

**USER ASSOCIATION IN HYBRID VLC/RF CELL-
FREE MASSIVE MIMO SYSTEMS VIA ITERATIVE
ALGORITHMS AND DEEP REINFORCEMENT
LEARNING**

BY

AHMED FUAD ALMEHDHAR

A Thesis Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

ELECTRICAL ENGINEERING

DECEMBER 2021

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **AHMED FUAD ALMEHDHAR** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN ELECTRICAL ENGINEERING**.

Thesis Committee

Dr. Salam A. Zummo (Adviser)

Dr. Ali H. Muqaihel (Member)

Dr. Ali A. Nasir (Member)

Dr. Abdallah S. Al-Ahmari

Department Chairman

Dr. Suliman Al-Homidan
Dean of Graduate Studies

Date



©Ahmed Fuad Almehdhar
2022

To my beloved family

ACKNOWLEDGMENTS

All praise be to Allah, the Most Compassionate, and the Most Merciful, for his ever-enduring blessings on me. May his peace be upon our prophet Muhammad, and his family.

I would like to express my deep and sincere appreciation to my thesis advisor Prof. Salam A. Zummo for his continuous guidance and encouragement. I would also like to express my respect and gratitude to Dr. Mohanad Obeed for his counselling and enlightenment, without whom this work would not have been delivered in this shape.

Thanks are also delivered to my thesis committee members, Prof. Ali Muqaibel and Dr. Ali Nasir for their valuable time and effort devoted to reviewing and providing their constructive criticism to the work of this thesis.

I would also like to extend my thanks to include everyone who has helped me in any aspect, whether technical or moral, throughout the course of this work, especially, Dr. Saleh Alawsh, and my friend, Adnan Hamida.

My heartfelt thanks go to my family, whom words cannot express how grateful I am for their continuous support and for always believing in me, may Allah reward them for their patience.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xiii
ABSTRACT (ENGLISH)	xvii
ABSTRACT (ARABIC)	xix
CHAPTER 1 INTRODUCTION	1
1.1 Theoretical Background	5
1.1.1 RF Channel and VLC Channel Models	5
1.1.2 Cell-Free Massive MIMO	7
1.2 Literature Review	9
1.2.1 Hybrid VLC/RF Systems	11
1.2.2 Cell-Free Massive MIMO Systems	16
1.2.3 Deep Reinforcement Learning in Wireless Communication	19
1.3 Thesis Work Motivation	22
1.4 Thesis Contribution and Outline	23
1.5 Notations	24
CHAPTER 2 HYBRID CELL-FREE MASSIVE MIMO SYSTEMS	27
2.1 Networks' Achievable Sum-rate	29

2.1.1	VLC Achievable Rate	30
2.1.2	RF Achievable Rate	31
2.2	Problem Formulation and Solution	33
2.2.1	Problem Formulation	33
2.2.2	Proposed Approach	34
2.3	Simulation Results	35
2.4	Conclusion	43

CHAPTER 3 USER-CENTRIC HYBRID CELL-FREE MASSIVE

	MIMO SYSTEMS	45
3.1	Networks' Achievable Sum-rate	46
3.1.1	VLC Achievable Rate	48
3.1.2	RF Achievable Rate	51
3.2	Problem Formulation and Solution	52
3.2.1	Problem Formulation	53
3.2.2	Gibbs Sampling-Based User Association	55
3.2.3	User Clustering	57
3.2.4	Constraints Satisfaction	61
3.3	Simulation Results	63
3.3.1	Hybrid CF-mMIMO without clustering	63
3.3.2	Hybrid CF-mMIMO with Clustering	67
3.4	Conclusion	71

CHAPTER 4 DEEP REINFORCEMENT LEARNING-BASED

	USER ASSOCIATION	73
4.1	Theoretical Foundation of DRL	74
4.1.1	Deep Learning	74
4.1.2	Reinforcement Learning	77
4.1.3	Deep Reinforcement Learning	83
4.2	User Association in Hybrid VLC/RF CF-mMIMO systems: A DRL Approach	85

4.2.1	Proposed DRL Framework	86
4.2.2	Training the Agent's DQN	91
4.3	Simulation Results	91
4.4	Conclusion	96
CHAPTER 5 THESIS CONCLUSION AND FUTURE RESEARCH		
	DIRECTIONS	97
5.1	Thesis Conclusion	97
5.2	Future Research Directions	98
	REFERENCES	101
	VITAE	115

LIST OF TABLES

2.1	Simulation Parameters [1]	38
2.2	Room Dimensions and Number of APs in Figure 2.4	39
3.1	Parameters for Clustering	67
4.1	DRL Parameters	87
4.2	DQN Structure	88
4.3	DQN Hyperparameters	91

LIST OF FIGURES

1.1	Illustration of the difference between (a) small cells system model where each user equipment (UE) is served by the AP in that cell, and (b) cell-free massive MIMO system model where all AP serve all UEs jointly at the same time.	8
1.2	Tree diagram highlighting where the work of interest is located among other topics	11
2.1	Proposed hybrid VLC/RF CF-mMIMO system model	28
2.2	Flowchart illustrating Algorithm 1	37
2.3	Sum-rate versus number of users in the systems	39
2.4	Sum-rate versus total number of APs	40
2.5	Sum-rate versus FoV	41
2.6	CDF of average per-user rate	42
2.7	Sum-rate versus total number of changes	42
3.1	Flowchart illustrating Algorithm 2	58
3.2	Illustration of $\mathbf{A}_1$ clustering	60
3.3	Illustration of $\mathbf{A}_2$ clustering	61
3.4	Average per-user rate versus FoV	64
3.5	CDF of average per-user rate	65
3.6	Number of required iterations for convergence versus number of users	66
3.7	CDF of average per-user rate using $\mathbf{A}_1$ clustering	68
3.8	CDF of average per-user rate using $\mathbf{A}_2$ clustering	69
3.9	Sum-rate versus number of users under $\mathbf{A}_1$ clustering	70

3.10 Sum-rate versus number of users under $\mathbf{A}_2$ clustering	71
4.1 A structure for an ANN with two hidden layers with 10 neurons each. The figure is taken from [2]	75
4.2 A single neuron operation	77
4.3 Steps of the interaction between an agent and its environment. Figure (h) represents a case of DRL instead of RL (the agent is a neural network).	80
4.4 Agent's DQN structure	88
4.5 Performance of DRL-based user association for $N_u = 10$	93
4.6 Performance of DRL-based user association for $N_u = 14$	94
4.7 Performance of DRL-based user association for $N_u = 18$	95
4.8 Sum-rate versus number of users using different association algo- rithms	95

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
AP	Access Point
APS	Access Point Selection
CB	Conjugate Beamforming
CF-mMIMO	Cell-Free massive MIMO
CNN	Convolutional Neural Network
Co-NOMA	Cooperative NOMA
CPU	Central Processing Unit
DCC	Dynamic Cooperation Clustering
DDPG	Deep Deterministic Policy Gradient
DL	Deep Learning
DNN	Deep Neural Network
DQL	Deep Q-Learning
DQN	Deep Q-Network
DRL	Deep Reinforcement Learning

DUD	Discrete Uniform Distribution
EE	Energy Efficiency
FoV	Field-of-View
IM	Intensity Modulation
IoT	Internet of Things
LB	Load Balancing
LED	Light-Emitting Diode
LoS	Line-of-Sight
LSFD	Large-Scale Fading Decoder
MDP	Markov Decision Process
MIMO	Multiple Input Multiple Output
ML	Machine Learning
MMSE	Minimum Mean Square Error
NOMA	Non-Orthogonal Multiple Access
PA	Power Allocation
PD	Photo Detector
PD-NOMA	Power Domain Non-Orthogonal Multiple Access
P-MMSE	Partial Minimum Mean Square Error
PPO	Proximal Policy Optimization
PPZF	Protective Partial Zero Forcing
PZF	Partial Zero Forcing

QoS	Quality-of-Service
RAM	Random Access Memory
ReLU	Rectified Linear Unit
RF	Radio Frequency
RL	Reinforcement Learning
RNN	Recurrent Neural Network
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-Noise Ratio
TRPO	Trust Region Policy Optimization
UC	User Centric
UE	User Equipment
VLC	Visible Light Communication
WiFi	Wireless Fidelity
ZF	Zero Forcing

THESIS ABSTRACT

NAME: Ahmed Fuad Almehdhar

TITLE OF STUDY: User Association in Hybrid VLC/RF Cell-Free Massive MIMO Systems via Iterative Algorithms and Deep Reinforcement Learning

MAJOR FIELD: Electrical Engineering

DATE OF DEGREE: DECEMBER 2021

Cell-free (CF) systems have emerged as a potential technology for future wireless communications and have proven to outperform the systems of small cells, which attracted many researchers to study the CF systems using either radio frequency (RF) or visible light communication (VLC).

In this thesis, a hybrid VLC/RF cell-free massive MIMO system (CF-mMIMO) is introduced. Expressions of the downlink achievable rates in both the VLC network and the RF network are derived. An optimization problem that governs the user association to one of the existing networks aiming to maximize the system's achievable rate is formulated.

This thesis also contributes towards the design of a user-centric (UC) hybrid

VLC/RF CF-mMIMO system, where user clustering is taken into account when deriving the expression of the downlink achievable rate. A joint optimization problem for both user association and user clustering with the objective of maximizing the sum-rate of the system is formulated. This problem is simplified by discarding the link between the two variables of optimization.

Two distinct algorithms are proposed to solve the user association problem in both systems, where each of them dominates in certain cases. The user clustering problem is solved by proposing two suboptimal clustering schemes and studying their performance with the user association algorithms.

Deep Reinforcement Learning (DRL) is also used to perform the task of user association in the hybrid VLC/RF CF-mMIMO system. A full DRL framework is developed to perform the task, with a deep Q-network (DQN) as an agent.

The performance of the three user association schemes is evaluated and compared. And the effect of having user clustering is studied while using the the first two association algorithms. Moreover, the overall performance of the proposed system is shown to outperform standalone VLC and standalone RF CF systems and a hybrid VLC/RF CF-mMIMO system with random user association.

The findings of the thesis are useful to the advancement of hybrid VLC/RF systems, CF systems, and the application of DRL in future communication technologies.

ملخص الرسالة

الاسم:

أحمد فؤاد المحضار

عنوان الدراسة:

اقتران المستخدمين في الأنظمة اللاخولية الهجينة المكونة من اتصالات الراديو اللاسلكية واتصالات الضوء المرئي

باستخدام خوارزميات تكرارية و بالتعلم المعزز العميق

التخصص:

الهندسة الكهربائية

تاريخ الدرجة العلمية: جمادى الأول 1443

ظهرت أنظمة الاتصالات اللاخولية مؤخراً كإتجاه بحثي جديد في مجال الاتصالات اللاسلكية، حيث أنها أثبتت تفوقها على أنظمة الخلايا الصغيرة المستخدمة حالياً مما أدى إلى جذب اهتمام العديد من الباحثين لدراسة هذه الأنظمة باستخدام تقنيات ترددات الراديو اللاسلكية أو اتصالات الضوء المرئي. في هذه الأطروحة، تم طرح فكرة النظام اللاخولي الهجين المكون من اتصالات الراديو اللاسلكية واتصالات الضوء المرئي. كما تم حل معضلة اقتران المستخدمين بإحدى شبكتي الاتصال (شبكة الراديو وشبكة الضوء المرئي) باستخدام ثلاث طرق مختلفة وبهدف زيادة معدل البيانات الممكن تحقيقها. الطريقة الأولى مبنية على خوارزمية تكرارية تعتمد على معدل البيانات عند كل مستخدم. الطريقة الثانية تعتمد أيضاً على معدل البيانات عند كل مستخدم ولكنها تتبع بعض قوانين الاحصاء لايجاد الشبكة المناسبة لخدمة جميع المستخدمين. الطريقة الثالثة مبنية على التعلم المعزز العميق (أحد فروع تعلم الآلة) والتي رغم تعقيدها تتميز عن الخوارزميات السابقة بتوفير نتائج فورية عند الحاجة. تمت مقارنة أداء هذا النظام الهجين المقترح مع أنظمة لاخلوية غير هجينة تستخدم إما تقنيات ترددات الراديو اللاسلكية أو اتصالات الضوء المرئي وتم التغلب عليها بفارق معتبر. أيضاً تمت المقارنة مع والتغلب على نفس النظام اللاخولي الهجين ولكن باستخدام توزيع عشوائي للمستخدمين على الشبكات.

CHAPTER 1

INTRODUCTION

We live in an era where communication, in general, is a vital component of living. The forms of communications are uncountable. Advancements in communications technologies are driven by the unprecedented requirements of data rates to satisfy expected future growth of data caused by advancements in other fields and technologies. Proposing advanced communication technologies is one way to meet the requirements of future generations of wireless communications.

A simple, but non-trivial, solution to increase the achievable data rate is to explore new frequency ranges other than the currently used radio frequency (RF) range, such as, the visible light range in the electromagnetic spectrum. This idea has lead to what is currently known as the visible light communication (VLC) technology which uses light as a medium to transmit data. Moreover, with the invention of energy-efficient light-emitting diodes (LEDs) and their spread as the main source of lighting, the focus in research was pushed towards using the visible light spectrum to transmit data. Furthermore, VLC has proven to provide data rates much higher than the RF technology [3]. This triggered more developments

in the area of VLC as a whole; transmitters, receivers, channel models, and modulation techniques were studied, analyzed, and designed. However, VLC systems inherently suffer from short range of coverage and only perform well when the line-of-sight (LoS) link is available (i.e., when there is no opaque objects between the transmitter and the receiver) [4]. After all, that link is just a beam of light. A common solution to overcome the aforementioned problems is to support the standalone VLC system with RF access points (APs) as they have larger coverage and better operation in non-LoS environments and when the blockage rate is high. In addition, users served by one network (RF or VLC) do not receive interference from the other network, which leads to mitigating the total received interference [5]. Such systems are called hybrid VLC/RF systems and they are suitable for areas with multiple blockages [6].

Focusing on the downlink communication for now, where APs transmit to some receiving users, a hybrid system may consist of a VLC AP (or multiple APs) and an RF AP (or multiple APs) as transmitters. This configuration solves in fact two problems, 1) users who do not have a LoS channel link with the VLC network can be easily served by the RF network, 2) when multiple users are served by the same APs in the VLC network their individual data rates may drop, and moving (offloading) one or more of them to the RF network may help them experience higher data rates and better quality-of-service (QoS).

Aside from exploring other frequency ranges of the electromagnetic spectrum, several ideas were proposed/studied to elevate the user's QoS. A cell-free massive MIMO (CF-mMIMO) system was proposed to increase the data rates of all users

in the network. In such a system, all APs are expected to transmit to all users simultaneously and using the same time and frequency resources. In such systems, the boundaries between cells are removed, hence the name ‘cell-free’ [7].

Even though these ideas are proven to enhance the performance, additional enhancement is achieved by allocating resources and powers according to the channel conditions of each user. Different optimization problems were formulated with different sets of constraints to maximize different utilities, such as, the system’s sum-rate and the users’ minimum rate. Multiple optimal and sub-optimal algorithms with different levels of computational complexity and time complexity for resource allocation and power control were proposed. Nevertheless, as wireless communication systems advance, the formulation of their optimization problems gets complicated and highly non-convex, and many variables get involved in the problem raising the need to jointly optimize the objective function with respect to all of these variables at once. In many cases, having a closed-form solution is only a dream; thus, to solve those problems, assumptions are sometimes made to simplify the problems. Convexification methods could also be applied to make the problem convex and easier to solve. However, iterative methods are usually proposed to solve such complicated problems. Despite their ability to solve the optimization problems, those iterative algorithms require a relatively significant amount of time to converge into an optimal (or a near-optimal) solution. However, in most of the wireless communication scenarios, these algorithms need to be run every period of time (coherence interval) to insure that the solution is applicable to the changes in the system (moving users, interfering devices, etc.).

Meanwhile, in the field of data science, Machine Learning (ML) algorithms are being developed, enhanced, and optimized. These algorithms have the advantage of instant inference, i.e. once they are deployed they can provide an instant output whenever the system requires an output; a desirable feature in wireless communication systems. Lately, ML algorithms received attention in wireless communication systems to perform tasks that usually consume time in traditional methods, as well as tasks that are not easy to formulate and solve otherwise. These include, but not limited to, power allocation and AP-user association. Different ML approaches, specifically Deep Learning (DL) and Reinforcement Learning (RL), underwent through stages of test and design to accomplish the desired tasks. Moreover, in the last couple of years Deep Reinforcement Learning (DRL) was shown to be the most practical form of ML algorithms for wireless communication purposes. Although they are pretty new and not easy to be built for a specific system design, once they are built and trained they are ready to perform the required tasks instantly.

In this thesis, we propose a hybrid VLC/RF CF-mMIMO system, the system is studied and investigated throughout the chapters of this thesis. In this chapter, we introduce the different topics of this thesis to the readers and provide them with the necessary theoretical background to invade and conquer the following chapters. A thorough literature review of the different directions used in this thesis is presented in this chapter as well. The chapter ends by motivating the work and pointing out the distinct contribution of this thesis.

1.1 Theoretical Background

This section provides the reader with the necessary knowledge and background information to get familiar with the different subjects discussed in this thesis. Only topics that have a direct relevance to the proposed work are presented in this section.

1.1.1 RF Channel and VLC Channel Models

In every wireless communication system, the study of the wireless channel is significantly important to, 1) evaluate the performance of the system, 2) mitigate the effects of the channel and enhance the performance. The following two subsections present the RF channel and VLC channel Models.

RF Channel

Systems that use RF in their communications, study the effect of large-scale and small-scale fading; the two components of the RF channel gain. In general, for systems with multiple APs and multiple users, the RF channel link between user j and the RF AP i , denoted by $h_{j,i}^{(r)}$, is given as in [8]

$$h_{j,i}^{(r)} = \underbrace{\sqrt{10^{-\frac{L(d)}{10}}}}_{\text{Large-Scale Fading}} \underbrace{\left(\sqrt{\frac{K}{K+1}} h_d + \sqrt{\frac{1}{K+1}} h_s \right)}_{\text{Small-Scale Fading}}, \quad (1.1)$$

where the superscript $(\cdot)^{(r)}$ specifies that the quantity here is an RF quantity.

In addition, $K = 10$ dB is the Rician factor for indoor 60 GHz mmWave links,

$h_d = \sqrt{0.5}(1 + j)$ is the LoS path fading channel, $h_s \sim \mathcal{CN}(0, 1)$ is the scattered path fading channel, and $L(d)$ is the corresponding path loss in dB at a distance d from the transmitter and is given by

$$L(d) = L(d_0) + 10v \log_{10} \left(\frac{d}{d_0} \right) + X, \quad (1.2)$$

where $L(d_0) = 68$ dB is the reference path loss at a reference distance of $d_0 = 1$ m, $v = 1.6$ is the path loss exponent, and X is the random variable representing the shadowing effect which is assumed to have a log-normal distribution of zero mean and a standard deviation of 1.8 dB [8].

VLC Channel

The transmitters in the VLC are LEDs, thus, each VLC AP is assumed to be equipped with a single LED that uses intensity modulation (IM) technique to transmit light signals to users. Each user is assumed to be equipped with a photo detector (PD) to receive these light signals. It has been shown that the LoS component of the VLC channel link between user j and the VLC AP i , denoted by $h_{j,i}^{(v)}$, is given as in [9] by

$$h_{j,i}^{(v)} = \frac{(m+1)A_p}{2\pi d_{j,i}^2} \cos^m(\phi) G_{of} f(\theta) \cos(\theta), \quad (1.3)$$

where the superscript $(\cdot)^{(v)}$ specifies that the link here is VLC, m is the Lambertian index which is given by

$$m = -\frac{1}{\log_2(\cos(\theta_{1/2}))}, \quad (1.4)$$

where $\theta_{1/2}$ is the half intensity radiation angle, A_p is the physical area of the PD at the receiver side, $d_{j,i}$ is the Euclidean distance from user j to the AP i , G_{of} is the gain of the optical filter, ϕ is the angle of radiance from the AP, θ is the angle of incidence at the PD as illustrated in [3], and $f(\theta)$ is the optical concentrator gain and is given by

$$f(\theta) = \begin{cases} \frac{n^2}{\sin^2(\Theta)}, & 0 \leq \theta \leq \Theta; \\ 0, & \theta > \Theta, \end{cases} \quad (1.5)$$

where n is the refractive index of the PD and Θ is the semi-angle of the field-of-view (FoV) of the PD. All VLC parameter are further illustrated in [3]. The non-LoS component of the VLC channel has a model and is present in the literature, but due to the significant deterioration of the non-LoS transmissions we work only with the LoS paths.

1.1.2 Cell-Free Massive MIMO

When they were first proposed, cell-free massive MIMO systems proposed that there are no cells or cell boundaries, that is, all APs are communicating with all users at the same time and frequency resources. Stated differently, there are no user-AP assignments and each single user receives services from all existing APs in the system. The focus in this work is on the downlink communication only,

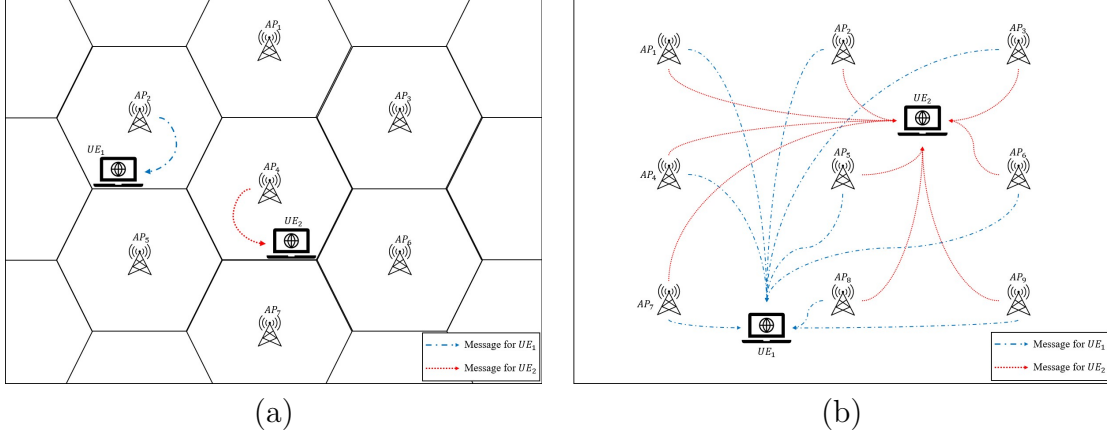


Figure 1.1: Illustration of the difference between (a) small cells system model where each user equipment (UE) is served by the AP in that cell, and (b) cell-free massive MIMO system model where all AP serve all UEs jointly at the same time.

hence, All APs receive data from a central processing unit (CPU) that contains messages to every user in the area. Each AP encodes all the message signals for all users (according to the channel link between that AP and all users). Figure 1.1 presents an illustration of both small cells and cell-free system models.

Mathematically expressed, the received signal by user j is given by

$$y_j = (\mathbf{h}_j)^T \mathbf{x} + n_j = \sum_{i=1}^{N_{ap}} h_{j,i} x_i + n_j, \quad (1.6)$$

where N_{ap} is the number of APs in the system, $\mathbf{h}_j$ is a channel vector of all APs to the j^{th} user, and $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_{N_{ap}}]^T$ is the column vector representing what each AP transmits, where AP i transmits the i^{th} entry that is given by

$$x_i = \sqrt{\rho_f} \sum_{j=1}^{N_u} \sqrt{P_j} w_{i,j} s_j, \quad (1.7)$$

where s_j is the message intended for user j , ρ_f is the total available power at each

AP, P_j is the power allocated to the messages of user j , $w_{j,i}$ is the weight assigned to symbol s_j that is transmitted by AP i .

