma}_s} \hat{f}_2\left(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right) = 0$, we can obtain the following series of equations, which can be expressed by

$$\sum_{n=1}^{N} \bar{\beta}_{m,n} h_{m,n}^2 \alpha_n^2 \sigma_{s,n}^2 = \left(\xi_m^*\right)^2, \quad \forall m = 1, 2, \cdots, M, \tag{46}$$

where $\xi_m^*$ can be explicitly calculated as

$$\xi_m^* = \frac{\sqrt{\varpi_m^*} \sum\limits_{n=1}^{N} \bar{\beta}_{m,n} h_{m,n}^2 \alpha_n^2}{\mu_m^* \left(\sum\limits_{l \neq m, l=1}^{M} \sum\limits_{j=1}^{N} \bar{\beta}_{l,j} h_{m,j}^2 C_j + \sum\limits_{n=1}^{N} \bar{\beta}_{m,n} h_{m,n}^2 \left(G_n + \alpha_n^2\right)\right)}. \tag{47}$$

Let $\boldsymbol{\xi} = \left[\left(\xi_1^*\right)^2, \cdots, \left(\xi_M^*\right)^2\right]^T$, $\mathbf{V}^a = \left[\mathbf{v}_1^a, \cdots, \mathbf{v}_N^a\right]$ and $\mathbf{v}_n^a = \left[\bar{\beta}_{1,n} h_{1,n}^2 \alpha_n^2, \cdots, \bar{\beta}_{M,n} h_{M,n}^2 \alpha_n^2\right]^T$, the solution of (46) is equivalent to the following

$$\boldsymbol{\sigma}_s^* = \arg\min_{\boldsymbol{\sigma}_s} \left\|\mathbf{V}^a \boldsymbol{\sigma}_s - \boldsymbol{\xi}\right\|_2 \tag{48a}$$

$$\text{s.t. } (25\text{f}), (25\text{g}), (25\text{i}). \tag{48b}$$

Note that, (48) is a constrained linear programming problem and it can be efficiently solved by the existing algorithms with polynomial-time computational complexity. In addition, the optimized $\boldsymbol{\sigma}_s^*$ obtained by the problem (48) exactly yields the same solution as that of the problem (45).

With fixed $\boldsymbol{\gamma}^*$, $\boldsymbol{\mu}^*$ and $\boldsymbol{\sigma}_s$, the optimization of $\boldsymbol{\zeta}$ is jointly related with $\Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\zeta}\right)$ and $\Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\zeta}\right)$, which are monotonically nondecreasing and concave over $\zeta_n$ as $\zeta_n \geq 1$.

As previously shown, the corresponding attenuation factors in terms of $\alpha_n^2$, $C_n$ and $G_n$ involved in $\Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\zeta}\right)$ and $\Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\zeta}\right)$ are monotonic and highly nonlinear functions for $\zeta_n \geq 1$. Although the optimal clipping ratio can be determined by using numerical methods such as the exhaustive or golden search for the associated APs that maximizes the $f_2\left(\boldsymbol{\sigma}_s^*, \boldsymbol{\zeta}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right)$. However, it will lead to prohibitive computational complexity and cannot guarantee the optimality for all UC-cells. Therefore, we could employ a lower bound approximation that can be analytically characterized to thereby reduce the computational complexity to some extent. To this end, we further simplify $\Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\zeta}\right)$ by combining (9), (11) and (14), shown as

$$\Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\zeta}\right) = \sum_{l \neq m, l=1}^{M} \sum_{j=1}^{N} \bar{\beta}_{l,j} h_{m,j}^2 C_j \left(\sigma_{s,j}^*\right)^2$$
$$+ \sum_{n=1}^{N} \bar{\beta}_{m,n} h_{m,n}^2 \left(G_n + \alpha_n^2\right) \left(\sigma_{s,n}^*\right)^2 + \sigma_{z_m}^2$$
$$= \sum_{l=1}^{M} \sum_{n=1}^{N} \bar{\beta}_{l,n} h_{m,n}^2 C_n \left(\sigma_{s,n}^*\right)^2$$
$$+ \sum_{n=1}^{N} \bar{\beta}_{m,n} h_{m,n}^2 \left(G_{\text{DAC},n} + G_{\text{ADC},n}\right) \left(\sigma_{s,n}^*\right)^2 + \sigma_{z_m}^2. \tag{49}$$

As $\zeta_n \geq 1$, it can be verified that $C_n$ is monotonically nondecreasing and concave, while the sum of $G_{\text{DAC},n}$ and $G_{\text{ADC},n}$ is monotonically increasing and convex. Moreover, as $\zeta_n$ is increasing, $C_n$ gradually approaches $\alpha_n^2$ and the average error between these two terms approximates an order of $10^{-6}$ for $\zeta_n \geq 3$. In addition, $C_n$ also shows much greater than the sum of $G_{\text{DAC},n}$ and $G_{\text{ADC},n}$, e.g., the former is $0.995$ whereas the latter is $1.4 \times 10^{-4}$ at $\zeta_n = 3$. Hence, we consider the upper band of $\Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\zeta}\right)$, shown as

$$\Omega_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\zeta}\right) \leq \tilde{\Omega}_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\zeta}\right)$$
$$\triangleq \sum_{l=1}^{M} \sum_{n=1}^{N} \bar{\beta}_{l,n} h_{m,n}^2 \alpha_n^2 \left(\sigma_{s,n}^*\right)^2$$
$$+ \sum_{n=1}^{N} \bar{\beta}_{m,n} h_{m,n}^2 \left(G_{\text{DAC},n}^{\max} + G_{\text{ADC},n}^{\max}\right) \left(\sigma_{s,n}^*\right)^2 + \sigma_{z_m}^2, \tag{50}$$

where $G_{\text{DAC},n}^{\max}$ and $G_{\text{ADC},n}^{\max}$ are the values considered at $\zeta_{\max}$. Due to $\alpha_n^2$ being monotonically increasing and concave over $\zeta_n$, we introduce the auxiliary variable $\chi_n$ to replace $\alpha_n^2$. Then, the optimization of the clipping ratio can be reduced to the following problem as

$$\boldsymbol{\chi}^* = \arg\max_{\boldsymbol{\chi}} \widehat{f}_2\left(\boldsymbol{\sigma}_s^*, \boldsymbol{\chi}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right), \tag{51}$$

where $\widehat{f}_2\left(\boldsymbol{\sigma}_s^*, \boldsymbol{\chi}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right)$ is shown by

$$\widehat{f}_2\left(\boldsymbol{\sigma}_s^*, \boldsymbol{\chi}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right) = \sum_{m=1}^{M} 2\mu_m^* \sqrt{\varpi_m^* \Omega_{m,1}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\chi}\right)}$$
$$- \sum_{m=1}^{M} \left(\mu_m^*\right)^2 \tilde{\Omega}_{m,3}\left(\bar{\boldsymbol{\beta}}, \boldsymbol{\sigma}_s^*, \boldsymbol{\chi}\right). \tag{52}$$

