



Adaptive beamforming and user association in heterogeneous cloud radio access networks: A mobility-aware performance-cost trade-off



Duc Thang Ha^{a,*}, Lila Boukhatem^a, Megumi Kaneko^b, Nhan Nguyen-Thanh^a, Steven Martin^a

^aLRI Laboratory, CNRS - Univ. Paris-Saclay - Univ. Paris-Sud, Orsay, France

^bNational Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, 101-8430 Tokyo, Japan

ARTICLE INFO

Article history:

Received 12 August 2018

Revised 25 March 2019

Accepted 12 May 2019

Available online 22 May 2019

Keywords:

H-CRAN
Beamforming
Clustering
User-to-RRH association
CSI Overhead

ABSTRACT

Heterogeneous Cloud Radio Access Network (H-CRAN) is a promising network architecture for the future 5G mobile communication system to address the increasing demand for mobile data traffic. In this work, we consider the design of efficient joint beamforming and user clustering (user-to-Remote Radio Head (RRH) association) in the downlink of a H-CRAN where users have different mobility profiles. Given the rapidly time-varying nature of such wireless environment, it becomes very challenging to enable optimized beamforming and user clustering without incurring large Channel State Information (CSI) and signaling overheads. The main objective of this work is to investigate and evaluate the trade-off between system throughput and the incurred costs in terms of complexity and signaling overhead, including the impact of different CSI feedback strategies given different user mobility profiles. We propose the Adaptive Beamforming and User Clustering (ABUC) algorithm which adapts its feedback parameters, namely the period of dynamic user clustering and the type of CSI feedback, in function of user mobility. Furthermore, we design a reinforcement-learning framework which enables the proposed ABUC algorithm to optimize its scheduling parameters on-the-fly, given each user mobility profile. Based on computer simulations, an analysis of the effect of mobility on system performance metrics is presented and conclusions are drawn regarding the algorithm's adequate parameter tuning for different mobility scenarios.¹

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Next generation of mobile and wireless communications system (5G) will revolutionize the way people communicate and extend the boundaries of the wireless industry. 5G will move beyond networks that are purpose-built for mobile broadband alone, toward systems that connect far more different types of devices at different speeds. The Internet of Things (IoT) is one of the primary contributors to global mobile traffic growth and this progression will lead to a huge mobile and wireless traffic volume predicted to increase a thousand-fold over the next decade [2]. Besides sustaining the tremendous growth of the traffic load, 5G system will be designed to fulfill diverse application requirements: far more stringent latency and reliability levels are expected to be necessary to

support applications related to healthcare, security, logistics, automotive applications, or mission-critical control; Network scalability and flexibility are required to support a large number of devices with very low complexity and to enable long battery lifetimes [3].

5G system is envisioned to meet such challenges thanks to the combination of several breakthroughs and technological advances such as ultra-dense small-cell deployments, intelligent multi-antenna, full duplex radios, millimeter wave transmissions, and cloud computing abilities. Particularly, the Cloud Radio Access Network (CRAN) is a network architecture based on cloud computing and centralized processing. It has been shown to provide high spectral and energy efficiencies while reducing both capital and operating expenditures [4]. At the same time, Heterogeneous Networks (HetNets) have emerged as another core feature for 5G network to enhance the capacity/coverage while saving energy consumption. HetNets are constituted by conventional macro cells and overlaying small cells. With small cells deployment, wireless links to end-users become shorter, thereby improving the link quality in terms of spectrum efficiency as well as energy efficiency. Therefore, combining both cloud computing and HetNet advantages results in the so-called Heterogeneous-Cloud Radio Access Networks

* Corresponding author.

E-mail addresses: duc-thang.ha@u-psud.fr, duc_thang.ha@lri.fr (D.T. Ha), lila.boukhatem@lri.fr (L. Boukhatem), megkaneko@nii.ac.jp (M. Kaneko), nhan.nguyen@lri.fr (N. Nguyen-Thanh), steven.martin@lri.fr (S. Martin).

¹ This is an extended version of the conference paper presented at IEEE PIMRC 2017 [1].

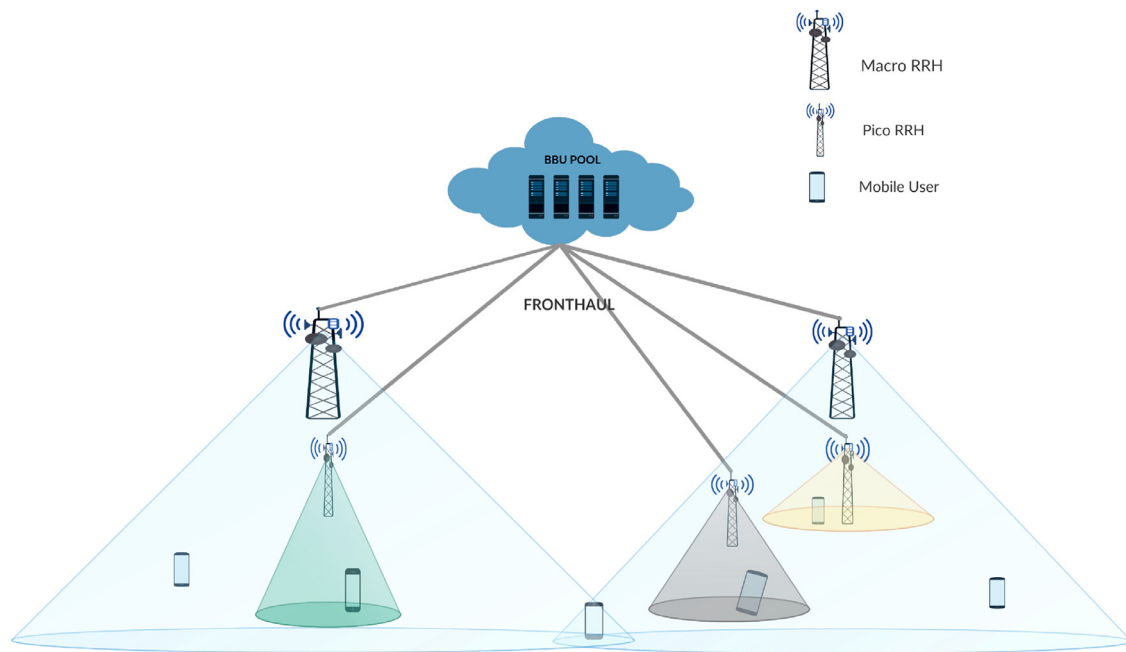


Fig. 1. H-CRAN system model.

(H-CRAN) depicted in Fig. 1 and which is regarded as one of the possible network architectures to meet 5G system requirements [5].

The technical challenges of 5G H-CRAN have been pointed out in a number of works, in particular regarding resource allocation, interference management and fronthaul constraint alleviation [6–8]. Particularly, the heterogeneous feature of access points in H-CRAN generalizes the problem of beamforming and user-to-RRH association compared to its CRAN counterpart. Many studies solved the scheduling problem by jointly optimizing the beamforming and user clustering in order to maximize network performance such as sum-rate, spectral efficiency and energy efficiency, etc. [9–11,11,12]. However, these solutions generate a large amount of control signaling and Channel State Information (CSI) overhead. Furthermore, most of these works did not consider the influence of user mobility and resulting time-varying wireless environment over the long-term scheduling performance.

To alleviate the problem of control signaling and CSI overhead costs, we have proposed in [1] a hybrid user clustering and beamforming algorithm aiming at weighted sum-rate maximization. This hybrid scheme is able to leverage the advantages of both dynamic and static user clusterings in CRAN, where the dynamic clustering performs optimally at the expense of maximum signaling overhead, while static clustering performs worse but drastically reduces the amount of overhead. The proposed hybrid algorithm was shown to achieve a good performance compared to dynamic clustering, while greatly reducing the required computational complexity, amount of CSI feedback and re-association signaling overhead over the long-term allocation process. However, in [1] we did not consider any user mobility issues.

In [13], we have considered the impact of more realistic channel variations due to different user velocities, but for a different optimization problem. Namely, we proposed a heuristic algorithm for minimizing a cost function in terms of computational complexity and CSI overhead for a given targeted sum-rate (minimum Quality-of-Service (QoS) requirement). However, the proposed algorithm required an initial empirical analysis for its feedback parameters' selection, and all users were assumed to have a homogeneous mobility profile.

Therefore in this paper, we investigate cost-efficient joint beamforming and user clustering methods for weighted sum-rate maximization in the downlink of a H-CRAN serving mobile users. More specifically, we investigate and evaluate the trade-off between network sum-rate and incurred costs in terms of complexity and signaling overhead, including the impact of different Channel State Information (CSI) feedback strategies. We propose an Adaptive Beamforming and User Clustering (ABUC) algorithm which extends the algorithm in [1] to cope with different types of user mobilities. The proposed algorithm is shown to be able to balance between the optimality of the beamforming and association solutions while being aware of practical system constraints, namely complexity and signaling overhead as well as the mobility behavior of users. Moreover, we identify the best feedback parameters such as type of CSI feedback and clustering period T , depending on the class of user mobility. Furthermore, we propose a reinforcement learning framework based on the Q-learning method for optimizing these feedback parameters on-the-fly, according to each user mobility profile. Indeed, reinforcement learning is especially suited for dealing with such intricate long-term resource allocation optimization, under dynamically varying wireless environments. Simulation results show the effectiveness of our proposed algorithm and approach, as it enables to select the best feedback parameters tailored to each user mobility profile, even in the difficult case where each user has a different mobility profile.

The rest of this paper is structured as follows: Section II presents the state of the art of CRAN, H-CRAN and most relevant works related to the beamforming and user scheduling issue. Section III describes the system model and Section IV gives the problem formulation and reference schemes. Next, we describe in Section V our proposed ABUC scheme and conduct its cost analysis in section VI. Section VII presents the Q-learning based clustering and beamforming framework. The simulation results are shown in section VIII. Finally, Section IX concludes the paper.

2. State of the art

To achieve the expected high performance, 5G system will rely on several advanced technologies such as heterogeneous small

cell deployment, millimeter wave communications, massive Multiple Input Multiple Output (MIMO), Network Function Virtualization (NFV), Software-Defined Networking (SDN), Device-to-Device (D2D) communications, and cloud computing concept [14].

The concept of CRAN was first proposed in [15] and described in details in [16]. In a CRAN system, all Baseband Units (BBUs) are shifted into the cloud to constitute a centralized processing pool. The radio frequency signals from geographically distributed users are collected by Remote Radio Heads (RRHs) and transmitted to the cloud platform through fronthaul links. This means that the system is able to adapt to non-uniform traffic and utilizes the resources more efficiently. Due to that fact that fewer BBUs are needed in CRAN compared to the traditional architecture, CRAN has also the potential to decrease the cost of network operation by reducing power and energy consumption. New BBUs can be added and upgraded easily, thereby improving scalability and easing network maintenance [17].

Different from CRANs, H-CRANs are proposed as a cost-efficient solution by incorporating the cloud computing into HetNets. The motivation behind H-CRANs is to enhance the capabilities of High Power Nodes (HPN, e.g., macro or micro base stations) with massive multiple antenna techniques and simplify the Low Power Nodes (LPN, e.g., small cells such as pico or femto cells) while connecting them to a “baseband signal processing cloud” with high speed optical fibers or RF (Radio Frequency) links [18]. As such, the baseband data processing as well as the radio resource control for LPNs are moved to the cloud server so as to take advantage of the cloud computing capabilities [6]. Similarly with the traditional CRAN, H-CRANs include a large number of RRHs with low energy consumption and which are coordinated with each other through the centralized BBU pool to achieve high cooperative gains.

