IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 1, 2018, accepted February 28, 2018, date of publication April 9, 2018, date of current version April 25, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2820898

User-Centric Cloud RAN: An Analytical Framework for Optimizing Area Spectral and Energy Efficiency

UMAIR SAJID HASHMI[1], (Graduate Student Member, IEEE), SYED ALI RAZA ZAIDI[2], (Member, IEEE), AND ALI IMRAN[1], (Member, IEEE)

[1]BSON Lab, School of Electrical and Computer Engineering, The University of Oklahoma, Tulsa, OK 74104, USA
[2]School of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT, U.K.

Corresponding author: Umair Sajid Hashmi (umair.hashmi@ou.edu)

The work of A. Imran was supported by the National Science Foundation under Grant 1619346, Grant 1559483, Grant 1718956, and Grant 1730650.

ABSTRACT In this article, we develop a statistical framework to quantify the area spectral efficiency (ASE) and the energy efficiency (EE) performance of a user-centric cloud based radio access network (UC-RAN) downlink. We propose a user-centric remote radio head (RRH) clustering mechanism, which: 1) provides significant improvement in the received signal-to-interference-ratio through selection diversity; 2) enables efficient interference protection by inducing repulsion among scheduled user-centric RRH clusters; and 3) can self-organize the cluster radius to deal with spatio–temporal variations in user densities. It is shown that under the proposed user-centric clustering mechanism, the ASE (bits/s/Hz/m^2) maximizes at an optimal cluster size. It is observed that this cluster size is sensitive to changes in both RRH and user densities and, hence, must be adapted with variations in these parameters. Next, we formulate the cost paid for the UC-RAN capacity gains in terms of power consumption, which is then translated into the EE (bits/s/Joule) of the UC-RAN. It is observed that the cluster radius which maximizes the EE of the UC-RAN is relatively larger as compared with that which yields maximum ASE. Consequently, we notice that the tradeoff between the ASE and the EE of UC-RAN manifests itself in terms of cluster radius selection. Such tradeoff can be exploited by leveraging a simple two player cooperative game. Numerical results show that the optimal cluster radius obtained from the Nash bargaining solution of the modeled bargaining problem may be adjusted through an exponential weightage parameter that offers a mechanism to utilize the inherent ASE-EE tradeoff in a UC-RAN. Furthermore, in comparison with existing state-of-the-art non user-centric network models, our proposed scheme, by virtue of selective RRH activation and non overlapping user-centric RRH clusters, offers higher and adjustable system ASE and EE, particularly in dense deployment scenarios.

INDEX TERMS User-centric architectures, cloud radio access networks, self organizing networks, area spectral efficiency, energy efficiency, Nash bargaining solution, Poisson point process.

I. INTRODUCTION

A. MOTIVATION

As the mobile data transmission is expected to grow 7-fold from 2016 to 2021 [1], network densification through a conglomeration of diverse technologies (HetNets) seems to be the viable way forward to 5G. Network operators are facing numerous challenges arising from the dense small deployment, high inter-cell interference being the primary culprit. Deployment wise, network densification increases the total cost of ownership (TCO) which includes capital and operational expenditures (CAPEX, OPEX). As the average revenue per user (ARPU) remains virtually flat, network operators are fearing a crisis situation where rising expenses may overcome the dwindling profit margins [2]. To aggravate things further, impromptu cell deployments by mobile users (MUs) render traditional cell planning strategies inept. The aforementioned idiosyncrasies of ultra-dense small cell networks call for a paradigm shift in network design.

Certain promising disruptive 5G technologies such as massive MIMO and mmWave are being considered for higher

average user throughput in 5G. However, both of these technologies, while offering higher network wide capacity, are likely to cost more in terms of energy efficiency and location-independent uniform user Quality of Experience (QoE). In this work, we investigate a User-centric Cloud Radio Access Network (UC-RAN) architecture that has the potential to address the aforementioned challenges. Conventional C-RAN allows centralizing and sharing of the baseband processing between several small cells in a virtual baseband processing unit (BBU) pool [3], [4]. By separating baseband units from the radio access units, the C-RAN architecture: (i) reduces the capital and operational expenditure [4]; (ii) provides huge energy saving (due to centralized air-conditioning etc.) and (iii) provisions implementation of sophisticated coordination mechanisms for reducing the co-channel interference [5]. However, there is one key 5G requirement that conventional C-RAN still fails to address, i.e. QoE. In conventional C-RAN, the Quality of Service (QoS) varies significantly from cell center to cell edge, same way it does in legacy networks, leading to poor QoE. UC-RAN on the other hand has potential to virtually remove cell edges by shifting the pivot of the cell design from the base station (BS) to the mobile user (MU) [6], [7]. The key distinct feature of UC-RAN is that, a cell is built around a user and not around the RRH or BS as in current networks [8]. This enables dynamic coverage as well as higher gains at the user terminals through spatial diversity from having several RRHs available to serve a user [5], [9].

The system design of user-centric architectures for small cell based networks has sparked interest for research in this area, which includes but is not limited to access point grouping mechanisms [6], [10], transmit power control strategies [11], interference alignment [12], RAN selection [13], dynamic load balancing [14] and optimal cluster dimensioning [5], [8], [9]. However, to the best of authors' knowledge, the analytical characterization of the area spectral and energy efficiencies and analysis of the impact different network parameters have on these efficiency metrics remains terra incognita. To this end, in this article, we address some fundamental design questions and propose a novel RRH clustering technique for designing efficient large scale UC-RANs. Furthermore, we present a game theoretic framework to trade-off between ASE and EE in dynamic fashion. This framework allows to retain a pareto-optimal performance while accommodating varying network load and operator's priority between ASE and EE.

B. USER-CENTRIC C-RAN: ARCHITECTURAL OVERVIEW

With the assumption of ultra-dense RRH deployment, which is a plausible scenario in future networks, inter-cell interference and network management issues due to close BS-user spatial proximity become prominent. UC-RAN addresses these by provisioning demand based baseband processing to RRHs and seamless coverage to the users. In a UC-RAN architecture, the BBU pool turns ON the RRHs that are required to serve a desired user at a certain QoS. The dynamic

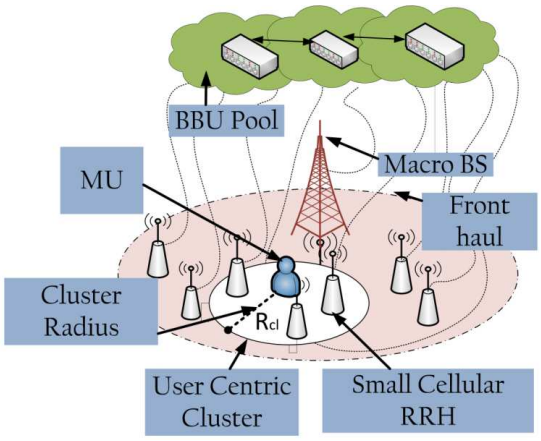

FIGURE 1. UC-RAN architecture with activation disc of radius R_{cl} for a served user.

clustering of RRHs allows users to be served seamlessly and provides uniform service experience regardless of user location and movement [5]. The user-centric cluster size serves as a proxy to the minimum spatial separation between an arbitrary user and its closest interfering RRH, thereby improving the average SINR at the user. Furthermore, contrary to the traditional small cell networks where the energy consumption and the handovers both increase with the density of the small cell RRHs, UC-RAN mitigates this problem by providing on demand coverage, i.e., by selectively turning ON as opposed to always turned ON RRHs. Additionally, the throughput gains provided by distributed diversity alleviate the overhead of cooperation.

Fig.1 provides a graphical illustration of a UC-RAN with RRH clustering. The RRHs are connected to the pool of BBUs via flexible front haul. The front haul is usually an optical fiber where signaling is done using radio-over-fiber (RoF) or common public radio interface (CPRI). Most of the signal processing at baseband level is delegated to the BBUs. The RRH deployment is expected to be very high density by leveraging the existing infrastructure (e.g. street lamp posts, poles, side of buildings etc.). The key idea here is to dynamically select the best RRH (in terms of SIR) within a circular area (virtual cell) of pre-defined radius around selected users (based on scheduling priority) during each scheduling interval. All other RRHs within the circle here after called cluster are kept OFF thereby minimizing the interference. The aforementioned UC-RAN architecture provides two-fold benefits: i) on-demand centralized processing at the BBU pools caters to non-uniform user traffic that subsequently enables OPEX reduction by as much as 30% [15], ii) user-centric RRH clustering reduces the number of nearby interfering RRHs and eliminates cell-edge coverage issues, hence improving the overall user experience.

C. DESIGN ISSUES & PROBLEM STATEMENT

UC-RAN functions on resource pooling and virtual cell formation around scheduled users. The centralized user-centric

RRH clustering not only reduces frequent handovers but also increases the signal-to-interference-plus-noise-ratio (SINR) that subsequently reduces the outage probability in dense networks. This in turn paves the way for increase in the system wide spectral efficiency. Additionally, selective RRH activation enables reduced power consumption, hence making the network more energy efficient. Effectively, the user-centric RRH clustering:

1) empowers UC-RAN to provide demand based coverage, i.e., the coverage can dynamically extend or shrink based on user density by intelligently switching RRHs ON/OFF;

2) enables energy savings as an RRH is only turned ON when required to serve a nearby user;

3) provides an efficient way to control the signal strength which is a function of RRH cluster size (through both the maximum path loss incurred and the number of serving RRHs providing selection diversity gain);

4) enables effective interference protection to an MU by inducing repulsion between RRH clusters (i.e., clusters are not allowed to overlap spatially which induces a natural guard-zone for a scheduled MU).

A critical design parameter in UC-RAN is the RRH cluster size around an arbitrary user. The cluster size may be defined in terms of number of nearby RRHs or an area covered by a circular disk around the user. In our model which is the latter case, the RRHs falling within the circular disk are designated to the corresponding user in a given time slot. Subsequent SIR based RRH activation is performed from amongst the RRHs within the cluster around the user. Increasing the cluster size offers following gains: i) larger distances between user and interfering RRHs results in larger link SINR and thus better link throughput; and ii) a larger cluster yields high macro diversity gain through selection among the larger number of RRHs in the cluster. However, the down side of a larger cluster is reduced spectrum reuse and a lower number of users that can be served simultaneously. This in turn reduces system level capacity. Hence with a larger cluster, there are fewer higher quality links as opposed to many low bit rate links (which occur with smaller cluster radius). Another dimension of the trade-off that cluster radius offers is the energy efficiency. Higher cluster radius keeps more RRHs off as compared to lower cluster radius. In the back drop of these insights the goal of this paper is to investigate following research questions:

- What is the optimal RRH cluster size that maximizes a key performance indicator of capacity, i.e., area spectral efficiency (ASE)?
- What is the cluster radius that yields optimal performance in terms of energy efficiency (EE)?
- What parameters are crucial in defining the optimal cluster sizes that maximize these system efficiencies (ASE, EE)? How sensitive are the efficiencies to variations in these parameters?
- Can we design a self-organizing framework to dynamically adjust the user-centric RRH cluster size and trade between ASE and EE in UC-RAN to cope with the spatio-temporal variations in user traffic?

In this paper, we take the first step towards analytical treatment of the above mentioned design issues and answering the key research question at hand, i.e., what is the optimal cluster size around a scheduled user? Amongst recent works, studies in [5] and [9] are most relevant. However, our analysis differs in three key aspects: 1) [5] and [9] leverage user-centric architectures to optimize virtual cluster radius that maximizes the system capacity. On the other hand, we present a framework to simultaneously analyze ASE and EE in a UC-RAN. 2) Unlike [5] where the proposed clustering is overlapping (scenarios where a single RRH may simultaneously serve multiple MUs), our model builds on non-overlapping user-centric clusters resulting in a one-to-one RRH-MU association during a given time slot. 3) Contrary to analysis in prior studies, we take into account variations in user density. By employing principles from stochastic geometry to model the thinned user and RRH densities in a particular time slot, we analyze the overall system efficiency more accurately. This allows investigation of relationship between key design parameters such as path loss exponent and SINR threshold on ASE and EE for given user and RRH densities.