It can be verified that (51) is a convex problem over $\boldsymbol{\chi}$. Let $\frac{\partial}{\partial \boldsymbol{\chi}} \widehat{f}_2\left(\boldsymbol{\sigma}_s^*, \boldsymbol{\chi}, \boldsymbol{\gamma}^*, \boldsymbol{\mu}^*\right) = 0$, we can obtain

$$\sum_{n=1}^{N} \bar{\beta}_{m,n} h_{m,n}^2 \left(\sigma_{s,n}^*\right)^2 \chi_n = \left(\upsilon_m^*\right)^2, \quad \forall m = 1, 2, \cdots, M, \quad (53)$$

where $\upsilon_m^*$ can be explicitly calculated by

$$\upsilon_m^* = \frac{\sqrt{\varpi_m^*} \sum\limits_{n=1}^{N} \bar{\beta}_{m,n} h_{m,n}^2 \left(\sigma_{s,n}^*\right)^2}{\mu_m^* \sum\limits_{l=1}^{M} \sum\limits_{j=1}^{N} \bar{\beta}_{l,j} h_{m,j}^2 \left(\sigma_{s,j}^*\right)^2}. \quad (54)$$

Let $\mathbf{v}_n^b = \left[\bar{\beta}_{1,n} h_{1,n}^2 \left(\sigma_{s,n}^*\right)^2, \cdots, \bar{\beta}_{M,n} h_{M,n}^2 \left(\sigma_{s,n}^*\right)^2\right]^T$, $\mathbf{V}^b = \left[\mathbf{v}_1^b, \cdots, \mathbf{v}_N^b\right]$, and $\boldsymbol{\upsilon} = \left[\left(\upsilon_1^*\right)^2, \cdots, \left(\upsilon_M^*\right)^2\right]^T$, then the optimized $\boldsymbol{\chi}^*$ can be determined by solving the following

$$\boldsymbol{\chi}^* = \arg\min_{\boldsymbol{\chi}} \ \left\|\mathbf{V}^b \boldsymbol{\chi} - \boldsymbol{\upsilon}\right\|_2 \quad (55a)$$

$$\text{s.t. } \operatorname{erf}^2\left(\frac{1}{\sqrt{2}}\right) \leq \chi_n \leq \operatorname{erf}^2\left(\frac{\zeta_{\max}}{\sqrt{2}}\right), \quad \forall n \quad (55b)$$

$$\operatorname{erf}^2\left(\frac{\vartheta_{\min}}{2\sqrt{2}\sigma_{s,n}^*}\right) \leq \chi_n \leq \operatorname{erf}^2\left(\frac{\vartheta_{\max}}{2\sqrt{2}\sigma_{s,n}^*}\right), \quad \forall n. \quad (55c)$$

After obtaining $\chi_n^*$, $\zeta_n^*$ can be obtained by back extrapolation, shown as

$$\zeta_n^* = \sqrt{2}\operatorname{erf}^{-1}\left(\sqrt{\chi_n^*}\right), \quad \forall n = 1, 2, \cdots, N, \quad (56)$$

where $\operatorname{erf}^{-1}\left(\cdot\right)$ is the inverse of $\operatorname{erf}\left(\cdot\right)$ and has the Maclaurin expression as

$$\operatorname{erf}^{-1}\left(z\right) = \sum_{k=0}^{\infty} \frac{b_k}{2k+1} \left(\frac{\sqrt{\pi}}{2} z\right)^{2k+1}, \quad (57)$$

where $b_0 = 1$ and

$$b_k = \sum_{k_0=0}^{k-1} \frac{b_{k_0} b_{k-1-k_0}}{(k_0+1)(2k_0+1)}. \quad (58)$$

The detailed description about the above iterative procedure is also shown in *Algorithm 1*. The rationale behind it is alternately optimizing $\boldsymbol{\gamma}$, $\boldsymbol{\mu}$, $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$ while fixing the others until the stationary points are found, which satisfy the KKT conditions of the problem $\mathcal{P}_3$.

*Lemma 2:* The *Algorithm 1* is guaranteed to be convergent and monotonically nondecreasing after each iteration, and the local optimal solutions can be achieved along with the convergence.

*Proof 2:* The proof is given in Appendix II. ∎

### C. Joint Optimization

After that, the binary variable $\boldsymbol{\beta}$ and continuous variables $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$ can be optimized in an alternating way, the details are shown in *Algorithm 2*. Firstly, the optimized $\boldsymbol{\beta}^{(t+1)}$ is obtained by solving the AP selection subproblem with fixed $\boldsymbol{\sigma}_s^{(t)}$ and $\boldsymbol{\zeta}^{(t)}$ at each iteration. Then, the optimized $\boldsymbol{\sigma}_s^{(t+1)}$ and $\boldsymbol{\zeta}^{(t+1)}$ are found via solving the subproblem (27) based on the given $\boldsymbol{\beta}^{(t+1)}$. Note that the optimizing procedure for the clipping ratio and IC power is just executed based on the fixed $\bar{\boldsymbol{\beta}}$, which

---

**Algorithm 1** Adaptive clipping and IC power allocation with the specified association factor.

**Input:**
1) Association decision $\bar{\beta}_{m,n}$
2) Weight $w_m$, channel gain $h_{m,n}^2$, background noise $\sigma_{z_m}^2$
3) Threshold $\Delta_{\min}$

1: **Initialization**
2: Set $\boldsymbol{\zeta}^{(1)}$ and $\boldsymbol{\sigma}_s^{(1)}$ with feasible values. Initialize $\left(\gamma_m^*\right)^{(1)} = \tilde{\mathrm{I}}_m$ and $t = 1$, respectively.
3: **while** 1 **do**
4:     Obtain $\left(\mu_m^*\right)^{(t+1)}$ via (43).
5:     Update $\left(\gamma_m^*\right)^{(t+1)}$ via (41).
6:     Update $\left(\xi_m^*\right)^{(t+1)}$ via (47).
7:     Update $\mathbf{V}^a$ and Obtain $\left(\boldsymbol{\sigma}_s^*\right)^{(t+1)}$ by solving the problem (48).
8:     Obtain $\left(\upsilon_m^*\right)^{(t+1)}$ via (54) and update $\mathbf{V}^b$.
9:     Obtain $\boldsymbol{\chi}^*$ by solving the problem (55).
10:    Update $\left(\zeta_n^*\right)^{(t+1)}$ via (56) based on $\boldsymbol{\chi}^*$.
11:    Check the halting criterion:
12:    **if** $\left|f_2^{(t+1)} - f_2^{(t)}\right| \leq \Delta_{\min}$ **then**
13:       Break.
14:    **end if**
15:    $t = t + 1$
16: **end while**
17: Obtain $\boldsymbol{\sigma}_s^* \Leftarrow \left(\boldsymbol{\sigma}_s^*\right)^{(t+1)}$ and $\boldsymbol{\zeta}^* \Leftarrow \left(\boldsymbol{\zeta}^*\right)^{(t+1)}$, respectively.