1.2 Literature Review

With the current pace of advancement in technology, we face an unprecedented data traffic loads that have been mentioned in [3]; this is quite a normal growth of data once we match it with the daily extensive use of multimedia applications and video streaming as well as the enormous numbers of devices being connected and added to the internet in a daily basis due to the advancements in the internet-of-things (IoT) [10], [11]. This has lead to a crowded electromagnetic spectrum for the frequency ranges of RF communication and wireless-fidelity (WiFi) networks [12], calling for more developed communication systems with a higher spectral efficiency and data rates, massive device connectivity, higher energy efficiency (EE), lower traffic fees, a more robust security, and an ultra-low latency [3], [13], [14], [15]. Several attempts have been made through the years to overcome these issues and to enhance existing systems, such as small cells or network densification [16], [17] which was aimed to, increase the system capacity, and enhance the networks' EE.

The idea of exploring other ranges in the electromagnetic spectrum was always an appealing idea. VLC has been introduced as a promising technology in 5G and beyond where the visible light spectrum is used for communication [18], [19]. Some problems that faces the VLC has been looked into in the literature such as the

short coverage of VLC transmitters and the fact that VLC functions properly only when LoS communication is feasible. Interested reader can find more details on the VLC system elements, VLC modulation scheme, bounds on the capacity of different VLC systems with different constraints in [3].

Despite the advantages of VLC systems, they have their own weaknesses compared to RF communication systems resulting in degradation in their performance. Those problems, which include small coverage area, non-LoS failure transmission, frequent handovers, and inter-cell interference, lead to poor service experienced by some users [3]. In fact, it was shown that existing objects in a room where a VLC system is used may block the LoS link between the user and the associated access point, and hence degrade the channel link by up to 90% of the LoS channel, causing a significant drop in the users' data rates [20]. Moreover, these blockages will probably block interfering links coming from another AP to a user causing a large fluctuation of the received signal-to-interference-plus-noise ratio (SINR) which affects the system fairness [3]. Another problem is the frequent handovers, which arise from the physical limits on LEDs that their coverage area is small, hence, a moving user needs to be served by another AP after short periods of time.

Nonetheless, such problems didn't stop researchers from studying the VLC network and improve its performance as in [21], [22], [23], [24], [25]. Several papers studied the effect of APs cooperation on the performance of VLC networks. In addition, many researchers included multiple-input multiple-output (MIMO) technique in their study to improve the VLC performance. Non-orthogonal mul-

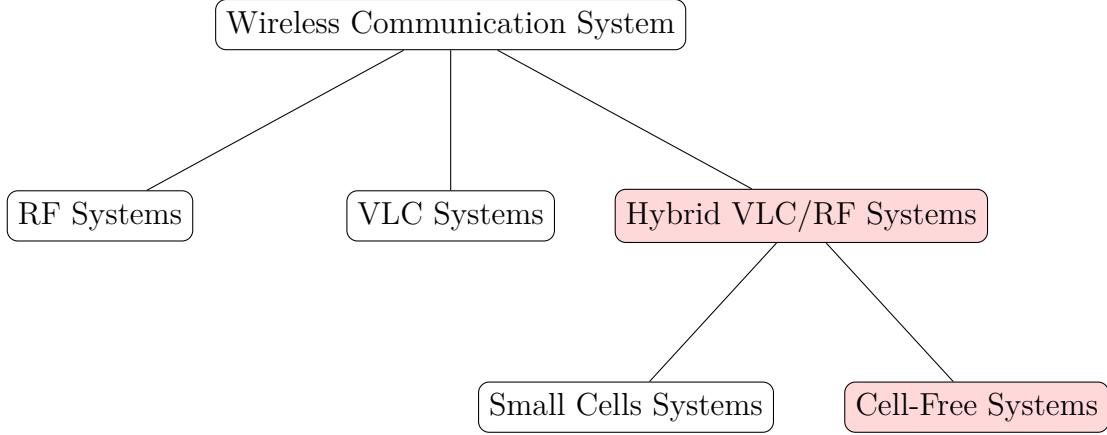


Figure 1.2: Tree diagram highlighting where the work of interest is located among other topics

multiple access (NOMA), specifically power domain NOMA (PD-NOMA), took its part in the advancement of VLC and was included in more than a few papers in the literature. However, since this part of the literature is not of deep relevance to our goal and work, it will not be expanded any further, interested readers might check the survey in [3] Sections IV and V for more details and sources on these topics. Figure 1.2 highlights the area of interest in wireless communications that the literature review (and hence our work) is focused on.

1.2.1 Hybrid VLC/RF Systems

As mentioned earlier, one of the problems in VLC networks is the small coverage area of the VLC transmitters resulting in not only the frequent handovers problem but also static users may be crowded in certain areas in the room causing them to be served by only few APs and not get served by other far APs, leading to having overloaded APs on one end and unloaded APs (or low-loaded) on the other end, a

reason for a poor performance experienced by the crowded users. However, having a small coverage area has its own advantage, that is, the whole frequency band can be reused by adjacent APs which eventually improves the spectral efficiency of the overall system [26]. Even though the overall spectral efficiency will increase, an inter-cell interference is guaranteed at the edges of the cells. With all the mentioned issues in VLC systems, users in motion, blocked by objects, located at the edge of cells, or connected to an overloaded AP will most likely receive a poor QoS compared to other users in the system. This raises questions regarding the performance and fairness of VLC systems.

A common solution to overcome the aforementioned issues and the negative aspects of the VLC systems is to support the standalone VLC network with an RF network. RF networks have larger coverage and better operation in non-LoS environments, making them perfectly suited for areas with multiple blockages. Thus, the complement between the VLC network and RF network is logical and such systems are called Hybrid VLC/RF systems. In addition, users served by the RF network do not interfere with or get interfered by users served by the VLC network [5]. Nevertheless, nothing is fully perfect. The advantages offered by having a hybrid system of VLC APs and RF APs working in harmony come with the price of the need of more complex algorithms to compromise between the large coverage and immunity of the RF network and the high capacity of the VLC network. Hence, more research was needed to fill this newly “at that time” generated gap. Stefan and Haas, in [27], used a hybrid VLC/RF system with multiple APs for both. They proposed a cell-user association based on minimum

distance between all users and APs in an uncoordinated manner and studied its effect on the performance of the system represented by the area spectral efficiency. They compared their proposed scheme with a maximum-SINR based association and a distributed association scheme that achieves a near-optima solution. The comparison was held with multiple FoVs and room sizes. The very simple, uncoordinated, proposed scheme performed well with only a very little loss compared to the near-optimal solution which required information exchange between users and APs. In [28], Wang *et al.* proposed a dynamic load balancing scheme in a dynamic hybrid system, that consists of a set of VLC APs and a single RF AP, taking into account the handover overhead. Also taking handover into account, the authors in [8] proposed another algorithm to balance the load in a hybrid system consisting of multiple VLC and RF APs. They proposed that at every time step all users are initially assigned to the VLC APs. However, all users that have an achievable data rate of less than a specified threshold γ should be transferred into the RF APs. Within each homogeneous network, users should be assigned to APs in a way that maximizes the system throughput. The authors in [29] considered a hybrid VLC/RF system where, depending on their channel conditions, users are assigned dynamically to both VLC and RF. Users connected to VLC network with lower achievable rate are migrated to the RF network. They also studied two variations of RF resource allocation in order to satisfy a good performance.

In the area of load balancing, Wang *et al.*, in [30], formulated two optimization problems for load balancing while considering handover overhead. The first

one is a joint optimization problem of access point assignment and resource allocation while the second one is a separate optimization problem for access point assignment followed by a proposed resource allocation scheme. Wu *et al.*, in [31] formulated another optimization problem in a dynamic hybrid system with handover control. In their work, they identified two different types of handover, a horizontal handover and a vertical handover. They formulated their optimization problem to only choose the best network for a specified user, then they assign an AP to the user based on the signal-to-noise ratio (SNR).

In [32], [33], [34], the authors considered load balancing in a hybrid VLC/RF system taking into consideration the handover cost and the fact that not every VLC user should be switched to the RF network whenever it experiences a link blockage. They proposed that 1) when a user has a high occurrence rate of channel blockage, the user should always connect to the RF network. 2) a user with no channel blockage should always be served by the VLC network. 3) a user with low occurrence rate of channel blockage but with high occupation rate, i.e. the blockage does not quickly vanish, should switch between networks when they face a channel blockage. According to that, they formulated an optimization problem that takes the performance degradation caused by channel blockage into account and they proposed a two-stage access point selection (APS) method. Taking into account both users' mobility and light-path blockages, they further formulated a joint optimization problem to determine the type of network access a user should be connected to.

Mohanad *et al.* in [35], [36] were the first to consider the joint load balancing

(LB) and power allocation (PA) for a hybrid VLC/RF network. They mathematically formulated an optimization problem for the PA sub-problem to allocate the power of each AP to its users in order to maximize the total achievable rate. They proposed an algorithm to find the optimal dual variables after writing them in terms of each other. An iterative algorithm was also proposed to balance the load with the help of the first algorithm that allocates the power. They also proposed a sup-optimal approach that does not require the exact interference information, though performed very similar to the optimal, however, with much lower complexity

As researchers seek more improvement in the spectral efficiency, NOMA might be a direction to go to. NOMA was investigated for the first time in a hybrid VLC/RF network in [37] where Papanikolaou *et al.* considered a system with multiple VLC APs and a single RF AP, all APs perform PD-NOMA. Their goal was to find an optimal grouping of the users that maximized a certain utility function that they defined based on the achievable data rate. They formulated the problem as Coalition Formation Game. After grouping, each group is assigned an AP and its users, as a group, perform NOMA.

Mohanad *et al.* in [38], [39] used the concept of cooperative NOMA (Co-NOMA) in a hybrid VLC/RF system. They also paired each weak user to a strong user and each weak user can be served either directly by the VLC AP or indirectly by the strong user via an RF link. Two mixed-integer non-convex optimization problems were formulated to maximize the system's weighted sum rate, one for a single VLC AP, and the second is for multiple VLC APs.

1.2.2 Cell-Free Massive MIMO Systems

The original idea of CF-mMIMO systems was first introduced in [7] and [40]. It was proposed that each user should be served by all APs simultaneously with the number of APs being much greater than the number of users. In [7], Nayebi *et al.* considered two different precoding schemes, conjugate beamforming (CB) and zero-forcing (ZF), in the downlink data transmission, and optimized transmit power to improve the performance. Whereas in [41], they proposed two different receivers in the uplink transmission, the partial minimum mean square error (P-MMSE) and a generalized large-scale fading decoding (LSFD) receivers, to maximize the SINR and the achievable rate. They showed that the proposed P-MMSE receiver has a near-optimal performance compared to minimum mean square error (MMSE) receiver. Furthermore, in [7], lower capacity bounds for CF-mMIMO systems utilizing CB and ZF precoders for downlink transmission were derived. A tight approximation for the SINR of the P-MMSE and the exact SINR of the LSFD receivers in the uplink transmission were derived in [41] as functions of large scale fading coefficients only. The work has further been extended in the downlink side where, in [42], power optimization problems to maximize the minimum SINR among all users (max-min problem) for both CB and ZF precoders were formulated. They proposed power allocation algorithms to find the optimal solutions to both problems. They also proposed suboptimal power allocation algorithms that have very low complexity, though very reasonable performance. Ngo *et al.*, in [40] and [43], adopted CB precoder in their work and derived closed-form

expressions for the achievable data rate for a finite number of APs in the downlink [40], and uplink [43] transmissions taking into account channel estimation error, power control, and non-orthogonality of pilot sequences. They also designed two pilot assignment schemes: random and greedy pilot assignment and they provided an algorithm for the greedy algorithm. They formulated (and proposed solutions for) a quasi-concave and a quasi-linear optimization problems to maximize the minimum downlink and uplink achievable rate of all users, respectively. Even though the idea of CF-mMIMO systems is appealing, it has two main problems: 1) all users, no matter how bad their channel links are, will be served by all APs, which means that each AP will allocate resources to all users even if they will not receive a good service. This prevents other users with better channel links from benefiting from these ‘wasted’ resources [44]; 2) the unscalability of CF-mMIMO systems [45]. S. Buzzi *et al.* in [44] proposed a modification on the original CF-mMIMO idea, where they proposed a user-centric (UC) approach where each AP should only serve (send messages to) the N users with the strongest channels in the downlink, and should only decode the messages of these N users in the uplink, where N is a system design parameter. This forms the basis to solve the scalability issue in the original CF-mMIMO systems since much less amount of data will propagate through the backhaul. The authors of [45] proposed a scalable CF-mMIMO where they grouped the APs into N pre-determined cell-clusters and each cluster is managed by an autonomous CPU, then they apply a UC approach to see what APs are best to serve a specific user. Then, all APs in the clusters that include those best serving APs will jointly serve that user.

Up to this end, all previous works dealt with single-antenna APs and single-antenna users in the CF-mMIMO systems. However, in the following works, authors started to include and get the benefits of adding multiple antennas at the APs side [46], [47], [48], [49], and at user side [46]. The authors in [46] utilized the UC CF idea presented in [44] but with multiple antennas in both APs and users sides, and designed a partial zero-forcing (PZF) precoding on the uplink transmission and a successive interference cancellation (SIC) in the uplink to improve the performance of the system.

Björnson *et al.* studied the scalability issue in the original CF-mMIMO system and provided a systematic procedure for determining whether a process/technique is scalable or not. Then, they developed a generalized dynamic cooperation clustering (DCC) framework for CF-mMIMO systems. They derived an achievable uplink and downlink data rate with general precoding vectors [47]. They proposed two more distributed precoding schemes, local PZF and local protective partial zero-forcing (PPZF), and derived closed form expressions for the achievable downlink data rate [48]. In there, they proposed an algorithm to globally solve the max-min fairness power control optimization problem. In [49], they proposed a new framework for scalable CF-mMIMO systems and provided an algorithm for joint initial access, pilot assignment, and cluster formation. Moreover, It was shown in [47] that the original CF-mMIMO and the UC CF-mMIMO are instances of the their generalized DCC framework.

Recently, few researchers utilized the dense nature of LED deployment to investigate having CF VLC systems. The authors in [50] and [51] formulated

optimization problems to improve the performance and to manage the downlink interference. A joint time-fraction user association and power allocation scheme in a CF VLC network was considered in [50] where a non-convex maximization problem was formulated. They proposed an iterative algorithm to find a local optimum by first separating the non-convex problem into two sub-problems then solving them alternately until convergence. While in [51] the authors formulated a non-convex optimization problem for joint AP grouping, AP-user association and AP power allocation. They divided the problem into sub-problems then provided algorithms to solve them. As a proof of concept, Beysens *et al.*, [52], [53], implemented and optimized a working CF VLC network with 36 VLC APs and 4 receivers and named it “DenseVLC”. Their DenseVLC was built with off-the-shelf devices and its performance was evaluated in extensive experiments.

1.2.3 Deep Reinforcement Learning in Wireless Communication

Artificial Intelligence (AI) refers to things that can be done via computers without the interference of human beings. The term AI is a “balloon” as everything starts to be enrolled under the AI umbrella. Machine Learning is a subsection of AI, however, it is a mathematically well-structured field. ML algorithms have proven to be useful in many applications in real life. Recently, research in wireless communications started to make use of ML algorithms in several aspects, from simple ML algorithms [54] to the new and more complicated ones [55], from simple

modulation-scheme recognition [56], [57], and direction-of-arrival estimation [2], to advanced and complicated association and power allocation problems [58]. Nevertheless, below are some of the research directions that have been investigated in the literature of Deep Learning, Reinforcement Learning, and Deep Reinforcement Learning in wireless communications. The authors of [59] and [60] used the concept of DL to perform user association and power allocation, respectively. In their problem formulation, [59] considered the sum-rate as the objective function to be maximized. They showed that their problem could be easily solved by the branch-and-cut algorithm, however, due to the large number of variables to be solved for, it is challenging to solve in real-time in typical Massive MIMO networks where a large number of users must be served. Thus, they built an artificial neural network (ANN) and trained it to perform the user association task. The authors in [60] formulated two optimization problems to maximize the minimum rate and the product rate instead of the sum rate, and they solved it using a deep neural network (DNN).

The Deep Learning methods are not very practical in terms of real-life application since they require a data set to train the model beforehand. This is generally inconvenient since we don't have data sets that ideally represent the solution to the problems we are trying to solve (user assignment, power allocation, ..., etc). Thus, researchers investigated the Reinforcement Learning in many wireless communication fields among which [61], [62], and [63] appeared. In [61], they considered the problem of downlink power control for cellular systems, while in [62] and [63] they tried to optimize the load balancing of Hybrid VLC/RF net-

works. In all of them, RL approaches were adopted and developed to approach the optimization problem, specifically, a Q-learning algorithm was used in [61], and a trust region policy optimization (TRPO) algorithm was used in [62] and [63].

In problems where the state space and the action space are large, which is usually the case in practical systems, the Q-learning algorithm may not be able to converge into an optimal policy and this is where the deep Q-learning (DQL), an algorithm of the DRL, comes to play. The agent of this DQL is a deep Q-network (DQN). Here, we use the terms DQL and DQN interchangeably. In the literature, variations of the DQN algorithm were adopted, modified and implemented to perform multiple tasks that include, but not limited to, allocating power in multi-cell networks to achieve high data rates [64], [65], selecting the best channel to access in dynamic multi-channel networks [66], controlling the interference in ultra-dense small cells [67], allocating resources for hybrid VLC/RF to maximize the system's achievable sum-rate [68].

Other works have investigated different types of agents in the areas of wireless communications. Since they are not of importance to us, interested readers are advised to check the references [69] for Deep deterministic policy gradient (DDPG) algorithms, and [70] for a proximal policy optimization (PPO) algorithm. Al-Eryani *et al.*, in [55], constructed a hybrid DDPG-DDQN (Deep Double Q-network) to maximize the sum-rate and the minimum rate in a cell-free system. The essence of their choice of the hybrid agents is the fact that DQNs are proven to be good candidates to perform the optimization when the action space is dis-

crete and the DDPG is a good candidate when the action space is continuous. Thus, the DQN was built to perform the clustering and the DDPG was built to take care of the power coefficients.

Surveys of many ML algorithms with different levels of complexity including DL, RL, and DRL along with their contribution to wireless communications are presented in [54], [71], [72], [73], [74], [75].

1.3 Thesis Work Motivation

The number of connected devices to the internet is rapidly increasing in an unprecedented rate, and the need for high data rates is trivially obvious. Using CF systems enhances the performance and achieves higher data rates than systems employing small cells. In addition, VLC enables the utilization of bandwidth of frequencies that are much larger than that used in the RF technology. This is due to the abundance of huge unlicensed ranges of frequency, unlike the limited frequency ranges used in different RF systems. Thus, having a VLC based CF system is a logical direction to go. Nonetheless, because VLC links are much more prone to failure and can easily be blocked by obstacles, it is reasonable to complement the VLC network with an RF network, which results in a system that exploits the advantages of both technologies. Embedding that with a CF system, we get our proposed hybrid VLC/RF CF-mMIMO system.

A Deep Reinforcement Learning framework is constructed to approach the user association problem in the proposed hybrid system. Such a problem can be

approached using some heuristic techniques. However, using the Deep Reinforcement Learning approach has three main advantages:

1. It has an almost instant time to take actions.
2. It has the potential to solve very complicated and highly non-convex problems.
3. It pushes to meet the worldwide research trends of developing and utilizing Machine Learning in different research fields.

1.4 Thesis Contribution and Outline

The contributions contained in this thesis are summarized as follows:

- In Chapter 2, a hybrid VLC/RF CF-mMIMO system is introduced. Expressions for the achievable data rate at each user are derived for both networks. A mathematical optimization problem is formulated to maximize the sum-rate in the proposed system by associating users to one of the networks in the system.
- An iterative algorithm is proposed in Chapter 2 to approach the optimization problem and provide a solution for the user association.
- In Chapter 3, a UC hybrid VLC/RF CF-mMIMO system is introduced. In this system, the user clustering is addressed along with the user association. Expressions for the achievable data rate at each user are derived for

both networks. A joint optimization problem is formulated to maximize the system's sum-rate, with user clustering and user association being the optimization variables. The joint optimization problem is broken into two sub-problems, one for each variable.

- The analysis of the system in Chapter 3 is generic and covers all possible types of user clustering, user association, and precoding schemes. This has been emphasized in the results section of Chapter 3.
- A Gibbs sampling-based algorithm is proposed in Chapter 3 to approach the user association sub-problem. This algorithm is compared with the iterative algorithm presented in Chapter 2.
- Two user clustering schemes that achieve sub-optimal performance are proposed in Chapter 3.
- In Chapter 4, a DRL framework is developed to approach the user association problem that is presented in Chapters 2 and 3. The performance of the proposed framework is compared with the performance achieved by the two classical algorithms proposed in Chapter 2 and Chapter 3.

1.5 Notations

Throughout this thesis, notations respect the following convention,

Vectors are denoted by boldface lowercase letters, whereas matrices are denoted by boldface uppercase letters. The transpose, Hermitian, and inverse op-

erators is represented by $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^{-1}$, respectively. The operator $\odot$ is the *Hadamard* product or element-wise product of matrices. The function $\text{diag}(\cdot)$ produced a diagonal square matrix with the diagonal elements being the elements of the input vector. In addition, the function $\max_i(\cdot)$ returns the maximum value of its input with respect to the variable i . The symbols $\mathbb{R}$, $\mathbb{R}^N$, and $\mathbb{R}^{N \times M}$ represent the set of real numbers, set of N -dimensional vectors of real numbers, and the set of $N \times M$ matrices of real numbers, respectively. The superscript $(\cdot)^{(v)}$ indicates that the quantity inside the parentheses is a VLC quantity while $(\cdot)^{(r)}$ mean that it is an RF quantity.

CHAPTER 2

HYBRID CELL-FREE MASSIVE MIMO SYSTEMS

In this chapter, we introduce the centralized hybrid VLC/RF CF-mMIMO system for indoor communication and study its performance. We first formulate an optimization problem that assigns the users to the VLC or RF networks aiming at maximizing the total achievable rate. An efficient iterative algorithm is then proposed to provide a solution for the formulated problem. We evaluate the hybrid CF-mMIMO system and show that it performs better than the VLC or RF standalone CF-mMIMO systems. We also show that the proposed algorithm significantly outperforms a random user association scheme.