D. CONTRIBUTIONS & ORGANIZATION

The contribution of this work is six-fold:

1) First, we introduce the user-centric RRH clustering mechanism. Borrowing from well established tools in stochastic geometry [16], we formulate a spatial model for a UC-RAN under the proposed clustering mechanism (Section III). We then characterize the mean and variance for the average aggregate interference (Section IV) experienced by a scheduled MU in a large scale UC-RAN. Our analysis considers both the geometric uncertainty due to the randomness in topology and the channel uncertainty due to small-scale multipath fading (see Section II).
2) We then derive a closed form expression for the lower bound on the link success (which also corresponds to the coverage probability) for a scheduled MU. The bound is employed to establish a lower-bound on the area spectral efficiency of the UC-RAN (Section IV).
3) Shifting our attention from the network level performance to the link level throughput, we consider the scenario where RRHs encode downlink (DL) transmission at maximum supportable rate for a certain reliability constraint. Under this consideration, we characterize the outage capacity of the scheduled MU under the proposed user-centric RRH clustering protocol. We then investigate the scaling behavior of the per user throughput with respect to the density of RRH. It is shown that the aggregate interference contributes to a loss in the distributed diversity gain which is obtained by the RRH selection (see Section V).

4) We then proceed to explore the overhead associated with discovering the best RRH under the proposed protocol in terms of the power consumption (see Section VI). The formulated power consumption is employed with the link level throughput to study the energy efficiency of UC-RAN.
5) We employ the developed analytical framework to investigate the design questions which were formulated in section IC. Specifically, we address the problems of optimal dimensioning of the cluster radius and selection of the RRH deployment density (Section VII). We investigate the impact of different parametric variations on these design issues and highlight the need for a self-organizing network (SON) features [17] to cope with the varying user densities. It is shown that there exists an optimal cluster radius which maximizes the energy efficiency of the network. However, such an optimal cluster size is not necessarily same as that which maximizes the area spectral efficiency. Consequently, there exists a trade-off between energy and area spectral efficiency of the UC-RAN.
6) Lastly, we model the inherent ASE-EE trade-off in UC-RANs through a bargaining problem [18] where the performance metrics are modeled as virtual game players and a Nash bargaining solution is found that corresponds to a unique optimal cluster radius for a given set of network parameters. Using an exponential weightage parameter in the optimization framework, we vary the bargaining powers of the players and show that the ASE-EE trade-off may be adjusted in real-time as a function of the network operator's spatio-temporal revenue model which may include traffic intensity, time of the day and hotspot locations (e.g. cafes, stadiums) [19] (see Section VIII).

E. NOTATION

Throughout this paper we use $\mathbb{E}_Z(.)$ to denote the expected value of a random variable Z. A particular value of random variable Z is denoted by z. The probability density function (PDF) of a random variable z is denoted by $f_Z(.)$. The bold face lower case letter (e.g. $\mathbf{x}$) is employed to denote a vector in $\mathbb{R}^2$. For sake of compactness, we employ $\mathbf{x}$ to refer to the vector itself and its location as well. The symbol $\setminus$ denotes set subtraction and $\|x\|$ denotes Euclidean norm of the vector x. The symbol b($\mathbf{x}$, r) denotes a ball of radius r centered at a point $\mathbf{x}$. The symbol $\in$ denotes set membership and Π is used to denote the point process. The point process is also used as a counting measure by using the notation $\Pi(\mathcal{A})$ which returns the number of points in Π which lie inside $\mathcal{A} \in \mathbb{R}^2$. The symbol $Z \sim \mathcal{U}(a, b)$ is used to denote a random variable which takes values between a and b with uniform probability. Similarly, $Z \sim \mathcal{E}(\mu)$ is used to denote an exponential random variable with mean μ. The symbol $\mathbb{1}(x > y)$ represents an indicator which is one if the condition $(x > y)$ is satisfied and 0 otherwise.

II. NETWORK MODEL

A. SPATIAL MODEL OF THE NETWORK

We consider a cloud radio access network under-laid within a large-scale cellular network. Both the small cell RRHs and MUs are assumed to be spatially distributed across the macrocells (see Fig. 1). The spatial distribution of the RRHs and the MUs is captured by two independent stationary Poisson point processes (SPPPs): $\Pi_{CLR} \in \mathbb{R}^2$ and $\Pi_{MU} \in \mathbb{R}^2$ with intensities λ_{CLR} and λ_{USR} respectively. Specifically, at an arbitrary time instant, the probability of finding $n_i \in \mathbb{N}, i \in \{RRH, MU\}$ RRHs/MUs inside a typical macrocell with area foot-print $\mathcal{A} \subseteq \mathbb{R}^2$ follows the Poisson law with mean measure $\Lambda_i(\mathcal{A}) = \lambda_i v_2(\mathcal{A})$. The mean measure is characterized by the average number of RRHs/MUs per unit area (i.e. λ_{CLR} and λ_{USR}) and the Lebesgue measure [16] $v_2(\mathcal{A}) = \int_{\mathcal{A}} d\boldsymbol{x}$ on $\mathbb{R}^2$, where if $\mathcal{A}$ is a disc of radius r then $v_2(\mathcal{A}) = \pi r^2$ is the area of the disc.

B. CHANNEL MODEL

The channel between a UC-RAN RRH $\boldsymbol{x} \in \Pi_{CLR}$ and an arbitrary MU $\boldsymbol{y} \in \Pi_{MU}$ is modeled by $h_{xy} l(||\boldsymbol{x} - \boldsymbol{y}||)$. Here $h_{xy} \in \mathcal{E}(1)$ is a unit mean exponential random variable which captures the impact of a Rayleigh fading channel between an RRH and an MU. The small-scale Rayleigh fading is complemented by a large-scale path loss modeled by $l(||\boldsymbol{x} - \boldsymbol{y}||) = K||\boldsymbol{x} - \boldsymbol{y}||^{-\alpha}$ power-law function. Here $||\boldsymbol{x} - \boldsymbol{y}||$ is the distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, K is a frequency dependent constant and $\alpha \geq 2$ is an environment/terrain dependent path loss exponent. The fading channel gains are assumed to be mutually independent and identically distributed (i.i.d.). Without any loss of generality, we will assume $K = 1$ for the rest of this discussion. It is assumed that the communication is interference limited and hence the thermal noise is negligible. Furthermore, we assume that all RRHs employ the same transmit power P_{CLR}.

III. USER-CENTRIC CLUSTERING IN UC-RAN

In this article, we propose a user-centric clustering mechanism for the UC-RANs. More specifically, we envision a scenario where out of the multitude of small cell RRHs deployed in close proximity of an intended MU, a single RRH that provides the best channel gain (and consequently the highest signal-to-interference ratio (SIR)) is activated to serve that MU. The proximity or neighborhood of an MU is characterized by the cluster radius R_{cl}. The proposed user-centric clustering mechanism (Algorithm 1, Fig.2) yields Π'_{MU} and Π'_{CLR} which is the set of scheduled MUs and activated RRHs during a particular time slot respectively.

As specified by Algorithm 1, the macro-cell or the BBU data center assigns a mark/tag $p_{USR} \sim \mathcal{U}(0, 1)$ to each MU. These marks correspond to the downlink scheduling priority of the MUs. More specifically, the lower the value of the mark, the higher is the priority of the user to be served by the RRHs. Effectively, these marks can be thought of as the timers corresponding to each MU which are decremented on

Algorithm 1 User-Centric Clustering for C-RAN

Data: Π_{MU},Π_{CLR},R_{cl}
Result: Π'_{MU},Π'_{CLR}
$\Pi'_{MU} \leftarrow \emptyset$,$\Pi'_{CLR} \leftarrow \emptyset$;
foreach $\boldsymbol{x} \in \Pi_{MU}$ **do**
  $F_{PRIO}[\boldsymbol{x}] \leftarrow \mathcal{U}(0,1)$
end
foreach $\boldsymbol{x} \in \Pi_{MU}$ **do**
  $p^{\{x\}}_{USR} \leftarrow F_{PRIO}[\boldsymbol{x}]$ $p^{\{x\}}_{SCH} \leftarrow 1$ **foreach** $\boldsymbol{y} \in \Pi_{MU}$ **do**
    if $\boldsymbol{y} \neq \boldsymbol{x}$ **then**
      if $\boldsymbol{y} \in b(\boldsymbol{x}, 2R_{cl})$ *and* $p^{\{y\}}_{USR} > p^{\{x\}}_{USR}$ **then**
        $p^{\{x\}}_{SCH} \leftarrow 0$
      else
        continue
      end
    else
      continue
    end
  end
  foreach $\boldsymbol{r} \in \Pi_{CLR}$ **do**
    if $\boldsymbol{r} \in b(\boldsymbol{x}, R_{cl})$ *and* $p^{\{x\}}_{SCH} == 1$ **then**
      if $h_{rx} l(||\boldsymbol{r} - \boldsymbol{x}||) > h_{r'x} l(||\boldsymbol{r'} - \boldsymbol{x}||), \forall \boldsymbol{r'} \in \Pi_{CLR}, \boldsymbol{r'} \in b(\boldsymbol{x}, R_{cl}), \boldsymbol{r'} \neq \boldsymbol{r}$ **then**
        $\Pi'_{CLR} \cup \{\boldsymbol{r}\}$
      end
    end
  end
  if $p^{\{x\}}_{SCH} == 1$ **then**
    $\Pi'_{MU} \cup \{\boldsymbol{x}\}$
  end
end

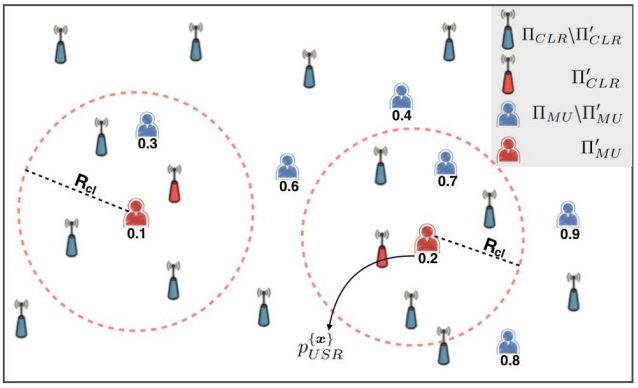

FIGURE 2. User-centric RRH clustering in a UC-RAN. Each scheduled user is served by a single RRH in its respective cluster that maximizes its received SIR. The user-centric scheduling is based on $p^{\{x\}}_{USR}$ values with lower marks corresponding to high scheduling priorities.

each time slot where service to this MU is deferred. A MU is scheduled for a downlink transmission iff it has highest scheduling priority in its neighborhood. In other words, there is no other MU in a disc of radius R_{cl} centered at MU with a higher priority. This round robin scheduling scheme ensures fair DL scheduling among MUs.[1] Notice that this disc also characterizes the size of the RRH cluster from which MU is being served. For a fixed R_{cl}, the percentage of MUs served in a given transmission time interval (TTI) is a function of relative RRH and MU PPP densities, i.e., if $\lambda_{USR} >> \lambda_{CLR}$, the average wait time before an arbitrary MU is served will be longer as compared to the scenario with same order MU and RRH densities.

The activation of RRHs is coupled with the user-centric scheduling mechanism (Algorithm 1). Only the RRHs which lie in the neighborhood of the scheduled users and provide the best propagation channel gain to their respective MUs are activated by the macro base station (MBS) (or BBU pool). This implies that each scheduled MU has a set of nearby RRHs that defines its user-centric RRH cluster. From this cluster of RRHs, only one that yields the highest SINR at the user is activated. Consequentially, there is at max one activated RRH that lies within a user-centric circular disk of area πR^2_{cl}. Effectively, activation of RRHs is on demand basis which provides UC-RAN capability of self-organizing the coverage to cope with the spatio-temporal variations of the user demography.

One might argue that such a non overlapping user-centric clustering scheme may result in service holes, i.e. there may exist MUs that are not associated with any RRHs due to empty RRH clusters around those MUs. Since we are considering dense small cell deployments with comparable λ_{CLR} and λ_{USR}, user-centric RRH clusters with realistic R_{cl} will hardly be void. In the unlikely scenario of a void cluster though, user clustering strategies [20] may be employed where nearby MUs are grouped together and optimization is performed on the MU clusters rather than individual MUs.[2] Furthermore, it is known that best RRH activation with a proximity constraint provides dual benefits of low outage probability and high power efficiency in dense deployment scenarios [21].