Many scientific challenges ahead for H-CRANs may be listed as resource allocation optimization, interference management and fronthaul constraint alleviation. In [7], resource allocation solutions for H-CRAN are investigated, and different schemes are proposed, namely, coordinated scheduling and multicloud association. In the coordinated scheduling, the main issue is to maximize the network-wide utility subject to user connectivity constraints. The search space is exponentially large and makes such exhaustive search clearly infeasible. Then, the authors proposed a graph-theoretical-based approach to solve the problem. In addition, multicloud H-CRAN is proposed to overcome the main limitations of single cloud RAN such as the distance separating BSs in the network and also the computation burden when connecting multiple BSs to the same cloud. In [8], the authors provided a framework for downlink resource allocation for D2D communications underlying H-CRAN to maximize the system performance while guaranteeing the QoS requirements. The resource allocation problem is formulated into a many-to-one matching game and is solved by constrained Deterministic Annealing (DA) algorithm to achieve low computational complexity. Another critical challenge in H-CRAN that may prevent the RRHs from fully utilizing available radio resources is the insufficient backhaul capacity. Therefore, efficient planning of the H-CRAN is important for its RRH deployment. Load balancing has been noticed as an efficient way to optimize factors such as resource utilization, fairness, waiting/processing delays, or throughput [19]. In [20], the authors proposed a novel method that allocates almost equal traffic load to each access point to minimize the number of activated RRHs and reduce the burden of backhauls in the C-RAN. The RRHs deployment strategy is able to minimize the system-wide power consumption while providing QoS-guaranteed performance with relative low capital expenditure (CAPEX) and operating expenditure (OPEX).

In H-CRANs architecture, users can benefit from the coverage diversity provided by several heterogeneous nodes leading to a user-centric architecture [7]. Determining the optimal user-to-RRH

association that offers the best performance becomes a combinatorial optimization problem of high complexity. Exhaustive search is infeasible for any reasonably sized network, even with very powerful processors at the cloud. Consequently, this problem has been addressed in many research works [21–23]. The problem is equivalent to the joint beamforming and user-to-RRH clustering problem which by solving the sparse beamforming issue indirectly decides which users should be served by which RRHs.

In [24], the authors studied the user clustering problem and evaluated the appropriate number of associated RRHs per user to balance throughput gain and implementation cost. In [9], the advantages of small cells clustering are evaluated in a dense heterogeneous network for downlink MIMO. It is shown that by giving reasonable cluster sizes, each cluster can form a virtual MIMO network wherein users are separated via spatial multiplexing using jointly designed downlink beamforming vectors. In [10,11], a Weighted Minimum Mean Square Error (WMMSE) method is used for solving joint beamforming and user clustering problems for sum-rate maximization. In [10], the authors implemented a greedy RRH clustering algorithm and compared it to two transmit precoding schemes, Zero-Forcing Beamforming (ZFBF) and WMMSE-based coordinated beamforming. They showed that WMMSE outperforms other reference beamforming approaches. In [11], the Weighted Sum-Rate (WSR) maximization problem is solved using similar WMMSE approach under dynamic and static user clusterings. It is shown that dynamic clustering significantly improves the system performance over other naive clustering schemes, while static clustering also achieves substantial performance gain. Due to the lack of convergence guarantee for the algorithm in [11], the same authors propose a new algorithm in [25] which is proved to converge to a local optimum. Different to previous works, whose aim is to solve the weighted sum rate maximization problem with backhaul constraints in CRANs, the paper [12] applies the generalized WMMSE approach to solve the average weighted energy efficiency (EE) utility objective function with each RRH's transmit power, individual fronthaul capacity, and inter-tier interference constraints. The majority of these previous works did not fully consider the required computational complexity and the incurred signaling costs over a long-term scheduling process. In particular, such beamforming techniques rely on accurate CSI feedback for all user-to-RRH links, which may create excessive burden on fronthaul links. Moreover, most of these works did not consider the cost in terms of computational complexity at the BBU side and user-to-RRH re-association cost during successive scheduling frames. To the best of our knowledge, there are very limited works considering the time-dimension and the induced complexity and signaling costs in their user clustering and beamforming design problems.

Unlike the above-mentioned works that presented algorithms optimizing a certain objective, such as power consumption and sum-rate, for the current time frame, some recent works began to consider the scheduling process in the long-term. Moreover, the scheduling in mobility environment is a substantial problem due to its strong stochastic characteristic and the non-deterministic variables. Therefore, machine learning is known as an effective solution for resource allocation in time-variant dynamic systems, such as wireless networks and cloud computing systems [26].

Reinforcement learning is an important branch of machine learning, in which an agent makes interactions with an environment trying to control the environment to its optimal states that receive the maximal rewards. Usually in reinforcement learning, the problem to resolve can be described as a Markov Decision Process (MDP) without mandatory requirement of state space, explicit transition probability and reward function [27]. Therefore, reinforcement learning is expected to handle tough situations that approach real-world complexity [28]. In [29], a reinforcement learning technique is applied to dynamic resource allocation in CRAN.

The authors present a Deep Reinforcement Learning (DRL) framework which is able to reach the objective of minimizing power consumption and meeting demands of wireless users over a long operational period. The model defines the state space, action space, and reward function for the DRL agent. The proposed framework not only achieves significant power savings while satisfying user demands but also well handles highly dynamic cases. The authors of AIQerm and Shihada [30] propose a centralized resource allocation scheme using online Q-learning, which guarantees interference mitigation and maximizes energy efficiency while maintaining QoS requirements for all users in 5G H-CRAN. Their simulation results confirm that the proposed Q-learning solution can mitigate interference, increase energy and spectral efficiencies significantly, and maintain users' QoS requirements. In [26], the authors consider the problem of cache-enabled opportunistic interference alignment (IA) in wireless networks. The finite-state Markov channel is used instead of block-fading channel or invariant channel which are not realistic in high mobility environments. To reduce the system complexity of finite-state Markov channel, the authors formulated the system as a DRL problem. Simulation results show that DRL is an effective approach to solve the optimization problem in cache-enabled opportunistic IA wireless networks. In [31], a DRL based algorithm is proposed for the joint mode selection and resource management problem in Fog Radio Access Network (Fog-RAN). Each user can operate either in C-RAN mode or in device-to-device mode, and the resource managed includes both radio and computation resources. The DRL agent makes smart decisions on user communication modes and processors' on-off states to minimize long-term system power consumption while considering the varying states of edge caches.

3. System model

We consider a H-CRAN model which consists of a BBU Pool, L macro and pico RRHs and K users (see Fig. 1). Each RRH and user are equipped with M and N antennas, respectively, and users are randomly located in the network area.

Let $\mathbb{L} = \{1, 2, \dots, L\}$ and $\mathbb{K} = \{1, 2, \dots, K\}$ be the sets of RRHs and of users, respectively. The propagation channel from all RRHs to the k th user is represented by matrix $\mathbf{H}_k \in \mathbb{C}^{N \times ML}$, $\forall k \in \mathbb{K}$ which includes the impacts of path loss and Rayleigh fading. We denote by h_{nq} , the (n, q) th element of matrix \mathbf{H}_k , where $q = (l-1)M + m$, where $l \in [1, L]$ and $m \in [1, M]$. Hence, h_{nq} is the channel gain between the m th antenna of the l th RRH and the n th antenna of user k . Given the user mobility profile, channel correlations will be assumed between the consecutive scheduling frames, as detailed later in Section 5.

Let $\mathbf{w}_k \in \mathbb{C}^{ML \times 1}$ be the transmit beamforming vector from all RRHs to the k th user,

$$\mathbf{w}_k = [\mathbf{w}_{1k}^H, \dots, \mathbf{w}_{lk}^H, \dots, \mathbf{w}_{Lk}^H]^H,$$

where $\mathbf{w}_{lk} \in \mathbb{C}^{M \times 1}$.

Let $s_k \in \mathbb{C}$ be the encoded information symbol for user k with $\mathbb{E}[|s_k|^2] = 1$.

The received signal at user k , $\mathbf{y}_k \in \mathbb{C}^{N \times 1}$, is expressed as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{w}_k s_k + \sum_{j=1, j \neq k}^K \mathbf{H}_k \mathbf{w}_j s_j + \mathbf{n}_k,$$

where $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma_k^2 \mathbf{I}_N)$ is the additive white Gaussian noise and \mathbf{I}_N is the identity matrix of size $N \times N$.

The Signal-to-Interference-plus-Noise Ratio (SINR) at user k can be expressed as

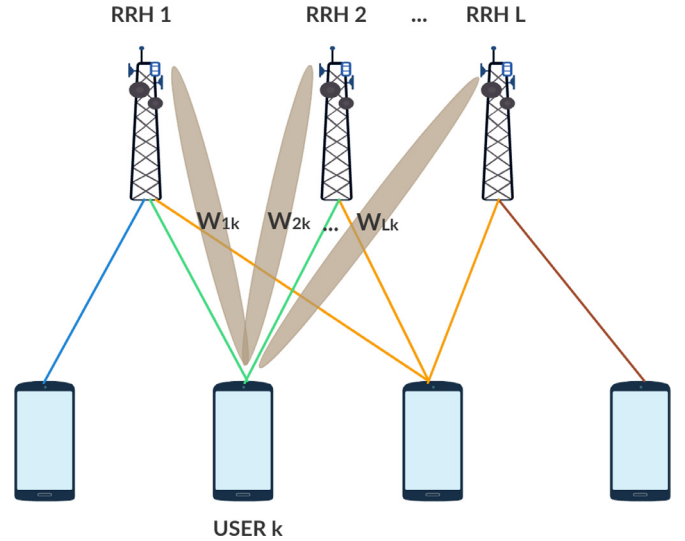


Fig. 2. System model.

$$\text{SINR}_k = \frac{|\mathbf{u}_k^H \mathbf{H}_k \mathbf{w}_k|^2}{\sum_{j=1, j \neq k}^K |\mathbf{u}_k^H \mathbf{H}_j \mathbf{w}_j|^2 + \sigma_k^2 \|\mathbf{u}_k\|_2^2}, \quad (1)$$

where $\mathbf{u}_k \in \mathbb{C}^{N \times 1}$ is the receive beamforming vector of user k .

Then, the achievable rate of user k under MMSE criterion can be expressed as [25]

$$r_k = \log_2 \left(1 + \mathbf{w}_k^H \mathbf{H}_k^H \left(\sum_{j=1, j \neq k}^K \mathbf{H}_k \mathbf{w}_j \mathbf{w}_j^H \mathbf{H}_k^H + \sigma_k^2 \mathbf{I}_N \right)^{-1} \mathbf{H}_k \mathbf{w}_k \right). \quad (2)$$

Fig. 2 illustrates an example of a user k clustering configuration with user-to-RRH associations determined by the derived \mathbf{w}_{lk} beamforming vector variables.

4. Problem formulation and reference schemes

We focus on the following WSR maximization problem [25], i.e., maximize the WSR of all users in the network under the fronthaul link capacity constraints and per-RRH power constraints. This problem is formulated as,

$$\max_{\{\mathbf{w}_{lk}, l \in \mathbb{L}, k \in \mathbb{K}\}} \sum_{k=1}^K \alpha_k r_k \quad (1a)$$

$$\text{s.t. } P_l = \sum_{k=1}^K \|\mathbf{w}_{lk}\|_2^2 \leq P_l^{\max} \quad (1b) \quad (\text{P1})$$

$$\sum_{k=1}^K \mathbb{1}\{\|\mathbf{w}_{lk}\|_2^2\} r_k \leq C_l^{\max} \quad (1c)$$

where α_k is the scheduling priority weight associated with user k . In (P1), the first constraint (1b) corresponds to the transmit power constraint of RRH l , i.e., P_l should be smaller than the maximum transmit power P_l^{\max} . The second constraint (1c) expresses that the sum-rate of users connected to RRH l should be smaller than its fronthaul link capacity C_l^{\max} . Problem (P1) is a non-convex Mixed-Integer Non-Linear Programming (MINLP) proven to be NP-hard [25], and hence cannot be solved in polynomial time.