Without loss of generality, we assume a room of size $M \times M \times height$ that contains N_{vap} VLC APs and N_{rap} RF APs, with a uniform deployment for both on the ceiling of the room. IM is assumed to be used by each VLC AP to transmit optical signals to the associated users. Each user equipment is assumed to be

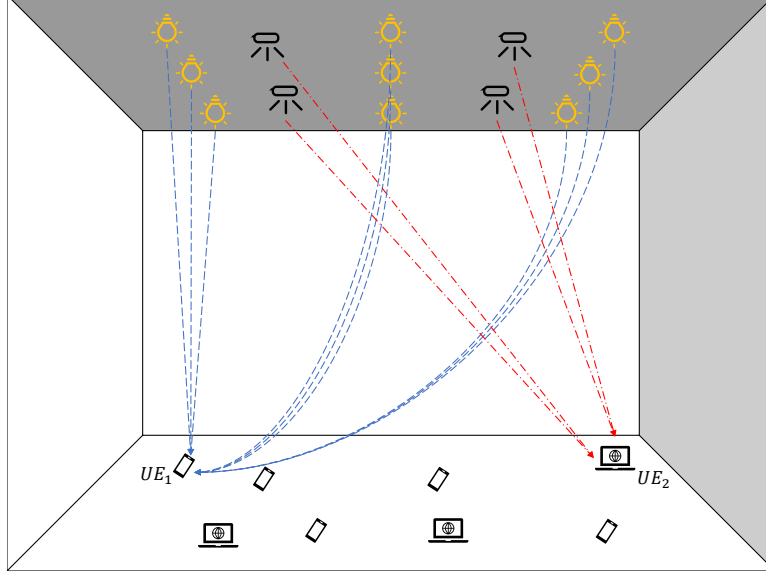


Figure 2.1: Proposed hybrid VLC/RF CF-mMIMO system model

equipped with a PD to directly detect the light intensity and a regular antenna to receive RF signals. Moreover, it is assumed that the room contains a total of N_u users that are randomly distributed over the floor according to the uniform distribution. We denote the number of users served by VLC APs and RF APs by $N_{u,v}$ and $N_{u,r}$, respectively. Furthermore, we assume that the channel coefficients remain constant during the transmission of one block of data. In this system, we consider the idea of a centralized hybrid VLC/RF cell-free system where multiple VLC (or RF) APs cooperate to serve each user. As implied by the word ‘centralized’, all APs are assumed to be connected to a central processing unit (CPU) that processes users’ signals and coordinates them. To eliminate confusion, the word ‘network’ is used only to indicate a network of only VLC APs or a network with only RF APs, and thus our system consists of two disjoint networks. In this system, a user would be served by all APs in one network (either VLC network or RF network). Moreover, all VLC (or RF) APs use the same bandwidth to

serve all their associated users as implied by the original definition of the *cell-free* systems. The system model is presented in Figure [2.1](#).

2.1 Networks' Achievable Sum-rate

In this section, we analyze the received signal at the receiver side and provide an expression for the achievable sum-rate. Let $\mathcal{U}^{(v)}$ be the set of users served by the VLC network and $\mathcal{U}^{(r)}$ be the set of users served by the RF network. Note that the two sets, $\mathcal{U}^{(v)}$ and $\mathcal{U}^{(r)}$, are mutually exclusive, which means that each user can only be served by either the VLC network or the RF network. We can see that $|\mathcal{U}^{(v)}| = N_{u,v} \leq N_u$, $|\mathcal{U}^{(r)}| = N_{u,r} \leq N_u$, and $\left|\bigcup_{m \in \{v,r\}} \mathcal{U}^{(m)}\right| = N_u$. In addition, we define a vector $\mathbf{b}$ of length N_u that corresponds to the users being connected to either the VLC network or the RF network, it can be written as

$$\mathbf{b} = \begin{bmatrix} b_1 & b_2 & \cdots & b_{N_u} \end{bmatrix}^T, \quad (2.1)$$

where $b_j \in \{0, 1\}$. Setting $b_j = 0$ means that user j is connected to the RF network, while $b_j = 1$ indicates that user j is connected to the VLC network. Vector $\mathbf{b}$ guarantees that a single user cannot be served simultaneously by both networks.

2.1.1 VLC Achievable Rate

Here, we provide an achievable rate expression that can be achieved at the users that are served by the VLC network. Each VLC AP transmits an entry of the vector $\mathbf{x}^{(v)} = \begin{bmatrix} x_1^{(v)} & x_2^{(v)} & \dots & x_{N_{vap}}^{(v)} \end{bmatrix}^T$, where AP i transmits the i^{th} entry that is given by

$$x_i^{(v)} = \sum_{j \in \mathcal{U}^{(v)}} \sqrt{\rho_f^{(v)} P_j^{(v)}} w_{i,j}^{(v)} s_j = \sum_{j=1}^{N_u} b_j \sqrt{\rho_f^{(v)} P_j^{(v)}} w_{i,j}^{(v)} s_j, \quad (2.2)$$

where s_j is the message intended for user j , $\rho_f^{(v)}$ is the total available power at each VLC AP, $P_j^{(v)}$ is the power allocated to the messages of user j in the VLC network, and $w_{j,i}^{(v)}$ is the weight assigned to symbol s_j that is transmitted by VLC AP i . We define $\mathbf{W}^{(v)} \in \mathbb{R}^{N_{vap} \times N_u}$ to be the precoding matrix, where its $(i, j)^{th}$ entry is given by $w_{i,j}^{(v)}$. Therefore, we can write the received signal at the VLC user j as

$$y_j^{(v)} = \rho_{oe} \rho_{eo} \sum_{i=1}^{N_{vap}} h_{j,i}^{(v)} x_i^{(v)} + n_j^{(v)}, \quad (2.3a)$$

$$= \rho_{oe} \rho_{eo} \sum_{i=1}^{N_{vap}} h_{j,i}^{(v)} \left(\sqrt{\rho_f^{(v)}} \sum_{l=1}^{N_u} b_l w_{i,l}^{(v)} \sqrt{P_l^{(v)}} s_l \right) + n_j^{(v)}, \quad (2.3b)$$

$$= \rho_{oe} \rho_{eo} \sqrt{\rho_f^{(v)}} \left(\sum_{l=1}^{N_u} b_l \left(\mathbf{h}_j^{(v)} \right)^T \mathbf{w}_l^{(v)} \sqrt{P_l^{(v)}} s_l \right) + n_j^{(v)}, \quad (2.3c)$$

where ρ_{oe} and ρ_{eo} are the optical-to-electrical and the electrical-to-optical conversion factors, respectively, and $n_j^{(v)} \sim \mathcal{N}(0, \sigma_v^2)$ is the additive white Gaussian noise at the PD of user j . $\mathbf{H}_v = \begin{bmatrix} \mathbf{h}_1^{(v)} & \mathbf{h}_2^{(v)} & \dots & \mathbf{h}_{N_u}^{(v)} \end{bmatrix}^T$ is the VLC channel matrix, where $\mathbf{h}_j^{(v)} = \begin{bmatrix} h_{j,1}^{(v)} & h_{j,2}^{(v)} & \dots & h_{j,N_{vap}}^{(v)} \end{bmatrix}^T$ is a channel vector of all VLC APs to the

j^{th} user, and $\mathbf{w}_l^{(v)}$ is the symbol used to denote the l^{th} column vector of $\mathbf{W}^{(v)}$. The achievable rate at the user j connected to the VLC network can be written as [76]

$$R_j^{(v)} = \frac{B^{(v)}}{2} \log_2 \left(1 + \gamma_j^{(v)} \right), \quad (2.4)$$

where $B^{(v)}$ is the total VLC bandwidth and $\gamma_j^{(v)}$ is given by

$$\gamma_j^{(v)} = \frac{\left| \rho_{oe}\rho_{eo}\sqrt{\rho_f^{(v)}}b_j \left(\mathbf{h}_j^{(v)} \right)^T \mathbf{w}_j^{(v)} \sqrt{P_j^{(v)}} \right|^2}{\sum_{\substack{l=1 \\ l \neq j}}^{N_u} \left| \rho_{oe}\rho_{eo}\sqrt{\rho_f^{(v)}}b_l \left(\mathbf{h}_j^{(v)} \right)^T \mathbf{w}_l^{(v)} \sqrt{P_l^{(v)}} \right|^2 + 9\sigma_v^2}. \quad (2.5)$$

Note that, compared to the Shannon's formula, equation (2.4) has the 1/2 factor.

This can be explained by the fact that VLC signals are always real and they do not utilize the imaginary dimension that is usually used in the RF signals. Equation (2.5) contains a factor of 9 multiplied by the noise power at the PD of a VLC receiver, which has been derived in [76].

2.1.2 RF Achievable Rate

Here, we provide an expression for the achievable rate at the users that are served by the RF network. The vector $\mathbf{x}^{(r)} = \begin{bmatrix} x_1^{(r)} & x_2^{(r)} & \dots & x_{N_{rap}}^{(r)} \end{bmatrix}^T$ represents the transmitted signals from RF APs, where $x_i^{(r)}$ is the transmitted signal from the

i^{th} RF AP and is given by

$$x_i^{(r)} = \sum_{j \in \mathcal{U}^{(r)}} \sqrt{\rho_f^{(r)} P_j^{(r)}} w_{i,j}^{(r)} s_j, \quad (2.6a)$$

$$= \sum_{j=1}^{N_u} (1 - b_j) \sqrt{\rho_f^{(r)} P_j^{(r)}} w_{i,j}^{(r)} s_j. \quad (2.6b)$$

where $\rho_f^{(r)}$ is the total available power at each RF AP, $P_j^{(v)}$ is the power allocated the messages of the j^{th} user in the RF network, and $w_{i,j}^{(r)}$ is the weight assigned to symbol s_j transmitted by the i^{th} RF AP. Let $\mathbf{W}^{(r)} \in \mathbb{R}^{N_{rap} \times N_u}$ be the precoding matrix, where its $(i, j)^{th}$ entry is given by $w_{i,j}^{(r)}$. The received signal by the user j can then be written as

$$y_j^{(r)} = \sum_{i=1}^{N_{rap}} h_{j,i}^{(r)} x_i^{(r)} + n_j^{(r)}, \quad (2.7a)$$

$$= \sum_{i=1}^{N_{rap}} h_{j,i}^{(r)} \left(\sqrt{\rho_f^{(r)}} \sum_{l=1}^{N_u} (1 - b_l) w_{i,l}^{(r)} \sqrt{P_l^{(r)}} s_l \right) + n_j^{(r)}, \quad (2.7b)$$

$$= \sqrt{\rho_f^{(r)}} \left(\sum_{l=1}^{N_u} (1 - b_l) \left(\mathbf{h}_j^{(r)} \right)^T \mathbf{w}_l^{(r)} \sqrt{P_l^{(r)}} s_l \right) + n_j^{(r)}. \quad (2.7c)$$

where $\mathbf{w}_l^{(r)}$ is the l^{th} column vector of $\mathbf{W}^{(r)}$, and $\mathbf{h}_j^{(r)} = \begin{bmatrix} h_{j,1}^{(r)} & h_{j,2}^{(r)} & \dots & h_{j,N_{rap}}^{(r)} \end{bmatrix}^T$ is a channel vector of all RF APs to the j^{th} user, and the RF channel matrix is given as $\mathbf{H}_r = \begin{bmatrix} \mathbf{h}_1^{(r)} & \mathbf{h}_2^{(r)} & \dots & \mathbf{h}_{N_u}^{(r)} \end{bmatrix}^T$, and $n_j^{(r)} \sim \mathcal{CN}(0, \sigma_r^2)$ is the additive white Gaussian noise at the antenna of user j . Thus, the achievable rate at the user j that is connected to the RF network is given by Shannon's formula as

$$R_j^{(r)} = B^{(r)} \log_2 \left(1 + \gamma_j^{(r)} \right) \quad (2.8)$$

where $B^{(r)}$ is the total available RF bandwidth and $\gamma_j^{(r)}$ is given by

$$\gamma_j^{(r)} = \frac{\left| \sqrt{\rho_f^{(r)}} (1 - b_j) \left(\mathbf{h}_j^{(r)} \right)^T \mathbf{w}_j^{(r)} \sqrt{P_j^{(r)}} \right|^2}{\sum_{\substack{l=1 \\ l \neq j}}^{N_u} \left| \sqrt{\rho_f^{(r)}} (1 - b_l) \left(\mathbf{h}_j^{(r)} \right)^T \mathbf{w}_l^{(r)} \sqrt{P_l^{(r)}} \right|^2 + \sigma_r^2}, \quad (2.9)$$

2.2 Problem Formulation and Solution

In this section, we are interested in the users association to the system networks aiming at maximizing the system's sum-rate. In other words, what is the best selection of the vector $\mathbf{b}$ to maximize the sum-rate. We first mathematically formulate the optimization problem and then propose a heuristic solution for the formulated problem.

2.2.1 Problem Formulation

Given that the association variable b_j of user j is either one or zero, we can express the achievable rate of user j as follows

$$R_j = R_j^{(v)} + R_j^{(r)}. \quad (2.10)$$

where $R_j^{(v)}$ and $R_j^{(r)}$ are defined in (2.4) and (2.8), respectively. Expression (2.10) implies that $R_j = R_j^{(v)}$ if $b_j = 1$ or $R_j = R_j^{(r)}$ if $b_j = 0$. In other words, if user j is connected to the VLC networks, $R_j = R_j^{(v)}$; otherwise, $R_j = R_j^{(r)}$. This emphasizes the fact that only one network, either VLC or RF, would serve user j at a time.

Our goal now is to optimize the vector $\mathbf{b}$ to maximize the system's sum-rate.

The problem can be formulated as

$$\max_{\mathbf{b}}. \quad \sum_j R_j \quad (2.11a)$$

$$\text{s.t.} \quad b_j \in \{0, 1\}, \quad \forall j, \quad (2.11b)$$

$$\sum_{j=1}^{N_u} \left(w_{i,j}^{(v)} \right)^2 P_j^{(v)} \leq 1, \quad \forall i, \quad (2.11c)$$

$$\sum_{j=1}^{N_u} \left(w_{i,j}^{(r)} \right)^2 P_j^{(r)} \leq 1, \quad \forall i. \quad (2.11d)$$

Constraint (2.11b) guarantees that each user is served by only one network at a time. Constraints (2.11c) and (2.11d) are imposed to make sure that the transmit power from the AP i is not greater than ρ_f (the total available power). Problem (2.11) is not an easy problem to tackle since it is a non-convex nonlinear integer problem. Hence, in the following, we propose a suboptimal solution to problem (2.11), where an iterative algorithm is proposed to efficiently distribute the users to the networks.

2.2.2 Proposed Approach

The optimal solution can be attained by using exhaustive search approach, where we go over every possible choice of the vector $\mathbf{b}$ and select the one that maximizes the objective function given in (2.11a). However, this approach is not practical since increasing the number of users in the system results in an exponential increase in the number of possibilities of $\mathbf{b}$. One simple and efficient way to approach

this problem is by iteratively considering the impact of switching each user from its current network to the other network. This starts by first sorting the users' rates in an ascending order then start switching the users with minimum rates first. If changing the user's network increases the sum-rate, the approach keeps that user in its new network; otherwise, brings it back to its previous network. This continues until convergence. This approach is presented in details in **Algorithm 1** and Figure 2.2. Since allocating the power is not in the scope of this work, we distribute the powers equally among users while guaranteeing that the constraints in (2.11d) and (2.11d) are satisfied.

2.3 Simulation Results

In this section, the performance of the proposed system is evaluated by studying the effect of changing several parameters on the sum-rate. Unless stated otherwise, the simulation parameters used to generate all the results are listed in Table 2.1. The precoding scheme used in calculating $\mathbf{W}$ for both networks is the typical ZF precoding, i.e. $\mathbf{W} = \mathbf{H}^H(\mathbf{H} * \mathbf{H}^H)^{-1}$. Monte-Carlo simulation is implemented to obtain more accurate and generic results, and every displayed point is the arithmetic mean of 1000 points resulted from 1000 independent uniform random distributions of users.

In Figure 2.3, the sum-rate of the proposed system is compared to three systems, a similar hybrid system that uses a random user association instead of the proposed algorithm, and two cell-free systems that consists of either RF APs or

Algorithm 1 Iterative User Association

Input: $N_u, \mathbf{H}_v, \mathbf{H}_r, \mathbf{W}^{(v)}, \mathbf{W}^{(r)}, \rho_f^{(v)}, \rho_f^{(r)}, B^{(v)}, B^{(r)}, \rho_{eo}, \rho_{oe}, \sigma_v, \sigma_r$.

Output: $\mathbf{b}$

Initialize: $b_j \leftarrow 1, \forall j \in [1, N_u]$

```
1: Convergence  $\leftarrow$  False
2: while Convergence  $\neq$  True do
3:   Changes  $\leftarrow$  0
4:   Calculate  $R_s \leftarrow \sum_j R_j$  using (2.4), (2.8) and (2.10)
5:   Let  $\bar{R}_j \leftarrow [R_j]_{j=1}^{N_u}$ 
      return index
6:    $I \xleftarrow{\text{sort}} \text{sort}(\bar{R}_j, \text{"ascending"})$ 
7:   for  $i \leftarrow 1$  to  $N_u$  do
8:      $j \leftarrow I(i)$ 
9:      $b_j \leftarrow 1 - b_j$  ▷ Switch user's network
10:    Calculate  $R_{new} \leftarrow \sum_j R_j$ 
11:    if  $R_{new} > R_s$  then
12:       $R_s \leftarrow R_{new}$ 
13:      Changes  $\leftarrow$  Changes + 1
14:    else
15:       $b_j \leftarrow 1 - b_j$  ▷ Return to previous network
16:    end if
17:  end for
18:  if Changes = 0 then
19:    Convergence  $\leftarrow$  True
20:  end if
21: end while
```

VLC APs (standalone CF-mMIMO systems). The comparison is held over different number of users in the system. To have a fair comparison, the number of APs in each of the standalone systems is set to be equal to the total number of APs in our hybrid system. Nonetheless, the proposed system shows a slightly better performance than the VLC standalone CF-mMIMO system until a point where it significantly surpasses it. This can be explained by the intervention of RF network to support the crowded VLC network by serving some of the users that are selected by Algorithm 1 to obtain a better sum-rate. VLC blockage effect is also

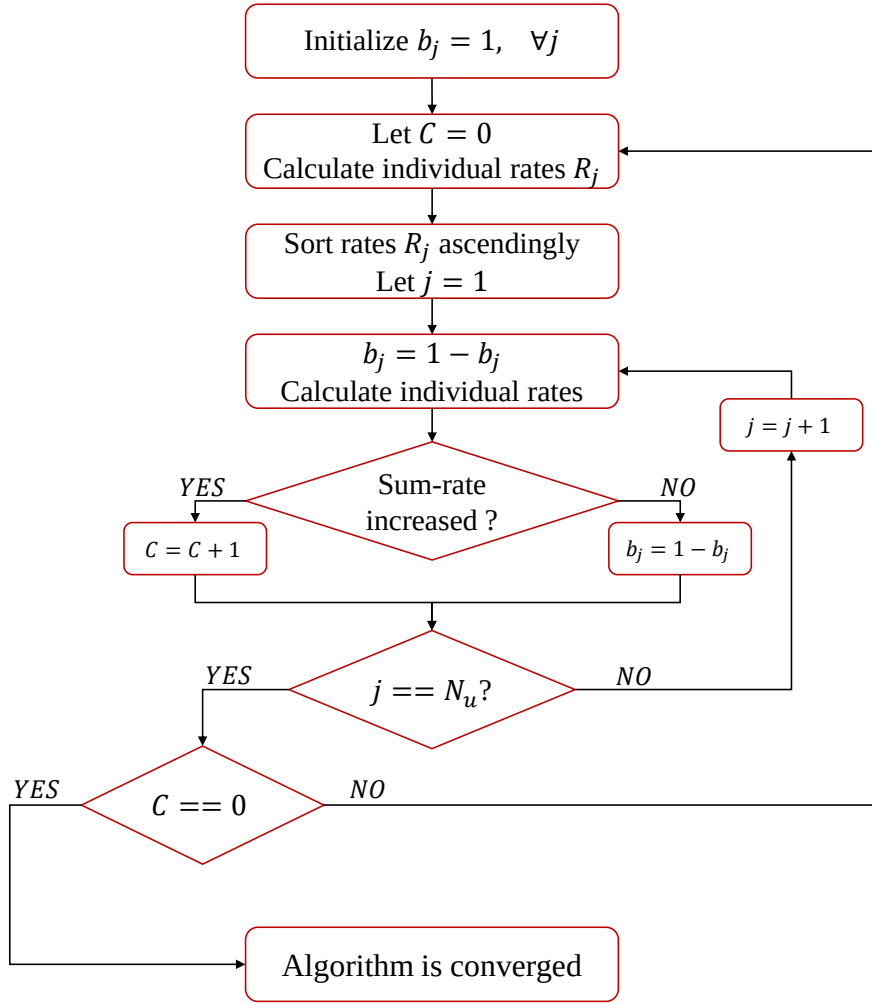


Figure 2.2: Flowchart illustrating Algorithm 1

included in the figure for both the VLC standalone and the Hybrid CF-mMIMO systems that uses the proposed user association mechanism. The figure shows that the proposed system does not deteriorate much when the blockage is introduced. Thus, blockage effect is assumed in all the following figures. Figure 2.3 also shows that the proposed system with Algorithm 1 dramatically outperforms the system with random user association, and the performance of the system with random user association is between the standalone CF-mMIMO systems, which indicates

Table 2.1: Simulation Parameters 1

Parameter	Value
Room length	10 m
Room width	10 m
Room height	3 m
Optical-to-electrical conversion factor ρ_{oe}	10 W/A
Electrical-to-optical conversion factor ρ_{eo}	0.53 A/W
Bandwidth of the VLC network $B^{(v)}$	40 MHz
Bandwidth of the RF network $B^{(r)}$	15 MHz
Noise power spectral density in VLC network	10^{-22} W/Hz
Noise power spectral density in RF network	10^{-19} W/Hz
Power limit for each VLC AP $\rho_f^{(v)}$	5 W
Power limit for each RF AP $\rho_f^{(r)}$	5 W
Number of VLC APs N_{vap}	16
Number of RF APs N_{rap}	9
User's altitude	0.85 m
Physical area of a PD A_p	1 cm
Semi-angle of the FoV of PDs Θ	60°
Half-intensity radiation angle $\theta_{\frac{1}{2}}$	60°
Gain of the optical filter G_{of}	1
Refractive index of PDs n	1.5
Monte-Carlo simulation	1000 scenarios
Blockage rate	10% of users

that the system does not take a good advantage of having both networks.

Figure 2.4 shows the steady performance of the proposed hybrid system using the proposed algorithm for an increasing number of the total APs in the room with a fixed number of users $N_u = 30$. In this figure, we adapt the room dimensions to accommodate all the APs in the system in order to maintain a reasonable deployment density of the APs. For instance, we preserve the minimum distance between any two adjacent VLC APs to be around 2.5 m. The exact numbers of VLC APs, RF APs and room dimensions are presented in Table 2.2. It is clear from Figure 2.3 and Figure 2.4 that the proposed system and algorithm work well

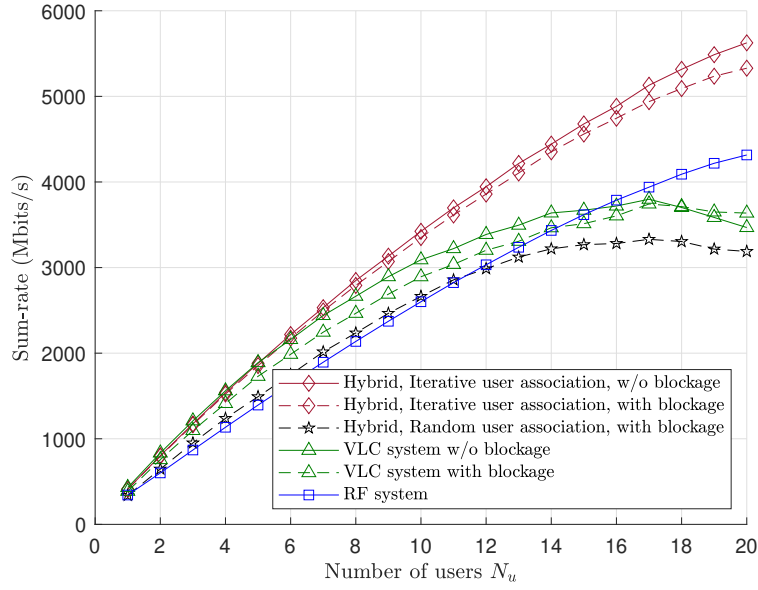


Figure 2.3: Sum-rate versus number of users in the systems

with increasing the number of users and the number of APs, respectively.