IV. QUANTIFYING THE AREA SPECTRAL EFFICIENCY OF A UC-RAN

In the previous section, we presented an outline of a user-centric clustering algorithm for a UC-RAN. As is obvious from the algorithm, the size of the cluster employed for scheduling is a critical system design parameter. Optimal dimensioning of the R_{cl} is necessitated by the fact that:

1) The cluster size determines the number of the active RRHs at any given time. In turn the density of active RRHs shapes the co-channel interference experienced by a scheduled MU.
2) The radius of the cluster, also characterizes the number of concurrently scheduled MUs per unit area.
3) The dimensions of a cluster also determine the number of RRHs serving a scheduled MUs. This in turn

[1]The case with MUs having non-uniform scheduling priorities will be covered in future extensions of this work.

[2]In the interest of space, detailed discussion and evaluation of MU clusters will be presented in future publications.

determines the diversity gain experienced due to spatially distributed RRHs.

The area spectral efficiency of a UC-RAN network is strongly coupled with these three factors. In this section, our objective is two fold:

- To characterize the area spectral efficiency of a large scale UC-RAN.
- To investigate the optimal dimensioning of the cluster radius for maximizing the throughput potential of the UC-RAN.

A. SIGNAL MODEL

Consider a scheduled user $\boldsymbol{x} \in \Pi'_{MU}$. Let $\mathcal{S}_{cop}(\boldsymbol{x}, R_{cl}) = \Pi'_{CLR} \cap (\boldsymbol{x}, R_{cl})$ be the unique RRH which is fed by the same BBU as $\boldsymbol{x}$ and selected to serve $\boldsymbol{x}$ on the basis of the scheduling criteria (Algorithm 1). Furthermore, let $\Pi_I = \Pi'_{CLR} \backslash \mathcal{S}_{cop}(\boldsymbol{x}, R_{cl})$ be the set of RRHs which are concurrently scheduled to serve $\boldsymbol{y} \neq \boldsymbol{x}, \forall \boldsymbol{y} \in \Pi'_{MU}$. In this article, we assume that the UC-RAN employs the RRH selection to serve its intended MU. Under RRH selection transmission the received signal at $\boldsymbol{x}$ can be written as (1), as shown at the top of the next page, where $\max_{\boldsymbol{i} \in \mathcal{S}_{cop}} h_{\boldsymbol{ix}} l(||\boldsymbol{x} - \boldsymbol{i}||)$ is the channel gain between the serving RRH and the MU $\boldsymbol{x}$, $\max_{\boldsymbol{j} \in \Pi'_{CLR} \cap (\boldsymbol{y}, R_{cl})} h_{\boldsymbol{xj}} l(||\boldsymbol{x} - \boldsymbol{j}||)$ is the interference experienced at $\boldsymbol{x}$ due to RRH $\boldsymbol{j}$ serving MU $\boldsymbol{y}$ and $s_{\boldsymbol{k}}$ is the message signal transmitted to MU $\boldsymbol{k}$ by its selected RRH. P_{CLR} is the transmit power employed by the RRH and $\varphi_{\boldsymbol{x}}$ is the additive white Gaussian noise (AWGN) at the receiver front end.[3]

B. PROBE CLUSTER

In order to characterize the area spectral efficiency of a UC-RAN, quantification of the success probability for a scheduled MU is essential. The key hurdle in characterizing the performance arises from the fact that unlike Π_{MU} the point process of the scheduled users Π'_{MU} is non-stationary. A closer inspection of Π'_{MU} reveals that it is indeed a modified version of a type II Matern hardcore process [16]. Fortunately, for such processes it is well established that they can be approximated by an equidense SPPP with appropriate modified intensity [22], [23].[4] Consequently, Π'_{MU} can be approximated by a SPPP with the intensity

$$\lambda_{USR}^{\{EF\}} = \frac{1 - \exp(-4\pi \lambda_{USR} R_{cl}^2)}{4\pi R_{cl}^2}. \tag{2}$$

Notice that since the user-centric clusters are considered non over lapping, the minimum distance between any two user-centric clusters should be $2R_{cl}$. Exploiting the stationary characteristics of the resultant SPPP, it is sufficient to focus on a typical MU. According to Silvnyak's theorem [16], the law of the SPPP does not change by addition of a single point. Hence we add a probe MU at origin. Moreover, the received signal $(r_{x(o)})$ in Eq. (1) can now be simplified with $r_i = ||\boldsymbol{i} - \boldsymbol{o}||$ and $r_y = ||\boldsymbol{y} - \boldsymbol{o}||$. For the sake of compactness, we will drop the index o for the rest of the discussion (e.g., $h_{io} = h_{\boldsymbol{i}}$).

C. LOWER-BOUND ON THE SUCCESS PROBABILITY OF SCHEDULED MU

From Eq. (1) the received SIR at the probe MU can be expressed as

$$\text{SIR} = \Gamma_{MU} = \frac{\max_{\boldsymbol{i} \in \mathcal{S}_{cop}} h_{\boldsymbol{i}} l(r_{\boldsymbol{i}})}{\sum_{\boldsymbol{j} \in \Pi_I} h_{\boldsymbol{j}} l(r_{\boldsymbol{j}})}. \tag{3}$$

Notice that $\mathcal{S}_{cop}$ is a function of the non-stationary Poisson point process Π'_{CLR}.

Proposition 1 (Moments of Aggregate Interference): The mean and variance of the aggregate interference experienced by a typical MU during a user-centric algorithm can be approximated as follows

$$\begin{aligned} \kappa_1 &= \mathbb{E}(I) \\ &= \frac{2\pi \lambda_{CLR}[1 - \exp(-[1 - \exp(-4\pi \lambda_{USR} R_{cl}^2]/4)]}{(\alpha - 2)(R_{cl})^{\alpha-2}(\lambda_{CLR} \pi R_{cl}^2)}, \\ \kappa_2 &= \mathbb{V}(I) \\ &= \frac{\pi \lambda_{CLR}[1 - \exp(-[1 - \exp(-4\pi \lambda_{USR} R_{cl}^2]/4)]}{(\alpha - 1)(R_{cl})^{2(\alpha-1)}(\lambda_{CLR} \pi R_{cl}^2)}, \end{aligned} \tag{4}$$

where λ_{CLR} is the density of the UC-RAN RRHs, α is the path loss exponent and R_{cl} is the radius of UC-RAN cluster.

Proof: Consider the SPPP Π_{CLR}, then under user-centric clustering algorithm, for each scheduled user, only a single RRH which resides in the vicinity as well as provides maximum channel gain to that user is activated by the macrocell. A natural implication of this policy is that the resulting PPP Π'_{CLR} is non-stationary. However, like Π'_{MU} it can be approximated with an equivalent SPPP with modified density $\lambda_{CLR} \cdot p_{ACT}$. Here p_{ACT} is the activation probability for the RRH and can be computed as (5), as shown at the top of the next page, where (a) follows from the fact that a BS is only activated if: i) there is a scheduled user within distance R_{cl}, and ii) there is no other BS within this distance of that user that can provide better channel gain. Now noticing that $\Pi_I = \Pi'_{CLR} \backslash \mathcal{S}_{cop}(\boldsymbol{o}, R_{cl})$, we can precisely describe $\Pi_I = \Pi'_{CLR} \backslash b(\boldsymbol{o}, R_{cl})$. Hence the mean and the variance can be computed using Campbell's theorem [16] as follows

$$\begin{aligned} \kappa_1 = \mathbb{E}(I) &= \mathbb{E}\left(\sum_{\boldsymbol{j} \in \Pi'_{CLR} \backslash b(\boldsymbol{o}, R_{cl})} h_{\boldsymbol{j}} l(r_{\boldsymbol{j}}) \right), \\ &= 2\pi \lambda_{CLR} p_{ACT} \int_{R_{cl}}^{\infty} \mathbb{E}(H) r^{1-\alpha} dr, \end{aligned} \tag{6}$$

[3] We notice that even with the induced spatial repulsion between MUs which consequently thins the PPPs of serving MUs and activated RRHs, the noise power at served MUs is negligible as compared to the aggregate interference. In other words, the network operates in a saturated conditions and the links are interference limited.

[4] This is validated by employing Ripley's K function [16] for both PPPs. It is observed that the K function of equi-dense PPP forms a lower bound on the Ripley's K function of scheduled MUs. The bound is very tight as the results are indistinguishable. In subsequent discussion, the equi-dense approximation is further validated by comparing theoretical area spectral efficiency of C-RAN against Monte-carlo simulation results (see Fig. 3).

$$r_x = \sqrt{P_{CLR} \max_{\boldsymbol{i}\in\mathcal{S}_{cop}} h_{\boldsymbol{i}\boldsymbol{x}} l(||\boldsymbol{x}-\boldsymbol{i}||)} s_{\boldsymbol{x}} + \sum_{\boldsymbol{y}\in\Pi'_{MU}, \boldsymbol{y}\neq\boldsymbol{x}} \sqrt{P_{CLR} \max_{\boldsymbol{j}\in\Pi'_{CLR}\cap(\boldsymbol{y},R_{cl})} h_{\boldsymbol{j}\boldsymbol{x}} l(||\boldsymbol{x}-\boldsymbol{j}||)} s_{\boldsymbol{y}} + \varphi_{\boldsymbol{x}}. \tag{1}$$

$$\begin{aligned} p_{ACT} &\overset{(a)}{=} \Pr\left[\Pi'_{MU} \cap b(\boldsymbol{r}, R_{cl}) \neq \emptyset | \boldsymbol{r} \in \Pi'_{CLR}\right].\{h_{\boldsymbol{r}} l(r_{\boldsymbol{r}}) > h_{\boldsymbol{j}} l(r_{\boldsymbol{j}}) | \boldsymbol{j} \in \Pi'_{CLR}, \boldsymbol{j} \neq \boldsymbol{r}\}\Big], \\ &= \left[1 - \Pr\{\Pi'_{MU} \cap b(\boldsymbol{r}, R_{cl}) = \emptyset | \boldsymbol{r} \in \Pi'_{CLR}\}\right].\Pr\{h_{\boldsymbol{r}} l(r_{\boldsymbol{r}}) > h_{\boldsymbol{j}} l(r_{\boldsymbol{j}}) | \boldsymbol{j} \in \Pi'_{CLR}, \boldsymbol{j} \neq \boldsymbol{r}\}, \\ &= \left[1 - \exp(-\lambda^{\{EF\}}_{USR} \pi R^2_{cl})\right].(1/[\lambda_{CLR}\pi R^2_{cl}]), \\ &= \frac{1 - \exp(-[1 - \exp(-4\pi\lambda_{USR} R^2_{cl}]/4)}{\lambda_{CLR}\pi R^2_{cl}}. \end{aligned} \tag{5}$$

and similarly

$$\kappa_2 = 2\pi\lambda_{CLR} p_{ACT} \int_{R_{cl}}^{\infty} \mathbb{E}(H^2) r^{1-2\alpha} dr. \tag{7}$$

Substituting $\mathbb{E}(H) = \mathbb{E}(H^2) = 1$ in the (6) and (7) concludes the proof. ■

Remarks:

1) From (4), we notice that the average aggregate interference experienced by an MU increases with an increase in the user density. For the fixed density of RRH, the only parameter that designer can adjust to compensate for the increase in the user density is to reduce the size of the cluster. While reducing the cluster size will increase the number of RRHs activated by accommodating more users, it also reduces the interference protection available to each MU link. More specifically, the small number of large clusters or large number of small clusters may lead to a similar co-channel interference environment.
2) The average interference experienced by an MU decreases with an increase in path loss. This follows from the fact that with an increase in path loss, signals attenuate more rapidly and hence the aggregate interference power is reduced. However, the signal strength is also reduced for the same reason.

Proposition 2 (Link Success Probability for a Scheduled MU): The link success probability of the probe MU served under the proposed user centric clustering and RRH selection scheme algorithm can be lower-bound as follows

$$\mathbb{P}_{suc}(\gamma_{th}, R^2_{cl}) \geq 1 - \exp\left(-\frac{\lambda_{CLR}\pi\delta}{\gamma^{\delta}_{th}\kappa^{\delta}_1} \gamma(\delta, \gamma_{th}\kappa_1 R^{\alpha}_{cl})\right), \tag{8}$$

where γ_{th} is the MU's desired SIR threshold, $\delta = \frac{2}{\alpha}$ and $\gamma(a,b) = \int_0^b t^{\alpha-1}\exp(-t)dt$ is the lower incomplete Gamma function.

