User-Centric Visible Light Communications for Energy-Efficient Scalable Video Streaming

Xuan Li, Yongkai Huo, Rong Zhang, Lajos Hanzo

Southampton Wireless, School of ECS, University of Southampton, UK

xl26g12,yh3g09,rz,lh@ecs.soton.ac.uk

Abstract—An energy-efficient indoor visible light communication (VLC) system relying on dynamic user-centric (UC) cluster formation is designed for scalable video streaming. Explicitly, the radically new UC cluster formation technique is based on an amorphous user-to-network association structure, which is ultimately the basis of our energy-efficient indoor VLC system. Furthermore, in order to optimise the system-level energy efficiency, our objective function is selected by jointly considering both the video quality and the power consumption. We then propose a 3-tier dynamic-programming-based algorithm for user/layer-level adaptive modulation mode assignment, for access-point-level power allocation and for cluster-level energy efficiency optimisation, respectively. Based on a scalable video coded sequence, our simulation results demonstrate the superior performance of our UC clusters compared to the conventional cell design in terms of its energy efficiency, throughput as well as video quality in most of the scenarios considered.

I. INTRODUCTION

1) Background: Following the launch of the global 5G research initiatives [1] conceived for tackling the explosive escalation of wireless tele-traffic, the horizon of communication spectrum has been expanded from the conventional radio frequency (RF) band both to the millimetre wave [2] and to the visible light [3] frequency band spanning from 400 to 700 THz. Hence, the pursuit of 'green' communications has motivated both the academic as well as the industrial community to continuously explore future technologies, in order to improve the achievable energy efficiency (EE) of the entire network infrastructure [4]. As a promising complimentary extension to the well-established indoor RF networks [5], [6], visible light communication (VLC) is becoming an additional promising enabler for providing indoor coverage, owing to its energy-efficient nature simultaneously supporting both communications and illumination. Hence, a careful design of both functions is required for fully realising the EE potential of VLC.

To elaborate, the light-emitting diode (LED) transmitters are primarily used for illumination with the aid of a constant DC power, which also provides sufficient forward biasing voltage across the LEDs for wireless communications. Hence, the additional communication function should not perturb the illumination specifications, nor should it violate the LEDs' physical limits. Ideally, the extra communication-related power

consumption should be as low as possible, while maintaining a minimum required quality of service (QoS). To this end, in addition to the extensive efforts invested in improving the attainable throughput of the VLC networks by utilising sophisticated optical devices [7], [8], by employing novel modulation schemes [9], [10], and by employing powerful multi-input-multi-output techniques [11], [12], valuable research has also been dedicated to the improvement of link-level EE, focusing on brightness and dimming control with the aid of both modulation-related [13], [14] as well as coding-related techniques [15], [16].

Whilst there are valuable link-level studies, there is a lack of system-level investigations on designing energy-efficient VLC networks supporting multiple users, which may require a radically new design approach. When considering VLC networks supporting multiple users, the performance degradation imposed by the escalating inter-cell interference at the cell edge may lead to dramatic reduction of the QoS as well as to the EE reduction of the VLC networks. As a result, careful VLC cell formation design becomes crucial, since it crucially influences the entire system design cycle. The authors of [17] proposed a joint link scheduling and brightness control scheme for energy-efficient multi-user VLC networks with the aid of a novel light source structure, which however still relies on the classic cellular design. By contrast, in [18] we conceived a novel user-centric (UC) design principle for VLC, where the resultant user-to-network association structure is based on amorphous cell-shapes. More explicitly, in our UC design, the cell formation is constructed by grouping the user equipment (UE) and associating access points (APs) to each UE-group based on the UEs' location information. Our recent work has demonstrated that the UC design principle leads to a higher system throughput and a more uniform user fairness [19] as well as to a reduced power consumption [20], when compared to the conventional VLC cell formation relying on either unity frequency reuse (UFR) or on a frequency reuse factor of two (FR2) [1].

---

[1]Conventional VLC cell formation follows the traditional cellular design principle, where each optical AP illuminates an individual cell and adopts UFR across all cells. As a result, inter-cell interference is imposed by the line-of-sight ray of neighbouring cells and the UEs may experience dramatic performance degradation at the cell edge. In order to reduce the inter-cell interference, appropriate frequency reuse patterns may be employed as an appealingly simple solution, while the system has to obey the classic trade-off between reduced bandwidth efficiency and improved cell-edge signal-to-interference-plus-noise-ratio, when using a frequency reuse factor higher than one.

2) Motivation: Owing to these exciting findings, we further develop the UC-VLC design philosophy for video streaming, which is believed to be the killer-application for VLC [21], since video services account for much of the tele-traffic in modern wireless communications systems [22], [23]. Given the rapid development of video coding techniques [24], [25], the advanced scalable extension of the so-called high-efficiency video coding (SHVC) techniques [26] has gained popularity, as a benefit of its scalable nature, because it is capable of offering diverse visual qualities by promptly adapting to the time-variant channel conditions of different UEs. The SHVC-based layered video stream is constituted by multiple unequal-importance layers, which are generated by using carefully designed source codecs [27], [28] as well as adaptive modulation and channel coding schemes [29], [30]. According to [25], if the subsets of the original video sequence may lead either to video reconstruction at a reduced picture size or at a reduced frame rate compared to the original one, then this video scheme exhibits either spatial or temporal scalability, respectively. Another popular scalability mode is the so-called quality-scalability, where the subsets of bits may provide a reduced video fidelity. Explicitly, the fidelity is often represented by the peak signal-to-noise ratio (PSNR). As a further advance, the afore-mentioned modes of scalability may be supported by a single scalable video sequence. The benefits of scalable video are multi-fold, amongst others, allowing for example the video decoder to progressively refine the reconstructed visual quality, as the channel-quality improves. The hierarchical structure of the video-stream also facilitates energy-efficient video communications by jointly considering the scalable video quality and the power consumption, as demonstrated for traditional RF networks in [31]–[33]. This motivated us to support energy-efficient scalable video streaming in the radically new context of UC-VLC networks.

3) Contribution: We design an energy-efficient scalable video system relying on dynamic UC cluster formation in VLC networks, while jointly considering adaptive modulation (AM) mode assignment and power allocation (PA). To be more specific,

i) we propose a distance-based UC cluster formation technique and employ two different joint transmission schemes within the clusters, which we refer to as combined transmission (CT) and vectored transmission (VT) [2]. The beneficial construction of UC clusters constitutes the basis of a structurally energy-efficient VLC network;

ii) we design an energy-efficient scalable video scheme and carefully formulate its EE maximisation problem by taking into account its unique video-related characteristics, such as the base- and enhancement video layers etc.;

iii) we propose a heuristic 3-tier dynamic-programming-based algorithm, including the user/layer-level AM mode assignment, the AP-level PA and the cluster-level EE opti-

mization, for maximising the system-level EE of our UC-VLC network;

iv) we evaluate the proposed EE scheme by transmitting a SHVC sequence and compare our UC design- in terms of its achievable EE, throughput, and video quality- to the conventional cells utilising both UFR and FR2. For our simulations results of the full video clips, please refer to http://www.ecs.soton.ac.uk/research/projects/924.

The rest of this paper is organized as follows. Our UC system model and the scalable video architecture considered are presented in Section II and III, respectively. Our problem formulation and its 3-tier dynamic-programming-based algorithm are described in Section IV and V, respectively. The simulation results are presented in Section VI. Finally, Section VII offers our conclusions.

II. SYSTEM MODEL

In contrast to the conventional cellular design leading to circular cells, in our UC-VLC network each UC cluster is served by a set of VLC APs, which simultaneously serve multiple UEs. More explicitly, a UC cluster includes a set of APs and UEs. In this section, we highlight the optical link characteristics and then discuss the UC cluster formation philosophy.

A. Link Characteristics

Since each UE has a limited field of view (FoV), it can only receive information from the optical APs, when one or more APs reside within the UE's FoV. According to [34], if the angle of incidence ψ from an AP a to a UE u is less than half of the UE's FoV ψ_F, the optical channel's direct current (DC) attenuation of the line-of-sight (LoS) path is given by

$$h_d[u,a] = \frac{(w+1)D_{PA}}{2\pi d_{a,u}^2} \cdot \cos^w(\phi) \cdot T_s(\psi) \cdot g(\psi) \cdot \cos(\psi), \quad (1)$$

where the Lambert index w depends on the semi-angle $\phi_{1/2}$ at half-illuminance of the source, which is given by $w = -1/\log_2(\cos\phi_{1/2})$. Furthermore, D_{PA} is the detection area of the detector's photo-diode (PD), $d_{a,u}$ is the distance between the VLC AP a and the UE u, while ϕ is the angle of irradiance. Still referring to (1), $T_s(\psi)$ and $g(\psi)$ denote the gain of the optical filter and of the optical concentrator employed, respectively, while $g(\psi)$ can be written as $g(\psi) = n_r^2/\sin^2\psi$ [34], where n_r is the refractive index of a lens at a PD. The optical channel of (1) may be widely adopted, when considering a Lambertian source operating in indoor optical wireless scenarios. Although some of the practical VLC channel characteristics have been simplified, our algorithm is a generic one, which may be readily adapted to other types of optical channels. Furthermore, according to [35], when the angle of incidence ψ is no larger than half of the FoV, the channel's DC attenuation on the first reflection path resulted from a small reflective area is given by

$$dh_r[u,a] = \frac{(w+1)D_{PA}}{2\pi^2 l_1^2 l_2^2} \cdot \rho \cdot dD_{wall} \cdot \cos^w(\phi) \cdot \cos(\beta_1) \cdot$$
$$\cos(\beta_2) \cdot T_s(\psi) \cdot g(\psi) \cdot \cos(\psi), \quad (2)$$

where l_1 denotes the distance between an AP and a reflective point, while l_2 is the distance between this point and a UE.

---

[2]In [5], relying on CT, each individual VLC AP of a multi-AP cell conveyed the same information on the same visible carrier frequency in their overlapping areas and served a single UE at a time. In order to eliminate the bandwidth efficiency reduction imposed by CT, zero-forcing-based VT techniques were employed for serving multiple users at the same time in the overlapping area, which will be exemplified in Section III-C.

The reflectance factor and the reflective area are denoted by ρ and dD_{wall}, respectively. Additionally, β_1 and β_2 represent the irradiance angles to the reflective point and to the UE, respectively. Our VLC parameter values are summarized in TABLE I in Section VI.

B. UC Cluster Formation

1) Preliminary: Let us introduce some notations first before constructing the UC clusters. Fig. 1 shows the down-link of a particular VLC network having ($N_A = 64$) optical APs and ($N_U = 25$) UEs, which are randomly distributed in the room of size 15m×15m×3m. Let $\mathcal{V}_A$ and $\mathcal{V}_U$ denote the AP set and UE set, respectively. Each UC cluster $\mathcal{C}_n$ is constituted by two subsets, i.e. the AP subset denoted by $\mathcal{V}_{A,n}$ as well as the UE subset denoted by $\mathcal{V}_{U,n}$, where $n = 1, \cdots, N$ and N denotes the total number of UC clusters. Hence, we have

$$\mathcal{C}_n = \mathcal{V}_{A,n} \cup \mathcal{V}_{U,n}, \ \mathcal{V}_{A,n} \cap \mathcal{V}_{A,n'} = \emptyset, \ \mathcal{V}_{U,n} \cap \mathcal{V}_{U,n'} = \emptyset, \quad (3)$$

$$\{\mathcal{V}_{A,n} : n = 1, \cdots, N\} \subseteq \mathcal{V}_A, \{\mathcal{V}_{U,n} : n = 1, \cdots, N\} \subseteq \mathcal{V}_U. \quad (4)$$

Note that N may not be determined before all UC clusters have been formed, because it is influenced by the various UEs' location, FoV, etc. Furthermore, since some UEs may not have information to transmit and some APs may not be active during the current cluster formation round, we have (4).