**Output:** $\boldsymbol{\sigma}_s^*$, and $\boldsymbol{\zeta}^*$

---

is sensitive to the initial condition. To guarantee the solving quality, random initial points are employed for $\boldsymbol{\beta}^{(1)}$, while the uniform power loading and clipping ratio are set for $\boldsymbol{\sigma}_s^{(1)}$ and $\boldsymbol{\zeta}^{(1)}$, respectively. At last, the solving results are averaged out. Here we set the empirical $\Delta_{\min} = 1 \times 10^{-3}$ for checking the halting conditions. By employing the proposed methodology, the UC-cells can be dynamically constructed according to the spatial distribution of all UEs, and the adjacent interference and nonlinear impairments would be consciously avoided or mitigated in a subtle way, which can effectively improve the achievable data rate and achieve a good trade-off between the system performance and computational complexity.

*Lemma 3:* The *Algorithm 2* always converges to a local optimum when a stationary point $(\boldsymbol{\beta}^\star, \boldsymbol{\sigma}_s^\star, \boldsymbol{\zeta}^\star)$ is obtained.

*Proof 3:* The proof is given in Appendix III. ∎

### D. Complexity Analysis

In *Algorithm 2*, the calculation of $\boldsymbol{\beta}^{(t+1)}$ is dominant in each iteration. For a given $N$ and $M$, the computational complexity of the optimization of $\boldsymbol{\beta}^{(t+1)}$ is mainly affected by branching and bounding rules since they determine the number of feasible solutions for each branching operation. With the help of the metric of participating eligibility, each UE has the limited potential AP subset $\mathcal{A}_m$, which gives $K_a = \prod\limits_{m=1}^{M} |\mathcal{A}_m|$ possible solutions to evaluate. However, some of them will turn out to be the infeasible solutions since the current assignment violates the constraints of (34b) and (34c), which should be pruned to further reduce the number

**Algorithm 2** Joint optimization of AP selection, clipping ratio and IC power.

**Input:**
1) Weight $w_m$, channel gain $h_{m,n}^2$, background noise $\sigma_{z_m}^2$.
2) Threshold $\Delta_{\min}$.

1: **Initialization**
2: Set $t = 1$, and initialize $\boldsymbol{\sigma}_s^{(1)}$, $\boldsymbol{\zeta}^{(1)}$ and $\boldsymbol{\beta}^{(1)}$ with feasible values, respectively.
3: **while** 1 **do**
4:     Obtain $\Gamma_{m,n}$ and $\tilde{\Gamma}_n$ via (28) and (29), respectively.
5:     Update $\boldsymbol{\beta}^{(t+1)}$ by solving the problem (34) based on the fixed $\boldsymbol{\sigma}_s^{(t)}$ and $\boldsymbol{\zeta}^{(t)}$.
6:     Update $\boldsymbol{\sigma}_s^{(t+1)} \Leftarrow \boldsymbol{\sigma}_s^*$ and $\boldsymbol{\zeta}^{(t+1)} \Leftarrow \boldsymbol{\zeta}^*$ via *Algorithm 1* based on the fixed $\boldsymbol{\beta}^{(t+1)}$.
7:     Calculate $\mathcal{F}\left(\boldsymbol{\beta}^{(t+1)}, \boldsymbol{\sigma}_s^{(t+1)}, \boldsymbol{\zeta}^{(t+1)}\right)$.
8:     Check the halting criterion:
9:     **if** $\left|\mathcal{F}^{(t+1)} - \mathcal{F}^{(t)}\right| \leq \Delta_{\min}$ **then**
10:         Break.
11:     **end if**
12:     $\boldsymbol{\beta}^\star \Leftarrow \boldsymbol{\beta}^{(t+1)}$, $\boldsymbol{\sigma}^\star \Leftarrow \boldsymbol{\sigma}^{(t+1)}$ and $\boldsymbol{\zeta}^\star \Leftarrow \boldsymbol{\zeta}^{(t+1)}$.
13:     $t = t + 1$.
14: **end while**
**Output:** $\boldsymbol{\beta}^\star$, $\boldsymbol{\sigma}_s^\star$ and $\boldsymbol{\zeta}^\star$

of evaluated solutions. Note that the remaining solutions for each UE are also related with the interference pattern of the whole system. We assume that there are remaining total $K_b$ feasible solutions, thus the computational complexity of the optimization $\boldsymbol{\beta}^{(t+1)}$ can be expressed as $\mathcal{O}\left(K_b \log K_b + K_a\right)$, where $K_b \log K_b$ represents the cost of sorting the feasible solutions based on their objective function values. As for the updating of $\boldsymbol{\sigma}_s^{(t+1)}$ and $\boldsymbol{\zeta}^{(t+1)}$, *Algorithm 1* consists of two main steps. In the first step, $\boldsymbol{\mu}^{(t+1)}$, $\boldsymbol{\gamma}^{(t+1)}$, $\boldsymbol{\xi}^{(t+1)}$ and $\boldsymbol{v}^{(t+1)}$ are successively calculated in a recursive manner, which needs to be run at most $M$ times for each iteration. In the second step, $\boldsymbol{\sigma}_s^{(t+1)}$ and $\boldsymbol{\zeta}^{(t+1)}$ can be obtained by solving the constrained linear programming (48) and (55), where the calculation of the Moore-Penrose matrix is needed for each iteration. The time complexity for solving this linear programming depends on the specific used algorithm. Considering the structure of the problem, we could invoke the interior point method where the time complexity is $\mathcal{O}\left(N^3\right)$. In summary, the computational complexity of the whole procedure can be denoted as $\mathcal{O}\left(K_b \log K_b + K_a + N^3 + M\right)$ approximately.

## V. NUMERICAL RESULTS AND DISCUSSIONS

In this section, numerical simulations are conducted to evaluate the capability of the proposed scheme for enhancing the achievable rate of a multi-user VLC network in the presence of nonlinear impairments. A room with a typical size of $6 \times 6 \times 3$ m$^3$ is considered here. A total of 16 APs are uniformly deployed on the ceiling with a horizontal interval of 1.5 m, which are used to provide down-link transmission and illumination. The UEs are randomly distributed in the room, and the receiving plane of each UE has a height of 0.85 m

| Parameters | Values |
|---|---|
| Room Size (m$^3$) | $6 \times 6 \times 3$ |
| Height of receiving plane (m) | 0.85 |
| Semi-angle $\phi_{1/2}$ (deg) | 60 |
| Active area $A_{\text{PD}}$ (cm$^2$) | 1 |
| Optical filter gain | 0.9 |
| Responsivity $R_{\text{PD}}$ (A/W) | 0.53 |
| Reflection factor $\rho$ | 0.8 |
| FOV $\psi_C$ (deg) | 45 |
| Maximum IC power $\sigma_{s,\max}^2$[1] | 1 |
| Total power budget $P_{\text{IC}}$[1] | 5 |
| Available LED range $[\vartheta_{\min}, \vartheta_{\max}]$[1] | [1,4] |
| Modulation bandwidth $B$ (MHz) | 50 |