It was shown in [10,11] that this WSR optimization problem is equivalent to a WMMSE problem with variables \mathbf{w}_k , \mathbf{u}_k and ρ_k defined below while being convex with respect to each of the variables. This enables to resolve the WMMSE problem through the block coordinate descent method by iteratively optimizing over

each variable [32] while keeping the others fixed. For finding the optimal beamformer \mathbf{w}_k under fixed \mathbf{u}_k and ρ_k , the following Quadratically Constrained Quadratic Programming (QCQP) problem is considered,

$$\begin{aligned} \min_{\{\mathbf{w}_k, l \in \mathbb{L}_k, k \in \mathbb{K}\}} & \sum_k \mathbf{w}_k^H (\sum_j \alpha_j \rho_j \mathbf{H}_j^H \mathbf{u}_j \mathbf{u}_j^H \mathbf{H}_j) \mathbf{w}_k \\ & - 2 \sum_k \alpha_k \rho_k \text{Re}\{\mathbf{u}_k^H \mathbf{H}_k \mathbf{w}_k\} \quad (2a) \\ \text{s.t.} & \sum_{k=1}^K \|\mathbf{w}_{lk}\|_2^2 \leq P_l^{\max} \quad (2b) \\ & \sum_{k=1}^K \beta_{lk} \hat{r}_k \|\mathbf{w}_{lk}\|_2^2 \leq C_l^{\max} \quad (2c) \end{aligned} \quad (\text{P2})$$

In QCQP problem (P2), \hat{r}_k is the achievable rate from the previous iteration and β_{lk} is a constant weight associated to RRH l and user k and is updated iteratively according to

$$\beta_{lk} = \frac{1}{\|\mathbf{w}_{lk}\|_2^2 + \tau}, \forall k, l, \quad (3)$$

where τ is a small constant regularization factor. e_k is the corresponding Mean Square Error (MSE),

$$e_k = \mathbf{u}_k^H \left(\sum_{j=1, j \neq k}^K \mathbf{H}_k \mathbf{w}_j \mathbf{w}_j^H \mathbf{H}_k^H + \sigma_k^2 \mathbf{I}_N \right) \mathbf{u}_k - 2 \text{Re}\{\mathbf{u}_k^H \mathbf{H}_k \mathbf{w}_k\} + 1, \quad (4)$$

and ρ_k is the MSE weight for user k ,

$$\rho_k = e_k^{-1}. \quad (5)$$

\mathbf{u}_k is the optimal receive beamforming vector under fixed \mathbf{w}_k and ρ_k ,

$$\mathbf{u}_k = \left(\sum_{j=1, j \neq k}^K \mathbf{H}_k \mathbf{w}_j \mathbf{w}_j^H \mathbf{H}_k^H + \sigma_k^2 \mathbf{I}_N \right)^{-1} \mathbf{H}_k \mathbf{w}_k. \quad (6)$$

Finally, the derived problem (P2) can be resolved using the Algorithm 1 given below [32].

Algorithm 1: Dynamic Algorithm.

```

initialize frame  $\beta_{lk}, \hat{r}_k, \mathbf{w}_k, \forall l, k$ 
repeat
  1) Fix  $\mathbf{w}_k$  and compute the MMSE receiver  $\mathbf{u}_k$  and the
  corresponding MSE  $e_k$  according to (8) and (6)
  2) Update MSE weight  $\rho_k$  according to (7)
  3) Find the optimal transmit beamformer  $\mathbf{w}_k$  under fixed
   $\mathbf{u}_k$  and  $\rho_k$ , by solving problem (P2)
  4) Compute the achievable rate  $r_k$ 
  5) Update  $\hat{r}_k = r_k$  and  $\beta_{kl}$  according to (5)
until convergence

```

To achieve the optimal performance, Algorithm 1 solves the beamforming problem by considering all possible user to RRH links, i.e., \mathbf{w}_k is of size LM in each scheduling frame, for each user k . The beamforming solution implicitly resolves the clustering problem since the user to RRH associations are updated according to the solution \mathbf{w}_k , i.e., the assignment of each link is identified by a non-zero element of the beamforming vector. This case, where \mathbf{w}_k is of size LM , is referred as the dynamic clustering algorithm. However, this global optimization requires intractable computational complexity and tremendous amounts of signaling and CSI overheads. Therefore, an alternative approach is to limit the computational burden by reducing the considered user to RRH links,

i.e., \mathbf{w}_k would be of size L_k where $L_k \leq L$ for each user k , at the expense of reduced sum-rate performance. This approach is referred as the static clustering algorithm.

Different to dynamic algorithm, in static scheduling, we consider only a fixed subset of RRHs in each cluster, i.e. $l \in \mathbb{L}_k$, where \mathbb{L}_k is the fixed cluster of RRHs serving user k . Likewise, we define \mathbb{K}_l as the subset of users associated with RRH l . The WSR maximization problem (P1) can now be re-formulated as

$$\begin{aligned} \max_{\{\mathbf{w}_k, l \in \mathbb{L}_k, k \in \mathbb{K}\}} & \sum_{k=1}^K \alpha_k r_k \quad (3a) \\ \text{s.t.} & P_l = \sum_{k \in \mathbb{K}_l} \|\mathbf{w}_{lk}\|_2^2 \leq P_l^{\max} \quad (3b) \\ & \sum_{k \in \mathbb{K}_l} r_k \leq C_l^{\max} \quad (3c) \end{aligned} \quad (\text{P3})$$

We can see that problem (P3) is much simplified as compared to problem (P1), as the constraints (3b) and (3c) consider only a fixed subset of users \mathbb{K}_l , while the variable \mathbf{w}_{lk} covers only the beamforming vectors from a subset of RRHs to each user since $\mathbf{w}_{lk} = 0$ for $l \notin \mathbb{L}_k$. Therefore, problem (P3) can be solved by applying a method similar to that of dynamic algorithm but under fixed \mathbb{L}_k instead of \mathbb{L} .

The WSR problem under static clustering can be resolved by Algorithm 2 given below [32]. The variables $\mathbf{w}_k^{\mathbb{L}_k}$ and $\mathbf{H}_k^{\mathbb{L}_k}$ denote the beamforming vector and the channel matrix to user k from the RRHs of its fixed cluster \mathbb{L}_k , respectively. They have the same sizes as \mathbf{w}_k and \mathbf{H}_k , respectively, but only their elements corresponding to RRHs within their cluster \mathbb{L}_k are non-zeros.

Algorithm 2: Static Algorithm.

```

initialize frame  $\mathbb{L}_k, \beta_k, \hat{r}_k, \mathbf{w}_k, \forall k$ 
repeat
  1) Compute (6), (7), (8) by replacing  $\mathbf{w}_k$  and  $\mathbf{H}_k$  by  $\mathbf{w}_k^{\mathbb{L}_k}$ 
  and  $\mathbf{H}_k^{\mathbb{L}_k}$ , respectively
  2) Fix  $\mathbb{L}_k$  during the whole process
  3) Call Dynamic Algorithm to solve (P3) under fixed  $\mathbb{L}_k$ 
until convergence

```

5. Proposed adaptive beamforming and user clustering (ABUC) algorithm

ABUC is based on the hybrid clustering concept that we introduced in [1] and in which we proposed to alternate dynamic and static clustering approaches in a periodic way. We define a period T and apply the dynamic algorithm at each T scheduled frames as illustrated in Fig. 3, while in the intermediate frames, a static algorithm is executed using the cluster subsets obtained from the previous dynamic frame. By doing so, we can narrow down the performance gap with the optimal dynamic solution, while reducing the required computational complexity, CSI feedback and re-association signaling overhead over the long-term allocation process. It is worth noting here that in the static intermediate frames the optimal beamforming vectors are updated upon receiving new CSI feedback information.

Our approach has the benefit to consider the temporal dimension of the allocation process, while being aware of the practical feasibility of the solution in terms of complexity and signaling costs. The motivation behind this approach lies in the fact that performing the optimal dynamic algorithm in each frame is not only computationally expensive but also unnecessary whenever the channel, user and network states stay more or less stable over few successive frames.

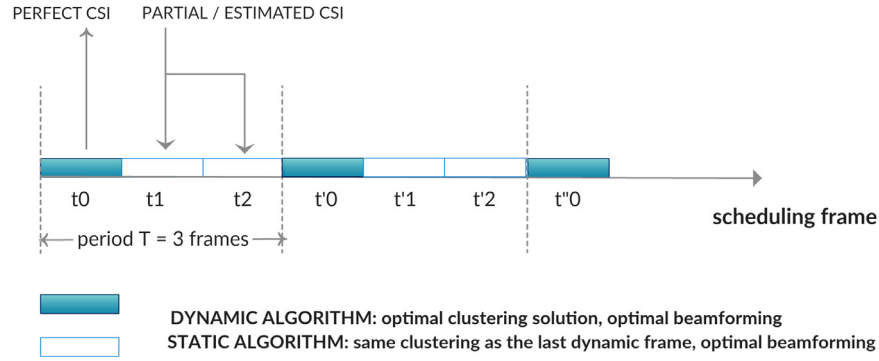


Fig. 3. ABUC clustering scheme with $T = 3$.

The goal of ABUC algorithm is to enable a high weighted sum-rate performance, while reducing the induced costs in terms of CSI feedback and signaling overhead, given channel time-variations. Compared to Thang et al. [1], in this paper we extend our algorithm to cope with different types of user mobility profiles. In consequence, the CSI estimation method is also improved such that it is tailored to time-varying channels reflecting user mobility.

User mobility causing inevitably CSI imperfectness at the cloud, we make use of a CSI estimation model which is aware of user mobility. The proposed approach is based on [33] for modeling the Estimated CSI matrix $\hat{\mathbf{H}}_k \in \mathbb{C}^{N \times ML}$, $\forall k \in \mathbb{K}$. We denote by \hat{h}_{nq} , the (n, q) th element of matrix $\hat{\mathbf{H}}_k$, where $q = (l-1)M + m$. Hence, \hat{h}_{nq} is the estimated channel gain between the m th antenna of the l th RRH and the n th antenna of user k , which is estimated as

$$\hat{h}_{nq} = \lambda h_{nq} + \left(\sqrt{1 - \lambda^2}\right) v_{nq}. \quad (7)$$

In (7), $v_{nq} \sim \mathcal{CN}(0, F_{lk})$ where F_{lk} is the large-scale fading gain of the downlink channel from RRH l to user k , and λ is the correlation coefficient between \hat{h}_{nq} and h_{nq} which is expressed as

$$\lambda = J_0(2\pi f_{d,ik} T_{dl}) \quad (8)$$

where $J_0(\cdot)$ is the (zero)th order Bessel function, T_{dl} is the fronthaul delay of the RRH l and $f_{d,ik}$ is the maximum Doppler frequency of the channel between the RRH l and user k . If the user moves at speed v (m/s), then the maximum Doppler frequency is calculated as $f_d = \frac{vf}{c}$, where f is the carrier frequency in Hertz and c is the speed of light. Therefore, we can express λ as function of v ,

$$\lambda_v = J_0\left(\frac{2\pi f T_{dl} v}{c}\right). \quad (9)$$

In ABUC, we consider different CSI feedback strategies during the scheduling process, namely: Full CSI, Partial CSI and Estimated CSI which are defined as follows:

- *Full CSI*: CSIs are assumed to be perfect at each frame and can be fully used in all scheduling frames.
- *Partial CSI*: CSIs are only known and perfect at dynamic frames (each period T), while at intermediate frames the channel gains are set equal to the last Full CSI received in the previous dynamic frame.
- *Estimated CSI*: CSIs are only known and perfect at dynamic frames (each period T) and at each intermediate frame CSIs are estimated according to the model given by Eq. (7).