Since the VLC channels are pre-dominantly static, the channel knowledge can be characterised by a single attenuation factor. Hence, the channel's impulse response can be readily estimated at the user side and then fed back to the AP side at the cost of a modest overhead. After acquiring the channel knowledge, the distances between APs and UEs may be inferred from (1). As a result, classical positioning and tracking may be used for determining the users' positions [36]–[38]. Thus the mutual distances $d_{u,u'}$ between any pair of UEs as well as the mutual distances $d_{a,u}$ between any AP-UE pair may be calculated. Since our UC cluster formation is ultimately based on the UEs' locations, in order to control the size of the clusters, the distance constraints are pre-set as d_μ for the UEs and d_α for the APs within a single cluster. In other words, the distance between the UE and its cluster centre is no more than d_μ and the distance between the AP and its cluster centre is no more than d_α. Let us now discuss, how to determine the cluster centre in the following detailed cluster formation steps. In this way, various forms of the UC clusters may be constructed by adjusting the value of d_μ or d_α. For example, a larger d_μ and d_α may improve the area spectral efficiency of the network according to [5], but as its price, the signal processing complexity within each cluster is increased. Hence, upon adjusting the value of d_μ or d_α, diverse requirements of the system design may be satisfied.

2) Implementation: Let us now carry out the UC cluster formation step by step. The UE subsets $\{\mathcal{V}_{U,n}\}$ are first constructed during Step 1 - Step 3, while the corresponding AP subsets $\{\mathcal{V}_{A,n}\}$ are determined during Steps 4 and 5. Finally, the UC clusters $\{\mathcal{C}_n\}$ are formed in Step 6.

Step 1) *Initial UE selection:* During this step, a UE $u \in \mathcal{V}'_U$ is randomly selected as the starting point for constructing a tentative single-UE subset denoted by $\mathcal{V}^t_{U,n}$ as part of the new

cluster $\mathcal{C}_n$, where $\mathcal{V}'_U$ is referred to as the idle UE set including all the UEs not belonging to any of the clusters constructed. Furthermore, the tentative centre c^t_n of $\mathcal{V}^t_{U,n}$ is the location of u denoted by (x_u, y_u), as shown in Fig. 1a.

Step 2) *Tentative UE-set expansion:* The initial tentative UE set $\mathcal{V}^t_{U,n}$ in Step 1 is expanded by including the nearby UEs within a certain range d_μ, which may be expressed as

$$\mathcal{V}^t_{U,n} = \left(\mathcal{V}^t_{U,n} \cup \{u' \in \mathcal{V}'_U : |(x_{u'}, y_{u'}) - c^t_n| \le d_\mu\}\right), \quad (5)$$

as also seen in Fig. 1a. Hence, the centre c^t_n of $\mathcal{V}^t_{U,n}$ should also be updated by averaging the locations of all UEs in $\mathcal{V}^t_{U,n}$, which may be calculated as

$$c^t_n = \left(\frac{\max_{u \in \mathcal{V}^t_{U,n}}\{x_u\} + \min_{u \in \mathcal{V}^t_{U,n}}\{x_u\}}{2},\right.$$
$$\left.\frac{\max_{u \in \mathcal{V}^t_{U,n}}\{y_u\} + \min_{u \in \mathcal{V}^t_{U,n}}\{y_u\}}{2}\right), \quad (6)$$

as explicitly shown in Fig. 1b. Then the idle UE set $\mathcal{V}'_U$ should exclude all the UEs satisfying (5), i.e. $\mathcal{V}'_U = (\mathcal{V}'_U \setminus \{u' \in \mathcal{V}'_U : |(x_{u'}, y_{u'}) - c^t_n| \le d_\mu\})$. Repeat checking the condition (5), until no more UEs are found in the proximity of the tentative UE subset $\mathcal{V}_{U,n}$.

Step 3) *UE set formation:* Following Step 2, the UE subset $\mathcal{V}_{U,n}$ is deemed to be determined, with its centre denoted as c_n, as long as no UE is close enough to the UE subset $\mathcal{V}_{U,n}$. Upon repeating Step 1 and 2, all the UE subsets $\{\mathcal{V}_{U,n}\}$ have been constructed, as shown in Fig. 1c.

Step 4) *Anchor AP selection:* Let us now find the corresponding AP subsets $\{\mathcal{V}_{A,n}\}$ for $\{\mathcal{V}_{U,n}\}$. In order to guarantee the number of APs is no less than the number of UEs within each UC cluster [3], each UE u first finds its closet AP denoted by $\text{AP}(u)$, which is referred to as the anchor AP, hence we have

$$\text{AP}(u) = \arg \min_{a \in \mathcal{V}_A} (d_{a,u}), \quad (7)$$

as shown in Fig. 1d. If more than one UEs select the same anchor AP, this AP selects its closest UE and the other UEs have to select their next nearest APs, until each UE selected a unique anchor AP satisfying $\text{AP}(u) \ne \text{AP}(u')$ for any two UEs u and u'. The unique anchor AP of the UE u is denoted as $\text{AP}^*(u)$. Hence, the tentative AP set for a specific cluster $\mathcal{C}_n$ may be written as

$$\mathcal{V}^t_{A,n} = \{\text{AP}^*(u) : u \in \mathcal{V}_{U,n}\}. \quad (8)$$

After all the tentative AP sets have been determined, similarly to the definition of the idle UE set $\mathcal{V}'_U$, let $\mathcal{V}'_A$ denote the idle AP set including all APs not belonging to any tentative AP set.

Step 5) *Tentative AP-set adjustment:* During this adjustment step, each tentative AP set is expanded first by including the

---

[3]The UC cluster formation process proposed in this paper should be operated after the multiple UE scheduling process, which should satisfy two requirements: i) the total number of the scheduled UEs does not exceed that of the VLC APs; ii) the scheduled UEs are not confined to a small area. This guarantees that the number of the APs is no less than that of the UEs within each cluster. However, the scheduling process is beyond our current scope.

(a) Initial UE selection.

(b) Tentative UE-set expansion.

(c) UE-set formation.

(d) Anchor AP selection.

(e) Tentative AP-set adjustment.
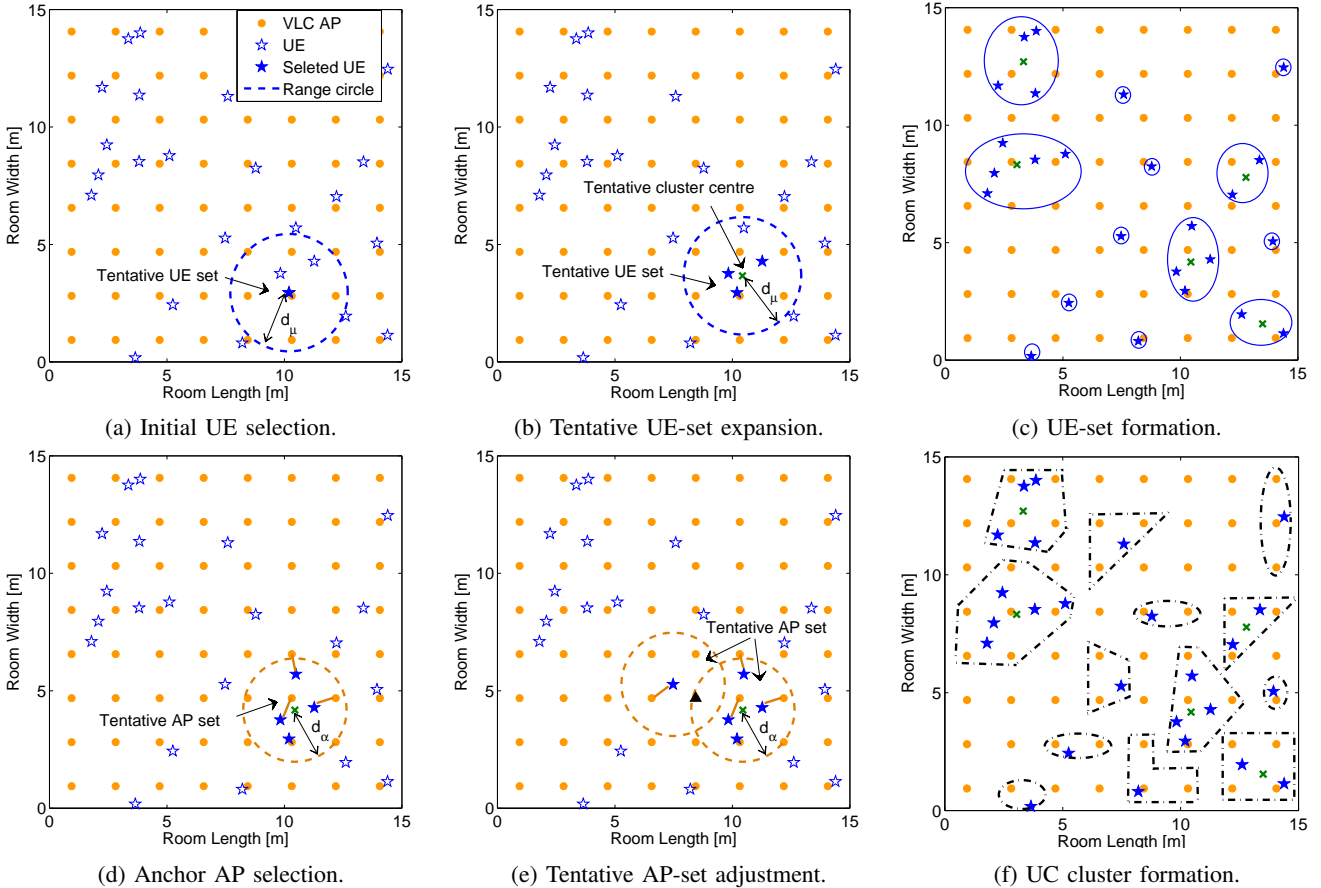
(f) UC cluster formation.

Fig. 1: Steps of UC cluster formation. The subfigures (a)-(c) illustrate the formation of the UE set for a specific UC cluster. (d) and (e) show the formation of the AP set for an example UC cluster, where a UE is connected by a short line with its anchor AP. Note that these steps are carried out after all UE sets are formed. In (f), multiple UC clusters are finally constructed. Note that the area of the closed irregular shapes do not represent the coverage of each cluster and the inter-cluster interference is not illustrated.

nearby idle APs within a certain range d_α, which may be expressed as

$$\mathcal{V}_{A,n}^t = \left(\mathcal{V}_{A,n}^t \cup \{a \in \mathcal{V}_A' : |(x_a, y_a) - c_n| \le d_\alpha\}\right), \quad (9)$$

where (x_a, y_a) denotes the position of the AP a. Furthermore, after gradually expanding all the tentative AP sets following the rules of (9), if a specific AP a was included in several $\mathcal{V}_{A,n}^t$ and in the meantime, a is not the anchor AP of any UE, i.e. if we have $a \in \left(\mathcal{V}_A' \cap \mathcal{V}_{A,n}^t \cap \mathcal{V}_{A,n'}^t\right)$, then AP a is set to its idle mode, which is shown as the solid triangle in Fig. 1e. This measure is taken for the sake of avoiding any extra interference as well as for saving energy. Thus, after the process of exclusive assignment based expansion, all tentative AP subsets $\{\mathcal{V}_{A,n}\}$ are determined.

Step 6) *UC cluster formation:* Upon combining the corresponding AP subset and UE subset, the UC clusters $\{\mathcal{C}_n\}$ are finally constructed, as shown in Fig. 1f. In order to distinguish the single-UE and the multi-UE clusters, the former is denoted as $\mathcal{C}_n^{\text{CT}}$, while the latter as $\mathcal{C}_n^{\text{VT}}$, corresponding to the link-level transmission techniques CT and VT, respectively, which will be introduced next in Section III-C.

III. SCALABLE VIDEO STREAMING

In this section, the scalable multiuser video streaming is introduced. The design aspects of the UC cluster formation and the video streaming are intricately inter-linked, where we adaptively assign AM modes to each video layer of the transmitted video stream and simultaneously allocate 'just sufficient' power to each optical AP.