[1] The IC power and the LED range are normalized with the reference 40 mW and 5 mA, respectively.

above the floor. Furthermore, the classic random waypoint mobility model is adopted to generate the spatial locations of all UEs in each time slot, where the minimum distance between UEs is not less than 1 m. The main simulation parameters of the VLC channel are listed in Table I. It should be noted that the normalized IC power and LED range are employed in this paper to facilitate the mathematical calculations in the optimization procedure[2]. Moreover, the receiver noise PSD with the value of $10^{-22}$ A$^2$/Hz is approximately equivalent to an effective SNR of 30.2 dB when the modulation bandwidth of 50 MHz and the IC power of 40 mW are employed. Therefore, the effective SNR of 30 dB is considered, and both $\sigma_{s,\max}^2$ and $\vartheta_{\min}$ can be set to 1, which are normalized with the reference 40 mW and 5 mA, respectively. In addition, the weight $w_m$ is set to be equal to $\frac{1}{M}$ while the objective function becomes the average rate of VLC network. For simplicity, $Q_{\text{DAC},n} = Q_{\text{ADC},m} = Q_{\text{bit}}$ is the adopted for each AP and UE.

### A. Convergence Performance

To show the feasibility of the proposed scheme, the convergence performance which can be evaluated by the variation of the objective function $\mathcal{F}(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta})$ versus the number of iterations is firstly investigated. As $Q_{\text{bit}} = 10$ and $M$ is varied from 3 to 7, $\mathcal{F}(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta})$ for each iteration is shown in Fig. 2. The results clearly demonstrate that the number of iterations required for reaching to the convergence state is gradually increased with the number of UEs. The main reason is that the computational complexity of the proposed scheme is directly related with the number of APs and UEs, and also with the interference patterns of the whole system derived from the specific distributions of UEs. When the room size and the number of APs are determined, the increase of $M$ will dominant the computational complexity of *Algorithm 1* and *Algorithm 2*. In addition, the interference strength of the whole system is also enhanced to a certain extent and the interference relations between users become more complex with the increase of UEs density. However,

---

[2]The average channel coefficients of indoor VLC are much smaller, which will encounter the calculation failure due to the accuracy limitations of some existing optimization solvers.
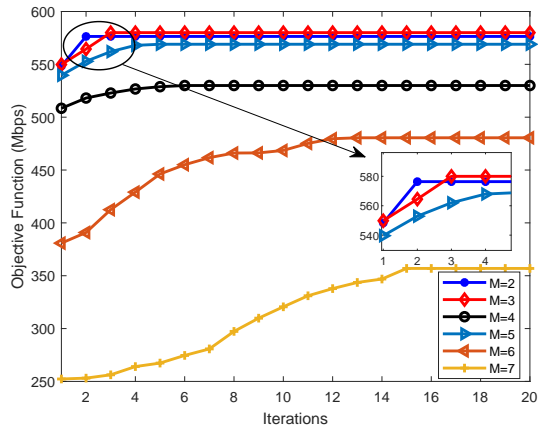
Fig. 2. Convergence comparison for different number of UEs as $Q_{\text{bit}} = 10$.



Fig. 3. Convergence comparison for different bit resolutions as $M = 4$.

the proposed scheme still can achieve favorable convergence performance under such conditions. Moreover, we can also observe that $\mathcal{F}(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta})$ declines as $M$ is increased. This is indeed as expected because more resources are required for simultaneously supporting their target UEs whereas the total resources are restricted.

Fig. 3 illustrates the convergence performance of the proposed scheme under different ADC/DAC bit resolutions as $M = 4$. We consider two scenarios, where *Scenario A* refers to the case of slight interference patterns and the coordinates of UEs are $(1.1, 1.5)$ m, $(4.5, 1.5)$ m, $(1.5, 4.6)$ m and $(4.7, 4.5)$ m, respectively. By comparison, *Scenario B* implies the complex interference case where the coordinates of UEs are $(1.1, 1.5)$ m, $(2.9, 2.1)$ m, $(1.5, 4.2)$ m and $(4.2, 2.9)$ m, respectively. The results demonstrate that all of the curves tend to become stable gradually with the increase of iteration number. In general, the proposed scheme involving slight interference patterns has a faster convergence rate under similar nonlinear quantization distortions. Additionally, the final optimized $\mathcal{F}(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta})$ is just slightly reduced as the lower $Q_{\text{bit}}$ is employed, e.g., less than 20 Mbps throughput reduction is introduced in $Q_{\text{bit}} = 8$ as compared with the case of $Q_{\text{bit}} = 10$ and $Q_{\text{bit}} = 12$, which indicates the effectiveness and reliability of the proposed scheme when facing different nonlinear quantization distortions. Therefore, the proposed scheme could achieve the convergence in a limited time and exhibit excellent interference and nonlinear impairments management capability.

### B. Optimized Stationary Points

Fig. 4 depicts the stationary points of the proposed scheme in terms of $\boldsymbol{\zeta}^{\star}$ and $\boldsymbol{\sigma}_s^{\star}$, which are collected at the convergence state under different $\vartheta_{\max}$. As the results show, the range of the stationary points is obviously distinct for different scenarios. In general, $\boldsymbol{\zeta}^{\star}$ and $\boldsymbol{\sigma}_s^{\star}$ in *Scenario A* fall within a relatively wider range than that of *Scenario B*. As shown in the case of $\vartheta_{\max} = 3.5$, the range of $\boldsymbol{\zeta}^{\star}$ and $\boldsymbol{\sigma}_s^{\star}$ in *Scenario A* are $[3.1006, 3.2484]$ and $[0.2902, 0.3186]$, respectively, where the range size is nearly 4 times wider than that of *Scenario B*,
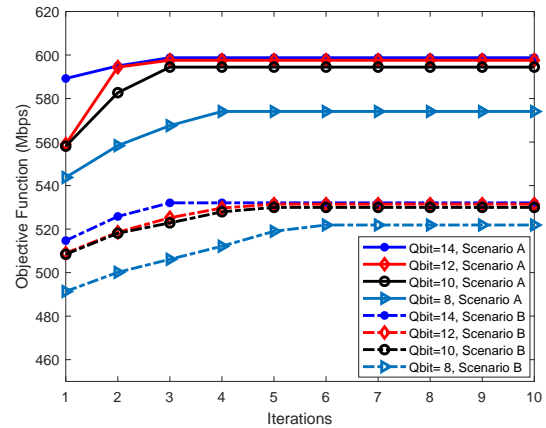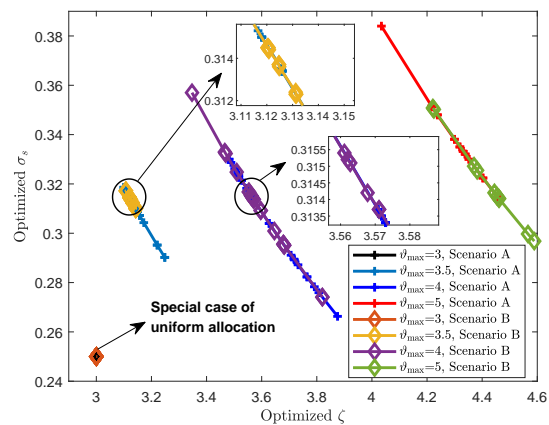