Table 2.2: Room Dimensions and Number of APs in Figure 2.4

Total number of APs	25	32	45	65	80	97	125
VLC APs	16	16	36	49	64	81	100
RF APs	9	16	9	16	16	16	25
Room dimensions	10 × 10	11 × 11	14 × 14	16 × 16	19 × 19	22 × 22	25 × 25

The FoV of the PDs is an important parameter that affects the VLC channel gain and hence the performance of VLC systems. Figure 2.5 shows the effect of changing the FoV of the VLC receivers on the average per-user data rate in both the proposed system and a VLC standalone CF-mMIMO system. In the VLC standalone system, we set that the number of APs to be equivalent to the total number of (VLC+RF) APs in the proposed hybrid system. This has been carried over three choices of number of users in the system. The figure shows that for large values of FoV, the performance diminishes, which should not be of any surprise

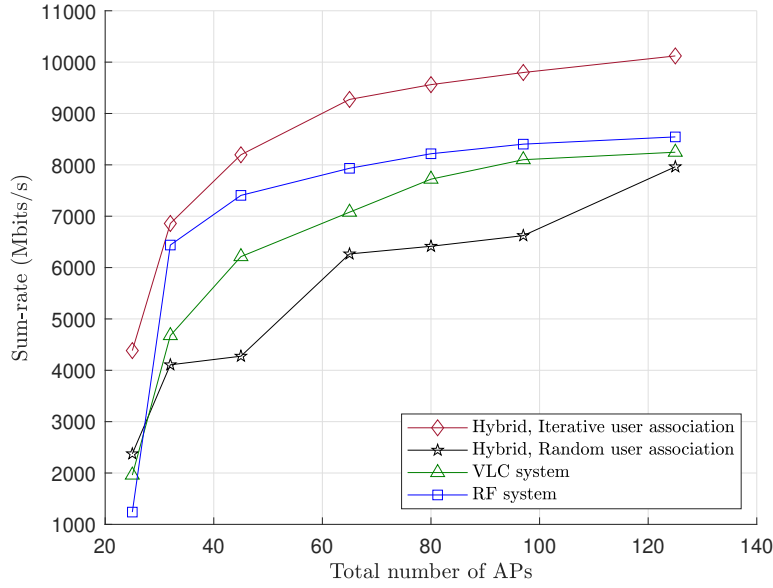


Figure 2.4: Sum-rate versus total number of APs

since the VLC channel gain drops as the angle of the FoV increases which can be seen in equation (1.3). Moreover, as the FoV decreases, the performance decreases as a result of more users not being able to be in the view of the VLC APs. VLC MIMO systems in general tend to push that limit to a very low FoV value since having more VLC APs to cover more ground area ensures that all users are able to connect. The proposed hybrid system, however, shows more resilience as the value of FoV drops which is a direct result of having the RF network as a support to the VLC network. Nevertheless, this resiliency is only visible when N_u is small since there could be cases where all VLC APs aren't in the FoV of any receiver. In addition, it is obvious from the figure that the performance gap between the proposed system and the VLC standalone widens as the number of user in the system increases.

Figure 2.6 shows the cumulative distribution function (CDF) of the average

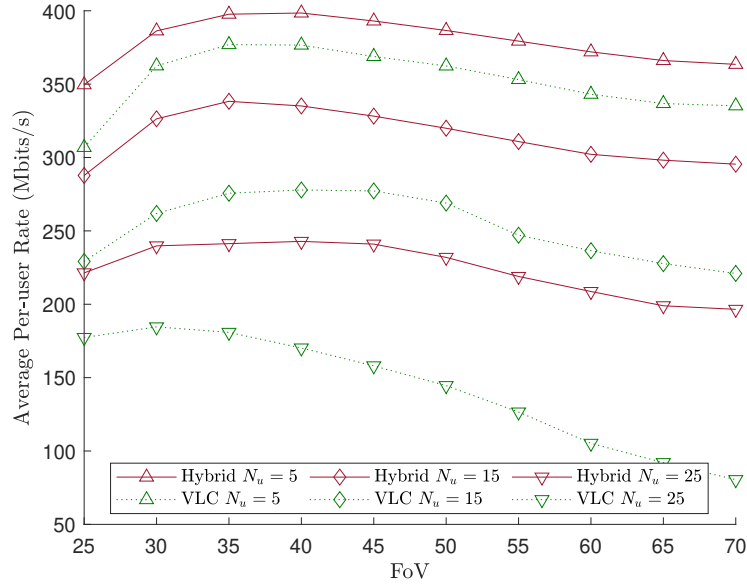


Figure 2.5: Sum-rate versus FoV

per-user rate of the four systems considered in Figure 2.3. All the systems are evaluated when $N_u = 10$ and $N_u = 20$. This figure emphasizes that the proposed system works best with relatively high number of users when the other systems fail to compete. It also emphasizes that the proposed user association (Algorithm 1) performs much better than the random user association.

Figure 2.7 studies the convergence of Algorithm 1, where it shows the total number of users' changes in CPU's internal processes and how any change affects the sum-rate of the system, where the overall performance is shown by the envelopes of the curves in which it succeeded in all four choices of N_u . The figure shows that Algorithm 1 converges with different number of users. However, increasing the number of users require Algorithm 1 to increase the number of changes to reach convergence.

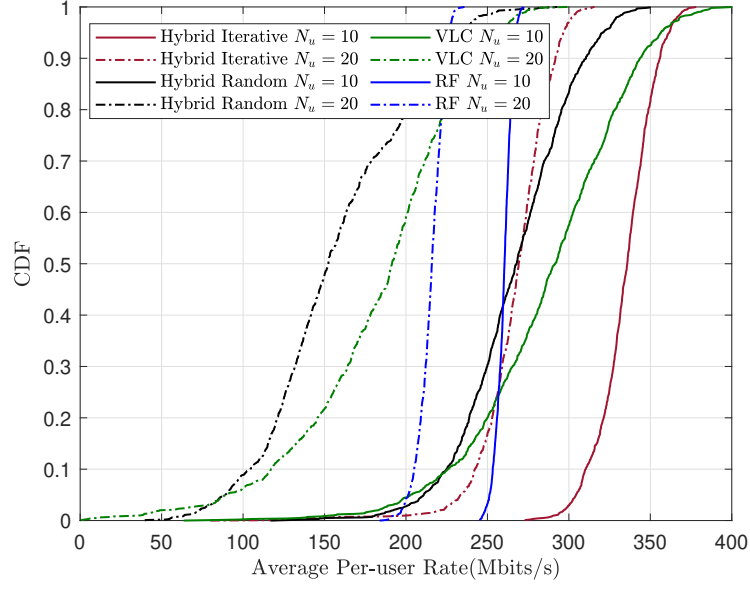


Figure 2.6: CDF of average per-user rate

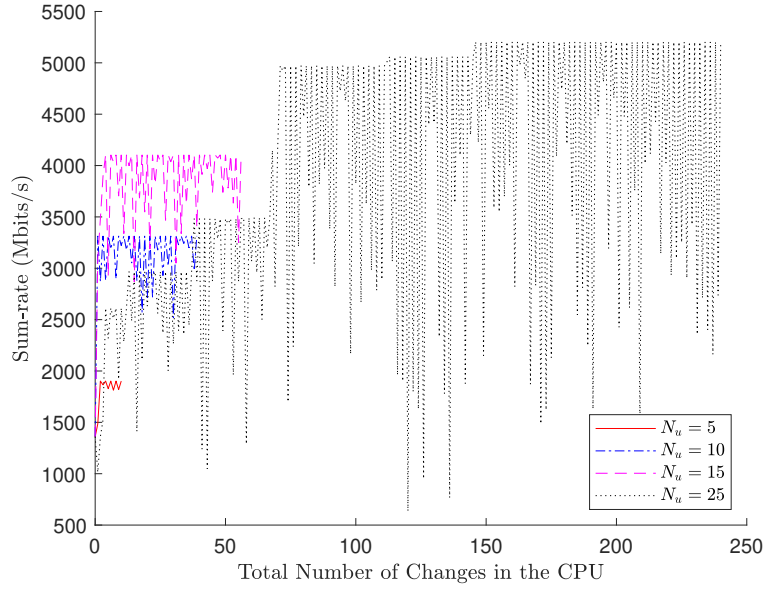


Figure 2.7: Sum-rate versus total number of changes

2.4 Conclusion

In this chapter, a hybrid VLC/RF cell-free massive MIMO system has been introduced and studied for the first time. Users are distributed over two existing cell-free networks (VLC and RF). This association problem is formulated as an integer optimization problem, and is approached via an iterative algorithm that converges to a suboptimal solution for the formulated optimization problem. The proposed system demonstrated better performance than the other cell-free systems with standalone VLC or RF networks in terms of both sum-rate and average per-user rate, especially at a relatively high number of users.

CHAPTER 3

USER-CENTRIC HYBRID CELL-FREE MASSIVE MIMO SYSTEMS

In this chapter, we propose a UC hybrid VLC/RF CF-mMIMO system for indoor communication. In such a system, users can be clustered using different, not necessarily mutually exclusive, clustering techniques. It is worth to point out the difference between the terms user association and user clustering; user association means the connection of the users to one of the existing networks, i.e. VLC or RF, while user clustering refers to the set of APs, within each network, that is serving each user. User clustering enables better performance when the backhaul capacity is limited, i.e. it limits the data exchange between the APs and the CPU. Generic derivations are carried to accept any form of clustering techniques and precoding schemes, with no conditions. The performance of the system is

studied for three clustering scenarios. We formulate a joint optimization problem for both users association and users clustering to maximize the sum-rate of the system. The joint optimization problem is then simplified into two independent sub-problems, where each is handled separately. The user association sub-problem is similar to the problem proposed in Chapter 2, however, it accounts for the clustering effect. A heuristic approach is then proposed to tackle the problem and obtain a sup-optimal solution. The proposed solution is compared to the algorithm provided in Chapter 2, and the strength of each algorithm is noted in the results. We evaluate the proposed user-centric hybrid CF-mMIMO system and show that it outperforms other systems that it is compared to. We also show that the proposed algorithm in this chapter significantly outperforms a random user association scheme.

The system model is exactly the same as the one in Chapter 2. Thus, we proceed to derive the performance metrics of interest in the system.

3.1 Networks' Achievable Sum-rate

In this section, we analyze the received signal at the receiver side and provide an expression for the achievable sum-rate. Let $\mathcal{U}_i^{(v)}$ be the set of users served by the VLC AP i , and $\mathcal{U}_i^{(r)}$ be the set of users served by the RF AP i . Moreover, let $\mathcal{A}_j^{(v)}$ be the set of VLC APs serving an arbitrary user j and let $\mathcal{A}_j^{(r)}$ be the set of RF APs serving an arbitrary user j . Note that the two sets $\mathcal{A}_j^{(v)}$ and $\mathcal{A}_j^{(r)}$ are mutually exclusive. In fact, for a specific user j one of the two sets will be

the empty set $\emptyset$, which means that a user can only be served by either the VLC network or the RF network but not both. Furthermore, the two sets $\mathcal{U}_i^{(v)}$ and $\mathcal{U}_i^{(r)}$ are mutually exclusive because a user cannot be served by both networks simultaneously. We can see that $\left| \bigcup_{i=1}^{N_{vap}} \mathcal{U}_i^{(v)} \right| = N_{u,v} \leq N_u$, $\left| \bigcup_{i=1}^{N_{rap}} \mathcal{U}_i^{(r)} \right| = N_{u,r} \leq N_u$, and $\left| \bigcup_{m \in \{v,r\}} \bigcup_{i=1}^{N_{\{m\}ap}} \mathcal{U}_i^{(m)} \right| = N_u$. In addition, we define a matrix $\mathbf{A}^{(v)} \in \mathbb{R}^{N_u \times N_{vap}}$ that represents the user clustering in the VLC network, such that

$$\mathbf{A}^{(v)} = \begin{bmatrix} a_{1,1}^{(v)} & a_{1,2}^{(v)} & \cdots & a_{1,N_{vap}}^{(v)} \\ a_{2,1}^{(v)} & a_{2,2}^{(v)} & \cdots & a_{2,N_{vap}}^{(v)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N_u,1}^{(v)} & a_{N_u,2}^{(v)} & \cdots & a_{N_u,N_{vap}}^{(v)} \end{bmatrix}, \quad (3.1)$$

where the element $a_{j,i}^{(v)}$ is a clustering factor between user j and VLC AP i , and is given by

$$a_{j,i}^{(v)} = \begin{cases} 1, & \text{if VLC AP } i \text{ serves user } j; \\ 0, & \text{if VLC AP } i \text{ does not serve user } j. \end{cases} \quad (3.2)$$

A similar matrix for RF users clustering $\mathbf{A}^{(r)} \in \mathbb{R}^{N_u \times N_{rap}}$ is also defined as

$$\mathbf{A}^{(r)} = \begin{bmatrix} a_{1,1}^{(r)} & a_{1,2}^{(r)} & \cdots & a_{1,N_{rap}}^{(r)} \\ a_{2,1}^{(r)} & a_{2,2}^{(r)} & \cdots & a_{2,N_{rap}}^{(r)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N_u,1}^{(r)} & a_{N_u,2}^{(r)} & \cdots & a_{N_u,N_{rap}}^{(r)} \end{bmatrix}, \quad (3.3)$$

where the element $a_{j,i}^{(r)}$ is a clustering factor between user j and RF AP i , and is given by

$$a_{j,i}^{(r)} = \begin{cases} 1, & \text{if RF AP } i \text{ chooses to serve user } j; \\ 0, & \text{if RF AP } i \text{ chooses not to serve user } j. \end{cases} \quad (3.4)$$

Finally, we define a vector $\mathbf{b}$ of length N_u that corresponds to the users being connected to either the VLC network or the RF network, it can be written as

$$\mathbf{b} = \begin{bmatrix} b_1 & b_2 & \cdots & b_{N_u} \end{bmatrix}^T, \quad (3.5)$$

where $b_j \in \{0, 1\}$. Setting $b_j = 0$ means that user j is connected to the RF network, while $b_j = 1$ indicates that user j is connected to the VLC network. Vector $\mathbf{b}$ guarantees that a single user cannot be served simultaneously by both networks.

3.1.1 VLC Achievable Rate

Here, we provide the achievable rate expression achieved at the users that are served by the VLC network. Each VLC AP transmits an entry of the vector $\mathbf{x}^{(v)} = \begin{bmatrix} x_1^{(v)} & x_2^{(v)} & \cdots & x_{N_{vap}}^{(v)} \end{bmatrix}^T$, where AP i transmits the i^{th} entry that is given

by

$$x_i^{(v)} = \sqrt{\rho_f^{(v)}} \sum_{j \in \mathcal{U}_i^{(v)}} w_{i,j}^{(v)} \sqrt{P_j^{(v)}} s_j \quad (3.6a)$$

$$= \sqrt{\rho_f^{(v)}} \sum_{j=1}^{N_u} b_j a_{j,i}^{(v)} w_{i,j}^{(v)} \sqrt{P_j^{(v)}} s_j, \quad (3.6b)$$

where s_j is the message intended for user j , $\rho_f^{(v)}$ is the total available power at each VLC AP, $P_j^{(v)}$ is the power allocated to the messages of user j in the VLC network, and $w_{j,i}^{(v)}$ is the weight assigned to symbol s_j that is transmitted by VLC AP i . We define $\mathbf{W}^{(v)} \in \mathbb{R}^{N_{vap} \times N_u}$ to be the precoding matrix, where its $(i, j)^{th}$ entry is given by $w_{i,j}^{(v)}$. Therefore, we can write the received signal at the VLC user j as

$$y_j^{(v)} = \rho \sum_{i=1}^{N_{vap}} h_{j,i}^{(v)} x_i^{(v)} + n_j^{(v)}, \quad (3.7a)$$

$$= \rho \sum_{i=1}^{N_{vap}} h_{j,i}^{(v)} \left(\sqrt{\rho_f^{(v)}} \sum_{l=1}^{N_u} b_l a_{l,i}^{(v)} w_{i,l}^{(v)} \sqrt{P_l^{(v)}} s_l \right) + n_j^{(v)}, \quad (3.7b)$$

$$= \rho \sqrt{\rho_f^{(v)}} \left(\sum_{l=1}^{N_u} b_l \left(\mathbf{h}_j^{(v)} \right)^T \mathbf{A}_{d,l}^{(v)} \mathbf{w}_l^{(v)} \sqrt{P_l^{(v)}} s_l \right) + n_j^{(v)}, \quad (3.7c)$$

where $\rho = \rho_{oe} \rho_{eo}$ is defined as the equivalent conversion factor, ρ_{oe} and ρ_{eo} are the optical-to-electrical and the electrical-to-optical conversion factors, respectively, and $n_j^{(v)} \sim \mathcal{N}(0, \sigma_v^2)$ is the additive white Gaussian noise at the PD of user j . $\mathbf{H}_v = \begin{bmatrix} \mathbf{h}_1^{(v)} & \mathbf{h}_2^{(v)} & \dots & \mathbf{h}_{N_u}^{(v)} \end{bmatrix}^T$ is the VLC channel matrix, where $\mathbf{h}_j^{(v)} = \begin{bmatrix} h_{j,1}^{(v)} & h_{j,2}^{(v)} & \dots & h_{j,N_{vap}}^{(v)} \end{bmatrix}^T$ is a channel vector of all VLC APs to the j^{th} user. $\mathbf{w}_l^{(v)}$ is the symbol used to denote the l^{th} column vector of $\mathbf{W}^{(v)}$. $\mathbf{A}_{d,l}^{(v)}$ is a

diagonal matrix with the diagonal elements being the elements in the l^{th} row of $\mathbf{A}^{(v)}$, and is mathematically written as

$$\mathbf{A}_{d,l}^{(v)} = \text{diag} \left(\mathbf{a}_{l,:}^{(v)} \right). \quad (3.8)$$

where $\mathbf{a}_{l,:}^{(v)}$ is the symbol used to indicate the l^{th} row vector of $\mathbf{A}^{(v)}$. Equation (3.7d) is further decomposed in equation (3.9) at the top of the next page where the desired signal at the receiver of VLC user j and the interference and noise are separated.

$$y_j^{(v)} = \rho \sqrt{\rho_f^{(v)}} \left(\overbrace{b_j \left(\mathbf{h}_j^{(v)} \right)^T \mathbf{A}_{d,j}^{(v)} \mathbf{w}_j^{(v)} s_j}^{\text{Desired Signal}} + \overbrace{\sum_{\substack{l=1 \\ l \neq j}}^{N_u} b_l \left(\mathbf{h}_j^{(v)} \right)^T \mathbf{A}_{d,l}^{(v)} \mathbf{w}_l^{(v)} s_l}^{\text{Interfering Signals}} \right) + \overbrace{n_j^{(v)}}^{\text{Noise}}, \quad (3.9)$$

The achievable rate at the user j connected to the VLC network can be written as [76]

$$R_j^{(v)} = \frac{B^{(v)}}{2} \log_2 \left(1 + \gamma_j^{(v)} \right) \quad (3.10)$$

where $B^{(v)}$ is the total VLC bandwidth and $\gamma_j^{(v)}$ is given by

$$\gamma_j^{(v)} = \frac{\left| \rho \sqrt{\rho_f^{(v)}} b_j \left(\mathbf{h}_j^{(v)} \right)^T \mathbf{A}_{d,j}^{(v)} \mathbf{w}_j^{(v)} \sqrt{P_j^{(v)}} \right|^2}{\sum_{\substack{l=1 \\ l \neq j}}^{N_u} \left| \rho \sqrt{\rho_f^{(v)}} b_l \left(\mathbf{h}_j^{(v)} \right)^T \mathbf{A}_{d,l}^{(v)} \mathbf{w}_l^{(v)} \sqrt{P_l^{(v)}} \right|^2 + 9\sigma_v^2}. \quad (3.11)$$

3.1.2 RF Achievable Rate

Here, we provide an expression for the achievable rate at the users that are served by the RF network. The vector $\mathbf{x}^{(r)} = \begin{bmatrix} x_1^{(r)} & x_2^{(r)} & \dots & x_{N_{rap}}^{(r)} \end{bmatrix}^T$ represents the transmitted signals from RF APs, where $x_i^{(r)}$ is the transmitted signal from the i^{th} RF AP and is given by

$$x_i^{(r)} = \sqrt{\rho_f^{(r)}} \sum_{j \in \mathcal{U}_i^{(r)}} w_{i,j}^{(r)} \sqrt{P_j^{(r)}} s_j \quad (3.12a)$$

$$= \sqrt{\rho_f^{(r)}} \sum_{j=1}^{N_u} (1 - b_j) a_{j,i}^{(r)} w_{i,j}^{(r)} \sqrt{P_j^{(r)}} s_j. \quad (3.12b)$$

where $\rho_f^{(r)}$ is the total available power at each RF AP, $P_j^{(r)}$ is the power allocated the messages of the j^{th} user in the RF network, and $w_{i,j}^{(r)}$ is the weight assigned to symbol s_j transmitted by the i^{th} RF AP. Let $\mathbf{W}^{(r)} \in \mathbb{R}^{N_{rap} \times N_u}$ be the precoding matrix, where its $(i, j)^{th}$ entry is given by $w_{i,j}^{(r)}$. The received signal by the user j can then be written as

$$y_j^{(r)} = \sum_{i=1}^{N_{rap}} h_{j,i}^{(r)} x_i^{(r)} + n_j^{(r)}, \quad (3.13a)$$

$$= \sum_{i=1}^{N_{rap}} h_{j,i}^{(r)} \left(\sqrt{\rho_f^{(r)}} \sum_{l=1}^{N_u} (1 - b_l) a_{l,i}^{(r)} w_{i,l}^{(r)} \sqrt{P_l^{(r)}} s_l \right) + n_j^{(r)}, \quad (3.13b)$$

$$= \sqrt{\rho_f^{(r)}} \left(\sum_{l=1}^{N_u} (1 - b_l) \left(\mathbf{h}_j^{(r)} \right)^T \mathbf{A}_{d,l}^{(r)} \mathbf{w}_l^{(r)} \sqrt{P_l^{(r)}} s_l \right) + n_j^{(r)}, \quad (3.13c)$$

where $\mathbf{w}_l^{(r)}$ is the l^{th} column vector of $\mathbf{W}^{(r)}$, and $\mathbf{h}_j^{(r)} = \begin{bmatrix} h_{j,1}^{(r)} & h_{j,2}^{(r)} & \dots & h_{j,N_{rap}}^{(r)} \end{bmatrix}^T$ is a channel vector of all RF APs to the j^{th} user, and the RF channel matrix is

given as $\mathbf{H}_r = \begin{bmatrix} \mathbf{h}_1^{(r)} & \mathbf{h}_2^{(r)} & \dots & \mathbf{h}_{N_u}^{(r)} \end{bmatrix}^T$, and $n_j^{(r)} \sim \mathcal{CN}(0, \sigma_r^2)$ is the additive white Gaussian noise at the antenna of user j . Equation (3.13c) is further decomposed in equation (3.14) at the top of this page where the desired signal at the receiver of RF user j and the interference and noise are separated.