Proof: Consider the probe MU scheduled under the proposed clustering mechanism, the link success probability for this MU is given by

$$\begin{aligned} \mathbb{P}_{suc}(\gamma_{th}, R^2_{cl}) &= \Pr\{\Gamma_{MU} > \gamma_{th}\}, \\ &= 1 - \Pr\{\Gamma_{MU} \leq \gamma_{th}\}, \\ &= 1 - \underbrace{\mathbb{E}_I[\Pr\{\max_{\boldsymbol{i}\in\mathcal{S}_{cop}} h_{\boldsymbol{i}} l(r_{\boldsymbol{i}}) \leq I\gamma_{th}\}]}_{A_1}. \end{aligned} \tag{9}$$

The term $A_1 = \Pr\{\max_{\boldsymbol{i}\in\mathcal{S}_{cop}} h_{\boldsymbol{i}} l(r_{\boldsymbol{i}}) \leq I\gamma_{th}\}$ can be computed by noticing the fact the $\mathcal{S}_{cop}$ is a SPPP inside a finite area $b(\boldsymbol{o}, R_{cl})$ and we can construct a Marked PPP by assigning the fading marks to each $\boldsymbol{i} \in \mathcal{S}_{cop}$.[5] Additional Bernoulli or indicator marks are assigned to the PPP such that the intensity of modified process[6] can be expressed as

$$\lambda_S(r, h) = \lambda_{CLR} 2\pi r \mathbb{1}(hl(r) \geq I\gamma_{th}) f_H(h). \tag{10}$$

Now A_1 can be computed by the void probability of the modified point process as

$$A_1 = \exp\left(-\underbrace{\int_0^{\infty}\int_0^{R_{cl}} \lambda_s(r,h) dr dh}_{\Lambda_s}\right), \tag{11}$$

The mean measure Λ_s can be evaluated by

$$\begin{aligned} \Lambda_S &= \lambda_{CLR} 2\pi \int_0^{R_{cl}}\int_0^{\infty} r\mathbb{1}(hl(r) \geq I\gamma_{th}) f_H(h) dr dh, \\ &\overset{(a)}{=} \lambda_{CLR} 2\pi \int_0^{R_{cl}} r \Pr(h \geq I\gamma_{th} r^{\alpha}) dr, \\ &= \lambda_{CLR} 2\pi \int_0^{R_{cl}} r \exp(-I\gamma_{th} r^{\alpha}) dr, \\ &= \frac{\lambda_{CLR}\pi\delta}{\gamma^{\delta}_{th} I^{\delta}} \gamma(\delta, \gamma_{th} I R^{\alpha}_{cl}), \end{aligned} \tag{12}$$

where (a) follows from the CDF of the exponential function. Employing (11) and (12), we obtain

$$\mathbb{P}_{suc}(\gamma_{th}, R^2_{cl}) = 1 - \mathbb{E}_I\left[\exp\left(-\frac{\lambda_{CLR}\pi\delta}{\gamma^{\delta}_{th} I^{\delta}} \gamma(\delta, \gamma_{th} I R^{\alpha}_{cl})\right)\right],$$

[5] A detailed discussion on the Marked PPP is beyond the scope of this article. Interested readers should refer to [16].

[6] The modified intensity corresponds to the dependently thinned point process.

$$\overset{(b)}{\geq} 1 - \exp\left(-\frac{\lambda_{CLR}\pi\delta}{\gamma_{th}^{\delta}\kappa_1^{\delta}}\gamma(\delta, \gamma_{th}\kappa_1 R_{cl}^{\alpha})\right).$$

where $\kappa_1 = \mathbb{E}_I(I)$ from (4) and (b) follows from the Jensen's inequality. ■

The area spectral efficiency of the large scale UC-RAN is defined as the number of bits/s which can be transmitted over a unit Hertz bandwidth per second in the area of 1 square meter. In other words, the area spectral efficiency measures the amount of information that is flowing through a unit area when one Hertz of bandwidth is employed. The lower bound on the link success probability (which is equivalent to the coverage probability) can be employed to establish a lower bound on the area spectral efficiency of the UC-RAN as

$$\mathcal{T}_{CLR} = \lambda_{USR}^{\{EF\}} \log_2(1+\gamma_{th})\mathbb{P}_{suc}(\gamma_{th}, R_{cl}^2), \tag{13}$$

where $\lambda_{USR}^{\{EF\}}$ is the effective density of the scheduled user defined in (2). As is clear from (13), the area spectral efficiency of the UC-RAN is strongly coupled with the cluster size. Intuitively, increasing the cluster size decreases the effective number of scheduled users. However it also increases both the SIR (due to lower number of nearby interfering RRHs) and the interference protection margin. Essentially, this implies that there exists an optimal radius for the cluster which will balance these two opposite effects to maximize the attainable area spectral efficiency.

V. OUTAGE CAPACITY AND IMPACT OF CO-CHANNEL INTERFERENCE

Until now we have focused on the scenario, where each MU has a certain desired QoS requirement which is reflected in their desired SIR threshold. We have developed a statistical framework to quantify the area spectral efficiency of the large scale UC-RAN under the proposed user-centric clustering mechanism. The area spectral efficiency is the measure of network wide performance of the UC-RAN. Nevertheless, consider a scenario where instead of a fixed desired SIR threshold, the MU's QoS is reflected by a reliability threshold ρ which upper-bounds the downlink outage probability. Then, under such a constraint, the scheduled RRH can encode the transmission at a maximum rate C_ρ for each MU. This rate measures the spectral performance on the downlink for an arbitrary MU and is known as the outage capacity [24]. Notice that the outage capacity is a link level performance metric. In this section, our aim is to:

1) derive a closed form expression for the bounds on the outage capacity of the user centric UC-RAN;
2) establish the scaling laws for per user throughput with respect to the RRH density;
3) explore the loss in diversity due to aggregate interference.

A. OUTAGE CAPACITY OF THE C-RAN DOWNLINK

The outage capacity (C_ρ) of the downlink between the probe MU and its serving RRH is defined as

$$C_\rho = \sup\{C_o : \mathbb{P}_{out}(c_o) = 1 - \mathbb{P}_{suc}(2^{C_o} - 1, R_{cl}^2) \leq \rho\}. \tag{14}$$

An upper-bound on the outage probability at a certain desired rate C_o can be obtained from (8) as follows,

$$\mathbb{P}_{out}(C_o) \leq \exp\left(-\frac{\lambda_{CLR}\pi\delta}{((2^{C_o}-1)\kappa_1)^{\delta}}\gamma(\delta, (2^{C_o}-1)\kappa_1 R_{cl}^{\alpha})\right). \tag{15}$$

Proposition 3 (Outage-Capacity of the Interference Limited MU Link): The ρ-outage capacity of the MU scheduled under the proposed scheme can be upper bounded as:

$$C_\rho \leq \log_2\left(1 + \frac{\lambda_{CLR}^{\frac{1}{\delta}-1}}{\bar{\kappa}_1 \ln(\rho^{-1})^{\frac{1}{\delta}}}\right) \text{(bits/s/Hz)} \tag{16}$$

where

$$\bar{\kappa}_1 = \frac{2\pi(\pi\Gamma(1+\delta))^{\frac{1}{\delta}}[1-\exp(-[1-\exp(-4\pi\lambda_{USR}R_{cl}^2)]/4)]}{(\alpha-2)-(R_{cl})^{\alpha-2}(\lambda_{CLR}\pi R_{cl}^2)}.$$

Proof: The outage probability expression in (15) can be upper-bounded using the fact that $\gamma(\delta, \gamma_{th}\kappa_1 R_{cl}^{\alpha}) \leq \Gamma(\delta)$[7] to give

$$\mathbb{P}_{out}(C_o) \leq \exp\left(-\frac{\lambda_{CLR}\pi\delta\Gamma(\delta)}{((2^{C_o}-1)\kappa_1)^{\delta}}\right) \leq \rho, \tag{17}$$

where $\delta\Gamma(\delta) = \Gamma(\delta+1)$. Bounding the outage probability by the desired reliability constraint and employing the definition in (14) along with some mathematical manipulations provides an upper-bound on C_ρ. ■

In order to gain further insights, let us define the effective SIR under the proposed user centric scheme as

$$\Gamma_{CRAN}^{I} = \frac{\lambda_{CLR}^{\frac{1}{\delta}-1}}{\bar{\kappa}_1 \ln(\rho^{-1})^{\frac{1}{\delta}}}. \tag{18}$$

The $C_\rho^I = log_2(1 + \Gamma_{CRAN}^I)$ is an increasing function of Γ_{CRAN}^I. Let us consider the case where each scheduled user can be assigned a separate frequency band. Effectively, the transmissions are noise limited rather than interference limited. In order to characterize the impact of co-channel interference, we need to quantify C_ρ for a noise limited user-centric C-RAN.

[7]Notice that for the reasonable parametric value of R_{cl}, the term R_{cl}^{α} is large and hence the bound is tight.

B. OUTAGE CAPACITY UNDER NOISE-LIMITED SCENARIO

Proposition 4 (Outage Capacity of a Noise-Limited MU Link): The ρ-outage capacity for a downlink MU transmission in a large scale interference free UC-RAN in the presence of additive white Gaussian noise at the receiver front-end is given by

$$C_{\rho}^{N} \leq \log_2\left(1+\frac{\lambda_{CLR}^{\frac{1}{\delta}}}{\kappa_2 \ln(\rho^{-1})^{\frac{1}{\delta}}}\right)(\text{bits/s/Hz}), \qquad (19)$$

where $\kappa_2 = (\pi\Gamma(\delta+1))^{\frac{1}{\delta}}P/\sigma^2$, σ^2 is noise variance and P is the transmit power employed by RRH.

Proof: The proof follows similar steps as in Proposition 3, with the only difference is that $\kappa_1 = \mathbb{E}(I)$ should be replaced with σ^2/P which can be interpreted as $\gamma^{-1} = \text{SNR}^{-1}$, i.e., the SNR in the absence of fading and path loss. ■

Similar to the interference limited case, we can define the effective SNR as $\Gamma_{CRAN}^{N} = \lambda_{CLR}^{1/\delta}/(\kappa_2 ln(\rho^{-1})^{1/\delta})$. Comparing Γ_{CRAN}^{N} with the Γ_{CRAN}^{I} reveals the impact of co-channel interference, as follows:

1) The power-gain obtained due to distributed diversity provided by the RRHs scales as $\mathcal{O}(\lambda_{CLR}^{1/\delta-1})$ for the interference limited scenario, while the scaling in the noise limited scenario follows $\mathcal{O}(\lambda_{CLR}^{1/\delta})$.
2) The effective SIR for the interference limited scenario is independent of the transmit power. Consequently, the number of bits transmitted with the desired reliability threshold ρ per Hertz usage of bandwidth cannot be increased by increasing the transmit power. This is contrary to the noise-limited scenario.

VI. THE COST OF DIVERSITY GAIN: ENERGY EFFICIENCY PERSPECTIVE

In previous sections, we focused on the spectral performance of the proposed UC-RAN. While the proposed user-centric mechanism exploits centralized processing in cloud to harness the distributed diversity gains, an important issue from network operator/designers perspective is the cost associated with these gains. More specifically, from an energy consumption perspective the cost-benefit analysis can be formulated in terms of energy efficiency. The network or link level energy efficiency characterizes the number of bits that can be transmitted per unit usage of available spectrum at the expense of one Joule in one second.

Due to a large spatio-temporal variance in user traffic, energy efficiency can be significantly improved in dense urban environment through efficient ON/OFF activation [25]. In order to quantify the energy consumption-throughput trade-off, our prime focus here is the energy consumption associated in discovering the best RRH for the association. To that end, we only focus on this additional energy which is required for the discovery purpose and can be considered as the overhead incurred for harnessing the diversity gain. Note that during the discovery process, each RRH is required to estimate the channel gain from the scheduled MU which comes at the expense of energy dissipation.