A. Multiuser Video Broadcast and SHVC

Since the UC clusters have now been constructed, let us introduce our energy-efficient scalable video system relying on the UC-VLC aided network. To elaborate, we consider a general video service scenario operating in a typical room having the dimensions of 15m×15m×3m, where each UE requests different video content. Supporting video multicast services is beyond the scope of this paper. Furthermore, we employ the appealing SHVC technique [24], [25], where the video sequence is encoded into multiple layers. This allows us to progressively refine the reconstructed video quality at the receiver, when the channel quality is improved. Given a specific UC cluster, where $N_{A,n}$ optical APs simultaneously support $N_{U,n}$ UEs, $N_{U,n}$ different scalable video sequences are requested, each of which is encoded into multiple layers at the transmitter side, as shown in the left subfigure of Fig. 2.
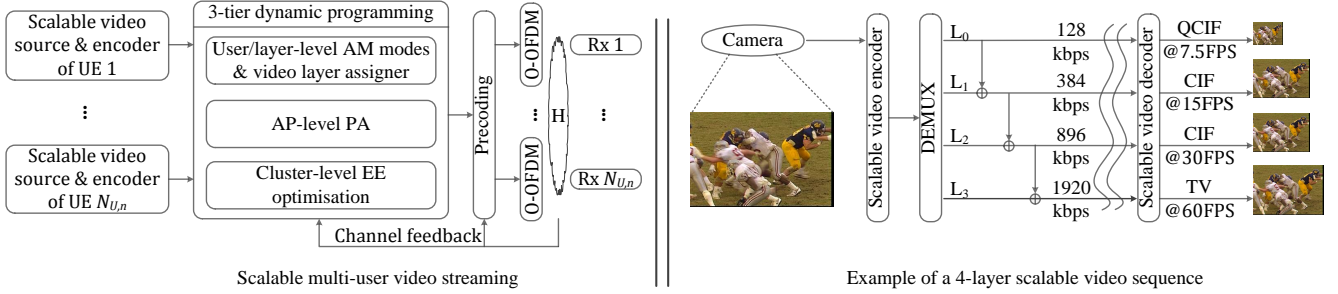
Fig. 2: Left: The scalable video streaming of a specific UC cluster $\mathcal{C}_n$, where $N_{A,n}$ optical APs supports $N_{U,n}$ UEs simultaneously. Right: The example of a 4-layer video sequence requested by a specific UE u in the UC cluster $\mathcal{C}_n$, where the video quality is gradually refined according to the video quality demand and the channel conditions of the UE u.

According to the channel conditions and to the specific video quality requirement of the UEs, the highest affordable-quality subset of transmitted video layers is decided for each video sequence and then it is assigned the appropriate AM modes by employing the 1st-tier of our proposed 3-tier dynamic-programming-based algorithm. Then, upon using the 2nd-tier PA and 3rd-tier EE optimisation algorithm, the EE of the system is maximised. In order to mitigate the intra-cluster interference, the video signals are preprocessed by exploiting the knowledge of the channel state information before transmission. At the receiver side, the UEs receive mutually interference-free video signals, which constitute a channel-quality-dependent subset of their requested video sequence, which leads to diverse visual qualities.

To elaborate a little further, the right subfigure of Fig. 2 shows an example of the 4-layer video sequence downloaded by a specific UE u within the above-mentioned UC cluster $\mathcal{C}_n$ in the left subfigure of Fig. 2, where the video sequence exhibits both the spatial as well as the temporal scalability. In general, the base layer L_0 is the most important one amongst them, whilst the less-important video layers are termed as enhancement layers. Gradually increased video qualities are associated with receiving the higher order layers L_1, L_2 and L_3, as seen in the stylised illustration of the right subfigure of Fig. 2. Each enhancement layer relies on both the base layer and on all the previous enhancement layers having lower orders. If either the base layer or any of the previous enhancement layers is lost or corrupted during their transmission, the dependent layers must be dropped by the decoder. For example, observe in the right subfigure of Fig. 2 that layer L_1 is dependent on the base layer L_0, while layer L_3 depends on both the base layer L_0 as well as on all the lower layers L_1 and L_2.

As far as the video streaming seen in the left subfigure of Fig. 2 is concerned, the video sequence of the UE u in the right subfigure is first encoded and a subset of its layers is pre-processed for transmission. Then the UE u may experience different visual qualities, depending both on its QoS requirements as well as on its channel conditions. For example, upon receiving only the base layer, the decoder is only capable of reconstructing the video at a quarter common intermediate format (QCIF) at 7.5 frame per second (FPS).

By contrast, a common intermediate format (CIF) based video sequence scanned at 15 FPS and 30 FPS can be reconstructed with the aid of the layers $\{L_0, L_1\}$ and $\{L_0, L_1, L_2\}$, respectively. In order to further improve the video quality for a high-resolution TV screen at 60 FPS, all the four layers $\{L_0, L_1, L_2, L_3\}$ must be flawlessly streamed. In practice, the different video broadcast scenarios mentioned above require different throughput, which rely on the assignment of different modulation schemes to different layers. Therefore, in order to provide the highest possible video quality for all UEs and at the same time to guarantee energy-efficient scalable video provision, the various modulation schemes have to be adaptively assigned and the power has to be appropriately allocated to each video layer requested by different UEs.

B. Adaptive Modulation

There are two popular techniques of constructing white LEDs, namely either by mixing the red-green-blue frequencies using three chips, or by using a single blue LED chip with a phosphor layer. We consider the latter one, which is the favoured commercial version. Hence, the modulation bandwidth B is typically around 20 MHz, albeit this measured bandwidth depends on the specific LED product used. Given this 20 MHz bandwidth, we are now ready to employ asymmetrically clipped optical OFDM (ACO-OFDM). To elaborate, ACO-OFDM is an energy-efficient scheme, which allows us to invoke AM modes relying on complex symbols, since the careful selection of a video-layer-specific AM mode guarantees scalable video provision.

A total of M different AM modes are adopted in our UC-VLC network, where the AM mode m provides a data rate of r^m, once its signal-to-interference-plus-noise-ratio (SINR) threshold of ξ_{thr}^m is satisfied. According to [39], given a target bit-error-rate (BER) and m-ary quadrature amplitude modulation (QAM), the minimum SINR required can be determined from:

$$\text{BER} = \frac{\sqrt{m}-1}{\sqrt{m}\log_2\sqrt{m}}\text{erfc}\left(\sqrt{\frac{3\xi_{\text{thr}}^m}{2(m-1)}}\right), \quad (10)$$

where erfc($\cdot$) denotes the complementary error function. Furthermore, we rank the M AM modes, such that if $m' \geq m$, we have $r^{m'} \geq r^m$ and $\xi_{\text{thr}}^{m'} \geq \xi_{\text{thr}}^m$.

C. Transmission

In order to simultaneously provide video services to multiple UEs within each UC cluster constructed in Section II-B, zero-forcing (ZF)-based VT techniques are introduced, where the intra-cluster interference is eliminated at the multiple AP transmitters, as indicated in the left subfigure of Fig. 2. Hence, all the UEs receive mutually interference-free signals.

To elaborate a little further, for a given UC cluster $\mathcal{C}_n^{\text{VT}}$, let $N_{A,n}$ and $N_{U,n}$ be the number of optical APs and the number of the UEs, respectively. Furthermore, let Z_t and Y_r denote the vectors of transmitted and received signals, respectively, where the entries of Z_t will be selected from a particular AM mode for the different video layers discussed in Section III-B. Upon employing VT techniques, the received signals may be obtained by

$$Y_r = \gamma \cdot H \cdot G \cdot \widetilde{\mathcal{P}}_n \cdot Z_t + \tau, \qquad (11)$$

where γ denotes the optical/electronic (O/E) conversion efficiency, while τ includes both the noise and the inter-cluster interference imposed by the neighbouring clusters. The channel matrix $H \in \mathbb{R}^{N_{U,n} \times N_{A,n}}$ hosts the channel attenuations between the $N_{A,n}$ APs and the $N_{U,n}$ UEs. In order to receive mutually interference-free signals at the receivers, the transmitted signals $Z_t = [z_1, z_2, \cdots, z_{N_{U,n}}]^T$, where each z_u entry satisfies that $\mathrm{E}[|z_u|^2] = 1$, are precoded as $(G \cdot Z_t)$, where the $(N_{A,n} \times N_{U,n})$-element matrix $G = H^H \cdot (H \cdot H^H)^{-1}$ obeys the ZF criterion for the sake of obtaining an interference-free identity matrix for $H \cdot G = I_{N_{U,n}}$. Furthermore, the PA matrix $\widetilde{\mathcal{P}}_n$ is a diagonal matrix and we have $\widetilde{\mathcal{P}}_n = \mathrm{diag}(p_1, \cdots, p_{N_{U,n}})$, where each diagonal entry p_u denotes the electronic power allocated to the signal z_u. Considering the AM mode m and the video layer l, where we have $l = 0, 1, \cdots, L - 1$, let us furthermore denote the power allocated to the l-th video layer of the UE u associated with the AM mode m in the electronic domain by $p_{u,l}^m$, which is a specific function of the power p_u.

Let us furthermore define the SINR as the aggregated electronic power over the noise power in a bandwidth of B [MHz] [40] plus the sum of the electronic power received from other optical sources in the vicinity. Still considering ACO-OFDM, we may express the SINR for the UE u within a multi-UE cluster $\mathcal{C}_n^{\text{VT}}$ as

$$\xi_{u,l}^m = \frac{(\gamma^2/2) \cdot p_{u,l}^m}{N_0 B + I_u}, \qquad u \in \mathcal{V}_{U,n^{\text{VT}}}, \qquad (12)$$

where $N_0 \sim 10^{-22}$ [A²/Hz] [40] denotes the noise power spectral density. I_u is the interference imposed by the reflected light as well as the LoS rays from the other clusters in the vicinity. Since the interference power received by the cluster under consideration is influenced by the PA within other clusters, for simplicity, we assume the interference imposed

is always equal to its maximum value formulated as [4]

$$I_u = \gamma^2 \pi \sum_{a \notin \mathcal{V}_{A,n}} h^2[u,a] \cdot (p_{\max}^{\text{tx}})^2, \qquad (13)$$

where $p_{\max}^{\text{tx}}$ is the maximum optical transmit power allowed for each optical AP and we have $h[u,a] = h_d[u,a] + \sum dh_r[u,a]$. (13) characterizes the worst-case situation in our VT cluster formation.

Furthermore, for the multi-UE VT cluster, the total electronic transmit power may be expressed as

$$p_{u,l}^{\text{tx}} = \sum_{a=1}^{N_{A,n}} g_{\text{VT}}^2[a,u] \cdot p_{u,l}^m \qquad u \in \mathcal{V}_{U,n^{\text{VT}}}, \qquad (14)$$

where $g_{\text{VT}}[a,u]$ denotes the $[a,u]$th entry of the transmit precoding (TPC) matrix G in (11). Given a specific UE u, the AM mode m may be assigned to it, provided that the SINR received is at least ξ_{thr}^m. Hence, the minimum power required by employing AM mode m for the UE u receiving the l-th video layer may be expressed as:

$$p_{u,l}^m = \xi_{\text{thr}}^m \cdot (N_0 B + I_u)/(\gamma^2/2). \qquad (15)$$

Having obtained the relationship between the power required and the AM mode of the video layers of different UEs, let us now formulate the PA problem in the UC-VLC network for our energy-efficient scalable video scheme in Section IV.