Note that unlike Partial and Estimated CSI approaches, the Full CSI assumption is impractical in real systems due to the fronthaul delays and the signaling burden, which justifies our assumption of imperfect CSI knowledge at the BBU pool.

Combining these CSI feedback strategies and the periodicity T results into the following variants of the proposed ABUC algorithm:

- $(T, \text{Full CSI})$: Full CSI feedback every frame for beamforming; dynamic clustering optimisation every T frames,
- $(T, \text{Partial CSI})$: CSI feedback every T frames, reused in all intermediate frames for beamforming; dynamic clustering optimisation every T frames,
- $(T, \text{Estimated CSI})$: CSI feedback every T frames, estimated for all intermediate frames for beamforming; dynamic clustering optimisation every T frames.

The detailed description of the proposed ABUC algorithm is given in Algorithm 3.

Algorithm 3: Proposed ABUC Scheme with different types of CSI.

```

initialize frame  $t = 0$ , user velocity  $v$ 
repeat
  if  $t \bmod T = 0$  then
    Get perfect CSI  $\mathbf{H}_k(t)$  for all users  $k$ 
    Call Dynamic Algorithm
  else
    if Full CSI then
      Get perfect CSI  $\mathbf{H}_k(t)$  for all users  $k$ 
    else if Partial CSI then
      Use imperfect CSI  $\hat{\mathbf{H}}_k(t) = \mathbf{H}_k(t - \text{mod}(t, T))$ , for all users  $k$ 
    else
      Estimate CSI  $\hat{\mathbf{H}}_k(t)$  following (12) for all users  $k$ 
      Call Static Algorithm
  Set clustering solution as the initial clusters for frame  $t+1$ 
  Move to next frame
until convergence

```

6. Cost analysis of the proposed ABUC algorithm

In this section, we elaborate a cost analysis of ABUC algorithm in terms of the previously mentioned cost parameters.

6.1. Signaling costs

6.1.1. CSI overhead

As aforementioned, we consider three cases for CSI feedback for the proposed ABUC algorithm. In the Full CSI case, the CSI is fed back by every user to each of their serving RRH for every frame. In the Partial CSI case, CSIs are returned only for dynamic frames, i.e., every T frames, and used for all successive intermediate static frames. In Estimated CSI case, CSIs are estimated from the Full CSIs by a correlation coefficient presented in Eq. (7). Based on these

considerations, we can easily derive the amount of CSI overhead for each period T in each of the following cases to be considered in the numerical evaluations:

- dynamic algorithm with Full CSI:

$$O_{dyn}^f = KLMNT, \quad (10)$$

- ABUC algorithm with Full CSI:

$$O_{ABUC}^f = \sum_k (L + (T - 1)L_k)MN, \quad (11)$$

- ABUC algorithm with Partial CSI and Estimated CSI:

$$O_{ABUC}^{p,e} = KLMN. \quad (12)$$

6.1.2. User-to-RRH re-association overhead

Signaling overhead is generally neglected in most of the related works, although it may be a serious issue in practical systems, especially for the dynamic clustering strategy as it continuously updates the user-to-RRH associations for each scheduling frame. In this work, we consider the total number of signaling messages generated by the re-association process following the derived beamforming and clustering solutions. More specifically, we denote by S this signaling cost. Re-associations are counted for all newly established user-to-RRH links, or for released links between two successive frames. A simple calculation of the re-association cost S averaged over all frames l is given as follows:

$$S = \frac{1}{l} \sum_i \sum_k S_k^i \quad (13)$$

where S_k^i is the re-association cost of user k at frame i and which can be formulated as

$$S_k^i = (|\{L_k^i \cup L_k^{i-1}\}| - |\{L_k^i \cap L_k^{i-1}\}|), \quad (14)$$

where L_k^i is the number of RRHs serving user k at frame i .

6.2. Computational complexity cost analysis

As in [11], we assume a typical network model where $K > L > M > N$. In the proposed hybrid algorithm, the complexity required for dynamic frames is given as $O(K^4 L^3 M^3)$ [11]. For static frames, the complexity of the algorithm is dominated by problem (P2). By solving the problem via interior point method, the complexity of this problem will be similar to that of a linear program. The complexity in practice is in order of $n^2 m$ (assuming $m \geq n$), in which n is the dimension of the solution and m is the dimension of the constraint. In our problem, the total number of variables \mathbf{w}_k and R_k is $(\sum_{k=1}^K L_k M + K)$, while the largest dimension of constraints is given by (2c), i.e., $K \sum_{k=1}^K L_k M$. Thus, the complexity of static frames may be expressed as

$$O \left(\left(\sum_{k=1}^K L_k M + K \right)^2 \left(K \sum_{k=1}^K L_k M \right) \right), \quad (15)$$

whose dominating term is given by

$$O \left(K^3 \sum_{k=1}^K L_k M \right). \quad (16)$$

7. Optimizing ABUC's feedback parameters using Q-learning

In this section, we design a reinforcement learning framework which enables the proposed ABUC algorithm to optimize its scheduling parameters on-the-fly, given each user mobility profile. Different to the initial ABUC presented above where the algorithm

parameters, i.e. period T and CSI feedback strategy, are fixed during the whole scheduling process, this Q-learning framework enables to activate dynamic and static clustering schemes under different CSI feedback strategies adaptively depending on the individual user mobility profile. To do that, we formulate the optimization problem as a Q-learning model in which an agent learns from the environment to manage it-self the feedback parameters, i.e. period T and CSI feedback type.

Q-learning is a reinforcement learning technique which goal is to learn a policy that informs an agent what action to take under what circumstances. It does not require a model of the environment and can handle problems with stochastic transitions and rewards, without requiring adaptations. The learning agent maximize its total (future) reward by adding the maximum reward attainable from future states to the reward for achieving its current state, effectively influencing the current action by the potential future reward. This potential reward is a weighted sum of the expected values of the rewards of all future steps starting from the current state.

In each decision epoch, the agent decides to make an action and observes the results from this action. Each action-state pair produces a Q-value that will be updated in a table in which the columns and the rows represent the actions and the states, respectively. The updated value of $Q^*(s_t, a_t)$ is computed by the Bellman function as presented in [34]:

$$Q^*(s_t, a_t) = Q(s_t, a_t) + \alpha [Rwd_t(s_t, a_t) + \gamma \max_{s_{t+1}} Q^{t+1}(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (17)$$

where $Q(s_t, a_t)$ denotes Q-value at state s_t when executing action a_t in time slot t , Rwd_t is the system reward at state s_t and action a_t . α and γ are learning rate and discount rate of the future expected reward, respectively.

In order to obtain the optimal policy, it is necessary to identify the actions, states and reward functions in the Q-learning model.

1) *System State*: The current system state s_t is jointly determined by the states of all K users. Due to the relationship of the received SINR and the channel coefficient, we model the channel coefficient, $|h_{lk}|^2$, as a Markov random variable. We partition and quantize the range of $|h_{lk}|^2$ into N levels. Each level corresponds to a state of the Markov channel. Each user k state is defined as the quantized CSI level n_k^t , where $1 \leq n_k^t \leq N$, $n_k^t \in \mathbf{N}$. The system state at time slot t is defined as,

$$s_t = \{n_1^t, n_2^t, \dots, n_K^t\}$$

2) *System Action*: In the system, the central scheduler has to decide which feedback parameters to be selected. Let \mathbf{T} and \mathbf{F} denote the set of possible values of T and CSI feedback schemes, respectively.

$$\mathbf{T} = \{T_1, \dots, T_p, \dots, T_P\}$$

$$\mathbf{F} = \{f_1, \dots, f_q, \dots, f_Q\}$$

where $P \in \mathbf{N}$ and Q represents the set dimension of all possible CSI feedback strategies.

The current composite action a_t is denoted by

$$a_t = \{a_1^t, a_2^t, \dots, a_K^t\}$$

where $a_k^t = (T_k^t, f_k^t)$ represents the feedback parameters of user k at time slot t , where period $T_k^t \in \mathbf{T}$ and CSI feedback type $f_k^t \in \mathbf{F}$.

3) *Reward Function*: The system reward needs to represent the optimization objective, that is to simultaneously reduce the system cost and satisfy the sum-rate demands. Here, we define the overall system reward at state s_t and action a_t as

$$Rwd_t(s_t, a_t) = \rho_1 \sum_{k=1}^K T_k(s_t, a_t) - \rho_2 \sum_{k=1}^K C_k(s_t, a_t) \quad (18)$$

where the first term is the system achieved sum-rate at state s_t and action a_t , and the second one denotes the CSI signaling overhead induced by the same state and action, ρ_1 and ρ_2 are weighting parameters representing the trade-off between the sum-rate and the cost, and $\rho_1 + \rho_2 = 1$.

The CSI overhead cost $C_k(T_k, f_k)$ of each user k is computed over T_k frames and can be expressed as follows:

- If Full CSI: $C_k(T_k, f) = \frac{1}{T_k} \left[L + (T_k - 1)L_k \right] MN$
- If partial or Estimated CSI : $C_k(T_k, p) = C_k(T_k, e) = \frac{LMN}{T_k}$

The Q-learning based framework uses an ϵ -greedy strategy [35] in which the amount of exploration is globally controlled by the parameter ϵ , that determines the randomness in action selections. In the ϵ -greedy method, the agent selects a random action with a fixed probability ϵ , $0 \leq \epsilon \leq 1$. At first, this rate must be initiated to its highest value, i.e. $\epsilon = 1$, as we don't have any knowledge about the values in the Q-table. At each time step, a uniform random number ξ is drawn, where $\xi \in [0, 1]$. If $\xi > \epsilon$, the action that gives the greatest value in the Q-table will be chosen, otherwise we select greedily one of the learned actions set. We reduce ϵ progressively as the agent becomes more confident at estimating Q-values. The Q-learning framework is detailed in Algorithm 4.

Algorithm 4: Proposed ABUC's Q-learning framework.

```

initialize user mobility profile set  $\mathbb{U}$ 
 $F_{\max}$ : number of learning episodes
 $F_0$ : number of frames for each learning episode
 $\epsilon$ : exploration rate,  $\epsilon \leftarrow 1$ 
for episode  $i = 1: F_{\max}$  do
  if  $i = 1$  then
    With probability  $\epsilon$ , randomly select an action
  else
    Randomly generate a probability  $\xi$ 
    if  $\xi \leq \epsilon$  then
      randomly select an action
    else
      choose action  $a_i = \arg \max Q(s_i, a_i)$ 
    end
  end
  for frame  $t=1:F_0$  do
    Execute ABUC with  $a_i^t$  parameter
    Obtain beamforming and clustering solutions for each frame  $t$ 
    Compute average sum-rate of episode  $i$  over all  $F_0$  frames
  end
  Compute the reward  $Rwd_i$  using (23) and observe the new state  $s_{i+1}$ 
  Update entry  $(s_i, a_i)$  of Q-table using (22)
  Update CSI quantization state  $\{n_k^i\}$ 
  Reduce exploration rate  $\epsilon$  by 0.1%
end

```

8. Simulation results

In this section, we numerically evaluate the performance of the proposed ABUC algorithm. We consider a 7-cell wrapped around two-tier H-CRAN. Each cell has a single macro-RRH and 3 pico-RRHs equally separated in space. The number of mobile users is varied between 5 and 30, uniformly distributed per macro-cell.