A. POWER CONSUMPTION MODEL

The power consumption of a standalone RRH was investigated in the project EARTH [26]. The proposed power consumption model provides accurate estimates of dissipated power in different building blocks such as antenna interface, cooling, power amplifiers and baseband processing. The model was extended by parameterization with the C-RAN efficiency in [27]. In this article, our primary interest is to compute the total power consumed in the discovery process in each user cluster. Thus, we propose a modified power consumption model which is inspired by [27] and [26]. Mathematically, the power consumption of the discovery process can be quantified as:

$$P_{CRAN} = \omega_{CRAN}(\text{M},\theta)P_0 + \Delta_u P_u + P_{ou}, \qquad (20)$$

where P_u is the transmit power employed by the MU, P_o is the fixed power consumption of the RRH in listening mode, Δ_u is the radio frequency dependent component of power consumption at the MU, $\omega_{CRAN}(\text{M},\theta)$ is the the UC-RAN coefficient and P_{ou} is the fixed circuit power consumed at the MU. The UC-RAN coefficient is coupled with the number of RRHs in each cluster (denoted by M) and a parameter θ which parameterizes the implementation efficiency. More specifically, $\omega_{CRAN}(\text{M},\theta) \leq \text{M}$ captures the performance gains due to consolidated architecture of UC-RAN. The lower the value of $\omega_{CRAN}(\text{M},\theta)$, the lesser is the amount of power dissipated in each cluster. A simple parameterization of the efficiency coefficient from can be obtained as follows:

$$\omega_{CRAN}(\text{M},\theta) = \theta\text{M}, \quad 0 \leq \theta \leq 1 \qquad (21)$$

where $\theta = 1$ captures the least efficient UC-RAN implementation. The mathematical expression for determining average number of RRHs in each cluster (M) is given in Lemma1.

Lemma I: The average number of RRHs within each user-centric cluster, i.e. M, is the complement of the void probability of the RRHs, i.e. $\text{M} = 1 - e^{-\lambda_{CLR}\pi R_{cl}^2}$.

Proof: Consider that Π_{CLR} is an SPPP with intensity λ_{CLR}, then under user-centric scheme, the average number of RRHs within a circular area of radius R_{cl} is given by $\lambda_{CLR}\pi R_{cl}^2$. Since each user-centric cluster can have at most one RRH, the average number of RRHs is the complement of the probability that an arbitrary cluster would at least one RRH within its foot-prints, i.e.

$$\begin{aligned}\text{M} &= \Pr\{\Pi_{CLR} \cap b(\boldsymbol{x}, R_{cl}) \neq \emptyset | \boldsymbol{x} \in \Pi'_{MU}\},\\ &= 1 - \Pr\{\Pi_{CLR} \cap b(\boldsymbol{x}, R_{cl}) = \emptyset | \boldsymbol{x} \in \Pi'_{MU}\},\\ &= 1 - \exp\{-\pi\lambda_{CLR}R_{cl}^2\}.\end{aligned}$$

■

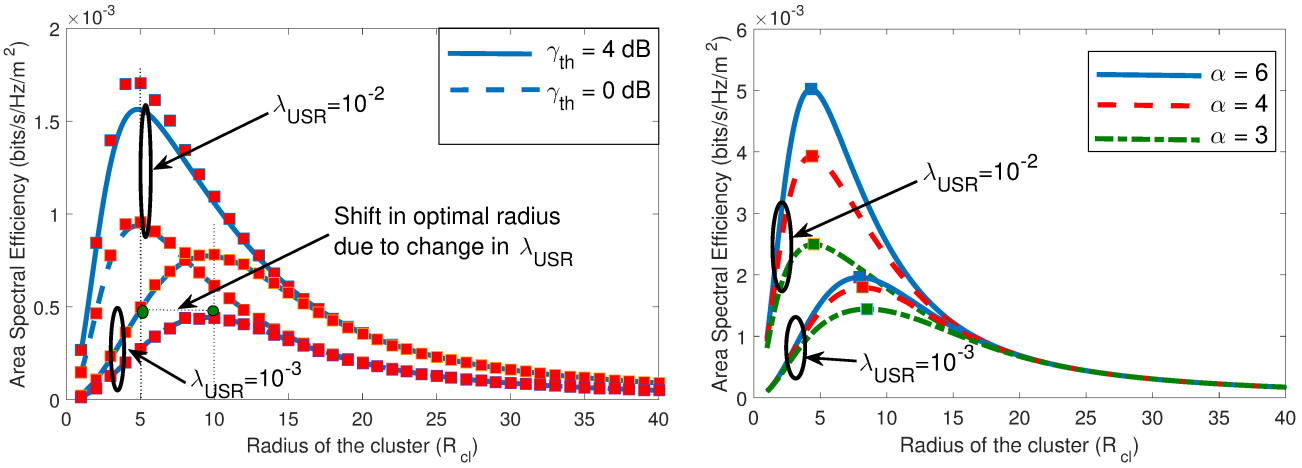

FIGURE 3. Impact of Parametric variations on the area spectral efficiency of the UC-RAN. The red ■ markers in (a) correspond to the Monte-carlo simulation results. (a) Area spectral efficiency of the UC-RAN with varying cluster radius, desired SIR threshold and user density for λ_{CLR}. (b) Area spectral efficiency of the UC-RAN with varying path loss exponent and user density for λ_{CLR}.

The average power consumption of each cluster can then be written as

$$P_{CRAN} = \omega_{CRAN}(1 - \exp\{-\pi\lambda_{CLR}R_{cl}^2\}, \theta)P_o + \Delta_u P_u + P_{ou}. \tag{22}$$

Notice that in this analysis we are mainly focusing on the power consumed at MU for re-broadcasting the pilot signal and the power consumed at RRHs to estimate channel from this pilot. We do not consider the power consumption at macro BS for initial transmission of pilot signal, since this cannot be regarded as an energy overhead. Such transmission is part of the macro BS operation even in the traditional cellular networks.

B. ENERGY EFFICIENCY

The network wide average energy efficiency is defined to be as the ratio of sustainable throughput for each scheduled MU and the average power consumption times the number of scheduled users. Mathematically

$$\eta_{EE} = \frac{\mathrm{B}\log_2(1 + \Gamma_{CRAN}^I)}{\omega_{CRAN}(1 - \exp[-\lambda_{CLR}\pi R_{cl}^2], \theta)P_o + \Delta_u P_u + P_{ou}}, \tag{23}$$

where B is the employed bandwidth (assumed unity for subsequent discussion) and Γ_{CRAN}^I is the effective SIR defined in (18).

Remarks:

1) The per user throughput scales as $\mathcal{O}(\lambda_{CLR}^{1/\delta-1})$ while the average power consumption of each cluster involved in discovery process scales as $\mathcal{O}(1-e^{-\lambda_{CLR}})$. This implies that both the user throughput and the power consumption are increasing functions of the RRH density. However, as λ_{CLR} increases, the power consumption quickly saturates to $P_o + \Delta_u P_u + P_{ou}$ as $\omega_{CRAN}(M, \theta) \to 1$.
2) Similar to RRH density, it is obvious that the throughput and power consumption are monotonically increasing functions of the cluster radius (R_{cl}). Due to the saturation of the power consumption though, the optimal cluster radius which maximizes the energy efficiency of the UC-RAN would be the maximum possible cluster size as per network operator's design specifications.
3) These two observations lead to an important design question, i.e., how different the EE-optimal cluster size is as compared to the cluster radius which maximizes the network wide area spectral efficiency? Also, since an energy efficient design would prefer a larger cluster size comprising on the area spectral efficiency, is there a way to work out a balance between these two parameters. The rest of our discussion will be formed across this design issue.

VII. RESULTS AND DISCUSSION

In this section, we discuss how the efficiency parameters ASE and EE are impacted by variations in user activity, deployment density and propagation environment. For ease of understanding, we denote the RRH cluster radii that maximize the ASE and EE in (13) and (23) as R_{cl}^* and $R_{cl,ee}^*$ respectively.

A. OPTIMAL CLUSTER RADIUS FOR ASE

Fig. 3 depicts the impact of different parametric variations on the area spectral efficiency of the UC-RAN. The solid lines correspond to the simulation of analytical expressions obtained in the previous section. Furthermore, the curves with "■" markers are obtained by performing Monte-Carlo simulations. The Monte-Carlo simulations employed 10^4 realizations of spatial and channel variations for each value of R_{cl} at a desired SIR threshold γ_{th} for each parametric value of

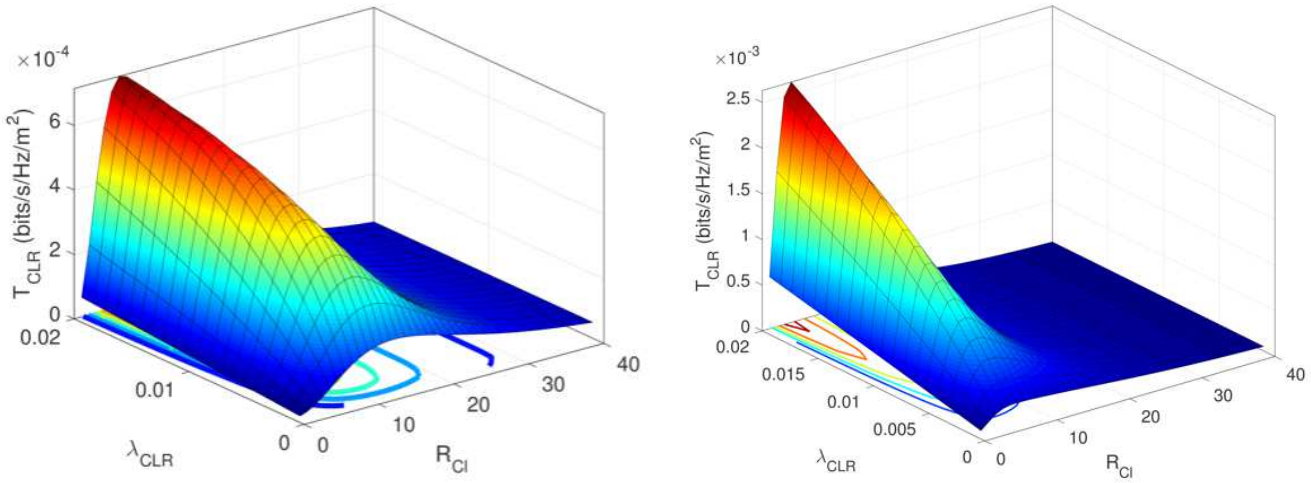

FIGURE 4. Area spectral efficiency of the UC-RAN with varying cluster radius and RRH density. (a) Area spectral efficiency of the UC-RAN with $\lambda_{USR} = 10^{-3}$, $\alpha = 4$ and $\gamma_{th} = 0$ dB. (b) Area spectral efficiency of the UC-RAN with $\lambda_{USR} = 10^{-2}$, $\alpha = 4$ and $\gamma_{th} = 0$.

the user density λ_{USR}. As shown in Fig. 3a, the lower-bound established in the previous sub-section is indeed extremely tight for all parametric variations.

Fig. 3a consolidates the observation which followed from (13), i.e., there exists an optimal cluster size which maximizes the area spectral efficiency of the UC-RAN. As indicated by Fig. 3a, the optimal cluster radius, R^*_{cl}, is not very sensitive to the changes in the desired SIR threshold. The impact of user density λ_{USR} on the R^*_{cl} is more pronounced as compared to the SIR threshold. With an increase in λ_{USR}, the optimal cluster radius decreases. Intuitively, with an increase in the user density, the cluster radius must be reduced so that signal strength can be improved. The gain in the signal strength offsets the loss due to increased interference. Since in real life the user density varies over the time, the UC-RAN must employ a self-organization mechanism (SON) by adapting the clustering radius. The optimal clustering radius can be established from the expression derived in the previous subsection. Such SON mechanism can be easily implemented on the BBU data centers or macro BS. At this juncture, it is worth highlighting that the SON algorithm will require the estimates of the path loss exponent and the user density. The estimation error in these parameters can lead to sub-optimal selection of the cluster radius, incurring significant penalty in terms of spectral efficiency.

Fig. 3b shows that the optimal cluster radius (R^*_{cl}) increases with an increase in the path loss exponent (α). Intuitively, a higher path loss exponent implies that the aggregate interference power is reduced. However the signal power is also reduced. To compensate for the signal power reduction, the cluster size can be increased to harness the spatial diversity gain for increasing the effective received power. It is clear from Fig. 3b that the ASE is an increasing function of α for a fixed cluster size.

B. OPTIMAL RRH DENSITY FOR MAXIMIZING ASE

Besides optimal selection of the cluster radius, from a network designer's perspective it is important to estimate the density of RRHs required to satisfy a certain desired QoS requirement. Fig. 4 plots the UC-RAN area spectral efficiency against the varying RRH density and cluster size.