$$y_j^{(r)} = \sqrt{\rho_f^{(r)}} \left(\overbrace{(1 - b_j) \left(\mathbf{h}_j^{(r)} \right)^T \mathbf{A}_{d,j}^{(r)} \mathbf{w}_j^{(r)} s_j}^{\text{Desired Signal}} + \overbrace{\sum_{\substack{l=1 \\ l \neq j}}^{N_u} (1 - b_l) \left(\mathbf{h}_j^{(r)} \right)^T \mathbf{A}_{d,l}^{(r)} \mathbf{w}_l^{(r)} s_l}^{\text{Interfering Signals}} \right) + \overbrace{n_j^{(r)}}^{\text{Noise}}, \quad (3.14)$$

The achievable rate at the user j that is connected to the RF network is given by Shannon's formula as

$$R_j^{(r)} = B^{(r)} \log_2 \left(1 + \gamma_j^{(r)} \right) \quad (3.15)$$

where $B^{(r)}$ is the total available RF bandwidth and $\gamma_j^{(r)}$ is given by

$$\gamma_j^{(r)} = \frac{\left| \sqrt{\rho_f^{(r)}} (1 - b_j) \left(\mathbf{h}_j^{(r)} \right)^T \mathbf{A}_{d,j}^{(r)} \mathbf{w}_j^{(r)} \sqrt{P_j^{(r)}} \right|^2}{\sum_{\substack{l=1 \\ l \neq j}}^{N_u} \left| \sqrt{\rho_f^{(r)}} (1 - b_l) \left(\mathbf{h}_j^{(r)} \right)^T \mathbf{A}_{d,l}^{(r)} \mathbf{w}_l^{(r)} \sqrt{P_l^{(r)}} \right|^2 + \sigma_r^2}. \quad (3.16)$$

3.2 Problem Formulation and Solution

In this section, we are interested in the users association to the system networks and users clustering within each network aiming at maximizing the system's sum-rate. In other words, what are the best selections of the vector $\mathbf{b}$ and the matrices $\mathbf{A}^{(v)}$ and $\mathbf{A}^{(r)}$ to maximize the sum-rate. Since the users association and clustering

problems are interlinked, we formulate a joint optimization problem for both users association and users clustering. The joint optimization problem is then simplified into two independent sub-problems, where each is handled separately

3.2.1 Problem Formulation

Given that the association variable b_j of user j is either one or zero, we can express the achievable rate at user j as follows

$$R_j = R_j^{(v)} + R_j^{(r)}. \quad (3.17)$$

where $R_j^{(v)}$ and $R_j^{(r)}$ are defined in (3.10) and (3.15), respectively. Expression (3.17) implies that $R_j = R_j^{(v)}$ if $b_j = 1$ or $R_j = R_j^{(r)}$ if $b_j = 0$. In other words, if user j is connected to the VLC networks, $R_j = R_j^{(v)}$; otherwise, $R_j = R_j^{(r)}$. This emphasizes the fact that only one network, either VLC or RF, would serve user j

at a time. Therefore, we can formulate our joint optimization problem as

$$\max_{\mathbf{A}^{(v)}, \mathbf{A}^{(r)}, \mathbf{b}} \sum_j R_j \quad (3.18a)$$

$$\text{s.t. } b_j \in \{0, 1\}, \quad \forall j, \quad (3.18b)$$

$$a_{j,i}^{(v)} \in \{0, 1\}, \quad \forall i, j, \quad (3.18c)$$

$$a_{j,i}^{(r)} \in \{0, 1\}, \quad \forall i, j, \quad (3.18d)$$

$$\sum_{j=1}^{N_u} \left(w_{i,j}^{(v)}\right)^2 P_j^{(v)} \leq 1, \quad \forall i, \quad (3.18e)$$

$$\sum_{j=1}^{N_u} \left(w_{i,j}^{(r)}\right)^2 P_j^{(r)} \leq 1, \quad \forall i. \quad (3.18f)$$

Constraint (3.18b) guarantees that each user is served by only one network at a time. Constraints (3.18c) and (3.18d) are responsible for clustering the users in each network. Constraints (3.18e) and (3.18f) are imposed to make sure that the transmit power from the AP i is not greater than ρ_f (the total available power). Problem (3.18) is not an easy problem to tackle since it is a non-convex nonlinear integer problem. However, by discarding the link between the three interlinked optimization variables ($\mathbf{b}, \mathbf{A}^{(v)}, \mathbf{A}^{(r)}$), we can simplify the problem in (3.18). Thus,

we can formulate a users association sub-problem as

$$\max_{\mathbf{b}}. \quad \sum_j R_j \quad (3.19a)$$

$$\text{s.t.} \quad b_j \in \{0, 1\}, \quad \forall j, \quad (3.19b)$$

$$\sum_{j=1}^{N_u} \left(w_{i,j}^{(v)} \right)^2 P_j^{(v)} \leq 1, \quad \forall i, \quad (3.19c)$$

$$\sum_{j=1}^{N_u} \left(w_{i,j}^{(r)} \right)^2 P_j^{(r)} \leq 1, \quad \forall i. \quad (3.19d)$$

This is a non-convex nonlinear integer problem which is still not easy to tackle. As mentioned in Subsection 2.2.2, the optimal solution can be reached by the impractical exhaustive search where we go over every possible choice of the vector $\mathbf{b}$. It is also possible to use the Algorithm 1 from Chapter 2 to solve (3.19). However, in the following we propose another iterative algorithm that is based on the Gibbs sampling algorithm. The advantages of this algorithm over Algorithm 1 would be clear as we evaluate and compare their performances later in this chapter.

3.2.2 Gibbs Sampling-Based User Association

Gibbs sampling is an algorithm that was mainly proposed to solve an optimization problem of the type

$$\min_{\mathbf{x}} \quad H(\mathbf{x}), \quad (3.20)$$

where $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_N \end{bmatrix}^T$ is an N -dimensional vector [77]. The iterative nature of Gibbs sampling is presented in the fact that the algorithm updates a single item in the vector $\mathbf{x}$ in each iteration of the algorithm based on the distribution $\Lambda_n(\mathbf{x}_{-n}) = (\Lambda_n(x_n|\mathbf{x}_{-n}), \forall x_n)$, where $\Lambda_n(x_n|\mathbf{x}_{-n})$ is given by

$$\Lambda_n(x_n|\mathbf{x}_{-n}) = \frac{\exp\left(-\beta H(x_n, \mathbf{x}_{-n})\right)}{\sum_{x'_n} \exp\left(-\beta H(x'_n, \mathbf{x}_{-n})\right)} \quad (3.21)$$

where n is the item to be updated, β is the temperature parameter, and $\mathbf{x}_{-n} = \begin{bmatrix} x_1 & \dots & x_{n-1} & x_{n+1}, \dots & x_N \end{bmatrix}^T$. At each iteration, the distribution $\Lambda_n(\mathbf{x}_{-n})$ varies according to the new $\mathbf{x}$. Due to this process, the value of x_n that results in a smaller value for the objective function $H(\mathbf{x})$ has higher probability of being selected.

The problem in (3.19) is a maximization problem. However, it can be turned into a minimization problem by using $U = 1/\sum_j R_j$ as the new objective function. The problem is now written as

$$\min_{\mathbf{b}} \quad U(\mathbf{b}) = \frac{1}{\sum_j R_j} \quad (3.22a)$$

$$\text{s.t.} \quad b_j \in \{0, 1\}, \quad \forall j, \quad (3.22b)$$

$$\sum_{j=1}^{N_u} \left(w_{i,j}^{(v)}\right)^2 P_j^{(v)} \leq 1, \quad \forall i, \quad (3.22c)$$

$$\sum_{j=1}^{N_u} \left(w_{i,j}^{(r)}\right)^2 P_j^{(r)} \leq 1, \quad \forall i. \quad (3.22d)$$

This alternation in the problem formulation enables us to make use of Gibbs

sampling with some other modifications.

A Gibbs-sampling based algorithm for user association is developed similar to [78]. At the beginning of each iteration t , we have the value of $\mathbf{b}^{t-1}$, which is the result of the previous iteration. We define the set $\mathcal{B}^{t-1}$ that contains all the vectors $\tilde{\mathbf{b}}_m^{t-1}$ that are different by only the m^{th} element from $\mathbf{b}^{t-1}$ and the vector $\mathbf{b}^{t-1}$ itself. For mathematical convenience $\tilde{\mathbf{b}}_0^{t-1} = \mathbf{b}^{t-1}$. Hence, the probability distribution function becomes

$$\Lambda_t(\mathbf{b}^t) = \frac{\exp\left(-\beta U(\mathbf{b}^t)\right)}{\sum_{\mathbf{b}' \in \mathcal{B}^{t-1}} \exp\left(-\beta U(\mathbf{b}')\right)}, \quad (3.23a)$$

$$= \frac{\exp\left(-\frac{\beta}{\sum_j R_j}\right)}{\sum_{\mathbf{b}' \in \mathcal{B}^{t-1}} \exp\left(-\frac{\beta}{\sum_j R_j}\right)}. \quad (3.23b)$$

The algorithm is presented in details in **Algorithm 2** and Figure 3.1. Notice that at the start of the algorithm we initialize the elements of the vector $\mathbf{b}$ from the set $\{0, 1\}$ according to the discrete uniform distribution (DUD).

3.2.3 User Clustering

User clustering has several advantages that include limiting backhaul transmission as well as enabling the scalability of the CF-mMIMO system. The aim here is to determine which APs in each network should serve each of the users in that network. Since the link between the optimization variables in (3.18) is discarded,

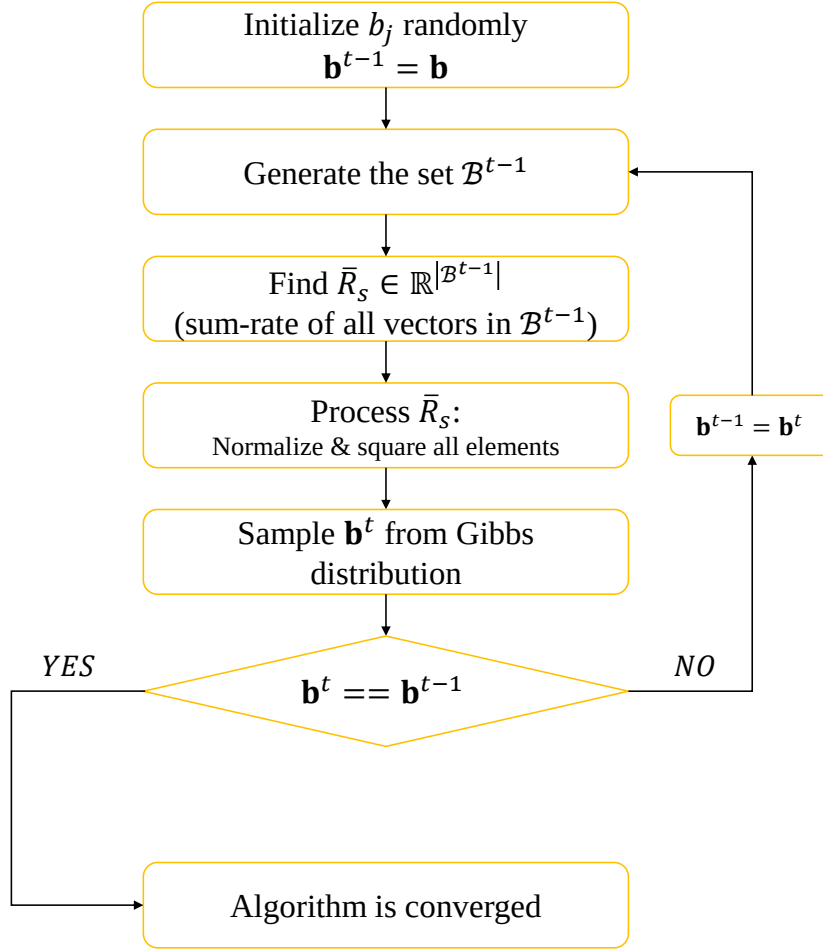


Figure 3.1: Flowchart illustrating Algorithm 2

we can formulate the user clustering sub-problem in any network $m \in \{v, r\}$ as

$$\max_{\mathbf{A}} \quad \sum_j R_j \quad (3.24a)$$

$$\text{s.t.} \quad a_{j,i} \in \{0, 1\}, \quad \forall i, j, \quad (3.24b)$$

$$\sum_{j=1}^{N_u} (w_{i,j})^2 P_j \leq 1, \quad \forall i. \quad (3.24c)$$

Algorithm 2 Gibbs Sampling-Based User Association

Input: $N_u, T_{max}, \beta, \mathbf{H}_v, \mathbf{H}_r, \mathbf{W}^{(v)}, \mathbf{W}^{(r)}, \rho_f^{(v)}, \rho_f^{(r)}, B^{(v)}, B^{(r)}, \rho_{eo}, \rho_{oe}, \sigma_v, \sigma_r$.

Output: $\mathbf{b}$

Initialize: Sample $\mathbf{b}$ according to DUD from $\{0, 1\}$

```

1: Convergence  $\leftarrow$  False
2:  $\mathbf{b}_{old} \leftarrow \mathbf{b}$ 
3: while  $t \leq T_{max}$  AND Convergence  $\neq$  True do
4:   Generate the set  $\mathcal{B}^{t-1} = \{\tilde{\mathbf{b}}_m^{t-1}\}_{m=0}^{N_u}$   $\triangleright$  Containing  $N_u + 1$  vectors
5:   for each  $\mathbf{b}' \in \mathcal{B}^{t-1}$  do
6:     From  $\mathbf{H}_v$  and  $\mathbf{H}_r$  Obtain  $\mathbf{A}_v$  and  $\mathbf{A}_r$ , respectively
7:     Calculate  $R_s \leftarrow \sum_j R_j$  using (3.10), (3.15) and (3.17)
8:   end for
9:    $\bar{R}_s \leftarrow [R_s] \in \mathbb{R}^{N_u+1}$   $\triangleright$  i.e. concatenate values of  $R_s$ 
10:  Normalize the vector  $\bar{R}_s$  to be in  $[0, 1]$ 
11:  Divide the elements of  $\bar{R}_s$  by  $\sum R_s$ 
12:  Square all the elements of  $\bar{R}_s$   $\triangleright$  To distinguish the values in  $\bar{R}_s$ 
13:  Sample  $\mathbf{b}^t$  from the distribution in (3.23)
14:  if  $\mathbf{b}^t = \mathbf{b}_{old}$  then
15:    Convergence  $\leftarrow$  True
16:  else
17:     $\mathbf{b}_{old} \leftarrow \mathbf{b}^t$ 
18:  end if
19: end while

```

Notice that the superscripts are omitted since the problem is applied to both networks. This is again a non-convex nonlinear integer problem and not easy to tackle. Therefore, without loss of generality, we propose to use two distinct sub-optimal clustering techniques, namely $\mathbf{A}_1$ and $\mathbf{A}_2$, to help evaluating the system. These clustering techniques certainly do not guarantee optimal solution but are justified selections in our problem.

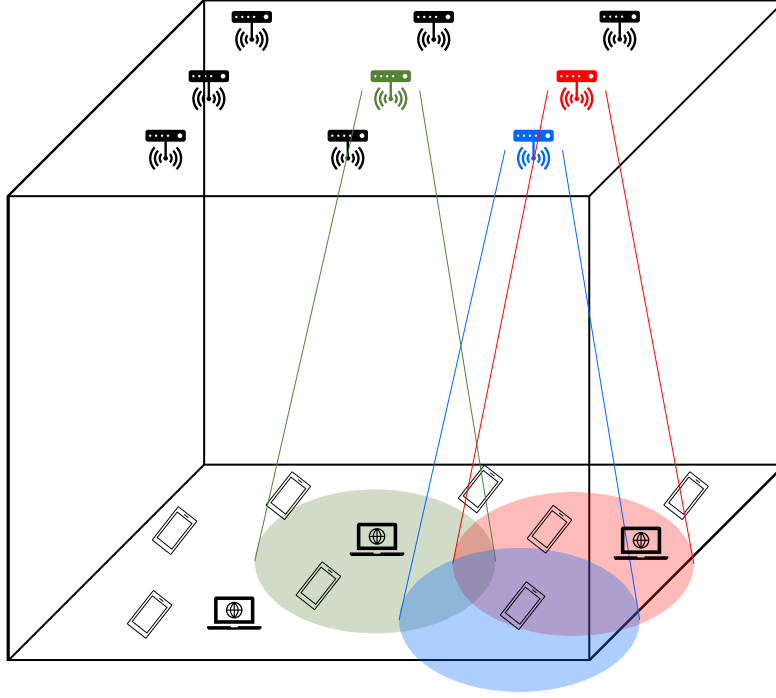


Figure 3.2: Illustration of $\mathbf{A}_1$ clustering

To start with, the first clustering technique, $\mathbf{A}_1 = [a_{1,j,i}]$, is given by

$$a_{1,j,i} = \begin{cases} 1, & d_{j,i} \leq d_{max}; \\ 0, & d_{j,i} > d_{max}, \end{cases} \quad (3.25)$$

where $d_{j,i}$ denotes the distance between user j and AP i , and d_{max} is a network design parameter representing the maximum distance a user could be from an AP and still be served by that AP. Figure [3.2](#) shows this clustering technique in action.

The second clustering technique, $\mathbf{A}_2 = [a_{2,j,i}]$, is based on the maximum number of served users per AP, i.e. each AP can serve up to N_{max} users. Those users are

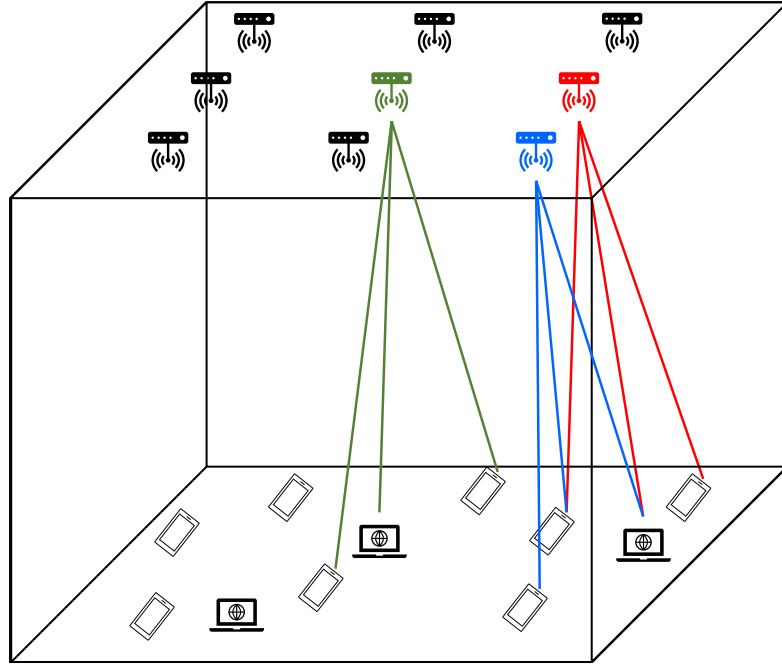


Figure 3.3: Illustration of $\mathbf{A}_2$ clustering

chosen based on their distance from that AP. This clustering is given by

$$a_{2j,i} = \begin{cases} 1, & \text{if } d_{j,i} \text{ is one of the } N_{max} \text{ smallest } d_{l,i}; \\ 0, & \text{otherwise.} \end{cases} \quad (3.26)$$

Figure 3.3 shows a simple scenario where this clustering is implemented with $N_{max} = 3$.

3.2.4 Constraints Satisfaction

In this section we convey the precoding schemes used in this paper. In addition, we provide the methodology that we follow to satisfy the constraints in (3.18e)

and (3.18f) which are the same as those in (3.19c), (3.19d), (3.22c), (3.22d), and the general one in (3.24c).

The precoding matrix $\mathbf{W}$ can be selected in multiple ways. However, we choose to use the ZF precoding scheme. This precoding scheme is meant to suppress the intra-network interference. The ZF precoding that incorporates the clustering is denoted by $\mathbf{W}_{ZF} = \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_{N_u} \end{bmatrix}$, where $\mathbf{w}_j$ is given as in [79] by

$$\mathbf{w}_j = \left(\sum_{l \in \mathcal{S}_j} \mathbf{A}_{d,j} \mathbf{h}_l (\mathbf{A}_{d,j} \mathbf{h}_l)^H \right)^{-1} \mathbf{A}_{d,j} \mathbf{h}_j, \quad (3.27)$$

where $\mathcal{S}_j$ represents the set of users that are served by partially the same APs as user j . Hence, this precoding scheme is usually referred to as partial ZF. Moreover, this set is given as $\mathcal{S}_j = \{i : \mathbf{A}_{d,i} \mathbf{A}_{d,j} \neq \mathbf{0}\}$, where $\mathbf{0}$ is the zero matrix [79].

Precoding schemes distribute the powers from APs to users. However, to meet the power constraints in our problem/sub-problems we need to allocate powers to users. Transmitted power from AP i in any network can be given in term of both user powers and the precoding matrix as

$$p_i = \sum_{j=1}^{N_u} (w_{i,j})^2 P_j, \quad (3.28)$$

where p_i is the power transmitted by AP i . To satisfy the constraints of our problem, we choose to allocate equal power to all users. Hence, users' powers are

given by

$$P_j = \frac{1}{\max_i \left(\left[\sum_{j=1}^{N_u} (w_{i,j})^2 \right] \right)}. \quad (3.29)$$

This guarantees that each AP transmits signal within its power budget.

3.3 Simulation Results

In this section, the performance of the proposed system along with the two proposed association algorithms are evaluated by studying the effect of changing several parameters on the sum-rate. Unless stated otherwise, the simulation parameters used in our setup to generate all the results are listed in Table 2.1 in the previous chapter. Monte-Carlo simulation has been implemented to obtain more accurate and generic results, and every displayed point is the arithmetic mean of 1000 points resulted from 1000 independent distribution of users. In addition, VLC blockage effect for both VLC systems and the Hybrid systems is assumed in all results since it was shown in the previous chapter that the performance of the proposed system does not deteriorate much when the blockage is introduced. The results in this section are divided into two subsections, where in the first subsection we evaluate the system with no clustering, and in the second subsection we evaluate it using the two proposed clustering techniques.

3.3.1 Hybrid CF-mMIMO without clustering

It has been already established that the clustering matrix $\mathbf{A}$ allows for generalized analysis. This means that it covers all possible connections between APs and

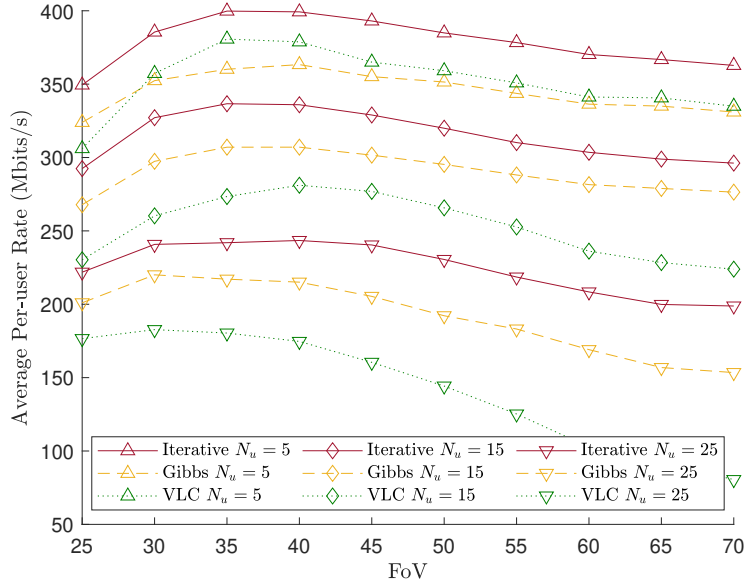


Figure 3.4: Average per-user rate versus FoV

users. Hence, the static (original) CF structure is nothing but a special case. In fact, setting $a_{j,i} = 1, \forall i, j$ leads to having all APs serve all users, a no-clustering scenario. In this subsection, we evaluate this special case of the system. However, the results of this part are greatly similar to the results of the previous chapter. Thus, we only provide the result of the average per-user rate as the FoV is changing and the CDF of the average per-user rate, using both Algorithm 1 and Algorithm 2.