Table 1
Parameter settings for simulation.

Simulation parameters	
Cellular layout	Hexagonal 7-cell wrapped-around two-tier model
Channel bandwidth	10 MHz
Inter-cell distance	0.8 km
TX power for macro/pico RRH	(43, 30) dBm
Antenna gain	15 dBi
Background noise	-169 dBm/Hz
Path-loss from macro RRH to user	$128.1 + 37.6 \log_{10}(d)$
Path-loss from pico RRH to user	$140.7 + 36.7 \log_{10}(d)$
Log-normal shadowing	8 dB
CSI error variance	-20 dB
User priority weights α_k	$1 \forall k$

We assume a Random Waypoint model to represent users' movements. The fronthaul constraints for macro-RRH and pico-RRH are 683.1 Mbps and 106.5 Mbps, respectively [25]. All channels undergo Rayleigh small scale fading and log-normal shadowing. The other parameter settings are presented in Table 1.

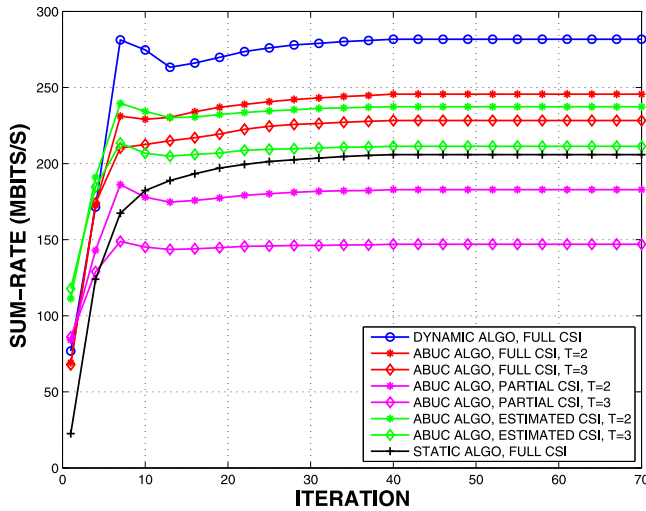
To evaluate the performance of the proposed ABUC algorithm and its behavior with regards to user mobility, we consider different mobility profiles represented by the parameter λ which is a function of velocity. In (9), we set the carrier frequency f and the fronthaul delay T_{dl} for all RRH l as 900 MHz and 2 ms [36], respectively.

8.1. ABUC Algorithm performance

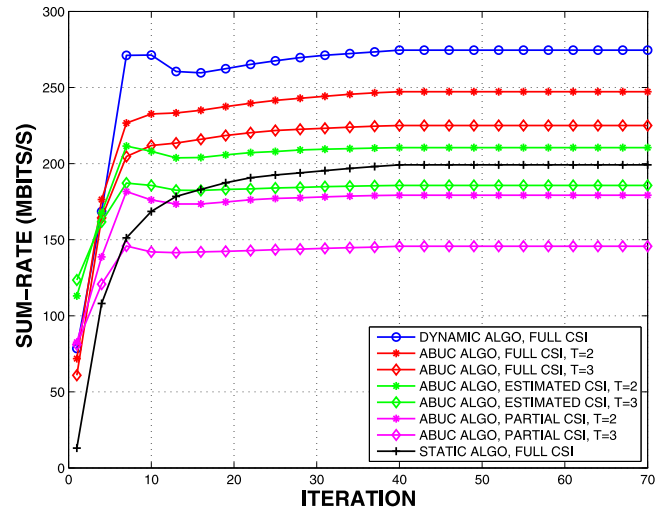
The performance of ABUC algorithm is compared to two baseline schemes, namely dynamic and static algorithms with Full CSI as in [11,25]. Note that considering Full CSI for static and dynamic algorithms guarantees their best sum-rate performance. For ABUC scheme, we varied both CSI feedback strategy and the period value T . We focus on three representative types of mobility profiles: low, medium and high with velocities 6 km/h, 36 km/h and 72 km/h, respectively.

In Fig. 4, we show the average sum-rate convergence of reference dynamic and static algorithms for $K = 10$ users per macrocell, and the proposed ABUC algorithm with periods $T = 2$ and $T = 3$ for all CSI feedback strategies. It is worth noting that ABUC with $T = 1$ is equivalent to the dynamic algorithm with Full CSI, and setting $T \geq 4$ induced excessive sum-rate degradation compared to baseline algorithms. This justifies our choice of limiting the appropriate period values to $T = 2$ and $T = 3$. As we can observe, the dynamic algorithm outperforms the static one, while ABUC algorithm with Full CSI is closer to the optimal sum-rate performance of the dynamic algorithm. As expected, the larger the period, the lower the performance of our proposed solution since it uses a more outdated clustering solution (derived in the last dynamic frame). In addition, ABUC scheme with Estimated CSI shows a very close performance to ABUC with Full CSI case under low mobility (see Fig. 4(a)). However, its performance degrades with higher velocities as the Estimated CSI quality loses accuracy as shown in Figs. 4(b) and (c). Under high mobility, Fig. 4(c) shows that the Estimated CSI strategy degrades even below the Partial CSI scheme. This behavior will be further discussed in the next subsection.

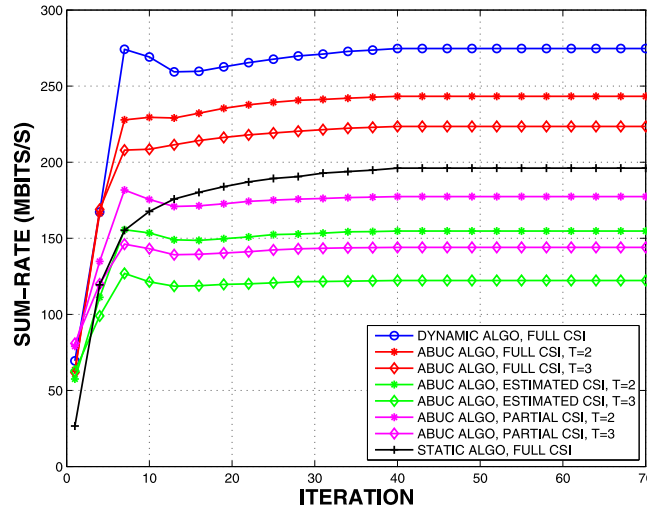
Fig. 5 evaluates the sum-rate performance as a function of the number of users and their mobility profile. We can observe the same tendency in performance with the increasing number of users. ABUC algorithm with Full CSI offers a close performance to the dynamic optimal solution for all mobility scenarios. Again, ABUC with Estimated CSI exhibits a good performance for lower mobility while Partial CSI outperforms it for high mobility. From these figures, we can also observe the effect of period T on the



(a) Low mobility (6 km/h)



(b) Medium mobility (36 km/h)



(c) High mobility (72 km/h)

Fig. 4. Average sum-rate convergence against number of iterations.

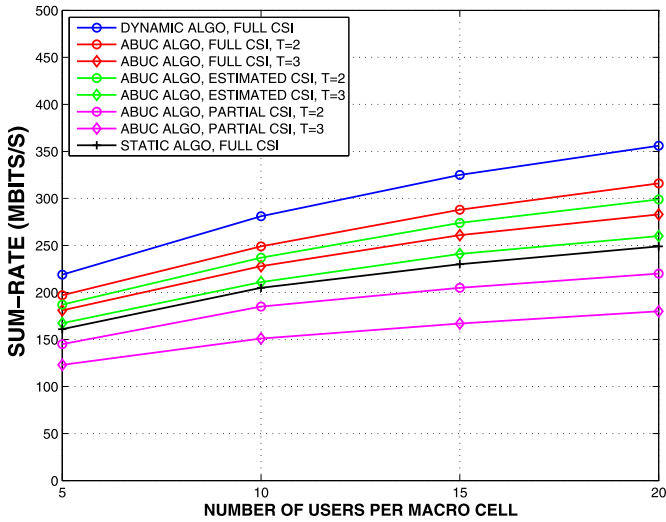
sum-rate: ABUC with $T = 3$ and Full CSI behaves better than with $T=2$ and Partial CSI. With $T = 3$ and Partial CSI, the proposed method is outperformed by the static algorithm with full (perfect) CSI, showing the importance of accurate CSIs.

In Fig. 6, we plot the cumulative distribution function (CDF) of per user data rates for proposed ABUC variants and baseline dynamic and static algorithms. The figure emphasizes the fairness among users in terms of instantaneous rate distribution. We notice that the ABUC algorithm shows a near performance to the dynamic algorithm in case of Full CSI, i.e. it reaches up to 91.3% and 86.8% of the optimal performance with $T = 2$ and $T = 3$ at the 50th percentile, respectively, for low mobility scenario. We can also point out the behavior of the proposed ABUC with Full CSI which tends to allocate more resources to low CSI quality users unlike the dynamic algorithm which concentrates the allocation towards best CSI users to maximize the global sum-rate. Therefore, ABUC scheme provides a better fairness for all users while approaching the optimal sum-rate.

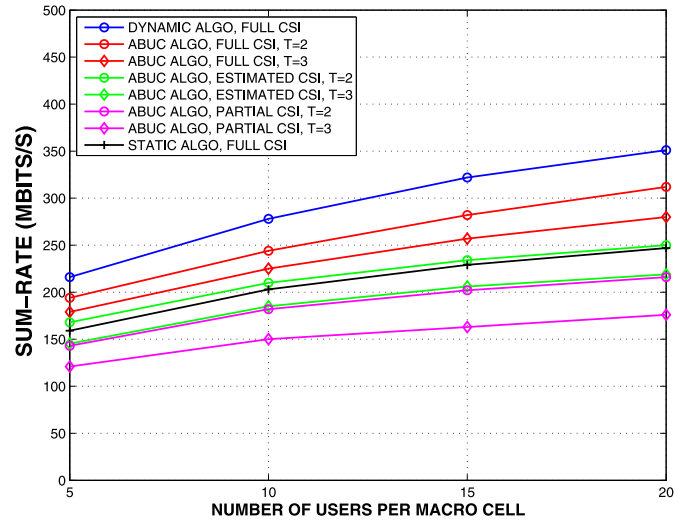
Another important analysis concerns the costs induced by all the compared schemes. We first plot in Fig. 7 the amount of generated CSI feedback overhead as a function of the number of users.

We can see that all variants of ABUC algorithm provide an important reduction in CSI overhead compared to the dynamic scheme (43–48% for $T = 2$ and 63–66% for $T = 3$). For the ABUC Full CSI case, this gain comes from smaller cluster sizes ($L_k \leq L$ (11)) in intermediate static frames. Moreover, the performance in terms of CSI overhead can be significantly improved by employing Partial CSI or Estimated CSI instead of Full CSI. Note that the amount of CSI feedback for partial and Estimated CSI is equal as shown in (12).