From Fig. 4, we can see that for a certain fixed cluster radius, the ASE increases with an increase in RRH density. This is naturally the consequence of the increased probability of presence of RRH within the user-centric clusters due to an increase in the RRH density. Notice that the density of RRH also impacts the optimal cluster size. Consequently, both the optimal radius which maximizes the ASE and the RRH density should be jointly selected to reap the full potential of UC-RAN.

As shown in the previous subsection, the optimal cluster radius is also a function of the user density. Hence from a SON perspective, the cluster radius must be adapted with any change in user density for a certain fixed RRH density. However, the RRH density at the selected cluster radius may not be optimal. This motivates the design where a certain density of the RRH is deployed as a baseline design. These RRHs can be turned ON/OFF depending on the density requirement. The SON algorithm then tracks the changes in the user density and optimizes the area spectral efficiency by re-configuring both the density of the RRHs and the optimal cluster radius.

C. OPTIMAL SIR THRESHOLD FOR MAXIMIZING ASE

From (13), it is obvious that there exists an optimal SIR threshold which maximizes the area spectral efficiency of the UC-RAN. This follows from the fact that the link rate is a logarithmically increasing function of the SIR threshold ($\log_2(1 + \gamma_{th})$) while the success probability is exponentially decreasing in terms of γ_{th}. Consequently both these effects

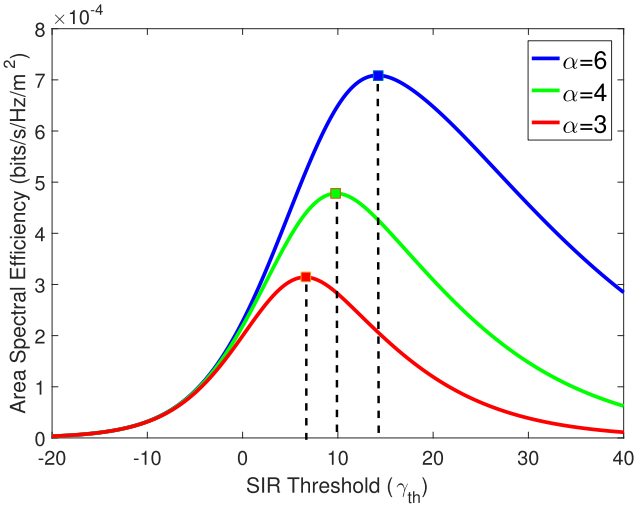

FIGURE 5. Area spectral efficiency of the UC-RAN with varying desired SIR threshold with user density $\lambda_{USR} = 10^{-2}$, $\lambda_{CLR} = 10^{-3}$ and $R_{cl} = R^*_{cl}$.

TABLE 1. Selection of SIR threshold γ_{th} for fixed BEP threshold P_b^{th}.

Modulation	SIR Threshold
BPSK/QPSK	$0.5(Q^{-1}(P_b^{th}))^2$
M-PSK	$\frac{0.5}{\log_2(\mathrm{M})}\left(\frac{Q^{-1}\left(\frac{P_b^{th}\log_2(\mathrm{M})}{2}\right)}{sin(\frac{\pi}{\mathrm{M}})}\right)^2$
M-QAM	$\frac{\mathrm{M}-1}{3\log_2(\mathrm{M})}\left(\frac{Q^{-1}\left(\frac{P_b^{th}\log_2(\mathrm{M})}{2}\right)}{4(1-1/\sqrt{\mathrm{M}})}\right)^2$

should balance out at a certain SIR threshold which will maximize the area spectral efficiency of the C-RAN. Notice that the rate at which the link success probability decreases depends on the path loss exponent α (through the term γ_{th}^{δ}). Hence the optimal SIR is coupled with the path loss exponent.

Fig. 5 plots the area spectral efficiency curve of the UC-RAN with a varying SIR threshold when an optimal cluster radius is employed, i.e., $R_{cl} = R^*_{cl}$. It corroborates our argument about the existence of an optimal SIR threshold which maximizes the spectral performance and its dependence on the path loss exponent. Consider a scenario where each UC-RAN user's QoS is determined by the maximum supportable rate subject to a certain bit error probability (BEP) experienced at a downlink. BEP of an arbitrary user is further coupled with the employed modulation scheme. Consequently, for a certain desired BEP threshold, the corresponding SIR threshold can be determined from the BEP expressions of the modulation scheme.

Table 1 summarizes the relation between the SIR threshold (γ_{th}) and desired BEP (P_b^{th}) for M-PSK and M-QAM schemes [28]. From a SON perspective, the optimal SIR threshold is coupled with the user density, cluster radius and density of the RRHs. Hence the SIR threshold or equivalently the modulation scheme must be adapted to cater for changes

TABLE 2. Power consumption parameters.

Parameter	Description	Value
P_u	Transmit power of MU	1 W
P_o	Fixed Power consumption of RRH	6.8 W
Δ_u	Radio Frequency dependent power consumption	4 W
P_{ou}	Circuit power consumed at MU during discovery	4.3 W

in these parameters. Effectively, this indicates the necessity of rate adaptation (by adapting value of the modulation index M for M-PSK/QAM) to maximize the attainable spectral performance at a certain BEP.

D. OPTIMAL CLUSTER RADIUS FOR ENERGY EFFICIENCY

Fig. 6 depicts the energy efficiency of the UC-RAN with varying cluster radii for the power consumption parameters summarized in Table 2. As shown in Fig. 6, the energy efficiency is maximized by selecting R_{cl} as large as possible. However, the area spectral efficiency is maximized by selecting a lower ASE-optimal R^*_{cl}. Hence there exists a trade-off between maximizing η_{EE} and $\mathcal{T}_{CLR}$. $R^*_{cl,ee} > R^*_{cl}$ implies that in order to maximize the energy efficiency, the area spectral efficiency must be sacrificed. The reason for this disparity in optimal cluster sizes can be attributed to the fact that η_{EE} scales differently, as compared to $\mathcal{T}_{CLR}$, with respect to R_{cl}. Hence the optimal cluster radii with respect to the energy and area spectral efficiencies captures the trade-off between these two design factors. The selection of optimal cluster-size is thus based on the network deployment objective. If the objective is to provide higher throughput without worrying about additional costs, the optimal cluster radius corresponding to the area spectral efficiency as a metric should be selected. On the other hand, if minimizing energy consumption across the network is the main objective, some throughput can be sacrificed by selecting an optimal cluster radius from the energy efficiency analysis. It is feasible for the operator to define different modes of operation, i.e., the energy efficient mode (for instance at night time) and the throughput efficient mode (for instance in day time) as proposed in [19]. The SON engine can then configure the optimal cluster radius in accordance with the desired mode.

Another observation from Fig.6 is the down scaling of η_{EE} under high θ values. This is due to the fact that $\omega_{CRAN}(\mathrm{M}, \theta) \approx \theta$ in (23) for $R_{cl} \geq 5$ m and the range of RRH densities considered for this work. The insensitivity of $\omega_{CRAN}(\mathrm{M}, \theta)$ with respect to R_{cl} and λ_{CLR} allows high θ values to increase the power consumption and consequently decrease η_{EE}.

E. QOE PERFORMANCE ANALYSIS IN A UC-RAN

Users' QoE analysis is conducted through SINR distribution between MUs in an LTE like simulation tool with network parameters: $\lambda_{USR} = 10^{-2}/\mathrm{m}^2$, $\lambda_{CLR} = 10^{-3}/\mathrm{m}^2$, $\alpha = 4, \theta = 0.5$, $\gamma_{th} = 4$ dB and bandwidth B = 1 Hz. Both the MU and RRH deployments are performed using uniform PPPs and average performance results are obtained via Monte

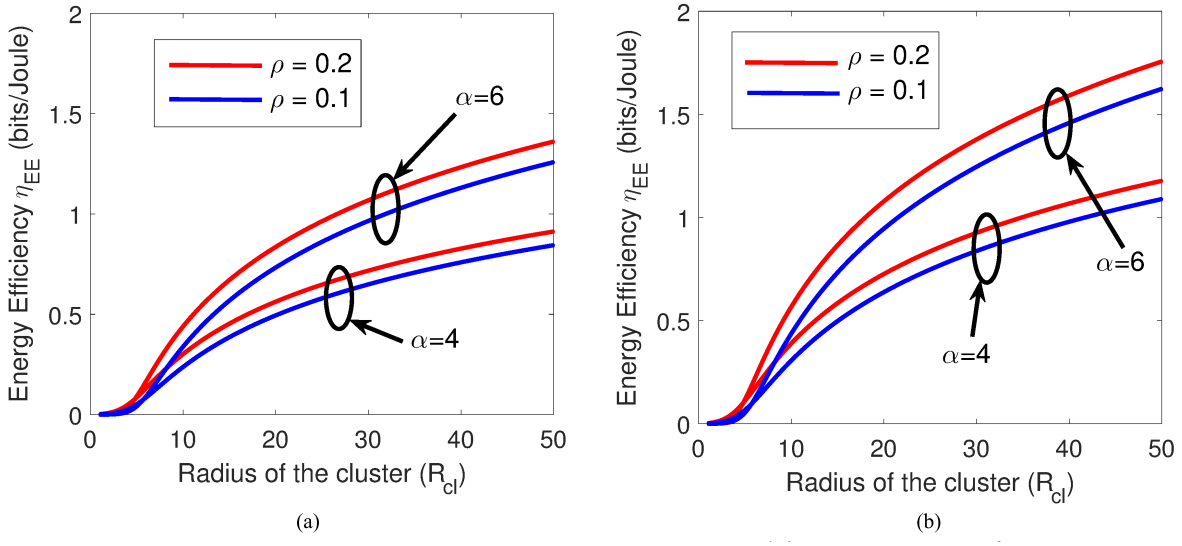

FIGURE 6. Energy Efficiency of the UC-RAN with varying cluster radius for $\lambda_{CLR} = 10^{-1}$ and $\lambda_{USR} = 10^{-2}$. (a) C-RAN efficiency $\theta = 1$. (b) C-RAN efficiency $\theta = 0.5$

Carlo simulations. We use two variants of the proposed user-centric approach: i) RRH cluster size deployment that maximizes ASE henceforth referred as UC(ASE), and ii) cluster size deployment that maximizes EE henceforth referred as UC(EE). To compare the performance with a standard non user-centric PPP deployment, we follow the approach in [29] and represent it as NUC. Results in Fig.7 show that even with the most data throughput efficient user-centric design, we obtain a SINR gain of over 20 dB for almost 50% of the users. The ruggedness in the CDF graph of UC(EE) in comparison to the other two CDFs is because of lower number of users in the thinned PPP Π'_{MU} which is a direct consequence of the larger cluster sizes in EE optimization. The 5 percentile SINR performance (for the cell-edge users with worst SINR in conventional networks) is also significantly improved for user-centric approaches with about 20 dB and 40 dB gain in UC(ASE) and UC(EE) respectively. Clearly, these results indicate that the user-centric approach eliminates cell-edge degradation and guaranteed high QoE for every user regardless of its physical location.

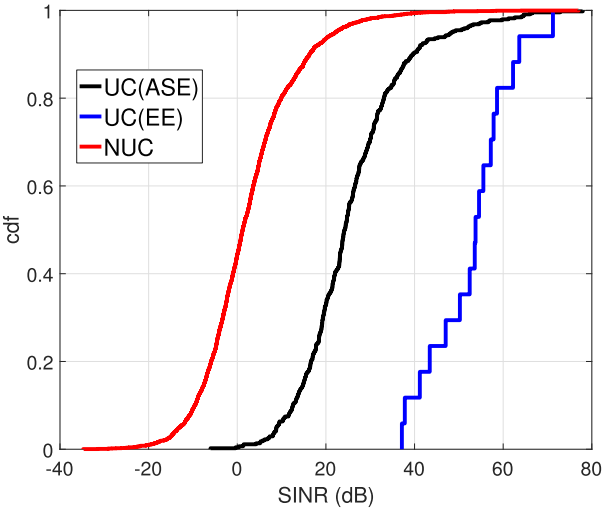

FIGURE 7. Downlink SINR CDF comparison between user-centric and non user-centric approaches.