Remark 1. When considering the single-UE cluster $\mathcal{C}_n^{\text{CT}}$ employing CT, each AP is assumed to emit the same amount of electronic power [5]. Hence the total electronic transmit power $p_{u,l}^{\text{tx}}$ of the UE u receiving the l-th video layer within the CT cluster $\mathcal{C}_n^{\text{CT}}$ may be written as

$$p_{u,l}^{\text{tx}} = \frac{N_{A,n} \cdot p_{u,l}^m}{\left(\sum_{a=1}^{N_{A,n}} h[u,a]\right)^2} = \sum_{a=1}^{N_{A,n}} \frac{p_{u,l}^m}{\left(\sum_{a=1}^{N_{A,n}} h[u,a]\right)^2}$$

$$= \sum_{a=1}^{N_{A,n}} g_{\text{CT}}^2[a,u] \cdot p_{u,l}^m, \qquad u \in \mathcal{V}_{U,n^{\text{CT}}}, \qquad (16)$$

where $h[u,a]$ denotes the $[u,a]$th entry of the channel matrix H of the single-UE CT cluster and $g_{\text{CT}}^2[a,u] = 1/\left(\sum_{a=1}^{N_{A,n}} h[u,a]\right)^2$. Hence, a unified expression

---

[4]Since our proposed cluster formation technique is distance-based, the distribution of the clusters constructed should be dispersed. Hence, for a specific UE, the distances to most of the APs in the neighbouring clusters should be sufficiently high compared to its anchor AP and consequently the interference imposed should be low due to the high-attenuation interference channels. Therefore, the differences between the maximum value of the interference and its actual value may be small. Furthermore, this assumption can support a more scalable and efficient solution within each cluster discussed in Section IV, hence avoiding the potentially excessive complexity imposed by a centralised algorithm, when true interference level is used.

[5]Since our objective is to allocate an appropriate power to each video layer of *multiple* UEs under the power constraint of each VLC AP, for the sake of maximising the achievable EE of the entire system, for simplicity, the equal power assumption was invoked in the single-UE scenarios. Hence, the optimisation in VT and the CT clusters can be operated in the same framework in Section IV and V.

derived for the CT and the VT clusters may be written as

$$p_{u,l}^{\text{tx}} = \sum_{a=1}^{N_{A,n}} g^2[a,u] \cdot p_{u,l}^m, \qquad (17)$$

where

$$g[a,u] = \begin{cases} g_{\text{CT}}[a,u], & u \in \mathcal{V}_{U,n^{\text{CT}}}, \\ g_{\text{VT}}[a,u], & u \in \mathcal{V}_{U,n^{\text{VT}}}, \end{cases} \qquad (18)$$

while the SINR received by the UE u may both be expressed as in (12).

IV. PROBLEM FORMULATION

In this section, the design problem of our energy-efficient scalable video scheme is formulated and both the video-related and power-related constraints are discussed.

A. Objectives

Let us now formulate the problem of energy-efficient scalable video streaming in our UC-VLC aided network. A general function $\mathcal{E}_{u,l}^m(\mathcal{P})$ is used for denoting the EE of the UE u receiving the video layer l associated with the AM mode m, where $\mathcal{P}$ represents the PA strategy of all UEs. For simplicity, we assume that the video chips requested by the different UEs have the same number of layers, but our algorithm can be readily applied, when considering different number of video layers for the different UEs. Formally, our goal is to maximise the sum EE of all UEs, which may be written as:

$$\max_{\mathcal{P}} \quad \Gamma = \max_{\mathcal{P}} \quad \sum_{\mathcal{C}_n} \sum_{u \in \mathcal{V}_{U,n}} \sum_{l=0}^{L-1} \sum_{m=1}^{M} \mathcal{E}_{u,l}^m(\mathcal{P})$$

$$= \max_{\mathcal{P}} \quad \sum_{\mathcal{C}_n} \sum_{u \in \mathcal{V}_{U,n}} \sum_{l=0}^{L-1} \sum_{m=1}^{M} \frac{\Delta_{u,l}^m(\mathcal{P})}{p_{u,l}^{\text{tx}}}, \qquad (19)$$

where $\Delta_{u,l}^m(\mathcal{P})$ is a general utility function assumed to be non-negative and non-decreasing in order to adequately reflect our QoS metric [30] and $p_{u,l}^{\text{tx}}$ denotes the total electronic power consumption as indicated in (17), when the UE u receives the lth video layer. Note that the expression of (19) supports a more scalable and efficient distributed solution, when compared to simply maximising the system-based global EE, which is usually defined as the aggregated throughput of all UEs divided by the aggregated power consumption. Furthermore, (19) may be maximised independently within a cluster, provided that the interference received is conservatively assumed to be always at its maximum value, which constitutes the worst-case scenario.

Moreover, we introduce a binary indicator $x_{u,l}^m$. If the AM mode m is assigned to transmit the video layer l of the UE u, we have $x_{u,l}^m = 1$; otherwise, $x_{u,l}^m = 0$. Therefore, our problem is a joint AM mode assignment and PA problem, where the objective function (OF) of (19) may be reformulated as

$$\max_{\mathcal{P}_n,\mathcal{X}_n} \quad \Gamma_n = \max_{\mathcal{P}_n,\mathcal{X}_n} \quad \sum_{u \in \mathcal{V}_{U,n}} \sum_{l=0}^{L-1} \sum_{m=1}^{M} \frac{\Delta_{u,l}^m(\mathcal{P}_n) \cdot x_{u,l}^m}{\sum_{a=1}^{N_{A,n}} g^2[a,u] \cdot p_{u,l}^m}, \qquad (20)$$

where we have $\mathcal{X}_n = \{x_{u,l}^m : u \in \mathcal{V}_{U,n}, m = 1, \cdots M, l = 0, \cdots, L-1\}$ and $\mathcal{P}_n$ denotes the PA strategy within the cluster

$\mathcal{C}_n$. Furthermore, $\Delta_{u,l}^m(\mathcal{P}_n)$ denotes the achievable QoS utility corresponding to the AM mode assignment m for the video layer l of the UE u, which is a function of $\mathcal{P}_n = \{p_{u,l}^m : u \in \mathcal{V}_{U,n}, m = 1, \cdots M, l = 0, \cdots, L-1\}$. Hence, after solving $\mathcal{P}_n$ in (19) within each cluster, the optimal solution of the EE maximisation problem in (19) is found by combining all $\mathcal{P}_n$.

B. Constraints

Let us now elaborate on the associated constraints, including both our video-related constraints and power-related constraints. During a video frame, the constraints imposed on video transmission may be formulated as follows:

$$\sum_{l=0}^{L-1} \sum_{m=1}^{M} t_{u,l}^m \cdot x_{u,l}^m \le T, \qquad \forall u \in \mathcal{V}_{U,n}; \quad (21)$$

$$\sum_{m=1}^{M} x_{u,l}^m = 1 : \prod_{l'=0}^{l-1} \sum_{m=1}^{M} x_{u,l'}^m = 1, \qquad \forall l > 0; \quad (22)$$

$$\sum_{m=1}^{M} x_{u,l}^m \le 1, \qquad \forall u \in \mathcal{V}_{U,n}, \forall l; \quad (23)$$

$$x_{u,l}^m \in \{0,1\}, \qquad \forall u \in \mathcal{V}_{U,n}, \forall l, \forall m, \quad (24)$$

where $t_{u,l}^m = S_{u,l}/r^m$ denotes the time required for transmitting the lth video layer constituted by $S_{u,l}$ bits to the UE u using the AM mode m. Note that ":" represents the condition. For example, $f(x) : g(x)$ means that $f(x)$ is true provided that $g(x)$ is true. The constraint in (21) guarantees that the time required for transmitting a video frame does not exceed its upper limit T. Note that our scalable video scheme only has a single-video-frame delay, but it can be generalised to various video services having different delay tolerance by loosening the constraint of (21). Furthermore, (22) states the nested dependence among the different layers of the SHVC-encoded video, which ensures once the $(l+1)$st video layer is received, all the previous lower video layers must have been received as well, provided that we assume a linear dependency model for the scalable video, where the video layers of the current frame are independent of those of the other frames. Moreover, (23) guarantees each video layer is assigned a specific AM mode.

In addition, since the transmit power of each AP should not exceed its maximum value of $p_{\text{max}}^{\text{tx}}$, the power-related constraints in the electronic domain of both the VT and CT clusters may be written as

$$\sqrt{\sum_{u=1}^{N_{U,n}} g^2[a,u]\bar{p}_u/2\pi} \le p_{\text{max}}^{\text{tx}}, \quad \forall a \in \mathcal{V}_{A,n}, \qquad (25)$$

where $\bar{p}_u$ denotes the average power allocated to the UE u, which may be formulated as

$$\bar{p}_u = \frac{\sum_{l=0}^{L-1} \sum_{m=1}^{M} t_{u,l}^m \cdot x_{u,l}^m \cdot p_{u,l}^m}{\sum_{l=0}^{L-1} \sum_{m=1}^{M} t_{u,l}^m \cdot x_{u,l}^m}. \qquad (26)$$

Note that we only consider the transmission power as the sole reason for power dissipation in (25). In our future work, other sources of power consumption, such as signal processing costs, back-haul power consumption etc. would also be taken into account in order to paint a holistic picture in terms of the
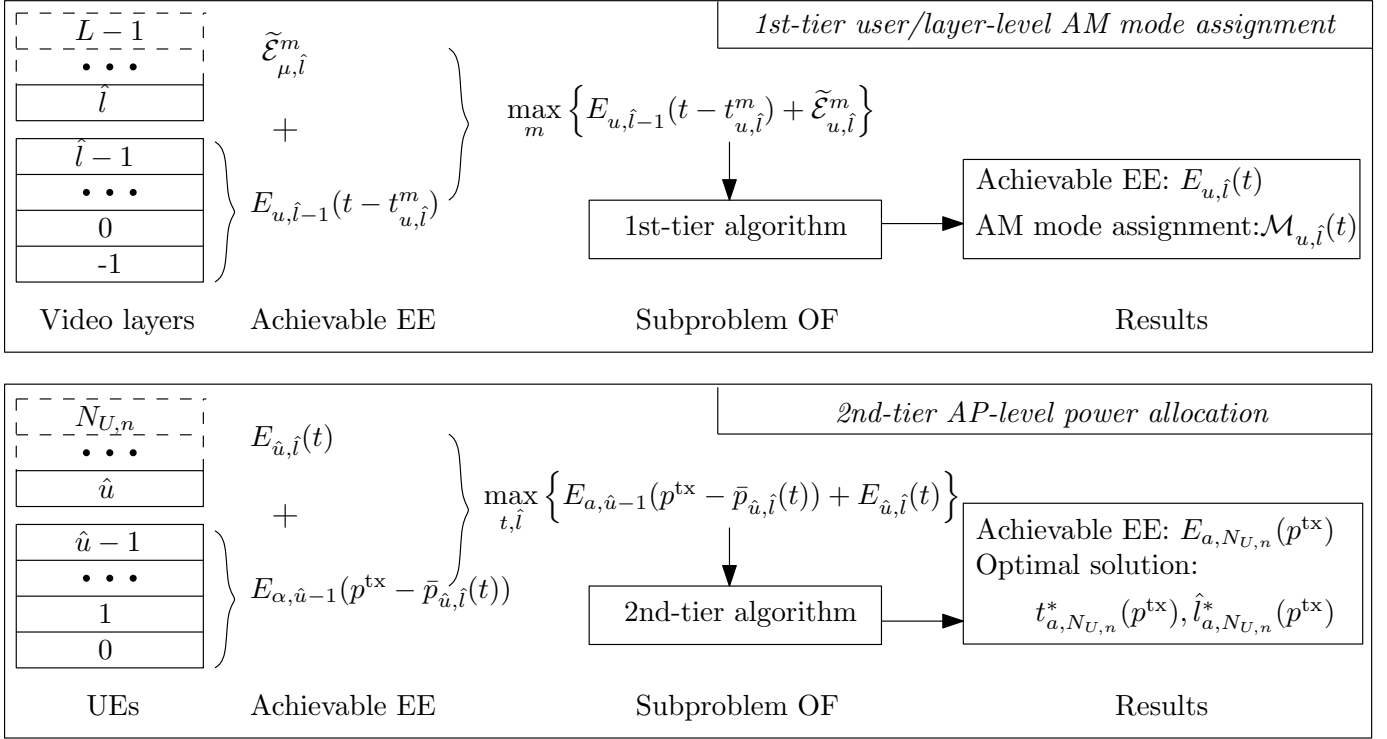
Fig. 3: The mind map of the 1st-tier user/layer-level AM mode assignment and the 2nd-tier AP-level PA.

overall network's power consumption. However, at the time of writing, quantifying the network's power consumption in case of VLC systems remains an open issue, since integrating our scheme with a certain VLC back-haul deserves dedicated treatment.