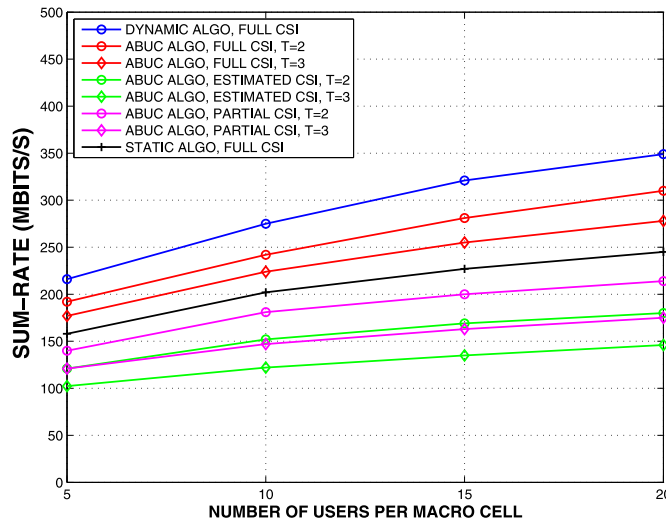
Furthermore, the ABUC algorithm enables a large reduction of the number of user-to-RRH association messages. We plot in Fig. 8 the re-association cost of proposed ABUC algorithm compared to the dynamic one for the three mobility profiles and different user loads (10 and 20 users per macro cell). In this figure, we observe a significant reduction of re-association messages for ABUC scheme compared to the dynamic one, which is even more under high user load. Note that as shown in Section 6.1.2, the re-association cost of the proposed algorithm depends only on period T and is unaffected by the CSI feedback strategy. With $T = 2$ and $T = 3$ ABUC achieves 54.3% and 74% reduction compared to the dynamic scheme, respectively. In addition, the results show that the mobility acts on



(a) Low mobility (6 km/h)



(b) Medium mobility (36 km/h)



(c) High mobility (72 km/h)

Fig. 5. Sum-rate performance as function of the number of users per macro cell.

the generated re-association cost as high user mobility requires more frequent changes of serving RRHs over successive scheduling frames. However, this increase in re-association cost is marginal proving the robustness of the proposed schemes against varying user loads and velocities.

8.2. Performance-cost tradeoff analysis: a mobility perspective

In this section, we discuss the performance-cost tradeoff under different mobility conditions. To characterize the different variants of ABUC algorithm, we define the parameters (T, f) , where T is the period of dynamic frames, and f is the CSI feedback strategy used for intermediate frames, i.e., $T \in \{2, 3\}$ and f may be chosen among Full CSI, Partial CSI and Estimated CSI.

Fig. 9 summarizes the sum-rate behavior of all algorithms against users velocities represented by the correlation coefficient λ . As we can see, λ approaching 1 corresponds to the static case and hence the Estimated CSI is very close to the real channel state, explaining why our ABUC scheme with Estimated CSI closely ap-

proaches its performance with Full CSI. On the other end, when λ tends to 0, the user velocity is very high leading to very dynamic channel variations. This can be observed on the performance of ABUC with Estimated CSI which degrades even lower than the case with Partial CSI given the poor quality of the CSI estimation. In addition, we observe that unlike ABUC with full and Partial CSI strategies, the performance of ABUC with Estimated CSI is strongly dependent on user mobility. In light of these considerations and the performance tendencies noticeable in Fig. 9, we identify three mobility regions: low mobility, medium mobility and high mobility corresponding to the velocity and λ ranges (0–18 km/h ; $0.96 \leq \lambda \leq 1$), (18–54 km/h ; $0.7 \leq \lambda \leq 0.96$), ($v \geq 54$ km/h ; $\lambda \leq 0.6$), respectively. For each range of velocity, we can find the most suitable algorithm with its parameters (T, f) that can balance the trade-off between sum-rate performance and the incurred costs.

- *Low mobility scenario:* from Fig. 9, ABUC scheme with Estimated CSI for both values of T ($T = 2$ and $T = 3$) closely approaches ABUC sum-rate performance with Full CSI, while largely reducing the CSI feedback overhead and

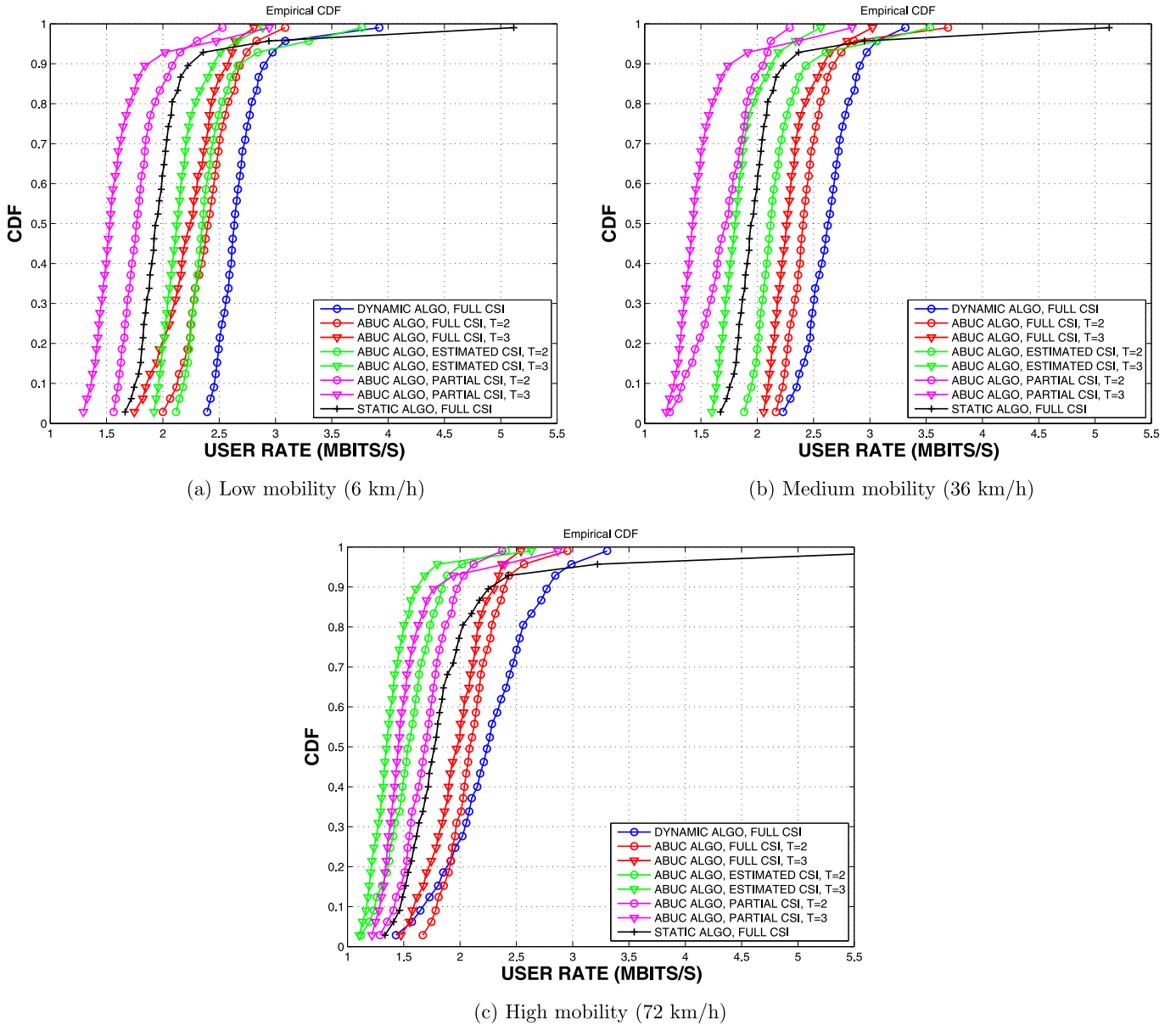


Fig. 6. Cumulative distribution function of user data rate.

Table 2

Computational complexity as function of the number of users per macro cell.

Number of users/macro cell	10	20	30
Dynamic algo	4.2×10^{12}	6.7×10^{13}	3.4×10^{14}
ABUC algo $T = 2$	5.4×10^7	4.7×10^8	1.8×10^9
ABUC algo $T = 3$	5.1×10^7	4.4×10^8	1.7×10^9

re-association costs compared to the baseline dynamic algorithm (as shown in Figs. 7 and 8). In addition, Fig. 6a reveals that ABUC with estimated and Full CSI achieve similar user fairness levels for each T . Moreover, regarding the computational complexity, we can see in Table 2 that ABUC with both values of T drastically decreases the computation complexity owing mainly to the smaller cluster sizes L_k used in intermediate static frames. Note that we can also observe that the reduction in complexity for ABUC is proportional to the period T . Therefore, we can conclude that ABUC with

($T = 3$; Estimated CSI) provides the best trade-off for low mobility users.

- *Medium mobility scenario:* for medium velocities, ABUC with Estimated CSI loses some performance compared to Full CSI but still outperforms Partial CSI case and performs close to baseline static algorithm with Full CSI. Here again, considering the balance between the loss of sum-rate performance and the gain in complexity and signaling costs (Figs. 7 and 8), we can infer that ABUC with ($T = 2$; Estimated CSI) provides the best trade-off performance for medium mobility. However, if the system can afford a higher overhead consumption, ABUC with ($T = 2$; Full CSI) offers the best sum-rate performance.
- *High mobility scenario:* In high mobility environments, it is clear that the Estimated CSI becomes obsolete and hence ABUC with Estimated CSI has no more benefits. As a consequence, ABUC with ($T = 2$; Partial CSI) is preferred as the best option to realize the sum-rate and cost trade-off.

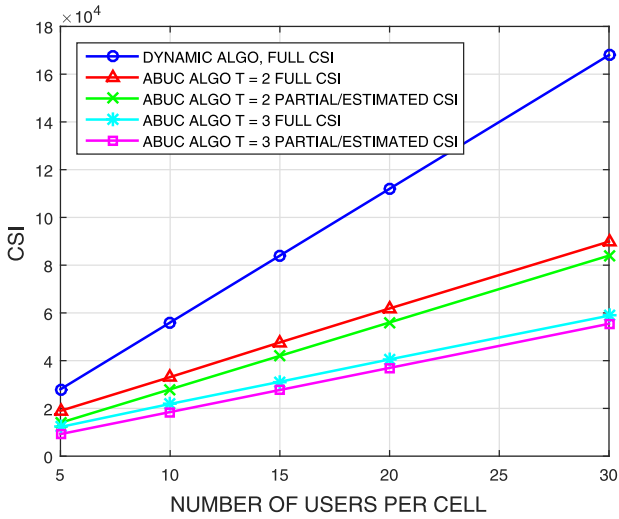


Fig. 7. CSI overhead as function of the number of users per macro cell.

From the above discussion, given that the set of parameters which provide the best performance-cost trade-off are highly dependent on the mobility profile of users, a Q-learning based method enabling the dynamic selection of parameters (T, f) is investigated in the next section.

8.3. Performance of ABUC with Q-learning

Our goal in this section is to derive some preliminary performance results of the proposed ABUC's Q-learning based feedback parameter selection framework on a simple network which consists of 1 macro-RRH and 3 pico-RRHs and 3 users. The agent will learn over 1200 episodes for a state space of 64 states, corresponding to 4 states of CSI quantization for each user. The action set for each user consists of 4 actions: $T = 1$ with Full CSI and $T = 2$ with Full CSI, Partial CSI and Estimated CSI, resulting in a Q-table of size 64×64 . The value of channel coefficients $|h_{lk}|^2$ are quantized into 4 ranges which are bounded by the following values: 6×10^{-6} , 1.8×10^{-5} and 8×10^{-5} . The other simulation parameters are the same as in Table 1.