F. EFFICIENCY PERFORMANCE EVALUATION OF SELECTIVE USER-CENTRIC RRH ACTIVATION IN A UC-RAN

Fig. 8 compares the system wide ASE and EE of the user-centric approaches with the baseline scheme at different RRH densities and $\lambda_{USR} = 10^{-2}/\text{m}^2$, $\alpha = 4$, $\theta = 0.5$ and $\gamma_{th} = 4$ dB. Fig. 8a reveals that as the RRH deployment density increases, UC(ASE) emerges as the most data efficient scheme. While NUC exhibits uniform ASE, UC(ASE) by virtue of increased $\mathbb{P}_{suc}$ exhibits highest system capacity. On the other hand, UC(EE), though not throughput efficient by any regards, yields more than 5 times power efficient network as compared to NUC approach (Fig. 8b). This observation highlights the inherent performance trade-off when the optimal cluster radius is adjusted according to objective function, i.e. when cluster size varies from R^*_{cl} to $R^*_{cl,ee}$ for maximizing the $\mathcal{T}_{CLR}$ and η_{EE} respectively.

VIII. SON FRAMEWORK FOR RRH CLUSTER SIZE OPTIMIZATION

Hitherto, we have demonstrated the superior efficiency performance in a UC-RAN with the user-centric RRH clustering. We also observed that the performance measures, i.e. ASE and EE require disparate RRH cluster sizes when maximized individually. Therefore, the key question of what should the optimal RRH cluster size be from a network design perspective remains unanswered. To address this research problem, we formulate a SON framework based on a two player

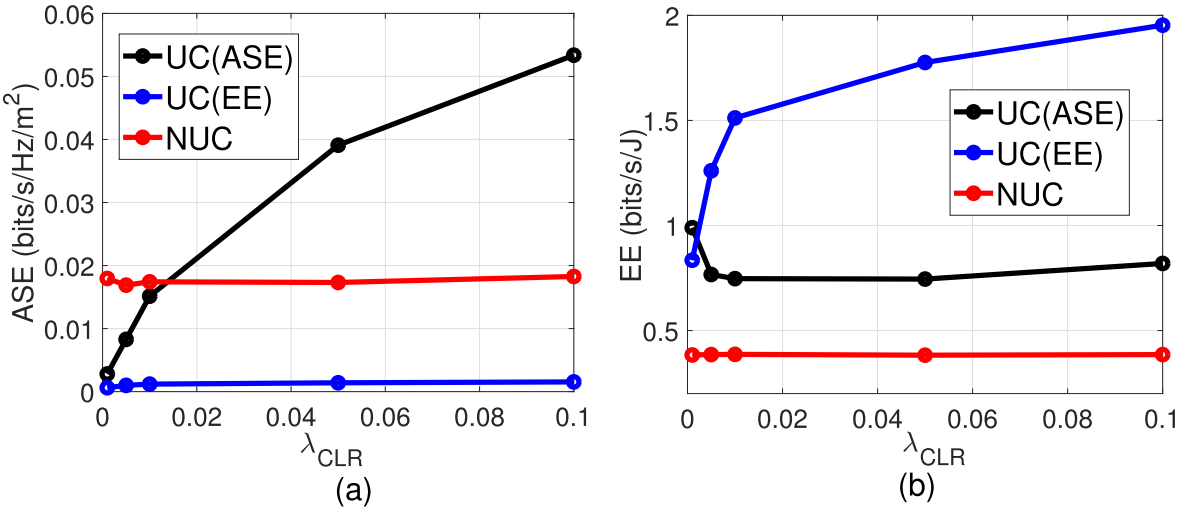

FIGURE 8. (a) ASE and (b) EE comparison of UC(ASE), UC(EE) and NUC approaches with different RRH densities.

cooperative bargaining game to investigate the cluster size estimation in a dynamic environment from the perspective of modeling the optimal trade-off between system wide ASE and EE. Both ASE and EE are modelled as virtual game players that independently estimate the best cluster size for maximizing their respective utility functions given by (13) and (23) respectively. In section VII, we observed the large dissimilarity in cluster size preferences of the players. However, both players can mutually benefit through the cooperative game where they negotiate for the R_{cl} that achieves optimal ASE-EE trade-off. Using Nashs axiomatic model, it is well known that the Nash bargaining solution (NBS) achieves a pareto-optimal solution, i.e. the optimal trade-off in the utilities of the players in such cooperative games [18].

Let $\mathbb{N} = \{1, 2\}$ be the set of the players, where player $i = 1$ denotes ASE (or $\mathcal{T}_{CLR}$), player $i = 2$ denotes EE (or η_{EE}) and $\mathbf{S}_i$ denotes the set of all feasible payoffs to an arbitrary MU i as

$$\boldsymbol{S}_i = \{U_i | U_i = U_i(R_{cl}), R_{cl} \in \mathbb{R} : R_{cl} > 0\}. \tag{24}$$

The achievable utilities for our virtual players can be represented by the space $\mathcal{S}$ which is the set of all feasible payoffs that players $i \in \mathbb{N}$ can achieve as they cooperate for cluster radius selection, i.e.

$$\mathcal{S} = \{U = (u_1, u_2) | u_1 \in S_1, u_2 \in S_2\}, \tag{25}$$

where $u_1(x_1)$ is the utility of the first player and $u_2(x_2)$ is the utility of the second player such that

$$s_1 = u_1(x_1) = [\mathcal{T}_{CLR}(R_{cl})]^{\beta}, \tag{26}$$
$$s_2 = u_2(x_2) = [\eta_{EE}(R_{cl})]^{1-\beta} \tag{27}$$

and $x_1 = x_2 = R_{cl} \in \mathbb{R} : R_{cl} > 0$. $\beta \in [0, 1]$ is the exponential bias factor in the NBS that defines the bargaining power (or the trade-off) division between the two players. We also define the disagreement space $\mathcal{D} \in \mathcal{S}$ as the set of the two disagreement points $d = (d_1, d_2)$ where $d_1 = u_1(\mathcal{D})$ and $d_2 = u_2(\mathcal{D})$ represent the payoffs for each player if the bargaining process fails and no outcome is reached. For our game, we set $d = (0, 0)$ thus giving both players uniform leeway to improve their utilities. Analysis in [30] has shown that the optimal trade-off in such parametric cooperative games can be obtained via Nash's axiomatic approach through the solution of a convex optimization problem which for our model can be given by (28), where P_{CRAN} is the average cluster power consumption expressed in (22) and R_{cl}^{opt} is the optimal cluster size that corresponds to the NBS, the unique solution (Nash equilibrium) that maximizes the Nash product in (28), as shown at the bottom of the next page.

Notice that the computational complexity of the SON optimization problem in (28) for determination of R_{cl}^{opt} is a function of the cluster size granularity, i.e. $\mathcal{O}(N_{CLR})$ where N_{CLR} denotes the number of distinct cluster sizes over which the optimization in (28) is performed. Since processing times are independent of MU or RRH densities, the optimal RRH cluster size estimation and subsequent deployment of user-centric scheduling through selective RRH activation is practically realizable and scalable throughout the network. Similar to other SON solutions [31], [32], the proposed optimization framework can be implemented in centralized BBU pools for dynamic adjustment of R_{cl}^{opt} with variation in network statistics and operator's preference (β).

A. SIMULATION RESULTS FOR SON OPTIMIZATION FRAMEWORK IN A UC-RAN

In this part of the paper, we will analyze and discuss the simulation results for determining the optimal cluster radii as network parameters and the exponential weightage parameter β fluctuate. We also look at the scheduling probability for an arbitrary MU under the RRH clustering mechanism explained in Algorithm 1. For simulation, we consider a two tier HetNet with a tri-sector hexagonal MBS of radius 500 m. We consider

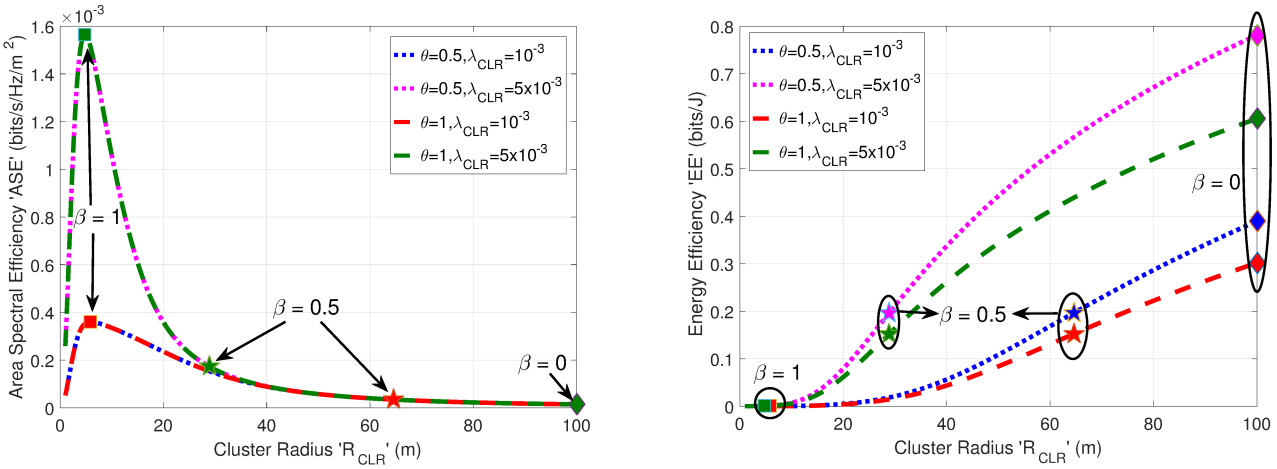

FIGURE 9. ASE and EE v/s R_{cl} when varying the λ_{CLR} and θ. The optimal RRH cluster size R_{cl}^{opt} for each use case is shown separately at β values of 0, 0.5 and 1 and denoted by ♦, ★ and ■ respectively.

a single sector of the MBS covering an area of 73850 m^2 where MUs and small cell RRHs are uniformly distributed according to their independent SPPPs. Without loss of generality, the channel power gains between all MUs and RRHs are assumed unity. We assume uniform transmit power of 30 dBm for all RRHs. Other power consumption parameters are taken from table 2. Simulation results are averaged over 1000 Monte Carlo trials.

1) IMPACT OF NETWORK PARAMETERS AND β ON ASE, EE IN A UC-RAN

Simulation results in Fig.9 depict how λ_{CLR} and θ impact the utilities of the virtual players with a set of exponential biases $\beta = [0, 0.5, 1]$ which denote the ASE-optimal, non-biased NBS and EE-optimal modes respectively. The results in Fig. 9 consolidate the established fact that there exist different cluster sizes that maximize the ASE and EE of the UC-RAN. As we would expect, the ASE is higher for dense RRH deployment and insensitive to θ. A non-biased ($\beta = 0.5$) trade-off yields mean losses of 89.5% and 62.5% to the utility functions output as compared to peak ASE ($\beta = 1$) and EE ($\beta = 0$) values respectively. Through appropriate β adjustment, the optimization framework yields an R_{cl}^{opt} between 5 m and 100 m (or the maximum allowable cluster size). Consequently, this allows the network operator to adjust the ASE and EE gains by 100x and 1000x respectively. Note that any gain in one efficiency parameter is accompanied by some degradation in the other.

As the operating point is shifted from EE to ASE regime ($\beta > 0.5$), we see that R_{cl}^{opt} reduces to allow for higher number of concurrent DL transmissions between RRHs and MUs. We also notice significant reduction in R_{cl}^{opt} as λ_{CLR} increases particularly at $\beta = 0.5$. Since in actual networks, the cluster density may dynamically vary over time, a SON placed within the centralized BBU pools can dynamically expand or shrink the RRH cluster sizes to maximize the utility in (28).

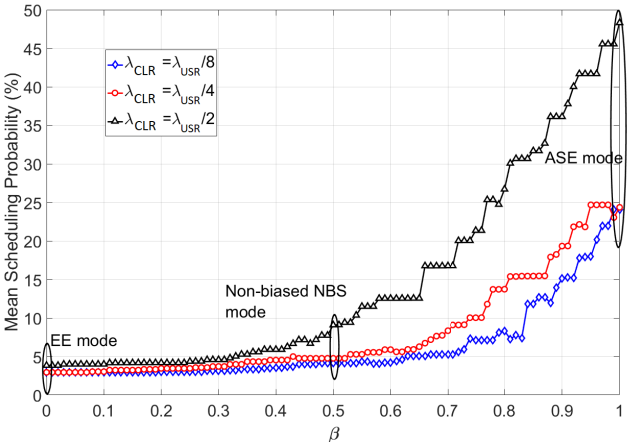

FIGURE 10. Mean scheduling success rate versus β under different $\lambda_{CLR}/\lambda_{USR}$.