Figure 3.4 shows the effect of changing the FoV of the VLC receivers on the average per-user data rate in both the proposed system (using both Algorithm 1 and Algorithm 2) and a VLC standalone CF system. In the VLC standalone system, we set the number of APs to be equivalent to the total number of (VLC+RF) APs in the proposed hybrid system. This has been carried over three choices of number of users in the system. The conclusion of this figure is similar to Figure

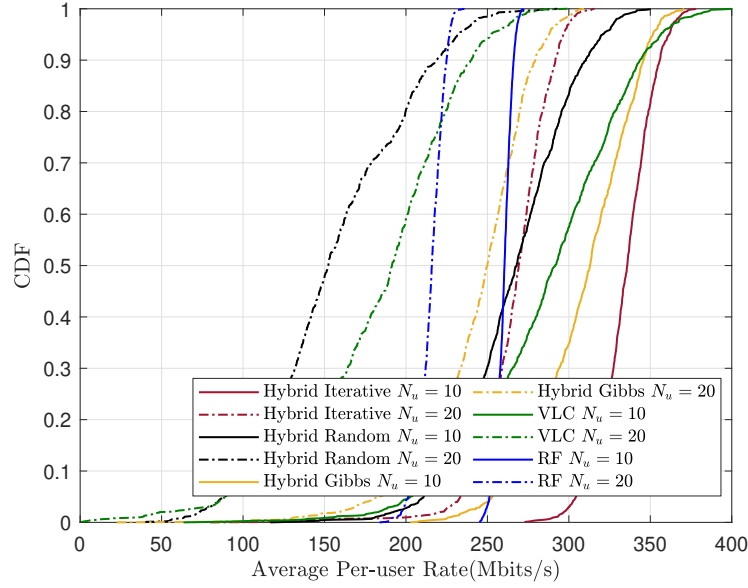


Figure 3.5: CDF of average per-user rate

2.5], where it tells us the proposed system outperforms the VLC standalone system. In addition, it is obvious from the figure that the performance gap between the proposed system and the VLC standalone widens as the number of user in the system increases. It came as surprise that when $N_u = 5$ the standalone VLC CF system performed better than the proposed system with Algorithm 2, which indicates that this algorithm may not always be suitable when the room has a small number of users.

Figure 3.5 shows the cumulative distribution function (CDF) of the average per-user rate of five systems, namely, the proposed hybrid system using Algorithm 1 and Algorithm 2 and a random user association scheme, as well as two cell-free systems that consists of either RF APs or VLC APs (standalone CF systems). All the systems are evaluated when $N_u = 10$ and $N_u = 20$. This figure emphasizes that the proposed system works best with relatively high number of users

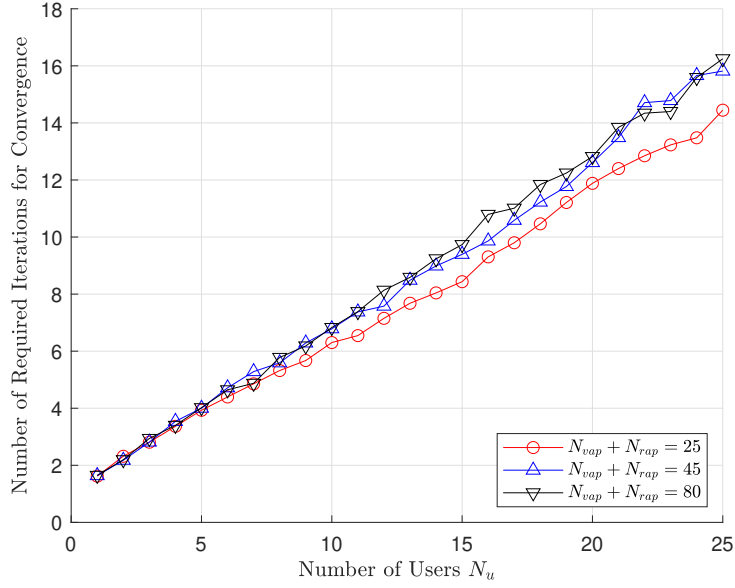


Figure 3.6: Number of required iterations for convergence versus number of users when the other systems fail to compete. It also emphasizes that the proposed user association algorithms [2](#) and [1](#) perform much better than the random user association.

Figure [3.6](#) shows the number of iterations (on average) required to reach convergence versus the number of users in the system. Stating the obvious, the number of iterations required to converge increases with the increase in the number of users. In other words, as the number of users increases, the algorithm require more iterations in order to converge to a solution. The figure numerically suggests a linear relationship between the number of users in the system and the number of required iterations for Algorithm [2](#) to converge. In addition, the figure indicates that even if the number of APs in the system increases (hence, the room dimensions increase as per Table [2.2](#)), it has no impact on the required number of iterations. However, this figure does not show the complexity within each it-

eration. It should not be difficult to notice that increasing N_u or the number of APs would require more computations within each iteration.

3.3.2 Hybrid CF-mMIMO with Clustering

In this subsection, the performance of the proposed system and algorithms along with the proposed clustering techniques are studied. It is worth to emphasize that despite the non-optimal selections of clustering techniques we adopt, they serve the purpose of the comparison held in this section. The main advantage of the proposed clustering techniques is that they are extremely simple to implement and do not require any complex computation or iterative search. Nonetheless, they achieve a satisfactory performance with the right choice of precoding schemes. Clustering parameters are presented in table 3.1

Table 3.1: Parameters for Clustering				
Clustering	$\mathbf{A}_1$		$\mathbf{A}_2$	
Network	VLC	RF	VLC	RF
d_{max}	4 m	6 m	—	—
N_{max}	—	—	3	5

Figure 3.7 shows the CDF of the proposed hybrid CF-mMIMO utilizing the first clustering technique $\mathbf{A}_1$ with three user association approaches, namely, Algorithm 2, Algorithm 1, and random association. Standalone CF systems were evaluated using the same clustering technique for comparison. In this figure, ZF precoding is implemented. We see the clear lead of the proposed hybrid system with both association algorithms over all other systems in the figure. It is worth

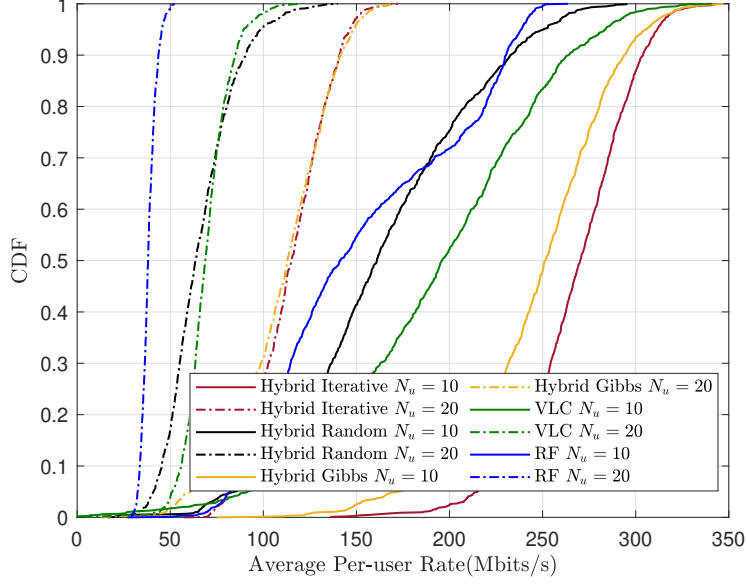


Figure 3.7: CDF of average per-user rate using $\mathbf{A}_1$ clustering

to notice that Algorithm [1](#) performs better than Algorithm [2](#), however, the performance gap between them is almost null when $N_u = 20$. It can be seen that using $\mathbf{A}_1$ for clustering reduces the sum-rate of all the systems in the figure. The drop in the performance is mostly due to the fact that each AP is limited to serve only users in d_{max} proximity, which removes the users that can receive some channel gain from any of the networks but are further than d_{max} from the AP, hence limiting the data rates.

The performance of the second clustering technique $\mathbf{A}_2$ using ZF precoding is shown in Figure [3.8](#). Multiple things can be concluded from this figure. The overall performance of the system using $\mathbf{A}_2$ clustering is better to the performance when using $\mathbf{A}_1$. The figure shows that with $\mathbf{A}_2$ clustering, less than 30% of the users experience an average rate of 300 Mbits/s, which is much better than that when using $\mathbf{A}_1$ where over 85% of the users receive an average rate of less than

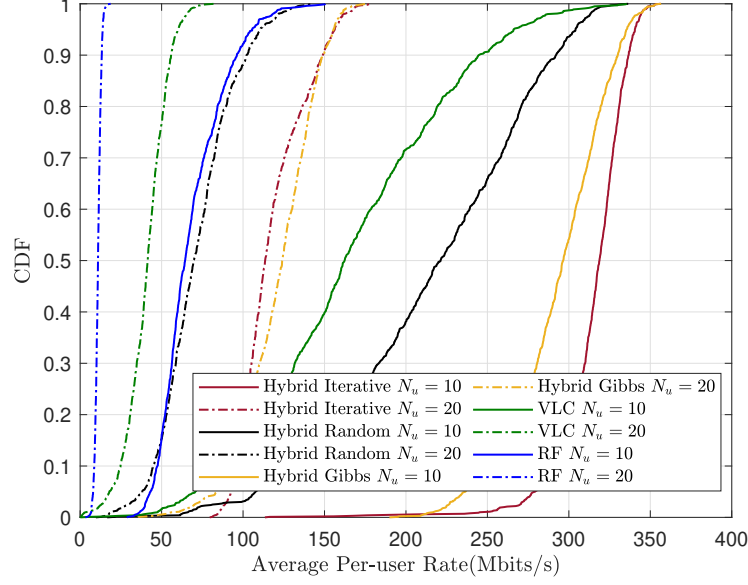


Figure 3.8: CDF of average per-user rate using $\mathbf{A}_2$ clustering

300 Mbits/s. Although one might notice that this specific clustering significantly lowers the performance of the VLC standalone CF system and kills the performance of the RF standalone CF system, it does not have such an impact on the hybrid system when using any of the proposed user association algorithms especially when the number of users is low. However, when the number of users is relatively high, $N_u = 20$ in this case, two things are noted; first, the performance when using $\mathbf{A}_2$ clustering is almost the same as when $\mathbf{A}_1$ clustering is used with a slight lead for the $\mathbf{A}_2$. Second, the performance of the gibbs sampling-based algorithm (Algorithm 2) is superior to that of Algorithm 1.

Figure 3.9 compares the two proposed algorithms in terms of the sum-rate they converge to with different number of users when $\mathbf{A}_1$ clustering is used. Figure 3.10 on the other hand presents the same comparison but with $\mathbf{A}_2$ clustering instead. Both figures show that regardless of what clustering technique is used,

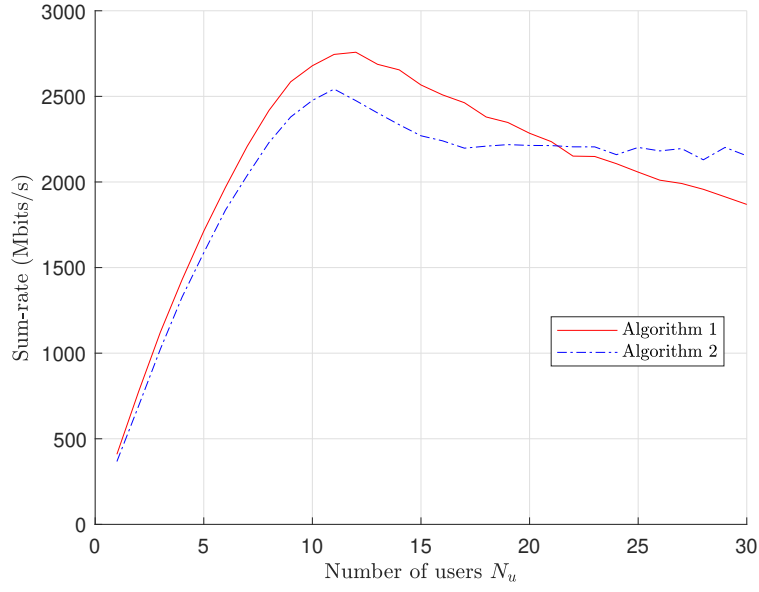


Figure 3.9: Sum-rate versus number of users under $\mathbf{A}_1$ clustering

it is better to adopt Algorithm 1 for user association when low number of users exists in the system. However, things turn around as the number of users increases above a certain number where Algorithm 2 becomes superior and its performance dominates. These figures suggest that according to the expected number of users in a given system/area we might select an algorithm to adopt.

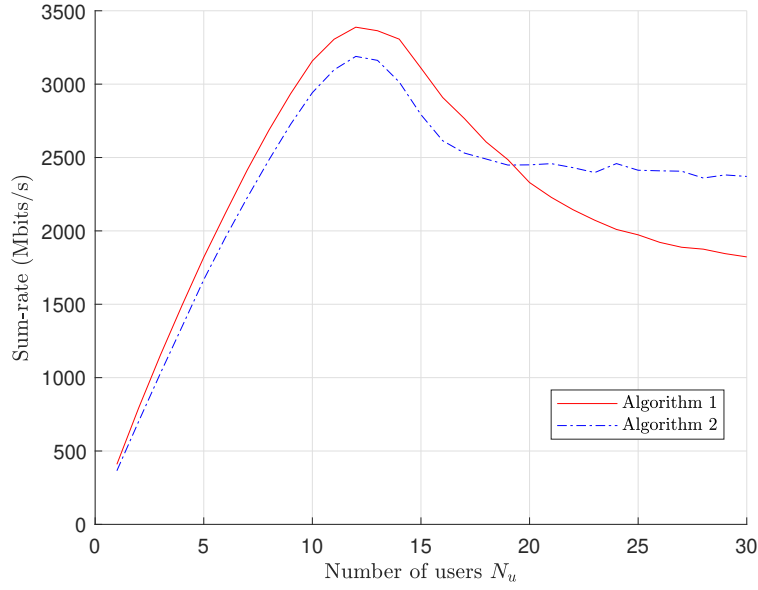


Figure 3.10: Sum-rate versus number of users under $\mathbf{A}_2$ clustering

3.4 Conclusion

In this chapter, a user-centric hybrid VLC/RF cell-free massive MIMO system has been introduced. User association and user clustering are two interlinked issues, and thus are formulated into a joint optimization problem that aims to maximize the sum-rate of the proposed system. The joint optimization problem is further simplified into two independent sub-problems, where each one is dealt with separately. A Gibbs sampling-based algorithm is proposed to find a sub-optimal solution for the user association sub-problem. This algorithm is compared to the iterative algorithm that was proposed in Chapter 2. Each algorithm showed high potential in certain cases. Two user clustering techniques are proposed and evaluated as well. The system outperformed other CF systems including a hy-

brid system with random user association, a VLC standalone system, and an RF standalone system.

CHAPTER 4

DEEP REINFORCEMENT LEARNING-BASED USER ASSOCIATION

Chapters 2 and 3 of this thesis studied the hybrid VLC/RF CF-mMIMO system with different user association algorithms and clustering techniques, and showed that they provide a significant enhancement compared to other systems. However, Algorithms 1 and 2 are iterative in nature and require a long time to run and converge, which is not exactly suitable for places with relatively fast varying channels. Meanwhile, in the field of data science, ML algorithms are known for their fast decision-making process. An ML model is usually trained in a process that might take several hours, or even days in some applications or model structures. However, a trained model can be fed with an appropriate input to predict an output in a very short inference time. This output prediction or decision mak-

ing is called inference, and ML algorithms are well known for their *near* instant inference.

In this chapter, we propose and develop a DRL framework to select the best network to serve each user in the hybrid VLC/RF CF-mMIMO system. That is, we solve the user association problem with the help of DRL. The user clustering problem provided in Chapter 3 is ignored here for reasons that will be discussed in the conclusion of this chapter.

First, we start by presenting a theoretical background of DRL and its basics as it is a relatively new field.

4.1 Theoretical Foundation of DRL

Machine Learning is a very broad field with many different algorithms. ML algorithms can be thought of as methods to optimize the performance of a variety of tasks by learning from a set of available data. All ML algorithms have a well mathematical structure and derivations, however, it is not the scope of this work to expand on that. Here, we are mostly interested in the DRL which is, in essence, a combination of both DL and RL. Therefore, we provide a brief background information on DL, RL, and DRL.

4.1.1 Deep Learning

Deep Learning is a sub-area of ML that is related to ANNs with more than a single hidden layer. An ANN is a brain-inspired ML algorithm that consists of layers

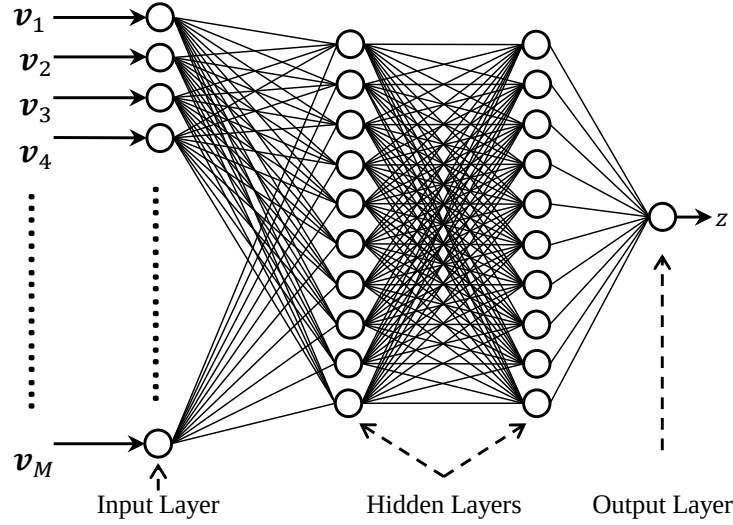


Figure 4.1: A structure for an ANN with two hidden layers with 10 neurons each. The figure is taken from [2]

of artificial neurons organized in a network structure similar to that provided in Figure 4.1. If an ANN contains more than one hidden layer it is called a deep neural network (DNN) which is then considered as a DL model. These DNNs are trained on data samples to estimate a model that fits these samples, which can then be used on any other set of similar data to provide estimations (or other tasks). This is generally useful when there is no direct equation that links the output to the input. To build a DNN model, it is required to have a training data set that consists of some optimal choices of data points that we need the system to learn from. The reader might have already guessed that this is of no use in practical wireless communication systems since having the optimal points means that we had a way to have reached it, which raises the question of why do we even need to use DL models in the first place.

Although that is partially true, there are two counter arguments that need

to be addressed. First, finding the optimal solutions sometimes requires a huge computational complexity, which is not always practical when a quick solution is needed. Thus, it could be beneficial to obtain the optimal solutions of only some data samples and then use these optimal solutions with their samples to train a DL that is then capable of providing a solution at any point in time. Second, the history of science and engineering showed us that things may not seem useful at first and then turn out to be powerful in the future with some developments (which is indeed the case here when the Deep Learning was involved in the Deep Reinforcement Learning, later in this section). Therefore, we dig a little bit in the Deep Learning or the neural networks.

A DNN is an ANN but with multiple hidden layers, each one holds a number of neurons. Neurons of different layers are connected via links that are called synapses which are usually called the weights. DNNs can come in different structures and types that are different from the basic structure provided in Figure 4.1, such as, convolutional neural networks (CNNs) and Recurrent neural networks (RNNs), which are beyond the scope of this work and will not be discussed.

To appreciate the working principle of giant DNNs, we first have a quick look on the operation of a single neuron. A neuron is a representation of an *activated* weighted sum, or a linear combination, of the features of its input layer (or the previous layer in a multi-layer network) as shown in Figure 4.2. The word ‘activated’ in the definition means that the weighted sum passes through an activation function $\varphi(\cdot)$ that activates the ability of the network to span other dimensions (non-linear dimensions). There is a large variety of activation functions, each is

suitable for certain applications, and all of them have a nonlinear relation between its input and its output.

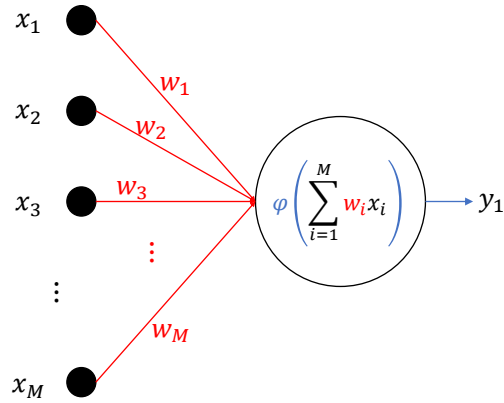


Figure 4.2: A single neuron operation

For instance, two well-known activation functions are the sigmoid activation function, which is given by

$$Sigmoid(x) = \sigma(x) = \frac{1}{1 + \exp(-x)}, \quad (4.1)$$

and the rectified linear unit (ReLU) that is given by

$$ReLU(x) = \max(0, x). \quad (4.2)$$

4.1.2 Reinforcement Learning

Reinforcement Learning is a sub-area of ML that deals with how something will act in a new environment given that it needs to reach a certain goal or maximize a specific function (or reward) via sequential decision making process. We can

think of RL as a technique that is inspired by the way human beings learn, i.e. by interacting with their surroundings. The sequential decision making process is crucial for achieving the maximum value of the reward and is then formalized using a Markov decision process (MDP). Therefore, MDPs are the basis for problems that are solved by RL. In this section, we explain the concept and components of an MDP, then we deviate to discuss the Q-learning, which is an algorithm that helps us solve for the optimal decision making policy.

Markov Decision Processes (MDPs):

There are five components that describe an MPD, an *agent*, an *environment*, a *state*, an *action*, and a *reward*. The *agent* is the decision maker. The *environment* is whatever the agent is interacting with. The *state* of the environment describes at what stage, level, or value it is in, and is brought to the agent and then the agent decides on what *action* within its set of all possible actions $\mathcal{A}_c$ is best to be chosen in order to maximize its *reward*. Once the agent selects its action based on the information it possesses, the environment transitions into a new state and delivers a reward value to the agent. This process occurs at each time step t , and the agent's goal, when choosing action, is to maximize the cumulative rewards over all time steps. Figure 4.3 shows the interaction between an agent and its environment. The expected return of rewards is the sum of future rewards. This is very useful since it is the goal of the agent to maximize the expected return. However, while this definition works fine when we have a limited number of time steps, it fails when we have a continuing task since the number of rewards

that needs to be added are infinite. This brings us to the definition of expected discounted return, and now this become the agent's objective to maximize. The expected discounted return can be written as

$$G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k}, \quad (4.3)$$

where G_t is the expected discounted return (expected return for short) at time step t , r_{t+k} is the reward at time step $t+k$, and γ is the discounting factor and is always non-negative and less than unity. This definition points out the fact that while the agent will always consider future rewards, it gives more weight (attention) to the immediate reward. We can further reformulate the equation as follows

$$\begin{aligned} G_t &= \sum_{k=0}^{\infty} \gamma^k r_{t+k}, \\ &= r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \gamma^3 r_{t+3} + \dots, \\ &= r_t + \gamma (r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots), \\ &= r_t + \gamma G_{t+1}, \end{aligned} \quad (4.4)$$

this shows that expected returns at different time steps are related to each other.

