IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 20 July 2022, accepted 21 August 2022, date of publication 25 August 2022, date of current version 6 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3201527

RESEARCH ARTICLE

Energy Efficient Resource Allocation for H-NOMA Assisted B5G HetNets

UMAR GHAFOOR[1], HUMAYUN ZUBAIR KHAN[1], (Senior Member, IEEE), MUDASSAR ALI[1,2], ADIL MASOOD SIDDIQUI[1], MUHAMMAD NAEEM[3], AND IMRAN RASHID[1]

[1]Department of Electrical Engineering, Military College of Signals, National University of Sciences and Technology, Islamabad 44000, Pakistan
[2]Department of Telecommunication Engineering, University of Engineering and Technology, Taxila 47050, Pakistan
[3]Department of Electrical and Computer Engineering, COMSATS University Islamabad, Wah Campus, Wah Cantt 47040, Pakistan

Corresponding author: Umar Ghafoor (ch.umar163@mcs.edu.pk)

ABSTRACT The resource allocation solution offered based on non-orthogonal multiple access (NOMA) and orthogonal multiple access (OMA) schemes are sub-optimal to address the challenging quality of service (QoS) and higher data rate viz-a-viz energy efficiency (EE) requirements in 5th generation (5G) cellular networks. In this work, we maximize the EE using user equipment (UE) clustering (UE-C) with downlink hybrid NOMA (H-NOMA) assisted beyond 5G (B5G) HetNets. We formulate an optimization problem incorporating UE admission in a cluster, UE association with a base station (BS), and power allocation assisted by H-NOMA, i.e., OMA and NOMA schemes in the macro base station (MBS) only and heterogeneous networks (HetNets) environments. The problem formulated is a type of non-linear concave fractional programming (CFP) problem. The Charnes-Cooper transformation (CCT) is applied to the formulated non-linear CFP problem to convert it into a concave optimization, i.e., mixed-integer non-linear programming (MINLP) problem. A two-phase ϵ-optimal outer approximation algorithm (OAA) is used to solve the MINLP problem. The simulation results show that H-NOMA with HetNets outperforms H-NOMA with MBS only in terms of UE admission, UE association, throughput, and EE.

INDEX TERMS UE-clustering, H-NOMA, fractional programming, MINLP, energy efficiency.

I. INTRODUCTION

The rapid growth in the number of mobile user equipment (UE), and heavy data-driven applications, i.e., live video gaming, video streaming, social networking, etc are imposing challenging requirements like minimum delay, higher data rates, spectrum efficiency (SE), and energy efficiency (EE) on the beyond 5th Generation (B5G) cellular networks. This exponential growth in mobile UEs viz-a-viz mobile data traffic is adding to a significant increase in the energy consumption in cellular networks. Information communication technology (ICT) is consuming almost 2% of the world's total energy. Energy consumption emits carbon which causes the greenhouse effect. Thus, ICT offering higher data rates, low carbon emission, and EE are the prime considerations in B5G cellular networks. Thus, academia and industry in the wireless communication field must divert their research towards future energy-efficient green cellular networks in B5G networks [1], [2], [3].

Energy-efficient radio resource management techniques are required to satisfy the needs for quality of service (QoS) and higher data rates with minimum energy consumption. EE is the ratio of data rate to the total energy consumed [4], [5].

EE can be improved using heterogeneous networks (HetNets), which include a macro base station (MBS) and small base stations (SBSs) [6], [7], [8]. HetNets cover more geographical areas and offer higher data rates and are energy efficient than the conventional MBS-only networks. In HetNets, MBS is large with high transmit power, and SBSs (i.e., femtocell, picocell, etc.) are of small size with low transmit power [9]. A low transmit-powered SBS will also preserve energy because additional energy will not be required for cooling purposes compared to a high transmit-powered MBS.

The associate editor coordinating the review of this manuscript and approving it for publication was Jie Tang.

Additionally, EE can also be improved by efficient reply to the question that how resources will access the network? The conventional multiple access (MA) schemes, i.e., orthogonal multiple access (OMA) includes time division multiple access (TDMA), frequency division multiple access (FDMA), and code division multiple access (CDMA) [10]. As an orthogonal FDMA scheme assigns a single subcarrier (SC) to a single UE which makes these conventional OMA schemes unsuitable due to the availability of limited spectrum, requirements of high data rate, and EE.