In order to solve our mixed integer non-linear non-convex problem of (20)-(25), we propose a heuristic 3-tier dynamic-programming-based algorithm, which is constituted by the 1st-tier user/layer-level AM mode assignment, the 2nd-tier AP-level PA and the 3rd-tier cluster-level EE optimisation. The user/layer-level AM mode assignment step mainly copes with the constraints in (21)-(24), which are determined by the video characteristics, regardless of the power-level budget of (25). All the solutions provided by the user/layer-level AM mode assignment step will be regarded as the input to the following AP-level PA step, where all the feasible solutions satisfying both the video characteristic constrains of (21)-(24) as well as the power constraint (25) can be found. Finally, during the cluster-level EE optimisation step, the solution of our problem (20)-(25) can be obtained based on our previous discussions.

V. DYNAMIC-PROGRAMMING-BASED ALGORITHM

In this section, a 3-tier dynamic-programming-based algorithm is proposed for solving our problem formulated in Section IV, including its user/layer-level AM mode assignment, the AP-level PA and the cluster-level EE optimisation.

A. User/layer-level AM Mode Assignment

In this step, we first deal with the constraints (21)-(24) imposed by the video characteristics, which are determined by each individual UE, but are transparent to how the clusters are formed. Thus, let us bypass the PA across the given cluster

and consider the problem on the basis of a single UE with the aid of the mind map of our proposed algorithm in Fig. 3.

1) Objectives: Given a specific UE u, the user/layer-level AM mode assignment problem may be formulated similar to a multi-choice knapsack problem [41], which is NP-hard. The classic multi-choice knapsack problem may be described as follows. Multiple groups of items are considered and each item has a value and a resource cost. The objective is to select exactly one item from each group into the knapsack and to maximise the total value of the selected items, whilst not exceeding the resource capacity available. In our problem, the L video layers may be viewed as the multiple groups, where the M AM modes are the items. The achievable EE and the corresponding transmission time costs may be interpreted as the value and the resource cost. Our goal is to select a specific one AM mode for each transmitted video layer in order to maximise the EE within the single-video-frame delay budget. We assume furthermore that if a video layer is not transmitted, no AM mode will be assigned to it. In contrast to the classic multi-choice knapsack problem, not all the video layers have to be transmitted by selecting an AM mode, but if the video layer l is transmitted, all the lower-order video layers must be transmitted and be assigned a specific AM mode as well. Note that the video layers' order requirement (22) is readily satisfied by employing our dynamic-programming-based algorithm as detailed below, since the proposed algorithm is implemented in accordance with the video layers' order, i.e. from layer 0 to layer $L-1$.

Before introducing our dynamic-programming-based algorithm, let us define an auxiliary function $E_{u,\hat{l}}(t)$ first, which may be interpreted as the maximum EE for the given UE u,

$$E_{u,\hat{l}}(t) = \max_{x_{u,l}^m} \left\{ \sum_{l=0}^{\hat{l}} \sum_{m=1}^{M} \frac{\Delta_{u,l}^m(p_{u,l}^m) \cdot x_{u,l}^m}{\sum_{a=1}^{N_{A,n}} g^2[a,u] \cdot p_{u,l}^m} : t_{u,\hat{l}} \le t, \ \sum_{m=1}^{M} x_{u,l}^m \le 1, \ x_{u,l}^m \in \{0,1\}, \ \forall l = 0,1,\dots,\hat{l} \right\}, \qquad (27)$$

when receiving the video layers from 0 to $\hat{l}$ and the transmission time duration is limited to t, where we have $0 \le \hat{l} \le L-1$ and $0 \le t \le T$. Since the time resource t has a continuous value, it is further decomposed into σ discrete values, which may be written as $t = 0, T/\sigma, 2T/\sigma, \cdots, T$. Following the basic principles of the dynamic programming approach, we consider the sub-problem under a specific transmission time limit t. Hence, $E_{u,\hat{l}}(t)$ may be written as (27), where $t_{u,\hat{l}}$ is the total transmission time required for transmitting all the video layers up to $\hat{l}$ and we have $t_{u,\hat{l}} = \sum_{l=0}^{\hat{l}} \sum_{m=1}^{M} t_{u,l}^m \cdot x_{u,l}^m$. Note that during this user/layer-level AM mode assignment procedure, we focus our attention on each single UE and the PA strategy $\mathcal{P}_n$ across the entire cluster may be treated in the next step. Hence, the achievable QoS utility has been written as $\Delta_{u,l}^m(p_{u,l}^m)$ in (27), instead of $\Delta_{u,l}^m(\mathcal{P}_n)$ as in (20).

2) Methodologies: Given a specific video layer having the length of $S_{u,l}$ bits, the lower the order of the AM mode, the longer the transmission time becomes, since the lower-order AM mode provides a lower data rate. We start the recursion of $E_{u,\hat{l}}(t)$ from the virtual base layer, which may be referred to as the '−1st' video layer. Since transmitting the virtual base layer represents no transmission, for any transmission time limit t, we have

$$E_{u,-1}(t) = 0, p_{u,-1}^m = 0, t_{u,-1}^m = 0, \forall t, \forall m, \qquad (28)$$

where $p_{u,-1}^m$ denotes the power allocated for the virtual base layer with the mth AM mode assigned to the UE u, which should be zero, since no power is consumed. Similarly, the corresponding time duration is also zero. In order to obtain a general recursion formulation for $E_{u,\hat{l}}(t)$, let us consider the base layer transmission as an explicit example. As mentioned above, if a certain video layer is transmitted, all the lower-order video layers must have been transmitted. Hence, when transmitting the base layer, as its lower video layer, the virtual '−1st' video layer must be transmitted. It is plausible that $(t_{u,0}^M + t_{u,-1}^M)$ is the least time required for transmitting the base layer by employing the highest-order AM mode M for the '−1st' and 0th video layer. If the time limit t is even shorter than $(t_{u,0}^M + t_{u,-1}^M)$, there will be no feasible solution for maximising $E_{u,0}(t)$, where $E_{u,0}(t)$ is assumed to be $-\infty$ and hence no AM mode will be assigned. Otherwise, for all $t \ge (t_{u,0}^M + t_{u,-1}^M)$, if the base layer is assigned an AM mode m, we need a duration of $t_{u,0}^m$ to transmit, when we have associated EE value of $\Delta_{u,0}^m(p_{u,0}^m)/p_{u,0}^m$. In this case, the time allocated to the virtual base layer is reduced to $(t - t_{u,0}^m)$. Hence, the problem of the AM mode assignment for transmitting the '−1st' and the 0th video layers may be formulated as (29). Similarly, for the other video layers $\hat{l} = 1,2,\dots,L-1$, as shown in the top subfigure of Fig. 3, a general formulation of $E_{u,\hat{l}}(t)$ may be written as (30), where the EE increment $\widetilde{\mathcal{E}}_{u,\hat{l}}^m = \Delta_{u,\hat{l}}^m(p_{u,\hat{l}}^m)/(\sum_{a=1}^{N_{A,n}} g^2[a,u] \cdot p_{u,\hat{l}}^m)$. Note that $(t - t_{u,\hat{l}}^m)$ may not be equal to any of the specific time

durations of $0, T/\sigma, 2T/\sigma, \cdots, T$. In this case, we assume that $(t - t_{u,\hat{l}}^m)$ corresponds to its nearest lower time-duration. For example, if the exact value of $(t - t_{u,\hat{l}}^m)$ is approximately $2.66T/\sigma$, we round it to $2T/\sigma$ and substitute $E_{u,\hat{l}-1}(2T/\sigma)$ into (30).

3) Results: Relying on the relationships formulated above, we obtain the AM mode assignment for each video layer of a specific UE u under various time limits. To elaborate with the aid of the top subfigure of Fig. 3, let $\mathcal{M}_{u,\hat{l}}(t)$ be an $(1 \times L)$-element vector of the AM mode assigned for each video layer under the time-resource t and the highest-order transmitted video layer $\hat{l}$. Since no AM mode is assigned to the video layers spanning from $\hat{l}+1$ to $L-1$, the corresponding elements in $\mathcal{M}_{u,\hat{l}}(t)$ are set to zero. Thus, the corresponding achievable EE value $E_{u,\hat{l}}(t)$ can be obtained from (30). Correspondingly, the power allocated for transmitting each video layer is known according to (15). Hence, corresponding to each $\mathcal{M}_{u,\hat{l}}(t)$ under the time limit t, the average PA $\bar{p}_{u,\hat{l}}(t)$ can be calculated according to (26).

Following the user/layer-level AM mode assignment step, we can obtain the $(\sigma \times L)$-element time-guaranteed EE matrix $E_u = \{E_{u,\hat{l}}(t) : t = T/\sigma \cdots, T, \hat{l} = 0, \cdots L-1\}$. Furthermore, each UE u has the time-guaranteed $(L \times \sigma \times L)$-element AM mode assignment matrix $\mathcal{M}_u$, which may be expressed as $\mathcal{M}_u = \{\mathcal{M}_{u,\hat{l}}(t) : t = T/\sigma, \cdots, T, \hat{l} = 0, \cdots L-1\}$, which is referred to as the time-guaranteed AM mode matrix of the UE u. Correspondingly, the $(\sigma \times L)$-element average PA matrix $\bar{p}_u = \{\bar{p}_{u,\hat{l}}(t) : t = T/\sigma, \cdots, T, \hat{l} = 0, \cdots L-1\}$ can also be obtained. Since all the results obtained so far are time-guaranteed and on a per-UE basis, we only have to find the highest achievable EE from E_u and its corresponding AM mode assignment, whilst avoiding the violation of the power constraint (25) imposed by the optical APs within each cluster.

B. AP-level PA

Let us now introduce the 2nd-tier AP-level PA algorithm, which is carried out within each UC cluster. Our objective in this step is to find the appropriate AM mode assignment solutions for each UE from its time-guaranteed AM mode assignment matrix $\mathcal{M}_u$, which further satisfies the power constraint (25) and maximises the achievable EE.