Fig. 4. The stationary points in terms of $\boldsymbol{\zeta}^{\star}$ and $\boldsymbol{\sigma}_s^{\star}$ under different $\vartheta_{\max}$.

i.e., $[3.1072, 3.1431]$ and $[0.3100, 0.3172]$, respectively. In addition, the range sizes is decreased as lower $\vartheta_{\max}$ is employed, and it will turn out to be one point for $\vartheta_{\max} = 3$. In this case, we conjecture that the available linear region of each AP is confined to be so small that the proposed algorithms converge to the same stationary points with high probability under the current power budget. Thus, all APs are implemented with the same clipping and power loading, which is a special case of uniform resource allocation. We know that a wider range of optimized variables enables the whole system to have a more flexible adjustment margin to achieve the desired performance, which can also indicate that the proposed scheme is superior to the uniform strategies. Furthermore, the smooth curves also mean that there exists a special functional relationship between the optimized $\boldsymbol{\zeta}^{\star}$ and $\boldsymbol{\sigma}_s^{\star}$ obtained by the proposed scheme.

In addition, a snapshot of the optimized $\boldsymbol{\beta}^{\star}$ with $\vartheta_{\max} = 4$ in *Scenario B* is shown in Fig. 5. According to their distribution, we know that some UEs are spatially close to each other, which will introduce complex and severe mutual interference when the conventional fixed-shape attocell was employed in the system. However, the proposed scheme can associate all UEs with their optimal AP candidates rather than the nearest one by comprehensively assessing the overall interference and
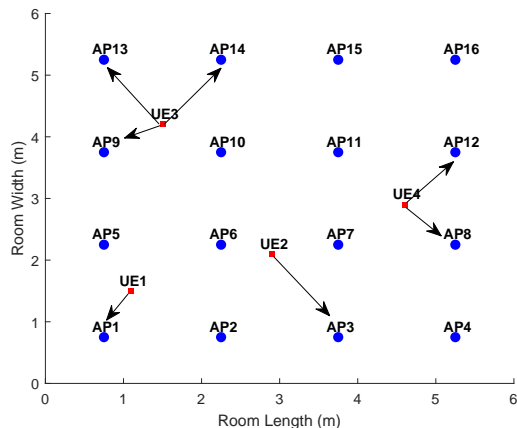
Fig. 5.  A snapshot of AP selection in Scenario B with $\vartheta_{\max} = 4$ .



Fig. 6.  The achieved sum rate comparison under different FOVs.

nonlinear impairments of the whole network. In addition, combined with adaptive clipping and power loading, the 4 UEs can obtain the achieved throughput as 532.1, 483.1, 554.7 and 562 Mbps, respectively, which indicates that the potential mutual interference of the whole system can be well managed and balanced to some extent for guaranteeing the system throughput, since the maximum performance gap among all UEs is less than 17%. Therefore, the UC-cells can be dynamically constructed by the proposed scheme with the optimized $\beta^\star$, $\zeta^\star$ and $\sigma_s^\star$ to make a good compromise among the interference and achieved throughput of the whole system.

### C. Achieved Sum Rate

The FOV and total IC power budget are important parameters in the VLC system which can directly affect the channel gain and also the interference patterns. In this section, the effect of various FOVs and total IC power budget on the achievable sum rate is investigated under the case of *Scenario B*. To validate the effectiveness of the proposed scheme, three traditional network formations are also considered here for comparison:

a) *Benchmark 1*: Multiple AP association based on distance metric, where the neighboring APs of one UE within a certain area are merged to be a UC-cell.

b) *Benchmark 2*: Single AP association based on fixed-shape cell, where the coverage of each AP is regarded as one attocell, and the UE will be connected to its nearest AP.

c) *Benchmark 3*: Multiple AP association based on fixed-shape cells, where the whole room is divided into $2 \times 2$ square cells and each cell contains 4 APs. In contrast to *Benchmark 1*, the corresponding combination of APs for each cell keep unchanged in *Benchmark 3*.

In addition, the uniform clipping and power loading in terms of $\zeta = 3$ and $\sigma_s = \frac{P_{\text{IC}}}{N}$ are used for the above three benchmark schemes.

Fig. 6 presents the sum rate performance comparison of the proposed scheme and three benchmark schemes, where $\psi_C$ is varied from 25 to 75 degrees and $P_{\text{IC}} = 5$. As shown in the figure, the sum rate of all curves is decreased when larger
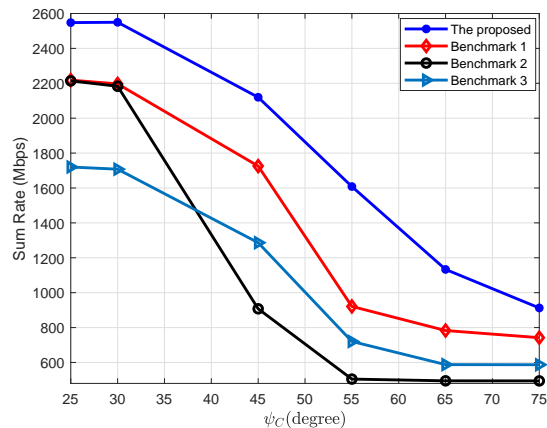
$\psi_C$ is employed, because the wide $\psi_C$ may drive a specific UE to receive more interference signals from the adjacent APs that serve the other UEs. However, the proposed scheme still exhibits an excellent sum rate performance for all benchmark schemes, and can achieve a significant gain especially for $\psi_C = 55$ degrees, which confirms its robustness against the inevitably interference at condition of larger FOVs.

Fig. 7 illustrates the corresponding sum rate performance comparison as $P_{\text{IC}}$ is varied from 2 to 10 and $\psi_C = 45$ degrees. As for the three benchmark schemes, their sum rate performance gradually increase and then sharply decline on the condition of $P_{\text{IC}} > 7$. Because the uniform power loading strategy is deployed in the three benchmark schemes, more power will be allocated to the modulated signal consequently as $P_{\text{IC}}$ is increasing. Thus, the majority part of the modulated signal will be clipped accordingly due to the limited available linear region of the LED, resulting in severe nonlinear distortions. It is a fact that the interference strength and nonlinear distortions are directly related to $P_{\text{IC}}$. The larger $P_{\text{IC}}$ employed, the higher the interference and the nonlinear distortions introduced. However, the proposed scheme just has a slight performance degradation as $P_{\text{IC}} > 7$, and its sum rate will gradually reach a stable state, which shows the robustness of the proposed scheme to the nonlinear distortions. Therefore, the proposed scheme can still work effectively to provide the robust sum rate performance even under severe interference and nonlinear impairments.