B. SCHEDULING RATE UNDER VARYING $\lambda_{CLR}/\lambda_{USR}$

In order to study the impact of dense RRH deployment, Fig. 10 depicts the mean scheduling success rate for the MUs under different β and deployment density ($\frac{\lambda_{CLR}}{\lambda_{USR}}$) scenarios. For each scenario, we simulate 1000 consecutive transmission time intervals (TTIs) for $\lambda_{USR} = 4\text{x}10^{-4}$. Using user-centric RRH clustering (Section II), we update the p_{USR} and RRH clusters for each TTI which allows us

$$R_{cl}^{opt} = \max_{R_{cl}} \left[\left(\lambda_{USR}^{\{EF\}} \log_2(1+\gamma_{th}) \mathbb{P}_{suc}(\gamma_{th}, R_{cl}^2) \right)^{\beta} \left(\frac{\log_2(1+\Gamma_{CRAN}^{I})}{P_{CRAN}} \right)^{1-\beta} \right]. \quad (28)$$

to estimate the mean number of MUs that are scheduled for DL transmission under each simulation scenario. The results in Fig. 10 show rapid increase in the scheduling probability as we move towards ASE regime. However, we notice that as RRH deployment density increases to $\lambda_{USR}/2$, the scheduling success doubles to 48% as compared to 24% in case of $\lambda_{USR}/4$ and $\lambda_{USR}/8$ for $\beta = 1$. This implies that increasing small cell deployment density in ASE mode allows higher number of concurrent DL transmissions that consequently reduces main scheduling delay for an arbitrary user in UC-RAN.

IX. CONCLUSION

In this article, we developed a comprehensive statistical framework for computation of the area spectral and energy efficiency of a large-scale user-centric cloud radio access network (UC-RAN). We proposed a user-centric RRH clustering algorithm which enables dynamic coverage extension and shrinkage by activating a single remote radio head within a specified demarcation around a scheduled user based on max SIR gain criteria. The user-centric architecture ensures uniform coverage and no cell-edge degradation for all the users irrespective of their physical location.We demonstrated that there exists an optimal cluster radius which maximizes the area spectral efficiency of the UC-RAN. It was also demonstrated that this optimal cluster radius is coupled with the user density and hence must be adapted by a self-organization mechanism.

The link level performance was then employed to perform cost-benefit analysis of the proposed protocol. More specifically, the amount of power dissipated in the association process under the proposed protocol is considered as the cost of obtaining the throughput gains. The throughput-cost ratio is hence essentially the energy efficiency of the UC-RAN. It was shown that there exists another optimal cluster radius which maximizes the energy efficiency of the UC-RAN. However, this is larger to the one obtained under area spectral efficiency criterion. Consequently, the throughput-energy consumption trade-off manifests itself in terms of dimensioning of the cluster radius in UC-RAN. Using a game theoretic framework, we demonstrated that a SON engine within the centralized BBU pools may be employed to dynamically configure the optimal cluster size. Simulation results indicated that: i) the SON mechanism allows more than 100x efficiency variation through real-time adjustment in the NBS bias parameter, ii) in comparison to current always-ON RRH deployments, selective RRH activation in UC-RAN offers high area spectral and energy efficiency gains, particularly when $\lambda_{CLR} > \lambda_{USR}$, and iii) significant SINR gains can be realized in both ASE and EE operating modes by virtue of interference-free RRH cluster zones around each scheduled user.

ACKNOWLEDGEMENT

A portion of this work has been accepted for presentation at IEEE International Conference on Communications (ICC) 2018. For more details, please visit www.bsonlab.com.

REFERENCES

[1] "Cisco visual networking index: Mobile data traffic forecast update, 2013 to 2018," CISCO, San Jose, CA, USA, Tech. Rep., 2014.
[2] J. Research. (Jun. 2011). *Press Release: Mobile Network Operator Revenues*. [Online]. Available: http://juniperresearch.com/viewpressrelease.php?pr=245
[3] R. Wang, H. Hu, and X. Yang, "Potentials and challenges of C-RAN supporting multi-RATs toward 5G mobile networks," *IEEE Access*, vol. 2, pp. 1187–1195, 2014.
[4] A. Checko *et al.*, "Cloud RAN for mobile networks—A technology overview," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 405–426, 1st Quart., 2014.
[5] Y. Zhang and Y. J. Zhang, "User-centric virtual cell design for cloud radio access networks," in *Proc. IEEE 15th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2014, pp. 249–253.
[6] S. Chen, F. Qin, B. Hu, X. Li, and Z. Chen, "User-centric ultra-dense networks for 5G: Challenges, methodologies, and directions," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 78–85, Apr. 2016.
[7] D. Huang, T. Xing, and H. Wu, "Mobile cloud computing service models: A user-centric approach," *IEEE Netw.*, vol. 27, no. 5, pp. 6–11, Sep./Oct. 2013.
[8] S. Zaidi, B. Hmidet, S. Affes, U. Vilaipornsawai, L. Zhang, "User-centric wireless access virtualization strategies for future 5G networks," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016, pp. 1–7.
[9] S. A. R. Zaidi, A. Imran, D. McLernon, and M. Ghogho, "Characterizing coverage and downlink throughput of cloud empowered HetNets," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 1013–1016, Jun. 2015.
[10] B. Dai and W. Yu, "Sparse beamforming and user-centric clustering for downlink cloud radio access network," *IEEE Access*, vol. 2, pp. 1326–1339, Oct. 2014.
[11] H. Zhang, Z. Yang, Y. Liu, and X. Zhang, "Power control for 5G user-centric network: Performance analysis and design insight," *IEEE Access*, vol. 4, pp. 7347–7355, 2016.
[12] C. Li, J. Zhang, M. Haenggi, and K. B. Letaief, "User-centric intercell interference nulling for downlink small cell networks," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1419–1431, Apr. 2015.
[13] A. Awad, A. Mohamed, and C. F. Chiasserini, "Dynamic network selection in heterogeneous wireless networks: A user-centric scheme for improved delivery," *IEEE Consum. Electron. Mag.*, vol. 6, no. 1, pp. 53–60, Jan. 2017.
[14] S. Bassoy, M. Jaber, M. A. Imran, and P. Xiao, "Load aware self-organising user-centric dynamic comp clustering for 5G networks," *IEEE Access*, vol. 4, pp. 2895–2906, 2016.
[15] "C-RAN: The road towards green RANs," C. Mobile, Hong Kong, White Paper, 2011.
[16] J. M. D. Stoyan, W. S. Kendall, and L. Ruschendorf, *Stochastic Geometry and its Applications*, vol. 2. Chichester, U.K.: Wiley, 1995.
[17] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 336–361, Feb. 2013.
[18] Z. Han, D. Niyato, W. Saad, T. Baar, and A. Hjrungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[19] A. Imran, M. A. Imran, and R. Tafazolli, "A novel self organizing framework for adaptive frequency reuse and deployment in future cellular networks," in *Proc. IEEE 21st Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2010, pp. 2354–2359.
[20] A. Imran, M. A. Imran, A. Abu-Dayya, and R. Tafazolli, "Self organization of tilts in relay enhanced networks: A distributed solution," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 764–779, Feb. 2014.
[21] H. He, J. Xue, T. Ratnarajah, F. A. Khan, and C. B. Papadias, "Modeling and analysis of cloud radio access networks using Matern hard-core point processes," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4074–4087, Jun. 2016.
[22] M. Haenggi, "Mean interference in hard-core wireless networks," *IEEE Commun. Lett.*, vol. 15, no. 8, pp. 792–794, Aug. 2011.
[23] Z. Chen *et al.*, "Aggregate interference modeling in cognitive radio networks with power and contention control," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 456–468, Feb. 2012.

[24] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[25] A. Taufique, M. Jaber, A. Imran, Z. Dawy, and E. Yacoub, "Planning wireless cellular networks of future: Outlook, challenges and opportunities," *IEEE Access*, vol. 5, pp. 4821–4845, 2017.
[26] G. Auer *et al.*, "How much energy is needed to run a wireless network?" *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.
[27] R. Gupta, E. C. Strinati, and D. Ktenas, "Energy efficient joint DTX and MIMO in cloud radio access networks," in *Proc. IEEE 1st Int. Conf. Cloud Network. (CLOUDNET)*, Nov. 2012, pp. 191–196.
[28] J. Proakis, *Digital Communications* in *Electrical Engineering Series*. New York, NY, USA: McGraw-Hill, 2001. [Online]. Available: https://books.google.com.pk/books?id=sbr8QwAACAAJ
[29] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.
[30] B. Romanous, N. Bitar, S. A. R. Zaidi, A. Imran, M. Ghogho, and H. H. Refai, "A game theoretic approach for optimizing density of remote radio heads in user centric cloud-based radio access network," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.
[31] A. Zoha, A. Saeed, A. Imran, M. A. Imran, and A. Abu-Dayya, "A learning-based approach for autonomous outage detection and coverage optimization," *Trans. Emerg. Telecommun. Technol.*, vol. 27, no. 3, pp. 439–450, 2016.
[32] U. S. Hashmi, A. Darbandi, and A. Imran, "Enabling proactive self-healing by data mining network failure logs," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Jan. 2017, pp. 511–517.

UMAIR SAJID HASHMI (GS'14) received the B.S. degree in electronic engineering from the Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan, in 2008, and the M.Sc. degree in advanced distributed systems from the University of Leicester, U.K., in 2010. He is currently pursuing the Ph.D. degree with the Big Data and Artificial Intelligence (AI) Enabled Self Organizing Network Laboratory, The University of Oklahoma, Tulsa, OK, USA. His research interests include stochastic geometry and the application of AI and game theory techniques toward optimization in fifth-generation cellular networks.

SYED ALI RAZA ZAIDI (M'09) received the Ph.D. degree from the School of Electronic and Electrical Engineering. From 2011 to 2013, he was associated with the International University of Rabat as a Research Associate. In 2013, he was also a Visiting Research Scientist with the Qatar Innovations and Mobility Centre involved in QNRF funded project QSON. From 2013 to 2015, he was associated with the SPCOM Research Group involved in US ARL funded project in the area of network science. He is currently a University Academic Fellow (Assistant Professor) of communication and sensing for robotics and autonomous systems with the University of Leeds. He has published over 90 papers in leading IEEE conferences and journals. He received the COST IC0902, Royal Academy of Engineering, EPSRC, and DAAD Grants to promote the research outputs. His current research interests include intersection ICT, applied mathematics, mobile computing, and embedded systems implementation. Specifically, his current research is geared toward the design and implementation of communication protocols to enable various applications (rehabilitation, health care, manufacturing, and surveillance) of future RAS and the design, implementation, and control of RAS for enabling future wireless networks (for e.g., autonomous deployment, management, and repair of future cellular networks). He received the G. W. and F. W. Carter Prize for best thesis and best research paper. From 2014 to 2015, he was an Editor of the IEEE Communication Letters and also a lead Guest Editor of *IET Signal Processing Journal's* special issue on signal processing for large-scale 5G wireless networks. He is also an Editor of *IET Access, Front haul and Backhaul* book. He is currently serving as an Associate Technical Editor for *IEEE Communication Magazine*.

ALI IMRAN (M'15) received the B.Sc. degree in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2005, and the M.Sc. degree (Hons.) in mobile and satellite communications and the Ph.D. degree from the University of Surrey, Guildford, U.K., in 2007 and 2011, respectively. He is the Founding Director of the *BSON Lab*, The University of Oklahoma, where he is currently an Assistant Professor with the School of Electrical and Computer Engineering. He is a principle investigator of several multinational research projects focused on next-generation wireless networks, for which he has secured research grants of over $3.5 million. His current research interests include big data and artificial intelligence (AI) enabled wireless networks, new RAN architectures for enabling IoT, and the vertical applications of big data and AI. On these topics, he has published over 70 refereed research articles. He has been an Invited Speaker and a Tutorial Presenter on these topics at numerous international conferences, including IEEE ICC, WF-IoT, PIMRC, and WCNC. He is an Associate Fellow of the Higher Education Academy, U.K., a member of the Advisory Board for Special Technical Community on Big Data at the IEEE Computer Society, a Board Member of ITERA, and an Associate Editor of the IEEE Access. He was a (co)-recipient of several awards that include the Best Paper Award IEEE CAMAD 2013 and the IEEE Green ICT Best Solution Award 2017.