1) Objectives: Similarly to the previous section, we first define an auxiliary function $E_{a,\hat{u}}(p^{tx})$, which may be interpreted as the maximum EE, when supporting a total of $\hat{u}$ UEs within a specific cluster $\mathcal{C}_n$ and the given AP $a \in \mathcal{V}_{A,n}$ with a maximum power limit p^{tx}, where we have $0 \le \hat{u} \le N_{U,n}$ and $0 \le p^{tx} \le p^{tx}_{max}$. Then we further discretize the transmit power into J levels, which may be written as $p^{tx} = p^{tx}_{max}/J, 2p^{tx}_{max}/J, \cdots, p^{tx}_{max}$. Hence, considering the cluster supporting $\hat{u}$ UEs, for a given AP a with a power limit p^{tx} within the cluster, $E_{a,\hat{u}}(p^{tx})$ may be formulated as (31), where $x_{u,\hat{l}}(t)$ is a binary indicator. If the AM mode assignment $\mathcal{M}_{u,\hat{l}}(t)$ is selected for the UE u, we have $x_{u,\hat{l}}(t) = 1$; otherwise, $x_{u,\hat{l}}(t) = 0$.

$$E_{u,0}(t) = \begin{cases} -\infty, & \text{for } 0 \leq t < t_{u,0}^M + t_{u,-1}^M; \\[2ex] \max_m \left\{ E_{u,-1}(t - t_{u,0}^m) + \dfrac{\Delta_{u,0}^m(p_{u,0}^m)}{\sum_{a=1}^{N_{A,n}} g^2[a,u] \cdot p_{u,0}^m} : t_{u,0}^m \leq t, 1 \leq m \leq M \right\}, & \\[2ex] & \text{for } t_{u,0}^M + t_{u,-1}^M \leq t \leq T. \end{cases} \quad (29)$$

$$E_{u,\hat{l}}(t) = \begin{cases} -\infty, & \text{for } 0 \leq t < \sum_{l=-1}^{\hat{l}} t_{u,l}^M; \\[2ex] \max_m \left\{ E_{u,\hat{l}-1}(t - t_{u,\hat{l}}^m) + \widetilde{\mathcal{E}}_{u,\hat{l}}^m : t_{u,\hat{l}}^m \leq t, 1 \leq m \leq M \right\}, & \text{for } \sum_{l=-1}^{\hat{l}} t_{u,l}^M \leq t \leq T, \end{cases} \quad (30)$$

$$E_{a,\hat{u}}(p^{\mathrm{tx}}) = \max_{t,\hat{l}} \left\{ \sum_{u=1}^{\hat{u}} \sum_{t=0}^{T} \sum_{\hat{l}=0}^{L-1} E_{u,\hat{l}}(t) \cdot x_{u,\hat{l}}(t) : \sum_{t=T/\sigma}^{T} \sum_{\hat{l}=0}^{L-1} x_{u,\hat{l}}(t) = 1, \forall u = 1, \cdots, \hat{u}, \right.$$
$$\left. \sqrt{\sum_{u=1}^{\hat{u}} \sum_{t=T/\sigma}^{T} \sum_{\hat{l}=0}^{L-1} \frac{g^2[a,u] \cdot \bar{p}_{u,\hat{l}}(t) \cdot x_{u,\hat{l}}(t)}{2\pi}} \leq p^{\mathrm{tx}}, x_{u,\hat{l}}(t) \in \{0,1\} \right\}, \quad (31)$$

$$E_{a,\hat{u}}(p^{\mathrm{tx}}) = \begin{cases} -\infty, & \text{for } 0 \leq p^{\mathrm{tx}} < \sqrt{\sum_{u=1}^{\hat{u}} g^2[a,u] \cdot \bar{p}_u^{\min}/(2\pi)}; \\[2ex] \max_{t,\hat{l}} \left\{ E_{a,\hat{u}-1}(p^{\mathrm{tx}} - \bar{p}_{\hat{u},\hat{l}}(t)) + E_{\hat{u},\hat{l}}(t) : \sqrt{g^2[a,u] \cdot \bar{p}_{\hat{u},\hat{l}}(t)/(2\pi)} \leq p^{\mathrm{tx}} \right\}, & \\[2ex] & \text{for } \sqrt{\sum_{u=1}^{\hat{u}} g^2[a,u] \cdot \bar{p}_u^{\min}/(2\pi)} \leq p^{\mathrm{tx}} \leq p_{\max}^{\mathrm{tx}}. \end{cases} \quad (33)$$

2) Methodologies: Next, we find the solution in an iterative manner. Let $\bar{p}_u^{\min}$ denote the minimum average power required, when transmitting video to the UE u, where we have $\bar{p}_u^{\min} = \min\{\min\{\bar{p}_u\}\}$. Similarly to the concept of the virtual base layer, we assume a virtual scenario, in which no UE is served, so that we have:

$$E_{a,0}(p^{\mathrm{tx}}) = 0, \bar{p}_{0,\hat{l}}(t) = 0, \quad \forall p^{\mathrm{tx}}, \forall t, \forall \hat{l}. \quad (32)$$

Hence, the basic principle of the general recursion for the AP-level PA step is similar to (30) of the previous section, as shown in the bottom subfigure of Fig. 3, which furthermore may be written as (33). Similar to (30) again, $(p^{\mathrm{tx}} - \bar{p}_{\hat{u},\hat{l}}(t))$ is rounded down to its nearest lower power level, when it is not equal to any of the legitimate power levels.

3) Results: By iteratively solving (33), the potential PA schemes are found, when separately considering the power limit of each AP. Since all the UEs within a single cluster must be fully supported, $E_{a,N_{U,n}}(p^{\mathrm{tx}})$ is the achievable EE corresponding to the AP a subjected to a power constraint p^{tx}, as shown in the bottom subfigure of Fig. 3. Correspondingly, the solution of (33) for achieving $E_{a,N_{U,n}}(p^{\mathrm{tx}})$ is denoted as $(t_{a,N_{U,n}}^*(p^{\mathrm{tx}}), \hat{l}_{a,N_{U,n}}^*(p^{\mathrm{tx}}))$, where $t_{a,N_{U,n}}^*(p^{\mathrm{tx}})$ and $\hat{l}_{a,N_{U,n}}^*(p^{\mathrm{tx}})$ are both $(1 \times N_{U,n})$-element vectors. Thus, the

PA scheme may be obtained by find the corresponding element $\bar{p}_{u,\hat{l}_{a,N_{U,n}}^*(p^{\mathrm{tx}})}[1,u](t_{a,N_{U,n}}^*(p^{\mathrm{tx}})[1,u])$ from the previous time-guaranteed PA matrix $\bar{p}_u$ for each of the UE u.

We further obtain the $(1 \times J)$-element vector $E_a = \{E_{a,N_{U,n}}(p^{\mathrm{tx}}) : p^{\mathrm{tx}} = p_{\max}^{\mathrm{tx}}/J, \cdots, p_{\max}^{\mathrm{tx}}\}$ to denote the achievable EE under various power limits and also the $(J \times N_{U,n})$-element PA matrix $\bar{p}_a$ at the AP a, where each row represents an PA scheme for the cluster $\mathcal{C}_n$ and we have J PA schemes. Considering all the $N_{A,n}$ APs, we now have a total of $(N_{A,n} \times J)$ PA schemes for a cluster $\mathcal{C}_n$. Next let us determine the final solution of the PA in the following section.

C. Cluster-level EE Optimisation

1) Solutions: Given a specific cluster $\mathcal{C}_n$, we have obtained both the $(N_{A,n} \times J)$ PA schemes denoted as a $(J \times N_{U,n} \times N_{A,n})$-element matrix $\{\bar{p}_a : a = 1, \cdots, N_{A,n}\}$ and the corresponding $(J \times N_{A,n})$-element EE matrix $\{E_a : a = 1, \cdots, N_{A,n}\}$. Then the solution of the PA problem formulated for achieving the maximum EE may be found by

$$\{a*, p^{\mathrm{tx}*}\} = \arg\max_{a,p^{\mathrm{tx}}} \{\{E_a\}\}, \quad (34)$$

where the solution $\{a*, p^{\mathrm{tx}*}\}$ corresponds to a row of the PA matrix $\bar{p}_{a*}$, which may be written as $\bar{p}_{a*}[j,:]$ and we have

TABLE I: Simulation parameters

VLC parameters	
Semi-angle at half power ($\phi_{1/2}$)	70°
Gain of an optical filter ($T_s(\psi)$)	1.0
Modulation bandwidth (B) [40]	20 [MHz]
Refractive index (n_r)	1.5
Detection area of a PD (D_{PA})	1.0 [cm^2]
O/E conversion efficiency (γ)	0.53 [A/W]
Video parameters	
Video sequence	RaceHorses
QPs/Frame	40,26,16
Resolution	416×240
Video codec	SHVC
FPS	30
No. of video layers	3
Bits per pixel	8
Error-free Y-PSNR	28.3, 36.9, 44.7 [dB]
Bitrate	5.1 [Mbits/s]
Representation	YUV 4:2:0
Error concealment	Frame-copy
No. of frames	30, 300
Coding structure	IPPPIPPP...
Algorithm parameters	
The range of each AP subset (d_α)	15 [m]
No. of discrete time values (σ)	100
No. of power level (J)	10

$j = J \cdot p^{\mathrm{tx}*}/p^{\mathrm{tx}}_{\max}$. It is plausible that the PA scheme $\bar{p}_{a*}[j,:]$ selected satisfies the power constraint (25) of the AP $a*$, but we have to check if it violates other APs' power limits. If $\bar{p}_{a*}[j,:]$ is feasible for all APs, it is the final solution of the PA for our energy-efficient video scheme. Otherwise, $\bar{p}_{a*}[j,:]$ and its corresponding EE $E_{a*,N_{U,n}}(p^{\mathrm{tx}*})$ are set to zero and we repeat (34) until we find a feasible PA scheme satisfying (25) for all APs within the cluster. Finally, each element of $\bar{p}_{a*}[j,:]$ corresponds to a specific layer-level PA $\bar{p}_{u,\hat{i}}(t)$ and to a specific AM mode assignment vector $\mathcal{M}_{u,\hat{i}}(t)$.

2) Complexity: The complexity imposed by our 3-tier dynamic-programming-based algorithm is much lower than that of the optimal exhaustive search, which is on the order of $O(L^{(M \cdot N_{U,n})})$ for a specific cluster $\mathcal{C}_n$. By contrast, during the 1st-tier AM mode assignment of the proposed algorithm, the complexity for each UE is on the order of $O(\sigma M L)$. Then in the PA step, the complexity imposed by each AP is $O(\sigma N_{U,n} J)$ and the last EE optimisation has a complexity of $O(N_{A,n} J)$. Furthermore, the complexity of our 1st-tier AM mode assignment algorithm may be reduced by using the following rule.

Remark 2. For any $\bar{p}_u$, if there exist two elements p_i and $p_{i'}$ such that $p_i \geq p_{i'}$ and the corresponding EE obeys $E_i < E_{i'}$, where E_i, $E_{i'} \in E_u$, then E_i is set to zero.

VI. PERFORMANCE EVALUATION

In this section, we will present our simulation results characterising the 3-tier dynamic-programming-based scheme conceived for energy-efficient scalable video transmission in the VLC system, with a special emphasis on our UC cluster formation. Before presenting our simulation results, let us elaborate further on the implementation of our energy-efficient video system. Relying on the positioning and tracking module, distance-based cluster formation is constructed. Within each cluster, our proposed optimisation algorithm is activated upon the UEs' requests. The solution gives the appropriate video layer selection, the adaptive mode assignment as well as the power allocation for the sake of maximising the EE. Thus different transmission strategies, i.e. VT and CT, are employed in multi-UE clusters and single-UE clusters correspondingly, based on the solution. Regarding the video implementation, at the transmitter side, only the carefully selected layers of the scalable SHVC stream are transmitted. At the decoder side, all the received layers will be decoded. Then, the error-infested layers will be discarded together with their dependent layers. This will result in a decodable bitstream, which is then passed to the SHVC decoder for reconstructing the final YUV clip.

A 15m$\times$15m$\times$3m room model is considered, which is covered by the VLC down-link, including (4$\times$4) uniformly distributed optical APs at a height of 2.5m. The parameters of the LED arrays are summarised in TABLE I. Our investigations include both the LoS and the first reflected light-path, where the channel's DC attenuation is given by (1) and (2). As for the AM mode, we consider binary phase-shift keying (BPSK), 4-QAM, 16-QAM, 64-QAM and 256-QAM in our simulations, while our algorithm is a general one, which is not limited to the specific AM modes considered in this paper. Furthermore, advanced SHVC techniques are employed in our VLC system, where the 3-layer video sequence "RaceHorses" is transmitted at 30 FPS. The error-free Y-PSNR is 28.3dB, 36.9dB and 44.7dB, respectively, when the first, second and third video layer is received. At this stage, we assume that if a video layer is allocated power and assigned a specific AM mode, it can be successfully received, provided that a sufficiently low BER requirement is satisfied in the indoor VLC environment considered. If a frame is not transmitted, the low-complexity error concealment technique of frame-copying is used, where the lost frame is replaced by the most recent successfully decoded frame. The general QoS metric of $\Delta_{u,l}^m(p_{u,l}^m)$ is assumed to be the achievable throughput corresponding to the specific AM mode assignment m for the video l of the UE u, i.e. we have $\Delta_{u,l}^m(p_{u,l}^m) = r^m$. The parameters of the video sequence are detailed in Table I. Each of our simulation results are averaged over 50 independent simulation runs, where the UEs are randomly distributed in the room.