However, based on what would an agent select an action while it is in a certain state? and how good an action is under a given state of the environment? Here, we introduce the concepts of a *policy* and *value functions*, respectively. A policy is a function that provides probabilities for selecting actions given the current state of the environment. It is a common practice to denote the policy by the symbol π . Thus, π is a probability distribution over the set of all possible actions given a

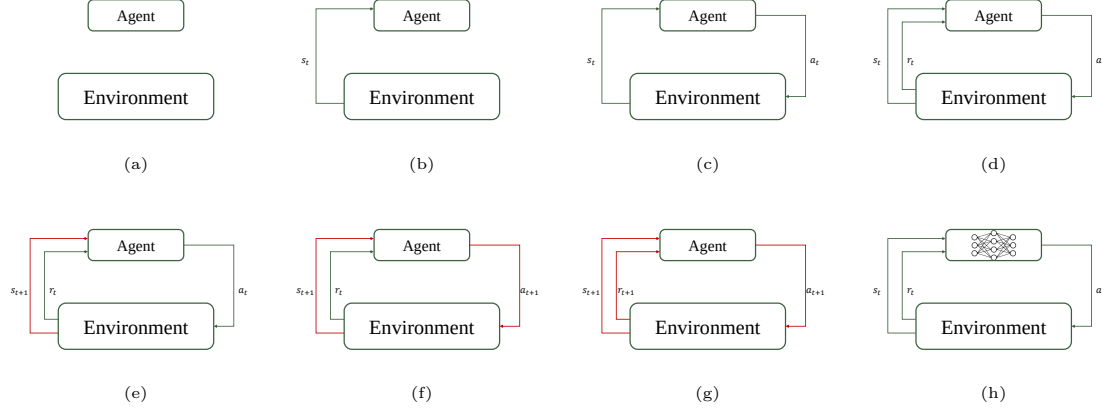


Figure 4.3: Steps of the interaction between an agent and its environment. Figure (h) represents a case of DRL instead of RL (the agent is a neural network).

specific state s . We can write the probability of selecting an action $a \in \mathcal{A}_c$ given the agent is in state s as $\pi(a|s)$.

Value functions tell us how good a state or a state-action pair is in terms of the expected return. However, we know that the expected return depends on the actions taken by the agent, and those actions are based on the policy π that the agent is following. Thus, value functions are defined for an agent with policy π . There are two definitions for value functions, state-value function and action-value functions. A state-value function provides the value of a state s for an agent that is following policy π and it is denoted by v_π . Whereas, an action-value function provides the value of an action taken by an agent (who is following policy π) given that the environment is at state s . It is denoted by q_π and is given as the expected return if an agent takes action a at time step t while it is at state s

$$q_\pi(s, a) = E_\pi [G_t | s_t = s, a_t = a] = E_\pi \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k} | s_t = s, a_t = a \right]. \quad (4.5)$$

This action-value function is conventionally known as the Q-function, and its

output as the Q-value. With that being established, it is now the goal of the Reinforcement Learning to find the policy that gives higher return than any other policy i.e. the *optimal policy*. Moreover, we say that a policy π is better than another policy π' , if the expected return from policy π is higher than that of π' , i.e.

$$\pi \geq \pi' \text{ if and only if } q_\pi(s, a) \geq q_{\pi'}(s, a), \forall s \in \mathcal{S}, a \in \mathcal{A}_c. \quad (4.6)$$

Associated with the optimal policy, we have an optimal action-value function (optimal Q-function), which provides the largest expected return for each state-action pair for an agent following policy π , and is given by

$$q_*(s, a) = \max_{\pi} q_\pi(s, a) \quad \forall s \in \mathcal{S}, a \in \mathcal{A}_c. \quad (4.7)$$

A fundamental and critical property of the optimal Q-function q_* is that it must satisfy the *Bellman optimality equation* that is given by

$$q_*(s, a) = E \left[r_t + \gamma \max_{a'} q_*(s', a') \right], \quad (4.8)$$

where s' is the next state of the environment reached by taking action a while the environment was in state s , and a' is the action that will be selected when the environment is currently at state s' if we follow the optimal policy. The Bellman equation states that the optimal Q-value of the current state action pair equals to the the expected return from that pair added to the discounted optimal Q-value of the next state-action pair, i.e. the maximum expected return.

Q-learning:

Q-learning is a Reinforcement Learning algorithm used to learn the optimal policy in an MDP. The essence of the Q-learning is to iteratively solve for the optimal q_* using the Bellman equation (4.8). Updating the Q-value aims at finding a q that is as close as possible to right hand side of Bellman equation, so that eventually it converges to the optimal q_* . In the action selection process, there is a trade-off between *exploration*, which means using random actions to explore the environment and check for possible better actions, and *exploitation*, which is utilizing the Q-values we have to select the best known action. For this trade-off, we use the Epsilon Greedy strategy which states that a random action will be chosen with a probability of ϵ and an action that have the maximum *known* corresponding Q-value will be chosen with probability of $(1 - \epsilon)$. A fairly common practice is to start with $\epsilon = 1$ since we do not have any information about the quality of any state-action pair, then we decrement the value of ϵ each time an action is taken to exploit the data that we have collected by the time.

We need to keep comparing the loss between the updated Q-value and the optimal Q-value and consequently update the Q-value to reduce the loss that is given by

$$\begin{aligned} loss &= q_*(s, a) - q(s, a), \\ &= E \left[r_t + \gamma \max_{a'} q_*(s', a') \right] - E \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k} \right]. \end{aligned} \tag{4.9}$$

In addition, in the learning process we can have a learning rate parameter α which allows us to control the influence of the learned value on the new Q-value. The

new Q-value is then calculated as

$$q^{new}(s, a) = (1 - \alpha) \overbrace{q(s, a)}^{\text{old value}} + \alpha \overbrace{\left(r_t + \gamma \max_{a'} q_*(s', a') \right)}^{\text{newly learned value}}. \quad (4.10)$$

Setting $\alpha = 1$ means that the new Q-value is the one we compute via the Bellman equation, which means throwing the old Q-value.

To summarize, the agent in a Reinforcement Learning learns a policy that maps each state to the best action with the goal of maximizing the expected return. However, learning this policy is nothing but learning the Q-values of all state-action pairs. In other words, an agent's policy can be represented by a table of states versus actions, and each entry of the table is the Q-value of the state-action pair that represent that cell.

4.1.3 Deep Reinforcement Learning

We can briefly define the Deep Reinforcement Learning as a Reinforcement Learning but with a DL-based agent. There are many algorithms that belong to the family of DRL, however, in this work, we only focus on the deep Q-learning (DQL).

Deep Q-Learning:

In problems where the action space is huge, the Q-table that represents the agent's policy become huge and updating it becomes computationally inefficient, which may lead to poor results and slow convergence. Thus, it is not recommended to update the Q-table in that same way, but it is preferred to estimate the Q-

table using a function estimator. A well-known function estimator is the ANNs discussed in Subsection 4.1.1. Therefore, the agent in the DQL employs a neural network, known as deep Q-network (DQN), that estimates the Q-values given a state and tries to reduce the loss and find the optimal Q-function¹. The DQN takes the state of the environment as input and outputs the Q-values of each action in the action space.

To help enhance the training, we use the *replay memory*. The replay memory is a buffer with finite capacity that stores some of the agent's experiences at each time step to use them later in the training. The agent's experience at time step t is given as the tuple

$$E_t = (s_t, a_t, r_t, s_{t+1}), \quad (4.11)$$

where s_t , a_t , r_t , and s_{t+1} represent the current state at time step t , the agent's action at time step t , the reward achieved when the state transitions from s_t as action a_t is performed, and the new state of the environment after applying a_t , respectively.

Training the DQN:

The process of training the DQN starts by initializing the weight of the DQN. Now for each time step in an episode², we explore/exploit the environment and select an action. Once an action is taken and executed, the environment returns a reward

¹The Q-function is another name for the Q-table, both of them represent a mapping between state-action pairs and the corresponding Q-values

²An episode is a full simulation length from the initial state to the final state or until convergence

to the agent and it transitions to a new state. The actual training then starts by sampling a batch from the replay memory and passing that batch through the DQN. The loss in (4.9) is then computed. The Q-values are computed by passing the state of the environment through the DQN which outputs the Q-value of each possible action. However, using that DQN to calculate both Q-values in (4.9) is impractical as updating the DQN's weights causes a change not only in the Q-values but in the target Q-values as well. An idea to overcome this issue is to introduce a second neural network that is responsible for the target Q-value in the loss i.e. $q_*(s, a)$. Specifically, this new separate network is meant to estimate the term " $\max_{a'} q_*(s', a')$ ". This network takes the next states from the sampled batch and estimates the Q-values of all possible actions to get the target Q-value in the loss expression. Thus, we call the original network the *policy network* since it approximates the policy, and this new network is called the *target network* since it is used to calculate the target Q-values. The weights of this target network are cloned from the policy network's weights every certain number of episodes.

4.2 User Association in Hybrid VLC/RF CF-mMIMO systems: A DRL Approach

After laying the foundation of the DRL, we now propose our own DRL framework to accomplish our goal of assigning users to one of the two networks, VLC or RF, in order to maximize the sum-rate of the system proposed in this thesis i.e. the hybrid VLC/RF CF-mMIMO system.

Let us first restate the optimization problem that we wish to solve

$$\max_{\mathbf{b}}. \quad \sum_j R_j \quad (4.12a)$$

$$\text{s.t.} \quad b_j \in \{0, 1\} \quad (4.12b)$$

$$\sum_{j=1}^{N_u} \left(w_{i,j}^{(v)} \right)^2 P_j^{(v)} \leq 1, \quad \forall i, \quad (4.12c)$$

$$\sum_{j=1}^{N_u} \left(w_{i,j}^{(r)} \right)^2 P_j^{(r)} \leq 1, \quad \forall i. \quad (4.12d)$$

which has already been mentioned in Chapter 2 and Chapter 3. In this part of the chapter we are only focused on proving that we can build a DRL framework that can solve this optimization problem; thus, we will only consider the case of a static CF-mMIMO system that was presented in Chapter 2, where the elements of the clustering matrices $\mathbf{A}^{(v)}$ and $\mathbf{A}^{(r)}$ are all ones. The constraints (4.12d) and (4.12d) are satisfied by allocating equal power to all users similar to what was described in Subsection 3.2.4. The following explains, in details, the design of our developed DRL framework that is built to solve (4.12).

4.2.1 Proposed DRL Framework

Before we present the details of the proposed framework, we draw the big picture of how the system behaves. Assume the same system model as in the previous chapters, we now try to assign users to one of the two networks using a DRL model. This system model, including the number of APs in each network and their locations in the room and all other parameters, is considered as our environment.

Table 4.1: DRL Parameters

Parameters	Value
Discounting factor γ	0.999
Number of episodes	10000
Time steps per episode	100
Initial exploration rate ϵ	1
Replay memory size	10000
Target network update period P_T	10 episodes

The Agent is the decision maker, which in our case is the responsible for deciding which one of the two networks serves each user; that is the CPU. An action is the user association vector. Thus, with a specific user distribution, the agent starts by randomly assigning users to one of the networks. Consequently, changes occur in the environment causing a change in its state, and a reward is then returned to the agent according to the selected action. This pattern continues as the agent is learning, and as it learns it starts to assign users to the networks based on the events that it experiences, i.e. the agent applies Epsilon-Greedy strategy. The details of the learning process are described in the previous section. All the required parameters for our DRL framework are presented in Table 4.1. In this subsection, we discuss the different elements of the proposed DRL framework and the design of its agent, states, actions, and reward.

Agent Design

The proposed DRL framework is based on the DQL algorithm where we design a DQN to estimate the Q-values of all state-action pairs. We design our DQN as three linear layers (a.k.a. fully connected layers) with an activation function

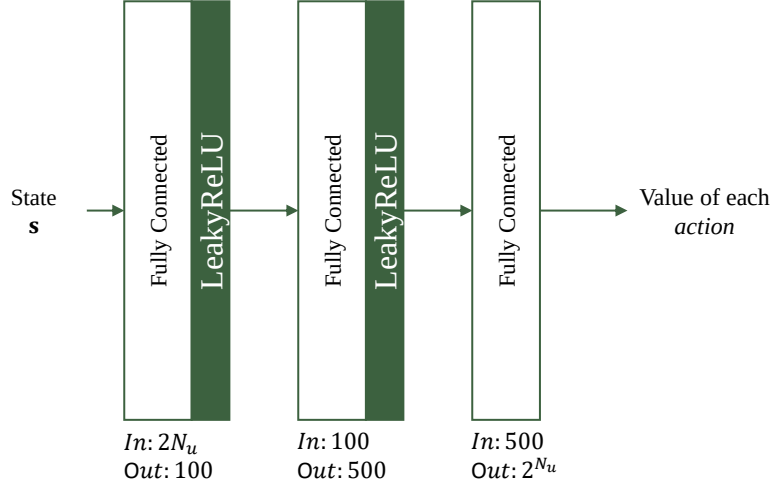


Figure 4.4: Agent's DQN structure

Table 4.2: DQN Structure

Input	$\mathbf{s} \in \mathbb{R}^{2N_u}$
Layer 1	Fully Connected (In: $2N_u$, Out: 100)
Activation	LeakyReLU
Layer 2	Fully Connected (In: 100, Out: 500)
Activation	LeakyReLU
Layer 3	Fully Connected (In: 500, Out: $2N_u$)

implemented after layers 1 and 2. The last layer is not followed by an activation function since it will provide us with the estimates of the Q-values, and we need these value with no changes applied to them. This structure is shown in Figure 4.4. The detailed structure of the network is provided in Table 4.2.

Action Space

The actions that the agent can take are the different combinations of users association to the networks. These are the different combinations of the elements b_j in the vector $\mathbf{b}$. In other words, in this work taking an action means selecting

a vector $\mathbf{b}$. We denote the set that contains all possible actions by $\mathcal{A}$, which we call the action space. Given that the number of users in the system is N_u , the number of possible actions is $|\mathcal{A}| = 2^{N_u}$. This is why the output layer of the agent's DQN contains 2^{N_u} neurons, each one outputs the Q-value of selecting that specific action given the state provided as input to the DQN.

State Space

As stated earlier, an agent takes an action based on the current state of the environment. In this context, the agent selects an association vector $\mathbf{b}$, we choose to make the SINRs³ of the users represent the state of the environment. This is done by first creating a vector $\mathbf{s}$ of length $2N_u$. This vector is a concatenation of two vectors each of length N_u . Each one of these two vectors contains the SINRs of all users in one network and are represented by $\mathbf{s}^{(v)}$ and $\mathbf{s}^{(r)}$ for the VLC and RF networks, respectively. This is shown in Equation 4.13.

³The term SINR is used broadly in this chapter to refer to the quantities in equations (2.5) and (2.9).

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}^{(v)} \\ \mathbf{s}^{(r)} \end{bmatrix} = \begin{bmatrix} SINR_1^{(v)} \\ SINR_2^{(v)} \\ \vdots \\ SINR_{N_u}^{(v)} \\ SINR_1^{(r)} \\ SINR_2^{(r)} \\ \vdots \\ SINR_{N_u}^{(r)} \end{bmatrix} = \begin{bmatrix} SINR_1^{(v)} \\ SINR_2^{(v)} \\ \dots \\ SINR_{N_u}^{(v)} \\ SINR_1^{(r)} \\ SINR_2^{(r)} \\ \dots \\ SINR_{N_u}^{(r)} \end{bmatrix} \Bigg\} 2N_u. \quad (4.13)$$

If a user is not connected to one network, the user's entry in the vector of that network is zero, which is shown in Equation 4.14. The set of all possible states in the system is denoted by $\mathcal{S}$, where $\mathbf{s} \in \mathcal{S}$. The values of SINR are continuous and thus there are infinite number of possible states, i.e. $|\mathcal{S}| = \infty$.

$$SINR_j^{(m)} = \begin{cases} \gamma_j^{(m)} & \text{if user } j \text{ is connected to network } m; \\ 0 & \text{otherwise.} \end{cases} \quad (4.14)$$

Rewards

After an agent monitors the state of the environment and selects an action it receives a reward that indicates whether the selected action was a good candidate action or not. Since we aim to maximize the sum-rate in equation (4.12a), we select the sum-rate of the system to indicate how good our action (or choice of user association vector) is. Therefore, the reward in our proposed framework is the system's sum-rate.

Table 4.3: DQN Hyperparameters

Hyperparameter	Value
Optimizer	Adam
DQN learning rate	0.001
Batch size	256

4.2.2 Training the Agent’s DQN

The elements presented in the previous subsection construct the DRL framework. We now train the DQN to select the best user association given the state of the environment, i.e. the SINR of each user. **Algorithm 3** describes the training procedure of the DQN. In the training we use another DQN as a target network to help in the training as discussed earlier. The chosen network’s hyperparameters are presented in Table 4.3.

4.3 Simulation Results

In this section, we evaluate the performance of the proposed DRL-based user association in the proposed hybrid VLC/RF CF-mMIMO system. Here, we implement ZF precoding as was done in Section 2.3. Table 2.1 contains the simulation parameters used to generate all the following results.

As the DQN network has a specific structure, once trained it can only accept inputs of the same size that it was trained with. In other words, if the DQN is trained on the existing N_u users in the room, it cannot (by itself) accommodate for a new entering user. Although not in the scope of this work, a possible solution for this is to use transfer learning process as proposed in [68]. Nonetheless, we

Algorithm 3 DQN Training Algorithm

```
1: Initialize policy network
2: Initialize replay memory
3: Initialize the environment
4: Clone policy network parameters to target network
5: Target network parameters  $\leftarrow$  policy network parameters
6: for  $episode \leftarrow 1$  to Number of episodes do
7:   for  $t \leftarrow 1$  to Number of time steps do
8:     Select an action  $a_t$  based on the epsilon-greedy strategy and the state
      of the environment and the policy network
9:     Execute the action and get the reward  $r_t$  and the next state  $s_{t+1}$ 
10:    Store the experience  $E_t = (s_t, a_t, r_t, s_{t+1})$  in the replay memory
11:    if Data in memory  $>$  batch size then
12:      Sample a batch from the memory
13:      Estimate current Q-values  $q(s, a)$  using the policy network
14:      Estimate the next Q-values using the target network
15:      Calculate the loss between the target Q-values and the current Q-
        values
16:      Update policy network parameters
17:      if  $\text{mod}(episode, P_T) = 0$  then
18:        Clone policy network parameters to target network
19:      end if
20:    end if
21:  end for
22: end for
```

perform a separate training process for different number of users.

In the following, we show the training curves of three different choices of the number of users. Specifically, the sum-rate versus the episode number in the training process when $N_u = 10$, $N_u = 14$, and $N_u = 18$ is presented in Figure 4.5, Figure 4.6, and Figure 4.7, respectively. In these figures, the black curves are the actual sum-rate achieved by the DRL, while the red curves are the filtered sum-rate. The filter used here is a moving average filter of length 100, which means that every point in the red curves represents the average of the last 100 points of

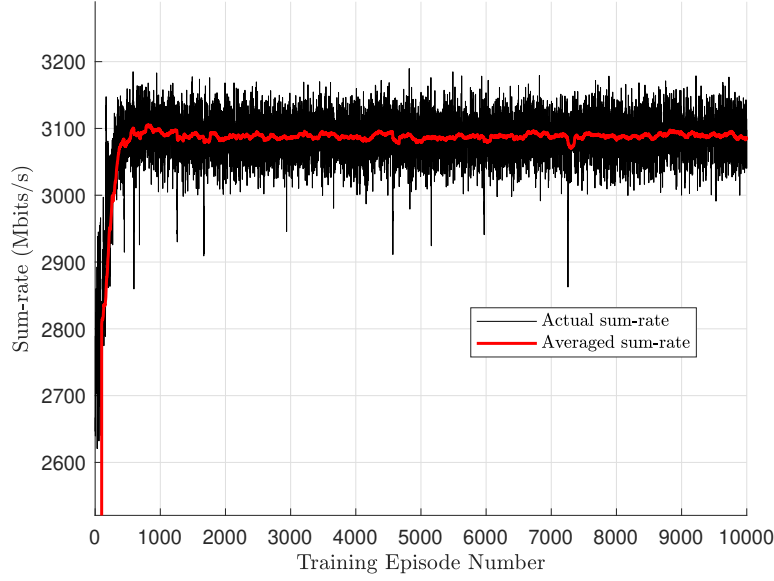


Figure 4.5: Performance of DRL-based user association for $N_u = 10$

the black curves. This moving average filter is given by

$$y[n] = \begin{cases} \frac{1}{100} \sum_{m=0}^{99} x[n-m], & n \geq 100; \\ 0, & otherwise, \end{cases} \quad (4.15)$$

where y is the output of the filter (red signal), and x is the filter's input (black signal).

It is important to compare the performance of the DRL-based user association with the classical association algorithms proposed in previous chapters. Thus, we plot a curve of the sum-rate vs the number of users for all of the association algorithms proposed in this thesis. Figure 4.8 presents this comparison for the hybrid system using the classical association algorithms proposed in Chapters 2 and 3 as well as the DRL-based user association proposed in this chapter. Note

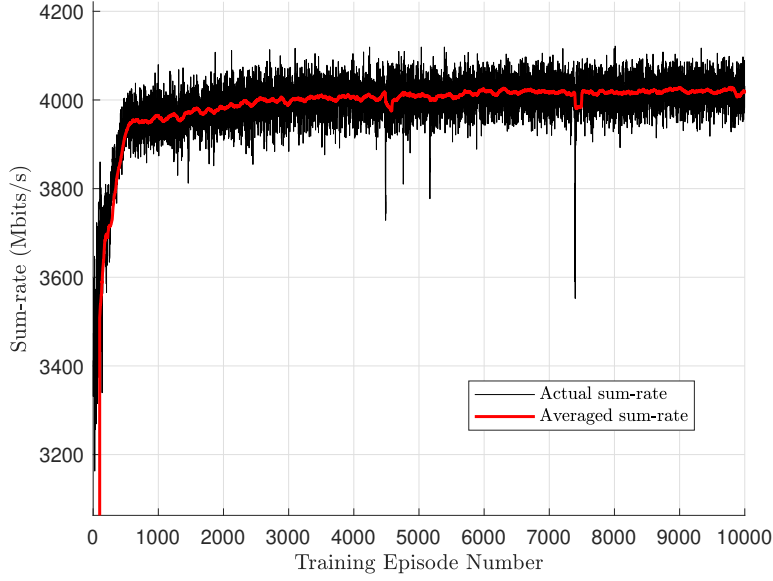


Figure 4.6: Performance of DRL-based user association for $N_u = 14$

that each point of the DRL curve is the mean of the performance of the last 1000 episode.

This comparison tells us that the DRL-based user association does indeed perform well even compared to the other proposed algorithms. In fact, with this specific framework and training conditions, it slightly outperforms the Gibbs sampling-based user association (Algorithm 2) as shown in Figure 4.8. There is an obvious lead when Algorithm 1 is used. However, in terms of sum-rate, the DRL-based approach performs pretty well. Moreover, it is important to notice that these trained DRL models have the potential to provide us with the association vector $\mathbf{b}$ instantly compared to the time it takes Algorithm 1 and 2 to provide their results.