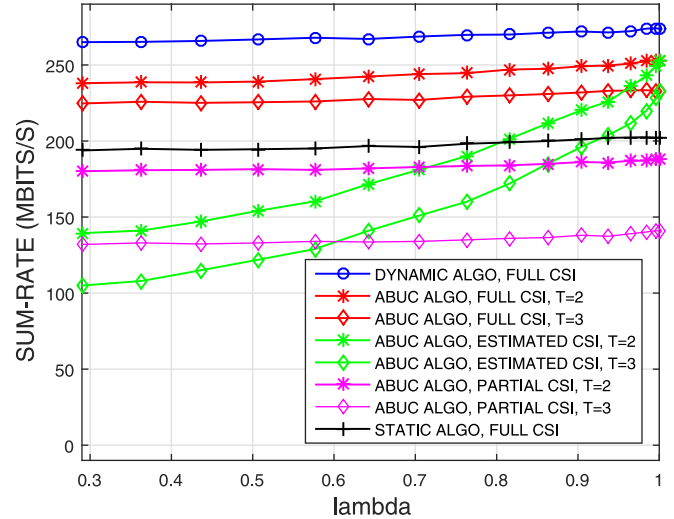
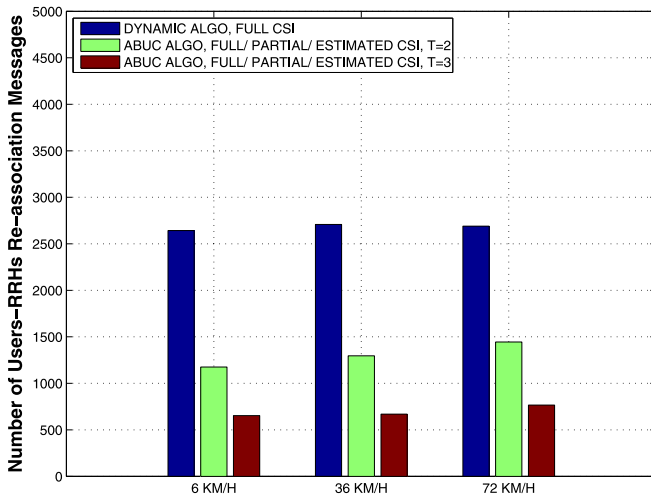


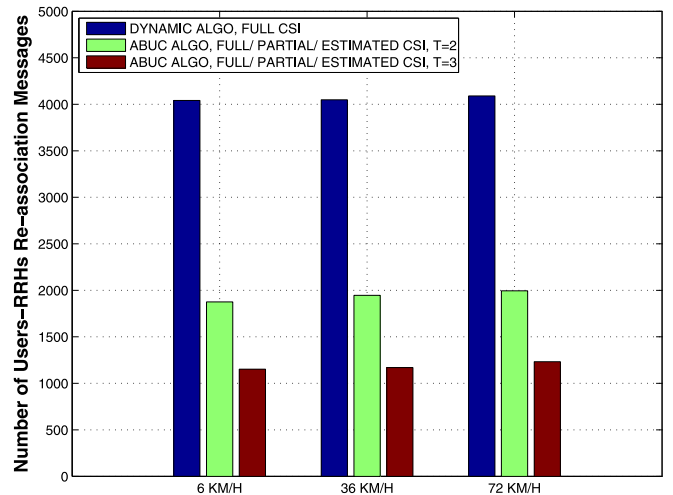
Fig. 9. Sumrate performance against lambda.

We consider two different scenarios: in the first one, all users have the same velocity while in the second, each user has its own individual velocity.

Firstly, we evaluated the proposed algorithm in the first scenario where all users undergo the same velocity which is varied over the three mobility profiles, namely: low, medium and high mobility corresponding respectively to 6 km/h, 36 km/h and 72 km/h. We plot in Fig. 10 converged value of the reward against different values of sum-rate weight ρ_1 which varies between 0 and 1. We observe that when ρ_1 approaches 1, the reward tends toward the sum-rate and all users whatever their velocity converge to take the same optimal action ($T = 1$, Full CSI). Inversely, when ρ_1 approaches 0, the CSI cost factor becomes more dominating on the reward value, then the users choose the optimal action that minimizes the CSI cost, i.e. ($T = 2$, Estimated CSI) for low and medium velocities, and ($T = 2$, Partial CSI) for high velocities. In addition, we also observe that the reward slightly decreases as the mobility become higher due to the CSI degradation. We varied the value of weights (ρ_1, ρ_2) to three cases (0.3, 0.7), (0.7, 0.3) and (0.9, 0.1) representing different situations of tradeoff between the sum-rate and the induced costs. We can see in Fig. 11 that the three users



(a) 10 users/macro cell



(b) 20 users/macro cell

Fig. 8. Re-association cost as function of number of users per macro cell.

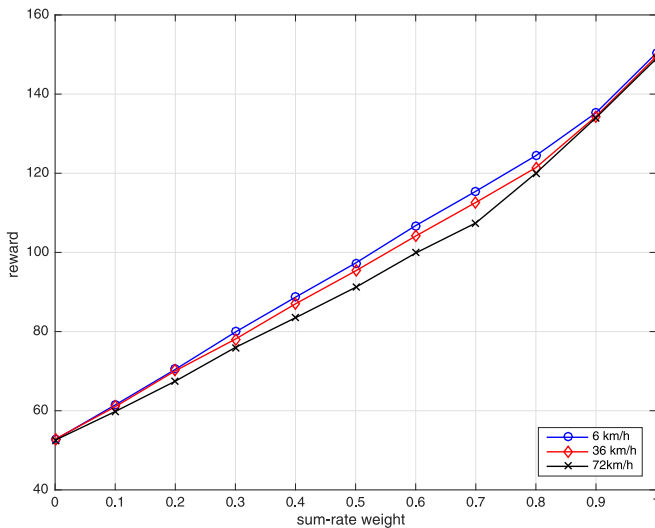


Fig. 10. Converged reward as function of weight ρ_1 .

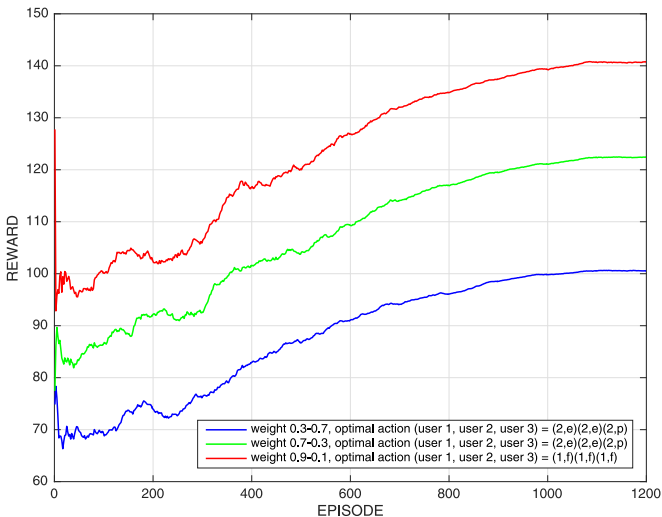


Fig. 11. Reward convergence over the learning episodes Individual velocities: user 1 = 6 km/h, user 2 = 36 km/h, user 3 = 72 km/h.

having different velocity converge individually to different optimal actions which are highly dependant on their respective velocities. We observe that for weight $\rho_1 = 0.3$ or $\rho_1 = 0.7$, the users with low and medium velocity take the action ($T = 2$, Estimated CSI), while the user with high mobility converges to ($T = 2$, Partial CSI). The selection of these actions confirms the effectiveness of the learning algorithm. Indeed, in case of low and medium mobility, the CSI estimation is accurate enough to be used when $\rho_1 = 0.3$ and $\rho_1 = 0.7$ for a balanced trade-off between sum-rate and cost, while the high mobility user no longer maintains an acceptable CSI estimation and then prefers selecting Partial CSI. Moreover, coupled with the fact that $T = 2$ is very effective for cost reduction, the users converge to $T = 2$ instead of $T = 1$. Finally, in case of $\rho_1 = 0.9$ which corresponds to a predominant sum-rate performance, the learning algorithm wisely converges to ($T = 1$, Full CSI) action regardless the user velocity as Full CSI and dynamic clustering at all frame presents the best action for sum-rate maximization.

9. Conclusion

In this paper, we investigated the trade-off between throughput gain and complexity/signaling overhead costs of the joint beamforming and user clustering problem in the downlink of a H-CRAN. We proposed ABUC, an algorithm that periodically activates dynamic and static clustering strategies for leveraging both the optimality of the dynamic solution and the low-complexity of its static counterpart. The key benefit of the proposed ABUC algorithm is to take into account the temporal dimension of the allocation process while being aware of practical system operation metrics, namely CSI feedback overhead, re-association cost, and computation complexity. We also propose to use a mobility-aware channel estimation model to better predict the channel variations when CSI feedback reduction is applied. The numerical results show that the proposed solution narrows significantly the performance gap with the optimal dynamic solution, while considerably reducing the required computational complexity, CSI feedback and re-association signaling overhead over the long-term allocation process. Moreover, we have identified the best parameter sets for different mobility scenarios. Furthermore, we designed a Q-learning framework which enables the proposed ABUC algorithm to optimize its scheduling parameters on-the-fly, given each user mobility profile. Our proposed framework is able to learn about the system dynamics and predict the best parameter set to apply for each user depending on its mobility behavior.

As a future work direction, the dynamic selection of algorithm parameters for a large scale H-CRAN network can be optimized through a more sophisticated machine learning techniques, e.g., deep reinforcement learning. The resource allocation and scheduling problem in Fog-RAN will be also considered.

Conflicts of Interest

None.

Acknowledgements

This work is supported by the CNRS PICS bilateral research fund between France and Japan, and by the Grant-in-Aid for Scientific Research (Kakenhi) no. 17K06453 from the Ministry of Education, Science, Sports and Culture of Japan.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.comnet.2019.05.005.

References

- [1] H.D. Thang, L. Boukhatem, M. Kaneko, S. Martin, Performance-cost trade-off of joint beamforming and user clustering in cloud radio access networks, in: 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), 2017, pp. 1–5, doi:10.1109/PIMRC.2017.8292411.
- [2] A. Gupta, R.K. Jha, A survey of 5G network: architecture and emerging technologies, IEEE Access 3 (2015) 1206–1232, doi:10.1109/ACCESS.2015.2461602.
- [3] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Que-seth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M.A. Uusitalo, B. Timus, M. Fallgren, Scenarios for 5G mobile and wireless communications: the vision of the metis project, IEEE Commun. Mag. 52 (5) (2014) 26–35, doi:10.1109/MCOM.2014.6815890.
- [4] M. Peng, H. Xiang, Y. Cheng, S. Yan, H.V. Poor, Inter-tier interference suppression in heterogeneous cloud radio access networks, IEEE Access 3 (2015) 2441–2455, doi:10.1109/ACCESS.2015.2497268.
- [5] M. Peng, Y. Li, Z. Zhao, C. Wang, System architecture and key technologies for 5G heterogeneous cloud radio access networks, IEEE Netw. 29 (2) (2015) 6–14, doi:10.1109/MNET.2015.7064897.
- [6] M. Peng, Y. Li, J. Jiang, J. Li, C. Wang, Heterogeneous cloud radio access networks: a new perspective for enhancing spectral and energy efficiencies, IEEE Wirel. Commun. 21 (6) (2014) 126–135, doi:10.1109/MWC.2014.7000980.