### D. Statistical Distribution

In this section, the statistical performance in terms of the cumulative distribution function (CDF) of the achieved sum rate are investigated. Taking $M = 4$ and $P_{\text{IC}} = 5$, the CDFs of the proposed and the other three benchmark schemes are illustrated in Fig. 8, where the position of randomly located UEs is employed and total 200 snapshots are used. It can be seen that the proposed method exhibits a better sum rate performance as expected because it outperforms all benchmark schemes for the majority of UE distribution. We know that the benchmark schemes employ the association structure with the fixed $\zeta$ and $\sigma_s$ without considering the relative spatial
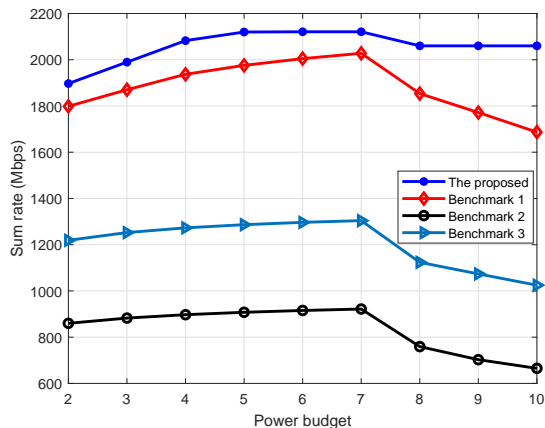
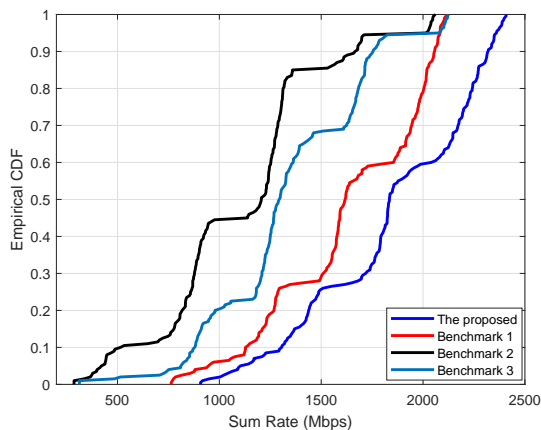Fig. 7. The achieved sum rate comparison under different power budgets.



Fig. 8. The CDF comparison of the proposed and the benchmark schemes.

distribution of all UEs. However, as for the proposed scheme, the candidate APs for each UE and the corresponding $\boldsymbol{\zeta}$ and $\boldsymbol{\sigma}_s$ are dynamically adjusted by considering the spatial distribution of all UEs comprehensively, which can alleviate the potential interference and nonlinear impairments effectively, resulting in an significant sum rate performance improvement. In summary, the proposed scheme has superiority over the traditional schemes and shows the validity in the UC-cells formation from the above simulations and discussions.

## VI. CONCLUSION

In this paper, a novel UC-cell formation based on the joint design of AP selection, clipping ratio and IC power allocation was proposed to achieve the sum rate maximization under the concerns of nonlinear impairments of multi-user VLC system. We formulated this designing as a mixed combinational and non-convex optimization problem and proposed an efficient method to obtain the local optimal solution. Specifically, the discrete variable and continuous variables were separated into two subproblems, and we developed a method to transform the complex subproblems to be binary nonlinear programming and constrained linear programming problems, respectively, which can be efficiently solved in relatively low-complexity manners.

Results demonstrated that the proposed scheme can effectively mitigate the interference and achieve significant performance gain. In addition, it also exhibited a relative robustness to the severe nonlinear impairments, outperforming the existing conventional methods when compared at the same conditions.

## APPENDIX I
## PROOF OF LEMMA 1

As $\boldsymbol{\sigma}_s$ is held to $\bar{\boldsymbol{\sigma}}_s$ and $\boldsymbol{\zeta}$ is fixed to $\bar{\boldsymbol{\zeta}}_s$, (27) can be reformulated by introducing the variable $\gamma_m$ to replace each ratio term in the logarithm function, shown as

$$\max_{\boldsymbol{\gamma}} \mathcal{F}(\boldsymbol{\gamma}) = \max_{\boldsymbol{\gamma}} \sum_{m=1}^{M} w_m \log_2(1+\gamma_m) \quad \text{(A.1a)}$$

$$\text{s.t.} \quad \gamma_m \leq \tilde{\mathrm{I}}_m(\bar{\boldsymbol{\sigma}}_s, \bar{\boldsymbol{\zeta}}), \quad \forall m. \quad \text{(A.1b)}$$

The subproblem (A.1) is a convex optimization in $\gamma_m$ with one global optimal solution. Therefore, the dual problem can be used to achieve the optimal solution since strong duality holds. By introducing the dual variables $\boldsymbol{\lambda} = \{\lambda_m \geq 0 : m = 1, \cdots, M\}$ to the inequality constraints (A.1b), the Lagrangian duality problem can be expressed as

$$\min_{\boldsymbol{\lambda}} \max_{\boldsymbol{\gamma}} \mathcal{L}(\boldsymbol{\gamma}, \boldsymbol{\lambda}), \quad \text{(A.2)}$$

where $\mathcal{L}(\boldsymbol{\gamma}, \boldsymbol{\lambda})$ is given by

$$\mathcal{L}(\boldsymbol{\gamma}, \boldsymbol{\lambda}) = \sum_{m=1}^{M} w_m \log_2(1+\gamma_m) - \sum_{m=1}^{M} \lambda_m \left(\gamma_m - \tilde{\mathrm{I}}_m(\bar{\boldsymbol{\sigma}}_s, \bar{\boldsymbol{\zeta}})\right). \quad \text{(A.3)}$$

From the KKT conditions, we have $\frac{\partial \mathcal{L}(\boldsymbol{\gamma}, \boldsymbol{\lambda})}{\partial \gamma_m} = 0$, then, the $\tilde{\lambda}_m$ can be obtained as

$$\tilde{\lambda}_m = \frac{\Omega_{m,2}(\bar{\boldsymbol{\beta}}, \bar{\boldsymbol{\sigma}}_s, \bar{\boldsymbol{\zeta}}) w_m}{\Omega_{m,3}(\bar{\boldsymbol{\beta}}, \bar{\boldsymbol{\sigma}}_s, \bar{\boldsymbol{\zeta}}) \ln 2}, \quad \forall m = 1, \cdots, M. \quad \text{(A.4)}$$