A. EE v.s. *Video Quality Investigations*

1) Investigations for Various FoV and UE Density: Since the FoV is one of the most influential parameters in VLC networks, we consider various FoVs and their effects on the
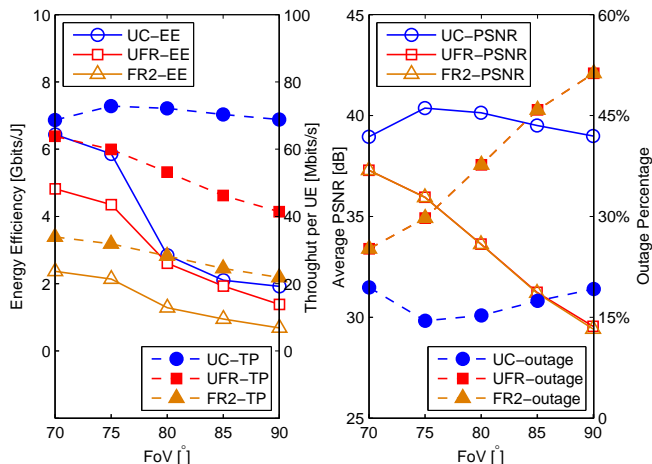
Fig. 4: Left: EE of the VLC network and average TP per UE for various FoVs. Right: Average PSNR and the UE outage percentage. The optical power is 0.3862W for maintaining the illuminance requirement of our room model as a typical work place.

attainable system performance, as shown in Fig. 4 [6]. The system's EE is reduced, when the FoV is increased, since a higher transmit power is required, when the interference is increased, while our UC clusters remain superior in all the scenarios considered. As for the average throughput (TP) and for the video PSNR per UE, which is used for quantifying the reconstructed video quality at the receiver side, they all reach a peak at FoV = 75° considering the UC design, due to having the lowest number of UEs in outage, according to the right subfigure of Fig. 4. When the FoV is further increased, the UEs suffer from more interference and some of them may experience an outage as a result. Note that in this case, increasing the transmitted power may not be an appealing solution, since according to (13) the interference is also increased proportionally. The traditional cell designs of UFR and FR2 have quite similar performances in terms of their video quality and the percentage of UEs in outage, although UFR provides a higher EE and TP. Hence using traditional frequency reuse may not be a desirable solution for interference mitigation in our VLC networks.

Fig. 5 shows the attainable system performance for various UE densities, ranging from a low-density scenario of 10 UEs to a high-density scenario of 40 UEs, where the FoV is 90° and the maximum transmitted optical power is 0.3862W. This power was calculated according to the illumination constraint of maintaining an average illuminance of 600lx, a minimum of 200lx and a maximum of 800lx, which are typical values for the work place. As expected, our proposed UC clusters are capable of offering the highest EE, TP and PSNR in most of the scenarios considered, whilst simultaneously supporting

[6]In Fig. 4, we observed the FoVs from 70° to 90°. With the FoV 70°/75°, the UE is capable of receiving data from two neighbouring APs and the area contaminated by potential interference is modest. When the FoV is increased to 80°/85°/90°, the UE is capable of receiving data from four APs and the potential interference contaminated area is also increased. Although the five FoVs considered have quite similar absolute values, but they correspond to different interference levels
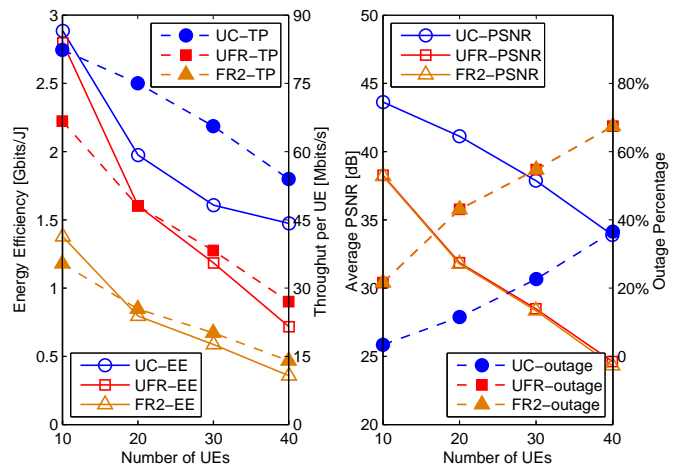

Fig. 5: Left: EE of the VLC network and average TP per UE for various UE densities. Right: Average PSNR and the UE outage percentage, where FoV = 90°. The optical power is 0.3862W for maintaining the illuminance requirement of our room model as a typical work place.
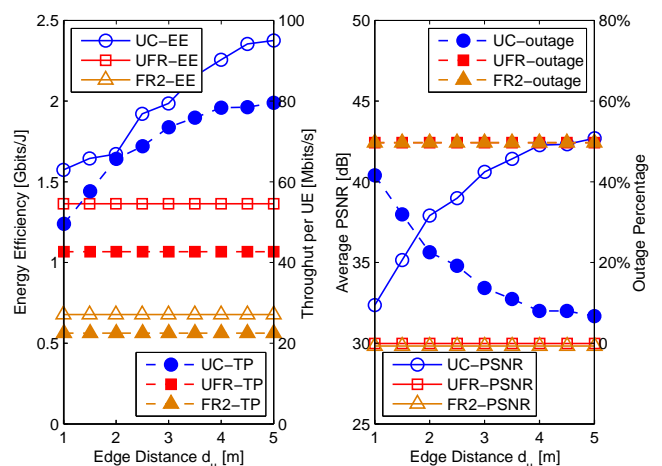

Fig. 6: Left: EE of the VLC network and average TP per UE for various UC cluster edge distance d_μ. Right: Average PSNR and the UE outage percentage, where FoV = 90°. The optical power is 0.3862W for maintaining the illuminance requirement of our room model as a typical work place.

more UEs at a given UE outage percentage, as seen in the right subfigure of Fig. 5.

2) Investigations for UC Cluster Edge Distance d_μ and Maximum Transmitted Optical Power: As mentioned in Section II-B, the various characteristics of the UC clusters may be satisfied upon adjusting the value of the atto-cell edge distance d_μ. For example, a large d_μ may improve the area spectral efficiency of the system according to our previous work [5]. Fig. 6 illustrates the system's performance for different d_μ values in terms of the EE, TP, PSNR and the outage percentage. As expected, upon increasing the value of d_μ, both the EE and TP are increased. Similarly, the average video quality is also enhanced for larger d_μ. Although a larger d_μ results in large-scale UC clusters, they potentially incur additional signal processing costs, such as the inversion of
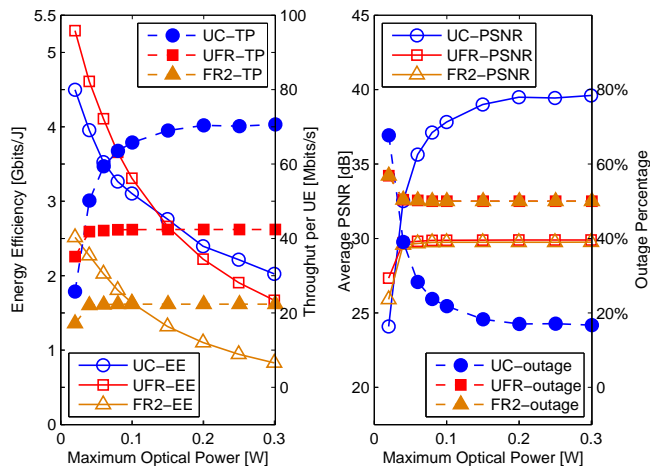
Fig. 7: Left: EE of the VLC network and average TP per UE for various transmitted optical power thresholds. Right: Average PSNR and the UE outage percentage, where FoV = $90°$.

a large matrix at the distributed APs as required by the ZF based transmit pre-coding. Furthermore, sharing data amongst the APs of large UC clusters may require a more capable back-haul. Hence, we have to determine the most beneficial edge-distance d_μ of our UC cluster in order to minimise the network's power consumption by additionally considering the power-dissipation of the signal processing as well as of the back-haul.

Fig. 7 shows the system's performance for different optical transmit power thresholds. Although the UC clusters are not as efficient as the traditional UFR design, when the maximum transmitted optical power is limited to 0.1W, it is capable of providing a higher TP as well as better video quality in most of the scenarios considered. Upon increasing the optical power threshold, the EE becomes lower, but the TP and video quality improvements of the system remain slow in all the scenarios considered. Note that the various optical transmit power thresholds may be realised by appropriately adjusting the LED array of each specific optical AP, in order to maintain the associated illuminance requirement.

B. Mobile UE Investigations

Our proposed UC clusters and the corresponding PA algorithm are now further investigated, when the UEs are mobile. Let the UE with index '1' move at 1.25m/s following a trajectory shown in Fig. 8a, where its UC cluster is dynamically constructed. The location of the UE is reported at the beginning of every video frame, i.e. every 1/30 s. The UC clusters for Frame 1, Frame 91, Frame 181 and Frame 271 are shown in Fig. 8a, respectively. Each frame's PSNR is shown in Fig. 8b for UE 1, where most frames have a high PSNR and a high subjective video quality, when UE 1 is moving and the UC clusters are dynamically updated. However, almost half of the 300 frames have very poor quality or have been lost, when employing the traditional UFR/FR2 design. Furthermore, the 5dB or so PSNR fluctuations on top of the curves seen in Fig. 8b are due to the coding structure of intra-coded (I) and
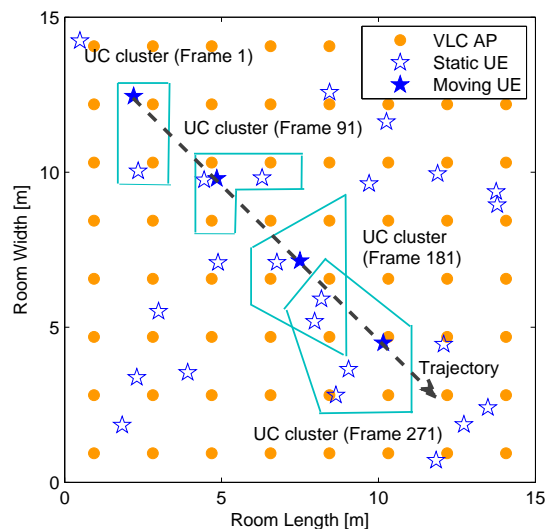
(a)
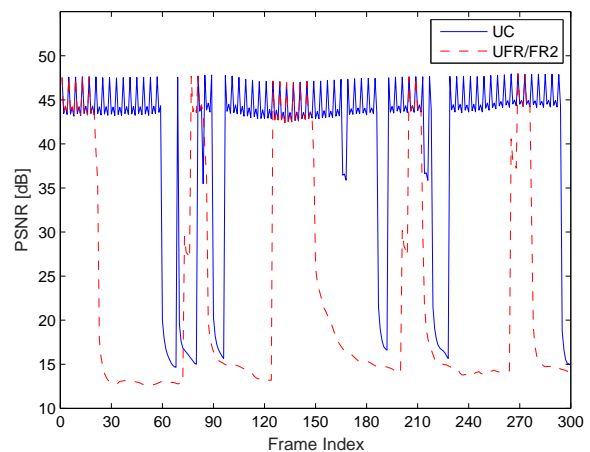
(b)
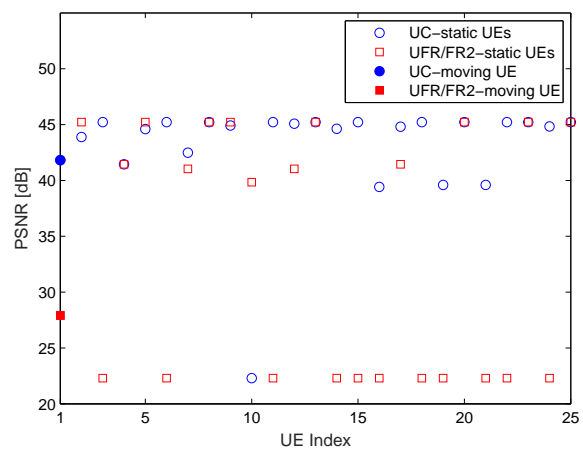
(c)