Compared to OMA, the non-orthogonal multiple access (NOMA) scheme can increase the SE and EE because more UEs can be accommodated on a single SC at the expense of multi-UE interference (MUEI). Power domain-NOMA (PD-NOMA) based on power domain multiplexing is the most widely used type of NOMA [11]. In downlink PD-NOMA, the base station (BS) at the transmitter side uses superposition coding to transmit information. Successive interference cancellation (SIC) is used to receive information at the receiver side by removing MUEI at the expense of increased complexity [12]. OMA and NOMA are combined as a hybrid-NOMA (H-NOMA) technique that could be a better choice to manage the complexity compared to NOMA. UE-clustering (UE-C) with H-NOMA can further reduce the MUEI. UE-C is defined as the group of UEs assigned an orthogonal SC. In H-NOMA based on UE-C, an SC assigned to a cluster that contains only one UE is defined as OMA-SC and otherwise NOMA-SC. Therefore, the data rate and EE can be improved using a UE-C-based downlink H-NOMA in B5G HetNets. Therefore, the optimal resource allocation for H-NOMA with UE-C in the downlink can enhance the EE.

The resource allocation for H-NOMA has been applied in many scenarios related to this paper. H-NOMA with UE-C in downlink is used for optimal resource allocation i.e. number of UEs admitted in clusters, number of UEs associated with BSs, power allocation to UEs for maximization of throughput [7]. H-NOMA including power domain and code domain NOMA is applied to connect low power trillion UEs and to achieve higher spectral efficiency in the uplink for next-generation internet of things (IoT) networks [13]. H-NOMA as a combination of NOMA and TDMA is used in the uplink to transmit information to clusters UEs and it is applied for reflecting beamforming of intelligent reflecting surface (IRS) and time allocation among BS's power transfer and different UEs clusters' information transmission to maximize the throughput of the IRS-assisted wireless powered communication network [14]. The next section provides the literature review.

A. LITERATURE REVIEW

Table 1 summarizes the past work on EE by using OMA, NOMA, H-NOMA, and UE-C techniques in HetNets and MBS-only networks.

The researchers have worked to maximize EE using NOMA in MBS-only networks from [15], [16], [17], [18]. In [15], the authors have maximized EE of cognitive radio (CR) inspired NOMA network using the suboptimal matching for sub-channel allocation (SOMSA) and difference of convex (DC) programming for power allocation with maximum transmit power and QoS constraints for each primary UE in MBS only networks. In [16], the authors proposed a sequential convex approximation (SCA) scheme to maximize the EE of CR-inspired NOMA network under each primary UE QoS constraint by solving the non-convex fractional programming (FP) problem. In [17], the authors proposed an iterative sub-optimal algorithm by adopting FP and DC programming to maximize EE under minimum rate and maximum transmit power constraints in downlink coordinated multipoint (CoMP) systems with NOMA. In [18], the authors proposed a novel energy-efficient matching scheme for sub-channel assignment and a Lyapunov optimization scheme for power allocation in a single downlink NOMA network.

The researchers have worked to maximize EE using OMA in HetNets from [19], [20], [21], [22]. In [19], the authors proposed a heuristic algorithm to improve the spectral efficiency and reduce the overall power consumption of a mobile communication network based on two-tier deployment of device-to-device (D2D) communication subject to maintaining a minimum signal threshold. In [20], the authors maximized the EE by using a stochastic geometry-based model with random discontinuous transmission (DTX) mode under the finite local delay constraint in the downlink HetNets. In [21], the authors maximized the EE using a heuristic algorithm based on the cross-entropy (CE) and Karush–Kuhn–Tucker (KKT) conditions-based method in a downlink HetNets system under maximum transmit power and minimum rate constraints. In [22], authors presented a joint EE resource allocation (JEERA) algorithm to improve EE and limit interference in HetNets subject to QoS constraints.