Substituting $\tilde{\lambda}_m$ into (A.3), $\mathcal{L}\left(\boldsymbol{\gamma}, \tilde{\boldsymbol{\lambda}}\right)$ can be further derived as

$$\mathcal{L}\left(\boldsymbol{\gamma}, \tilde{\boldsymbol{\lambda}}\right) = \sum_{m=1}^{M} w_m \log_2(1+\gamma_m) - \frac{1}{\ln 2} \sum_{m=1}^{M} w_m \gamma_m + \frac{1}{\ln 2} \sum_{m=1}^{M} w_m(\gamma_m + 1)\Theta_m, \quad \text{(A.5)}$$

Let $f_1(\bar{\boldsymbol{\sigma}}_s, \bar{\boldsymbol{\zeta}}, \boldsymbol{\gamma}) = \mathcal{L}\left(\boldsymbol{\gamma}, \tilde{\boldsymbol{\lambda}}\right)$, for any $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$, we define

$$\tilde{\boldsymbol{\gamma}}(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}) = \arg\max_{\boldsymbol{\gamma}} f_1(\bar{\boldsymbol{\sigma}}_s, \bar{\boldsymbol{\zeta}}, \boldsymbol{\gamma})\big|_{\bar{\boldsymbol{\sigma}}_s = \boldsymbol{\sigma}_s, \bar{\boldsymbol{\zeta}} = \boldsymbol{\zeta}}. \quad \text{(A.6)}$$

From the trivial solution of (A.1), we know that each of $\tilde{\gamma}_m$ satisfy $\tilde{\gamma}_m = \tilde{\mathrm{I}}_m(\boldsymbol{\sigma}_s, \boldsymbol{\zeta})$. Then, substituting $\tilde{\gamma}_m$ back in (A.5), we can obtain

$$\max_{\boldsymbol{\sigma}_s, \boldsymbol{\zeta}} f_1(\boldsymbol{\sigma}_s, \boldsymbol{\zeta},, \boldsymbol{\gamma}) = \max_{\boldsymbol{\sigma}_s, \boldsymbol{\zeta}} \mathcal{F}(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}). \quad \text{(A.7)}$$

*Proof* completed.

## APPENDIX II
## PROOF OF LEMMA 2

The updating of $\boldsymbol{\sigma}_s$ and $\boldsymbol{\zeta}$ are obtained by solving (48) and (55), respectively. However, the parameters $\mathbf{V}^a$, $\boldsymbol{\xi}$, $\mathbf{V}^b$ and $\boldsymbol{v}$ related to the linear programming with constraints are calculated based on the updating of $\boldsymbol{\gamma}$ and $\boldsymbol{\mu}$, which are directly optimized from the maximizing of $f_2(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu})$. Although the $\boldsymbol{\zeta}$ is optimally updated by maximizing the lower band of $f_2(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu})$, the solution would converge to the point which satisfies the KKT conditions of the original problem. Therefore, the solving of (48) and (55) are equivalent to the maximizing of $f_2(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu})$ since all of variables can be approximately regarded as the optimal updating based on the same gradient function. Therefore, *Algorithm 1* is essentially block coordinate ascent.

In addition, as we change one variable while fixing the others, it can be verified that $f_2(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu})$ is a monotonically nondecreasing function with a finite bound. After each iteration, thus, $f_2(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu})$ satisfies

$$f_2\left(\boldsymbol{\sigma}_s^{(t+1)}, \boldsymbol{\zeta}^{(t+1)}, \boldsymbol{\gamma}^{(t+1)}, \boldsymbol{\mu}^{(t+1)}\right) \geq f_2\left(\boldsymbol{\sigma}_s^{(t)}, \boldsymbol{\zeta}^{(t)}, \boldsymbol{\gamma}^{(t)}, \boldsymbol{\mu}^{(t)}\right). \tag{B.1}$$

Therefore, *Algorithm 1* will converge to the stationary point satisfying the KKT conditions of the problem (27).

*Proof* completed.

## APPENDIX III
## PROOF OF LEMMA 3

As aforementioned, the branch and bound algorithm is sufficient to guarantee the local optimal solution of $\boldsymbol{\beta}^{(t+1)}$ when bounded by the given $\boldsymbol{\sigma}_s^{(t)}$ and $\boldsymbol{\zeta}^{(t)}$, satisfying

$$\mathcal{F}\left(\boldsymbol{\beta}^{(t+1)}, \boldsymbol{\sigma}_s^{(t)}, \boldsymbol{\zeta}^{(t)}\right) \geq \mathcal{F}\left(\boldsymbol{\beta}^{(t)}, \boldsymbol{\sigma}_s^{(t)}, \boldsymbol{\zeta}^{(t)}\right). \tag{C.1}$$

According to *Lemma 2*, we know that $f_2(\boldsymbol{\sigma}_s, \boldsymbol{\zeta}, \boldsymbol{\gamma}, \boldsymbol{\mu})$ is monotonically nondecreasing and *Algorithm 1* is always convergent at a stationary point satisfying the KKT conditions once the optimal $\boldsymbol{\beta}^{(t+1)}$ is given, thus we have

$$\mathcal{F}\left(\boldsymbol{\beta}^{(t+1)}, \boldsymbol{\sigma}_s^{(t+1)}, \boldsymbol{\zeta}^{(t+1)}\right) \geq \mathcal{F}\left(\boldsymbol{\beta}^{(t+1)}, \boldsymbol{\sigma}_s^{(t)}, \boldsymbol{\zeta}^{(t)}\right). \tag{C.2}$$

Therefore, $\mathcal{F}(\boldsymbol{\beta}, \boldsymbol{\sigma}_s, \boldsymbol{\zeta})$ is monotonically nondecreasing in each iteration and *Algorithm 2* always converges to a local optimum when a stationary point is obtained. Note that, even though $\boldsymbol{\beta}$ is a binary variable, the convergence result of *Algorithm 2* is the best solution that one could hope for since the optimality of $\boldsymbol{\beta}^\star$ can be easily asserted at the condition of $\boldsymbol{\sigma}_s^\star$ and $\boldsymbol{\zeta}^\star$.

*Proof* completed.

## REFERENCES

[1] G. Gui, M. Liu, F. Tang, N. Kato, and F. Adachi, "6G: Opening new horizons for integration of comfort, security, and intelligence," *IEEE Wireless Communications*, vol. 27, no. 5, pp. 126–132, Oct. 2020.

[2] S. Aboagye, A. R. Ndjiongue, T. M. Ngatched, O. A. Dobre, and H. V. Poor, "RIS-assisted visible light communication systems: A tutorial," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 1, pp. 251–288, Dec. 2022.

[3] N. Chi, Y. Zhou, Y. Wei, and F. Hu, "Visible light communication in 6G: Advances, challenges, and prospects," *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 93–102, Sept. 2020.

[4] F. Xing, S. He, V. C. Leung, and H. Yin, "Energy efficiency optimization for rate-splitting multiple access-based indoor visible light communication networks," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 5, pp. 1706–1720, Jan. 2022.

[5] R. Bai, S. Hranilovic, and Z. Wang, "Low-complexity layered ACO-OFDM for power-efficient visible light communications," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 3, pp. 1780–1792, Sept. 2022.