Fig. 8: (a) The trajectory of the moving UE and dynamic UC cluster formation. (b) PSNR of each frame for the moving UE. (c) Average PSNR of each UE over 300 frames.
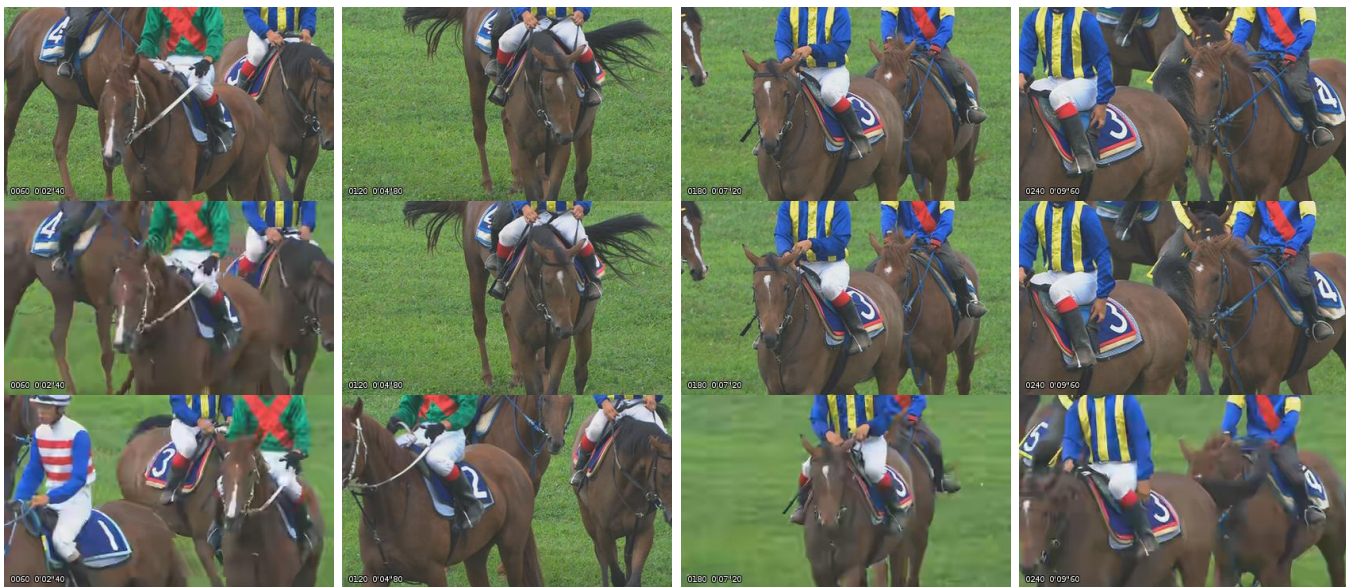
Fig. 9: Four video frames of the Racehorses video sequence transmitted in the UC-VLC network, where the 60th, 120th, 180th and 240th video frames have been extracted from the original video sequence (first row), UC (second row) and UFR (third row) schemes, respectively.

predictively-coded (P) 'IPPPIPPP...' frames in our simulations. Explicitly, the I frames are inserted every 4 frames, which are capable of providing a higher PSNR at the cost of an increased TP requirement. Moreover, Fig. 8c portrays the PSNR averaged over 300 frames for each UE, where the UC design remains capable of serving more UEs at a higher video quality compared to the traditional cell designs.

Fig. 9 shows the video frames reconstructed by the moving UE, where the 60th, 120th, 180th and 240th video frames have been extracted from the original video sequence (first row), UC (second row) and UFR (third row) schemes, respectively. Observe from the second row that the quality of the video frames is only slightly degraded compared to the original lossless video frames (first row). However, the visual quality provided by the UFR design is very poor, as shown in the third row. Three out of four frames have been lost and replaced by the previously successfully decoded frames. Hence, the video frames in the third row are quite different from the original frames. For our simulation results of the full video clips, please refer to http://www.ecs.soton.ac.uk/research/projects/924.

VII. Conclusions

In order to conceive an energy-efficient VLC system, a beneficial UC cluster formation was designed by conceiving a radically new perspective. Then a 3-tier dynamic-programming-based algorithm was proposed for transmitting the advanced scalable H.265-encoded video sequence, whilst relying on user/layer-level AM mode assignment, AP-level PA and cluster-level EE optimisation. Our numerical results demonstrated that the UC design constitutes an energy-efficient VLC system solution, by providing a higher system EE, average TP and video quality compared to the conventional cells. Despite the promise of the UC cluster formation, naturally, some open challenges arise, including the acquisition of accurate location information, the research of robustness under channel estimation uncertainties, the provision of up-link support, just to name a few.

References

[1] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[2] S. Rangan, T. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, March 2014.

[3] L. Hanzo, H. Haas, S. Imre, D. O'Brien, M. Rupp, and L. Gyongyosi, "Wireless myths, realities, and futures: From 3G/4G to optical and quantum wireless," in *Proceedings of the IEEE*, vol. 100, May 2012, pp. 1853–1888.

[4] C.-L. I, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: a 5G perspective," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 66–73, February 2014.

[5] 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.

[6] F. Jin, R. Zhang, and L. Hanzo, "Resource allocation under delay-guarantee constraints for heterogeneous visible-light and RF femtocell," *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, pp. 1020–1034, Feb 2015.

[7] P. Haigh, F. Bausi, H. L. Minh, I. Papakonstantinou, W. Popoola, A. Burton, and F. Cacialli, "Wavelength-multiplexed polymer LEDs: Towards 55 Mb/s organic visible light communications," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 9, pp. 1819–1828, Sept 2015.

[8] S. Zhang, D. Tsonev, S. Videv, S. Ghosh, G. A. Turnbull, I. D. W. Samuel, and H. Haas, "Organic solar cells as high-speed data detectors for visible light communication," *Optica*, vol. 2, no. 7, pp. 607–610, Jul 2015.

[9] P. Haigh, A. Burton, K. Werfli, H. Le Minh, E. Bentley, P. Chvojka, W. Popoola, I. Papakonstantinou, and S. Zvanovec, "A multi-CAP visible-light communications system with 4.85-b/s/Hz spectral efficiency," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 9, pp. 1771–1779, Sept 2015.

[10] A. Yang, Y. Wu, M. Kavehrad, and G. Ni, "Grouped modulation scheme for LED array module in a visible light communication system," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 24–28, April 2015.

[11] A. Nuwanpriya, S.-W. Ho, and C. S. Chen, "Indoor MIMO visible light communications: Novel angle diversity receivers for mobile users," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 9, pp. 1780–1792, Sept 2015.

[12] B. Li, J. Wang, R. Zhang, H. Shen, C. Zhao, and L. Hanzo, "Multiuser MISO transceiver design for indoor downlink visible light communication under per-LED optical power constraints," *IEEE Photonics Journal*, vol. 7, no. 4, pp. 1–15, Aug 2015.

[13] I. Din and H. Kim, "Energy-efficient brightness control and data transmission for visible light communication," *IEEE Photonics Technology Letters*, vol. 26, no. 8, pp. 781–784, April 2014.

[14] K. Lee and H. Park, "Modulations for visible light communications with dimming control," *IEEE Photonics Technology Letters*, vol. 23, no. 16, pp. 1136–1138, Aug 2011.

[15] S. Kim and S.-Y. Jung, "Modified Reed-Muller coding scheme made from the bent function for dimmable visible light communications," *IEEE Photonics Technology Letters*, vol. 25, no. 1, pp. 11–13, Jan 2013.

[16] Y. Suh, C.-H. Ahn, and J. K. Kwon, "Dual-codeword allocation scheme for dimmable visible light communications," *IEEE Photonics Technology Letters*, vol. 25, no. 13, pp. 1274–1277, July 2013.

[17] S. Shao, A. Khreishah, and I. Khalil, "Joint link scheduling and brightness control for greening VLC-based indoor access networks," *arXiv preprint arXiv:1510.00026*, 2015.

[18] 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.

[19] 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 2016.

[20] 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 (Accepted)*, 2015.

[21] J. Jiang, Y. Huo, F. Jin, P. Zhang, Z. Wang, Z. Xu, H. Haas, and L. Hanzo, "Video streaming in the multiuser indoor visible light downlink," *IEEE Access*, vol. 3, pp. 2959–2986, Dec 2015.

[22] L. Hanzo, P. J. Cherriman, and J. Streit, *Wireless video communications: second to third generation and beyond*. IEEE Press-John Wiley & Sons, 2001.

[23] Y. Huo, C. Hellge, T. Wiegand, and L. Hanzo, "A tutorial and review on inter-layer FEC coded layered video streaming," *IEEE Communications Surveys Tutorials*, vol. 17, no. 2, pp. 1166–1207, Secondquarter 2015.

[24] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, Dec 2012.

[25] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103–1120, Sept 2007.

[26] J. Boyce, Y. Ye, J. Chen, and A. Ramasubramonian, "Overview of SHVC: Scalable extensions of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. PP, no. 99, pp. 1–1, July 2015.

[27] Y. Huo, T. Wang, R. Maunder, and L. Hanzo, "Motion-aware mesh-structured trellis for correlation modelling aided distributed multi-view video coding," *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 319–331, Jan 2014.

[28] D. Vukobratovic, V. Stankovic, D. Sejdinovic, L. Stankovic, and Z. Xiong, "Scalable video multicast using expanding window fountain codes," *IEEE Transactions on Multimedia*, vol. 11, no. 6, pp. 1094–1104, Oct 2009.

[29] P. Li, H. Zhang, B. Zhao, and S. Rangarajan, "Scalable video multicast with adaptive modulation and coding in broadband wireless data systems," *IEEE/ACM Transactions on Networking*, vol. 20, no. 1, pp. 57–68, Feb 2012.

[30] H. Zhou, Y. Ji, X. Wang, and B. Zhao, "Joint resource allocation and user association for SVC multicast over heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3673–3684, July 2015.

[31] C. Singhal, S. De, R. Trestian, and G.-M. Muntean, "Joint optimization of user-experience and energy-efficiency in wireless multimedia broadcast," *IEEE Transactions on Mobile Computing*, vol. 13, no. 7, pp. 1522–1535, July 2014.

[32] M. Ismail, W. Zhuang, and S. Elhedhli, "Energy and content aware multi-homing video transmission in heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3600–3610, July 2013.

[33] S.-P. Chuah, Z. Chen, and Y.-P. Tan, "Energy-efficient resource allocation and scheduling for multicast of scalable video over wireless networks," *IEEE Transactions on Multimedia*, vol. 14, no. 4, pp. 1324–1336, Aug 2012.

[34] J. Kahn and J. Barry, "Wireless infrared communications," in *Proceedings of the IEEE*, vol. 85, no. 2, Feb. 1997, pp. 265–298.

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

[36] U. Hammes, E. Wolsztynski, and A. M. Zoubir, "Robust tracking and geolocation for wireless networks in NLOS environments," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 5, pp. 889–901, Oct 2009.

[37] A. Kushki, K. N. Plataniotis, and A. N. Venetsanopoulos, "Intelligent dynamic radio tracking in indoor wireless local area networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 3, pp. 405–419, March 2010.

[38] S. Feng, X. Li, R. Zhang, M. Jiang, and L. Hanzo, "Hybrid positioning aided amorphous-cell assisted user-centric visible light downlink techniques," *IEEE Access (Accepted)*, 2016.

[39] R. Mesleh, H. Elgala, and H. Haas, "Performance analysis of indoor OFDM optical wireless communication systems," in *2012 IEEE WCNC*, April 2012, pp. 1005–1010.

[40] J. Grubor, S. Randel, K.-D. Langer, and J. Walewski, "Broadband information broadcasting using LED-based interior lighting," *Journal of Lightwave Technology*, vol. 26, no. 24, pp. 3883–3892, Dec. 2008.

[41] S. Martello and P. Toth, *Knapsack problems: algorithms and computer implementations*. John Wiley & Sons, Inc., 1990.