- [7] H. Dahrouj, A. Douik, O. Dhimallah, T.Y. Al-Naffouri, M.S. Alouini, Resource allocation in heterogeneous cloud radio access networks: advances and challenges, *IEEE Wirel. Commun.* 22 (3) (2015) 66–73, doi:[10.1109/MWC.2015.7143328](https://doi.org/10.1109/MWC.2015.7143328).
- [8] X. Mao, B. Zhang, Y. Chen, J. Yu, Z. Han, Matching game based resource allocation for 5G H-CRAN networks with device-to-device communication, in: 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), 2017, pp. 1–6, doi:[10.1109/PIMRC.2017.8292250](https://doi.org/10.1109/PIMRC.2017.8292250).
- [9] K. Hosseini, W. Yu, R.S. Adve, Cluster based coordinated beamforming and power allocation for mimo heterogeneous networks, in: 2013 13th Canadian Workshop on Information Theory, 2013, pp. 96–101, doi:[10.1109/CWIT.2013.6621600](https://doi.org/10.1109/CWIT.2013.6621600).
- [10] M.M.U. Rahman, H. Ghauch, S. Imtiaz, J. Gross, RRH clustering and transmit precoding for interference-limited 5G CRAN downlink, in: 2015 IEEE Globecom Workshops (GC Wkshps), 2015, pp. 1–7, doi:[10.1109/GLOCOMW.2015.7414198](https://doi.org/10.1109/GLOCOMW.2015.7414198).
- [11] B. Dai, W. Yu, Sparse beamforming and user-centric clustering for downlink cloud radio access network, *IEEE Access* 2 (2014) 1326–1339, doi:[10.1109/ACCESS.2014.2362860](https://doi.org/10.1109/ACCESS.2014.2362860).
- [12] M. Peng, Y. Yu, H. Xiang, H.V. Poor, Energy-efficient resource allocation optimization for multimedia heterogeneous cloud radio access networks, *IEEE Trans. Multimed.* 18 (5) (2016) 879–892, doi:[10.1109/TMM.2016.2535722](https://doi.org/10.1109/TMM.2016.2535722).
- [13] D.T. Ha, L. Boukhatem, M. Kaneko, S. Martin, An advanced mobility-aware algorithm for joint beamforming and clustering in heterogeneous cloud radio access network, in: Proceedings of the 21st ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, in: MSWIM '18, ACM, New York, NY, USA, 2018, pp. 199–206, doi:[10.1145/3242102.3242120](https://doi.org/10.1145/3242102.3242120).
- [14] Z. Guizani, N. Hamdi, CRAN, H-CRAN, And F-RAN for 5G systems: key capabilities and recent advances, *Int. J. Netw. Manage.* 27 (2017).
- [15] Y. Lin, L. Shao, Z. Zhu, Q. Wang, R.K. Sathikhi, Wireless network cloud: architecture and system requirements, *IBM J. Res. Dev.* 54 (1) (2010) 4:1–4:12, doi:[10.1147/JRD.2009.2037680](https://doi.org/10.1147/JRD.2009.2037680).
- [16] C. Mobile, C-RAN: The Road Towards Green RAN, White Paper, ver 2 (2011).
- [17] A. Checko, H.L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M.S. Berger, L. Dittmann, Cloud RAN for mobile networks – a technology overview, *IEEE Communications Surveys Tutorials* 17 (1) (2015) 405–426, doi:[10.1109/COMST.2014.2355255](https://doi.org/10.1109/COMST.2014.2355255).
- [18] H. Dahrouj, A. Douik, F. Rayal, T.Y. Al-Naffouri, M. Alouini, Cost-effective hybrid RF/FSO backhaul solution for next generation wireless systems, *IEEE Wirel. Commun.* 22 (5) (2015) 98–104, doi:[10.1109/MWC.2015.7306543](https://doi.org/10.1109/MWC.2015.7306543).
- [19] C. Ran, S. Wang, C. Wang, Balancing backhaul load in heterogeneous cloud radio access networks, *IEEE Wirel. Commun.* 22 (3) (2015) 42–48, doi:[10.1109/MWC.2015.7143325](https://doi.org/10.1109/MWC.2015.7143325).
- [20] Q. Shen, Z. Ma, S. Wang, Deploying C-RAN in cellular radio networks: an efficient way to meet future traffic demands, *IEEE Trans. Veh. Technol.* 67 (8) (2018) 7887–7891, doi:[10.1109/TVT.2018.2838133](https://doi.org/10.1109/TVT.2018.2838133).
- [21] Y. Shi, J. Zhang, K.B. Letaief, Group sparse beamforming for green cloud-RAN, *CoRR abs/1310.0234* (2013).
- [22] H. Dahrouj, W. Yu, T. Tang, S. Beaudin, Power spectrum optimization for interference mitigation via iterative function evaluation, in: 2011 IEEE GLOBECOM Workshops (GC Wkshps), 2011, pp. 162–166, doi:[10.1109/GLOCOMW.2011.6162428](https://doi.org/10.1109/GLOCOMW.2011.6162428).
- [23] S.H. Park, O. Simeone, O. Sahin, S. Shamai, Joint precoding and multivariate backhaul compression for the downlink of cloud radio access networks, *IEEE Trans. Signal Process.* 61 (22) (2013) 5646–5658, doi:[10.1109/TSP.2013.2280111](https://doi.org/10.1109/TSP.2013.2280111).
- [24] M. Peng, S. Yan, H.V. Poor, Ergodic capacity analysis of remote radio head associations in cloud radio access networks, *IEEE Wirel. Commun. Lett.* 3 (4) (2014) 365–368, doi:[10.1109/LWC.2014.2317476](https://doi.org/10.1109/LWC.2014.2317476).
- [25] B. Dai, W. Yu, Backhaul-aware multicell beamforming for downlink cloud radio access network, in: 2015 IEEE International Conference on Communication Workshop (ICCW), 2015, pp. 2689–2694, doi:[10.1109/ICCW.2015.7247585](https://doi.org/10.1109/ICCW.2015.7247585).
- [26] Y. He, Z. Zhang, F.R. Yu, N. Zhao, H. Yin, V.C.M. Leung, Y. Zhang, Deep-reinforcement-learning-based optimization for cache-enabled opportunistic interference alignment wireless networks, *IEEE Trans. Veh. Technol.* 66 (11) (2017) 10433–10445, doi:[10.1109/TVT.2017.2751641](https://doi.org/10.1109/TVT.2017.2751641).
- [27] H.Y. Ong, K. Chavez, A. Hong, Distributed deep q-learning, *CoRR abs/1508.04186* (2015).
- [28] V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, D. Hassabis, Human-level control through deep reinforcement learning, *Nature* 518 (7540) (2015) 529–533.
- [29] Z. Xu, Y. Wang, J. Tang, J. Wang, M.C. Gursoy, A deep reinforcement learning based framework for power-efficient resource allocation in cloud RANs, in: 2017 IEEE International Conference on Communications (ICC), 2017, pp. 1–6, doi:[10.1109/ICC.2017.7997286](https://doi.org/10.1109/ICC.2017.7997286).
- [30] I. AlQerm, B. Shihada, Enhanced machine learning scheme for energy efficient resource allocation in 5G heterogeneous cloud radio access networks, in: 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), 2017, pp. 1–7, doi:[10.1109/PIMRC.2017.8292227](https://doi.org/10.1109/PIMRC.2017.8292227).
- [31] Y. Sun, M. Peng, S. Mao, Deep reinforcement learning based mode selection and resource management for Green Fog Radio Access networks, *IEEE Internet Things J.* (2019), doi:[10.1109/JIOT.2018.2871020](https://doi.org/10.1109/JIOT.2018.2871020). 1–1
- [32] M. Razaviyayn, M. Hong, Z.-Q. Luo, A unified convergence analysis of block successive minimization methods for nonsmooth optimization, *SIAM J. Optim.* 23 (2) (2013) 1126–1153, doi:[10.1137/120891009](https://doi.org/10.1137/120891009).
- [33] D.S. Michalopoulos, H.A. Suraweera, G.K. Karagiannidis, R. Schober, Amplify-and-forward relay selection with outdated channel estimates, *IEEE Trans. Commun.* 60 (5) (2012) 1278–1290, doi:[10.1109/TCOMM.2012.032012.110430](https://doi.org/10.1109/TCOMM.2012.032012.110430).
- [34] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, M.A. Riedmiller, Playing Atari with deep reinforcement learning, *CoRR abs/1312.5602* (2013).
- [35] M. Rovcanin, E.D. Poorter, I. Moerman, P. Demeester, A reinforcement learning based solution for cognitive network cooperation between co-located, heterogeneous wireless sensor networks, *Ad Hoc Netw.* 17 (2014) 98–113.
- [36] 3GPP, Technical Specification Group Services and System Aspects, Quality of Service (QoS) Concept and Architecture (Release 12), Report TS 23.107, V12.0.0 (2014).



Duc Thang Ha received the B.E. degree in electrical engineering from Hanoi University of Science and Technology, Hanoi, Vietnam, in 2014, and the M.Sc. degree in computer science from the University of Paris-Sud, Orsay, Ile-de-France, France, in 2015, where he is currently pursuing the Ph.D. degree with the Computer Science Department. His research interests include wireless communications and convex optimization.



Lila Boukhatem is associate professor in the ROCS (Networks and combinatorial and stochastic optimization) team of LRI laboratory (CNRS – University of Paris Saclay – University of Paris-Sud). She received a computer science engineering degree in 1997 from INI Institute, Algeria, MSc degree in 1998 from the University of Versailles Saint Quentin-en-Yvelines, France, and Ph.D. degree in 2001 from the University of Pierre et Marie Curie (Paris 6), France. In 2009, she obtained her HDR degree from Paris-Sud University, France. She joined the LRI laboratory in 2002 where she is developing various research works in mobile and wireless networks including: Cross-layer design, modeling and performance evaluation, radio resource allocation and optimization, mobility management, routing and energy in vehicular networks. She has lead and actively participated to several national and international projects and programs such as PHC Sakura, FP7 - NoE NEWCOM#, EIT ICT-Labs Digital Cities. Her current research interests are radio and energy resource allocation for edge computing enabled 5G systems.



Megumi Kaneko received her B.S. and MSc. degrees in communication engineering in 2003 and 2004 from Institut National des Télécommunications (Télécom SudParis), France, jointly with a MSc. from Aalborg University, Denmark, where she received her Ph.D. degree in 2007. In May 2017, she obtained her HDR degree (French Doctoral Habilitation for Directing Researches at Professor position) from Paris-Sud University, France. From September 2010 to March 2016, she was an Assistant Professor in the Graduate School of Informatics, Kyoto University. She is currently an Associate Professor at the National Institute of Informatics, Tokyo, Japan. Her research interests include wireless communications, radio resource and interference management, wireless sensing and cross-layer network protocols. She received the 2009 Ericsson Young Scientist Award, the IEEE Globecom 2009 Best Paper Award, the 2011 Funai Young Researcher's Award, the WPMC 2011 Best Paper Award, the 2012 Telecom System Technology Award and the 2016 Inamori Foundation Research Grant.



Nhan Nguyen-Thanh received the B.E. and M.Eng. degrees in electrical engineering from the Ho Chi Minh City University of Technology in 2003 and 2007, respectively, and the Ph.D. degree in electrical engineering from the University of Ulsan, South Korea, in 2012. He is currently with the LRI Laboratory, University of Paris Sud, France, as a Post-Doctoral Researcher. His current research interests include deep learning, game theory, IoT, and the next-generation wireless communication systems.