[6] Y. He, M. Jiang, X. Ling, and C. Zhao, "Robust BICM design for the LDPC coded DCO-OFDM: A deep learning approach," *IEEE Transactions on Communications*, vol. 68, no. 2, pp. 713–727, Feb. 2019.

[7] C. Wang, Y. Yang, J. Cheng, C. Guo, and C. Feng, "A dimmable OFDM scheme with dynamic subcarrier activation for VLC," *IEEE Photonics Journal*, vol. 12, no. 1, pp. 1–12, Feb. 2020.

[8] Z. Gao, Y. Wang, X. Liu, F. Zhou, and K.-K. Wong, "FFDNet-based channel estimation for massive MIMO visible light communication systems," *IEEE Wireless Communications Letters*, vol. 9, no. 3, pp. 340–343, March 2020.

[9] P. Miao, G. Chen, K. Cumanan, Y. Yao, and J. A. Chambers, "Deep hybrid neural network-based channel equalization in visible light communication," *IEEE Communications Letters*, vol. 26, no. 7, pp. 1593–1597, Jul. 2022.

[10] M. R. Ghaderi, "LiFi and hybrid WiFi/LiFi indoor networking: From theory to practice," *Optical Switching and Networking*, vol. 47, p. 100699, Feb. 2023.

[11] X. Li, R. Zhang, and L. Hanzo, "Cooperative load balancing in hybrid visible light communications and WiFi," *IEEE Transactions on Communications*, vol. 63, no. 4, pp. 1319–1329, April 2015.

[12] J. Beysens, A. Galisteo, Q. Wang, D. Juara, D. Giustiniano, and S. Pollin, "DenseVLC: A cell-free massive MIMO system with distributed LEDs," in *Proceedings of the 14th International Conference on Emerging Networking EXperiments and Technologies*. New York, USA: Association for Computing Machinery, Dec. 2018, pp. 320–332.

[13] H. Zhang, N. Liu, K. Long, J. Cheng, V. C. Leung, and L. Hanzo, "Energy efficient subchannel and power allocation for software-defined heterogeneous VLC and RF networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 3, pp. 658–670, March 2018.

[14] A. Ibrahim, T. Ismail, K. F. Elsayed, M. S. Darweesh, and J. Prat, "Resource allocation and interference management techniques for OFDM-based VLC atto-cells," *IEEE access*, vol. 8, pp. 127 431–127 439, July 2020.

[15] G. Ma, R. Parthiban, and N. Karmakar, "An adaptive handover scheme for hybrid LiFi and WiFi networks," *IEEE Access*, vol. 10, pp. 18 955–18 965, Feb. 2022.

[16] R. Zhang, J. Wang, Z. Wang, Z. Xu, C. Zhao, and L. Hanzo, "Visible light communications in heterogeneous networks: Paving the way for user-centric design," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 8–16, April 2015.

[17] H. A. Ammar, R. Adve, S. Shahbazpanahi, G. Boudreau, and K. V. Srinivas, "User-centric cell-free massive MIMO networks: A survey of opportunities, challenges and solutions," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 611–652, Dec. 2021.

[18] X. Li, F. Jin, R. Zhang, J. Wang, Z. Xu, and L. Hanzo, "Users first: User-centric cluster formation for interference-mitigation in visible-light networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 39–53, Jan. 2015.

[19] R. Zhang, H. Claussen, H. Haas, and L. Hanzo, "Energy efficient visible light communications relying on amorphous cells," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 894–906, April 2016.

[20] M. Obeed, A. M. Salhab, S. A. Zummo, and M.-S. Alouini, "New algorithms for energy-efficient VLC networks with user-centric cell formation," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 1, pp. 108–121, March 2019.

[21] R. Jiang, Q. Wang, H. Haas, and Z. Wang, "Joint user association and power allocation for cell-free visible light communication networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 1, pp. 136–148, Jan. 2018.

[22] J. Chen, Z. Wang, and R. Jiang, "Downlink interference management in cell-free VLC network," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 9007–9017, Jan. 2019.

[23] H. Liu, X. Gong, M. Huang, Y. Chen, X. Yuan, K. Chen, and S. Yang, "Joint AP grouping and user clustering for interference management in cell-free VLC network," *Optics & Laser Technology*, vol. 156, p. 108465, Dec. 2022.

[24] H. Yang, W.-D. Zhong, C. Chen, A. Alphones, and P. Du, "Qos-driven optimized design-based integrated visible light communication

and positioning for indoor IoT networks," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 269–283, Jan. 2020.

[25] R. Zhang, Y. Cui, H. Claussen, H. Haas, and L. Hanzo, "Anticipatory association for indoor visible light communications: Light, follow me!" *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2499–2510, April 2018.

[26] O. B. H. Belkacem, R. Dinis, and M. L. Ammari, "Nonlinear effects in NOMA-OFDM systems: Analytical signal characterization and receiver design," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 3, pp. 3739–3750, March 2023.

[27] X. Ling, X. Zhang, J. Sun, P. Ge, J. Wang, C. Zhao, and X. Gao, "When hammerstein meets wiener: Nonlinearity modeling for end-to-end visible light communication links," *IEEE Transactions on Communications*, vol. 71, no. 1, pp. 310–323, Jan. 2023.

[28] M. Laakso, A. A. Dowhuszko, and R. Wichman, "Predistortion of OFDM signals for VLC systems using phosphor-converted LEDs," in *2022 IEEE 23rd International Workshop on Signal Processing Advances in Wireless Communication (SPAWC)*.  Oulu, Finland: IEEE, Jul. 2022, pp. 1–5.

[29] L. Xia, X. Wang, Z. Sun, Z. Cheng, J. Jin, Y. Yuan, G. Liu, T. Jiang, and Y. Huang, "Signal clipping at transmitter and receiver of O-OFDM for VLC under optical power constraint," *China Communications*, vol. 19, no. 6, pp. 154–168, June 2022.

[30] X. Ling, J. Wang, X. Liang, Z. Ding, and C. Zhao, "Offset and power optimization for DCO-OFDM in visible light communication systems," *IEEE Transactions on Signal Processing*, vol. 64, no. 2, pp. 349–363, Jan. 2016.

[31] S. Mardanikorani, X. Deng, J.-P. M. Linnartz, and A. Khalid, "Compensating dynamic nonlinearities in LED photon emission to enhance optical wireless communication," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1317–1331, Feb. 2021.

[32] O. T. Demir and E. Bjornson, "The bussgang decomposition of nonlinear systems: Basic theory and MIMO extensions," *IEEE Signal Processing Magazine*, vol. 38, no. 1, pp. 131–136, Jan. 2021.

[33] J. Zhang, C. Liu, X. Li, H.-L. Zhen, M. Yuan, Y. Li, and J. Yan, "A survey for solving mixed integer programming via machine learning," *Neurocomputing*, vol. 519, pp. 205–217, Jan. 2023.