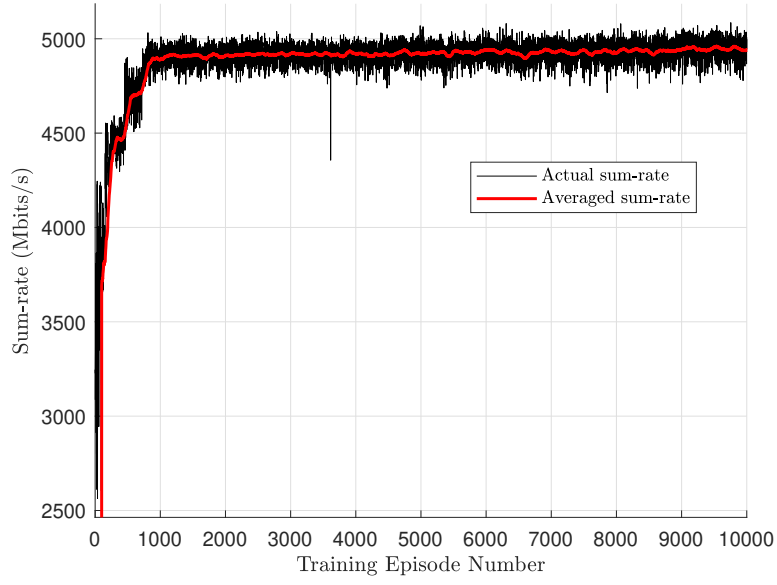


Figure 4.7: Performance of DRL-based user association for $N_u = 18$

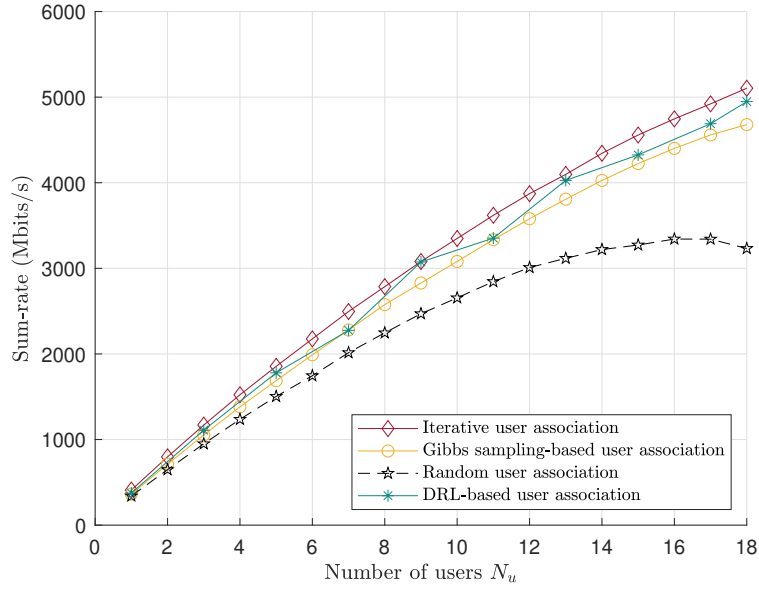


Figure 4.8: Sum-rate versus number of users using different association algorithms

4.4 Conclusion

In this chapter, a DRL framework was developed to solve the user association problem in the hybrid VLC/RF CF-mMIMO system proposed earlier in this thesis. We designed the agent, states, actions and rewards of the proposed framework. An algorithm to train the proposed agent's DQN is provided and the training results show great potential compared to the classical algorithms provided in previous chapters. DRL-based user association performed slightly better than the Gibbs sampling-based user association (Algorithm 2). However, Algorithm 1 performs the best amongst all of them. Nevertheless, it is noteworthy that the DRL-based user association has the advantage of providing an instant output once required, while Algorithms 1 and 2 take a noticeable amount of time to provide their output. Despite their advantages and performance, DRL models often consume a noticeable space in the random access memory (RAM) to generate, store, and use their neural networks. The size of these networks grow with the number of possible actions in the designed framework. In our proposed framework, the size of the action space is $|\mathcal{A}| = 2^{N_u}$. This indicates that, with the current setup, problems will arise as N_u increases. In fact, this is the main reason for ignoring the user clustering problem. If our framework was built with the intention of dealing with both user association and user clustering, i.e. to solve the joint problem in (3.18), the number of possible actions would be $|\mathcal{A}| = 2^{(N_u + N_{u,v} \times N_{vap} + N_{u,r} \times N_{rap})}$. This number is tremendously huge and can not be dealt with using the current setup.

CHAPTER 5

THESIS CONCLUSION AND FUTURE RESEARCH DIRECTIONS

5.1 Thesis Conclusion

In this Thesis, a hybrid VLC/RF CF-mMIMO system and a UC hybrid VLC/RF CF-mMIMO are proposed, analysed and evaluated in Chapter 2 and Chapter 3, respectively. Generic expressions for the achievable rate in both system are derived where any precoding scheme and any clustering scheme can be used in the final expressions. Two optimization problems are formulated, a user association optimization problem and a joint user association and user clustering optimization problem. Moreover, the joint optimization problem is simplified into a user association sub-problem and a user clustering sub-problem.

Three distinct approaches are proposed to solve the user association problem. An iterative search algorithm (Algorithm 1), a Gibbs sampling-based algorithm (Algorithm 2), and a DRL-based algorithm are proposed. The performance of the system using each of these algorithms is studied and numerical results are provided in their respective chapters.

Using Algorithm 1, we studied the sum-rate of the proposed system and how it increases with the number of users. We have also shown that the proposed system is resilient to blockage. The convergence of Algorithm 1 was studied as well.

Algorithm 2 was shown to provide a sub-optimal solution which is ever lower than what Algorithm 1 provides. Nonetheless, when clustering is introduced, the performance of Algorithm 2 dominates as the number of users in the system increases beyond a certain number.

The performance of the DRL-based association was shown to be around the performance of Algorithm 2, however, with the advantage of instant inference once required.

Moreover, a comparison between the three proposed algorithms is carried out for three different choices of the number of users in the system, which shows that Algorithm 1 performs the best.

5.2 Future Research Directions

There are several possible directions to continue and extend the work presented in this thesis, some of these ideas are mentioned in the following:

- In a static hybrid VLC/RF CF-mMIMO system, it is possible to relax the integer constraint on the user association variable. This allows the value of b_j to be in the range $[0, 1]$. Furthermore, with the assumption of ZF precoding with equal power distribution, the problem becomes convex and can be solved using Lagrange duality.
- In a static hybrid VLC/RF CF-mMIMO system, if we assume ZF precoding we can formulate a joint optimization problem of both user association and power allocation. This can be approached by relaxing the constraint on the user association variable and then jointly solved for both variables.
- In the UC hybrid VLC/RF CF-mMIMO system, we may try to approach the user clustering sup-problem with a heuristic approach, by first limiting the number of possible clustering combinations to lower the complexity of the algorithm.
- A DRL-based clustering technique may be formulated to perform the user clustering task. However, memory requirements issue has to be addressed, and a different framework needs to be developed to lower the possible number of actions and to increase the efficiency of the current implementation.

REFERENCES

- [1] M. Obeed, H. Dahrouj, A. M. Salhab, S. A. Zummo, and M.-S. Alouini, “DC-Bias and Power Allocation in Cooperative VLC Networks for Joint Information and Energy Transfer,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 5486–5499, 2019.
- [2] A. Almehdhar, A. Hamida, K. Aliyu, S. Alawsh, A. Muqaibel, S. Al-Dharrab, W. Mesbah, and G. L. Stüber, “DOA Estimation in Wireless Seismic Surveys Using Deep Learning,” in *2020 International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, 2021, pp. 1–6.
- [3] M. Obeed, A. M. Salhab, M. S. Alouini, and S. A. Zummo, “On Optimizing VLC Networks for Downlink Multi-User Transmission: A Survey,” *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2947–2976, 2019.
- [4] 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 Trans. Green Commun. Netw.*, vol. 3, no. 1, pp. 108–121, 2019.
- [5] M. Ayyash, H. Elgala, A. Khreishah, V. Jungnickel, T. Little, S. Shao, M. Rahaim, D. Schulz, J. Hilt, and R. Freund, “Coexistence of WiFi and LiFi to-

- ward 5G: concepts, opportunities, and challenges,” *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 64–71, 2016.
- [6] T. Rakia, H.-C. Yang, F. Gebali, and M.-S. Alouini, “Dual-Hop VLC/RF Transmission System with Energy Harvesting Relay under Delay Constraint,” in *2016 IEEE Globecom Workshops (GC Wkshps)*, 2016, pp. 1–6.
- [7] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and H. Yang, “Cell-Free Massive MIMO systems,” in *2015 49th Asilomar Conference on Signals, Systems and Computers*, 2015, pp. 695–699.
- [8] Y. Wang, D. A. Basnayaka, and H. Haas, “Dynamic load balancing for hybrid Li-Fi and RF indoor networks,” in *2015 IEEE International Conference on Communication Workshop (ICCW)*, 2015, pp. 1422–1427.
- [9] J. Kahn and J. Barry, “Wireless infrared communications,” *Proc. IEEE*, vol. 85, no. 2, pp. 265–298, 1997.
- [10] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, “Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications,” *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.
- [11] M. R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, and L. Ladid, “Internet of Things in the 5G Era: Enablers, Architecture, and Business Models,” *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 510–527, 2016.

- [12] D. Tsonev, S. Videv, and H. Haas, “Towards a 100 Gb/s visible light wireless access network,” *Opt. Express*, vol. 23, no. 2, pp. 1627–1637, Jan 2015. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-23-2-1627>
- [13] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, “What Will 5G Be?” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [14] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, “5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, 2017.
- [15] P. Pirinen, “A brief overview of 5G research activities,” in *1st International Conference on 5G for Ubiquitous Connectivity*, 2014, pp. 17–22.
- [16] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer, “Network densification: the dominant theme for wireless evolution into 5G,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 82–89, 2014.
- [17] X. Ge, S. Tu, G. Mao, C. Wang, and T. Han, “5G Ultra-Dense Cellular Networks,” *IEEE Wireless Commun.*, vol. 23, no. 1, pp. 72–79, 2016.
- [18] Y. Tanaka, S. Haruyama, and M. Nakagawa, “Wireless optical transmissions with white colored LED for wireless home links,” *11th IEEE Int. Symp. Pers.*

- Indoor Mob. Radio Commun. PIMRC. Proceedings (Cat. No.00TH8525)*, vol. 2, pp. 1325–1329 vol.2, 2000.
- [19] G. A. Mapunda, R. Ramogomana, L. Marata, B. Basutli, A. S. Khan, and J. M. Chuma, “Indoor Visible Light Communication: A Tutorial and Survey,” *Wirel. Commun. Mob. Comput.*, vol. 2020, pp. 8 881 305:1–8 881 305:46, 2020.
- [20] C. Chen, S. Videv, D. Tsonev, and H. Haas, “Fractional Frequency Reuse in DCO-OFDM-Based Optical Attocell Networks,” *Journal of Lightwave Technology*, vol. 33, no. 19, pp. 3986–4000, 2015.
- [21] M. Dehghani Soltani, X. Wu, M. Safari, and H. Haas, “Access point selection in Li-Fi cellular networks with arbitrary receiver orientation,” in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2016, pp. 1–6.
- [22] X. Wu, M. Safari, and H. Haas, “Bidirectional Allocation Game in Visible Light Communications,” in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, 2016, pp. 1–5.
- [23] M. Obeed, A. M. Salhab, S. A. Zummo, and M. Alouini, “Joint Power Allocation and Cell Formation for Energy-Efficient VLC Networks,” in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–6.
- [24] Y. Wang, X. Wu, and H. Haas, “Resource Allocation in LiFi OFDMA Systems,” in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6.

- [25] X. Ling, J. Wang, Z. Ding, C. Zhao, and X. Gao, “Efficient OFDMA for LiFi Downlink,” *Journal of Lightwave Technology*, vol. 36, no. 10, pp. 1928–1943, 2018.
- [26] H. Haas, L. Yin, Y. Wang, and C. Chen, “What is LiFi?” *Journal of Lightwave Technology*, vol. 34, no. 6, pp. 1533–1544, 2016.
- [27] I. Stefan and H. Haas, “Hybrid Visible Light and Radio Frequency Communication Systems,” in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, 2014, pp. 1–5.
- [28] Y. Wang, S. Videv, and H. Haas, “Dynamic load balancing with handover in hybrid Li-Fi and Wi-Fi networks,” in *2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*, 2014, pp. 575–579.
- [29] D. A. Basnayaka and H. Haas, “Hybrid RF and VLC Systems: Improving User Data Rate Performance of VLC Systems,” in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, 2015, pp. 1–5.
- [30] Y. Wang, D. A. Basnayaka, X. Wu, and H. Haas, “Optimization of Load Balancing in Hybrid LiFi/RF Networks,” *IEEE Transactions on Communications*, vol. 65, no. 4, pp. 1708–1720, 2017.
- [31] X. Wu, M. Safari, and H. Haas, “Joint Optimisation of Load Balancing and Handover for Hybrid LiFi and WiFi Networks,” in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017, pp. 1–5.

- [32] X. Wu and H. Haas, “Access point assignment in hybrid LiFi and WiFi networks in consideration of LiFi channel blockage,” in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2017, pp. 1–5.
- [33] X. Wu, M. Safari, and H. Haas, “Access Point Selection for Hybrid Li-Fi and Wi-Fi Networks,” *IEEE Transactions on Communications*, vol. 65, no. 12, pp. 5375–5385, 2017.
- [34] X. Wu and H. Haas, “Load Balancing for Hybrid LiFi and WiFi Networks: To Tackle User Mobility and Light-Path Blockage,” *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1675–1683, 2020.
- [35] M. Obeed, A. M. Salhab, S. A. Zummo, and M. Alouini, “Joint Load Balancing and Power Allocation for Hybrid VLC/RF Networks,” in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6.
- [36] —, “Joint optimization of power allocation and load balancing for hybrid VLC/RF networks,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 5, pp. 553–562, 2018.
- [37] V. K. Papanikolaou, P. D. Diamantoulakis, Z. Ding, S. Muhaidat, and G. K. Karagiannidis, “Hybrid VLC/RF Networks with Non-Orthogonal Multiple Access,” in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–6.

- [38] M. Obeed, H. Dahrouj, A. M. Salhab, S. A. Zummo, and M. S. Alouini, "User Pairing, Link Selection, and Power Allocation for Cooperative NOMA Hybrid VLC/RF Systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 1785–1800, 2021.
- [39] M. Obeed, H. Dahrouj, A. M. Salhab, A. Chaaban, S. A. Zummo, and M. S. Alouini, "Power Allocation and Link Selection for Multicell Cooperative NOMA Hybrid VLC/RF Systems," *IEEE Communications Letters*, vol. 25, no. 2, pp. 560–564, 2021.
- [40] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-Free Massive MIMO: Uniformly great service for everyone," in *IEEE Workshop Signal Process. Adv. Wirel. Commun. SPAWC*, 2015, pp. 201–205.
- [41] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and B. D. Rao, "Performance of cell-free massive MIMO systems with MMSE and LSFD receivers," in *2016 50th Asilomar Conference on Signals, Systems and Computers*, 2016, pp. 203–207.
- [42] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Pre-coding and Power Optimization in Cell-Free Massive MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4445–4459, 2017.
- [43] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-Free Massive MIMO Versus Small Cells," *EEE Trans. Wireless Commun.*,

- vol. 16, no. 3, pp. 1834–1850, 2017.
- [44] S. Buzzi and C. D’Andrea, “Cell-Free Massive MIMO: User-Centric Approach,” *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 706–709, 2017.
 - [45] G. Interdonato, P. Frenger, and E. G. Larsson, “Scalability Aspects of Cell-Free Massive MIMO,” in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, 2019, pp. 1–6.
 - [46] S. Buzzi, C. D’Andrea, and C. D’Elia, “User-Centric Cell-Free Massive MIMO with Interference Cancellation and Local ZF Downlink Precoding,” in *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, 2018, pp. 1–5.
 - [47] E. Björnson and L. Sanguinetti, “A New Look at Cell-Free Massive MIMO: Making It Practical With Dynamic Cooperation,” in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2019, pp. 1–6.
 - [48] G. Interdonato, M. Karlsson, E. Björnson, and E. G. Larsson, “Local Partial Zero-Forcing Precoding for Cell-Free Massive MIMO,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 7, pp. 4758–4774, 2020.
 - [49] E. Björnson and L. Sanguinetti, “Scalable Cell-Free Massive MIMO Systems,” *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.

- [50] R. Jiang, Q. Wang, H. Haas, and Z. Wang, “Joint User Association and Power Allocation for Cell-Free Visible Light Communication Networks,” *IEEE J. Sel. Areas Commun.*, vol. 36, no. 1, pp. 136–148, 2018.
- [51] J. Chen, Z. Wang, and R. Jiang, “Downlink interference management in cell-free vlc network,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 9007–9017, 2019.
- [52] J. Beysens, A. Galisteo, Q. Wang, D. Juara, D. Giustiniano, and S. Pollin, “DenseVLC: A Cell-Free Massive MIMO System with Distributed LEDs,” in *Proc. 14th Int. Conf. Emerg. Netw. EXp. Technol.*, ser. CoNEXT ’18. New York, NY, USA: Association for Computing Machinery, 2018, p. 320–332. [Online]. Available: <https://doi.org/10.1145/3281411.3281423>
- [53] J. Beysens, Q. Wang, A. Galisteo, D. Giustiniano, and S. Pollin, “A Cell-Free Networking System With Visible Light,” *IEEE/ACM Trans. Netw.*, vol. 28, no. 2, pp. 461–476, 2020.
- [54] A. Basu and B. Bhattacharyya, “A Study on the Integration of Machine Learning in Wireless Communication,” in *2018 International Conference on Communication and Signal Processing (ICCSP)*, 2018, pp. 0974–0978.
- [55] Y. Al-Eryani, M. Akroun, and E. Hossain, “Multiple Access in Cell-Free Networks: Outage Performance, Dynamic Clustering, and Deep Reinforcement Learning-Based Design,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1028–1042, 2021.

- [56] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *Engineering Applications of Neural Networks*, C. Jayne and L. Iliadis, Eds. Cham: Springer International Publishing, 2016, pp. 213–226.
- [57] N. E. West and T. O'Shea, "Deep architectures for modulation recognition," in *2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2017, pp. 1–6.
- [58] Y. S. Nasir and D. Guo, "Multi-Agent Deep Reinforcement Learning for Dynamic Power Allocation in Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2239–2250, 2019.
- [59] A. Zappone, L. Sanguinetti, and M. Debbah, "User Association and Load Balancing for Massive MIMO through Deep Learning," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, 2018, pp. 1262–1266.
- [60] L. Sanguinetti, A. Zappone, and M. Debbah, "Deep Learning Power Allocation in Massive MIMO," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, 2018, pp. 1257–1261.
- [61] E. Ghadimi, F. Davide Calabrese, G. Peters, and P. Soldati, "A reinforcement learning approach to power control and rate adaptation in cellular networks," in *2017 IEEE International Conference on Communications (ICC)*, 2017, pp. 1–7.

- [62] R. Ahmad, M. D. Soltani, M. Safari, and A. Srivastava, "Load Balancing of Hybrid LiFi WiFi Networks Using Reinforcement learning," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, 2020, pp. 1–6.
- [63] R. Ahmad, M. D. Soltani, M. Safari, A. Srivastava, and A. Das, "Reinforcement Learning Based Load Balancing for Hybrid LiFi WiFi Networks," *IEEE Access*, vol. 8, pp. 132 273–132 284, 2020.
- [64] Y. Zhang, C. Kang, T. Ma, Y. Teng, and D. Guo, "Power Allocation in Multi-Cell Networks Using Deep Reinforcement Learning," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, 2018, pp. 1–6.
- [65] S. Saeidian, S. Tayamon, and E. Ghadimi, "Downlink Power Control in Dense 5G Radio Access Networks Through Deep Reinforcement Learning," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.
- [66] S. Wang, H. Liu, P. H. Gomes, and B. Krishnamachari, "Deep Reinforcement Learning for Dynamic Multichannel Access in Wireless Networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 2, pp. 257–265, 2018.
- [67] L. Xiao, H. Zhang, Y. Xiao, X. Wan, S. Liu, L. Wang, and H. V. Poor, "Reinforcement Learning-Based Downlink Interference Control for Ultra-Dense

- Small Cells,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 423–434, 2020.
- [68] S. Shrivastava, B. Chen, C. Chen, H. Wang, and M. Dai, “Deep Q-Network Learning Based Downlink Resource Allocation for Hybrid RF/VLC Systems,” *IEEE Access*, vol. 8, pp. 149 412–149 434, 2020.
- [69] C. Qiu, Y. Hu, Y. Chen, and B. Zeng, “Deep Deterministic Policy Gradient (DDPG)-Based Energy Harvesting Wireless Communications,” *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8577–8588, 2019.
- [70] J. J. G. Luis, M. Guerster, I. del Portillo, E. Crawley, and B. Cameron, “Deep Reinforcement Learning for Continuous Power Allocation in Flexible High Throughput Satellites,” in *2019 IEEE Cognitive Communications for Aerospace Applications Workshop (CCA AW)*, 2019, pp. 1–4.
- [71] O. Simeone, “A Very Brief Introduction to Machine Learning With Applications to Communication Systems,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 648–664, 2018.
- [72] Z. Qin, H. Ye, G. Y. Li, and B. F. Juang, “Deep Learning in Physical Layer Communications,” *IEEE Wireless Communications*, vol. 26, no. 2, pp. 93–99, 2019.
- [73] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y. Liang, and D. I. Kim, “Applications of Deep Reinforcement Learning in Communications and

- Networking: A Survey,” *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3133–3174, 2019.
- [74] A. Feriani and E. Hossain, “Single and Multi-Agent Deep Reinforcement Learning for AI-Enabled Wireless Networks: A Tutorial,” 2020.
- [75] Y. L. Lee and D. Qin, “A Survey on Applications of Deep Reinforcement Learning in Resource Management for 5G Heterogeneous Networks,” in *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2019, pp. 1856–1862.
- [76] A. Chaaban, Z. Rezki, and M.-S. Alouini, “On the Capacity of the Intensity-Modulation Direct-Detection Optical Broadcast Channel,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3114–3130, 2016.
- [77] L. P. Qian, Y. J. A. Zhang, and M. Chiang, “Distributed Nonconvex Power Control using Gibbs Sampling,” *IEEE Transactions on Communications*, vol. 60, no. 12, pp. 3886–3898, 2012.
- [78] H. Liu, X. Yuan, and Y.-J. A. Zhang, “Reconfigurable Intelligent Surface Enabled Federated Learning: A Unified Communication-Learning Design Approach,” *IEEE Transactions on Wireless Communications*, pp. 1–1, 2021.
- [79] Özlem Tugfe Demir, E. Björnson, and L. Sanguinetti, “Foundations of User-Centric Cell-Free Massive MIMO,” *Foundations and Trends® in Signal Processing*, vol. 14, no. 3-4, pp. 162–472, 2021. [Online]. Available: <http://dx.doi.org/10.1561/2000000109>

VITAE

- Name: Ahmed Fuad Almehdhar
- Nationality: Yemen
- Date of Birth: 13 July 1995
- Email: *ahmed1521416@hotmail.com*
- Permenant Address: JEDDAH, KSA. ZIP code: 23341

Education

- B.S. in Electrical Engineering, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia in 2019.
- M.Sc. in Electrical Engineering, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia in 2021.

Publications

- Ahmed Almehdhar, Mohanad Obeed, Anas Chaaban, and Salam Zummo, “A Hybrid VLC/RF Cell-Free Massive MIMO System”, accepted to *IEEE International Conference on Communication 2022 (ICC 2022)*

- A. Almehdhar, A. Hamida et al., “DOA Estimation in Wireless Seismic Surveys Using Deep Learning,” in *2020 International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, 2021, pp. 1-6, doi: 10.1109/ICCSPA49915.2021.9385716.