The researchers have worked to maximize EE using NOMA in HetNets from [23], [24], [25], [26]. In [23], the authors maximized the EE by using a low-complexity sub-channel matching algorithm and lagrangian duality algorithm for power allocation under the constraint of the limited supply of energy in NOMA HetNets with simultaneous wireless information and power transfer (SWIPT). In [24], the authors considered perfect and imperfect channel state information (CSI) to maximize EE by using a heuristic algorithm in 5G NOMA HetNets subject to QoS and maximum transmit power constraints. In [25], authors presented a spectrum-efficient, energy-efficient, and delay-constrained heuristic algorithm to reduce energy consumption subject to the constraints of delay and QoS for full-duplex self-backhauled (FS) HetNets. In [26], the authors proposed a fair power allocation (FPA) approach to maximize the EE subject to the constraint of transmit powers for UEs in a two-tier downlink ultra-dense HetNet with NOMA.

The researchers have worked little to maximize EE using H-NOMA in MBS only networks from [27], [28], [29], [30], [31], [32]. In [27], a heuristic algorithm is proposed by collective optimization of the power allocation, and channel assignment that maximizes the EE in an MBS-only network

TABLE 1. Literature review.

Ref	Objective	Constraints	Optimization Type	UE	OMA	NOMA	UE-C	MBS-only	HetNets	Algorithm
[15]	maximize EE	Max. transmit power	Non-convex	✓		✓		✓		SOMSA and DC schemes
[16]	EE	QoS	Non-convex FP	✓		✓		✓		SCA
[17]	EE	QoS and transmit power	Non-convex	✓		✓		✓		Iterative scheme
[18]	EE	QoS and max. transmit power	Mix integer programming	✓		✓		✓		Matching and Lyapunov schemes
[19]	EE	Maintain signal threshold	Non-convex	✓	✓				✓	Heuristic algorithm
[20]	EE	Finite local delay	Non-convex	✓	✓				✓	Random DTx scheme
[21]	EE	Transmit power and QoS	NP-hard	✓	✓				✓	Heuristic and KKT based schemes
[22]	EE	QoS	Non-convex programming	✓	✓				✓	JEERA algorithm
[23]	EE	Limited energy	Non-convex	✓		✓			✓	Matching scheme
[24]	EE	Power and QoS	Non-convex programming	✓		✓			✓	Heuristic algorithm
[25]	EE	Delay and QoS	Non-linear	✓		✓			✓	Heuristic algorithm
[26]	EE	Transmit power	Non-linear	✓		✓			✓	FPA approach
[27]	EE	QoS and transmit power	Pseudo-concave	✓	✓	✓		✓		Heuristic algorithm
[28]	EE	QoS	Non-convex	✓	✓	✓		✓		Dinkelbach's algorithm
[29]	EE	QoS and min. transmit power	MINLP	✓	✓		✓	✓		Heuristic algorithm
[30]	Throughput	QoS	Non-linear	✓	✓	✓		✓		FPA
[31]	EE	QoS	Nash-equilibrium	✓	✓	✓	✓	✓		Heuristic algorithm
[32]	SE and EE	QoS	MOO	✓	✓	✓		✓		SCA and SOC algorithms
Paper	EE	Admission, association, power, QoS	MINLP	✓	✓	✓	✓	✓	✓	OAA

subject to QoS and maximum transmit power constraints. In [28], the authors proposed Dinkelbach's algorithm and SCA method to maximize the EE with a minimum transmission data rate in a cooperative CR network. In [29], EE is maximized in downlink NOMA systems by using the heuristic resource allocation (Heur-RA) algorithm and optimal resource allocation (Opt-RA) algorithm to solve mixed integer non-linear programming (MINLP) problem under QoS and minimum transmit power constraint. In [30], the authors have proposed an H-NOMA scheme for throughput improvement considering desired user traffic volume as well as channel conditions. In [31], a game theory-based joint energy-efficient optimization of resources in cognitive radio networks (CRNs) consisting of secondary networks and primary networks through H-NOMA is studied. In [32], a SE-EE trade-off-based technique for H-NOMA, i.e., a combination of TDMA and NOMA systems is proposed.

B. MOTIVATIONS AND CONTRIBUTIONS

After going through the past literature [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32] and looking at Table. 1, to the best of the author's knowledge, B5G challenges, i.e., maximum UE accommodation, higher data rate, and EE, etc using an H-NOMA, i.e., NOMA and OMA, etc and UE-C modeled in a HetNets has not been explored in the past. We consider these challenges and formulate an optimization problem incorporating UE admission in a cluster, UE association with a BS, power allocation, and H-NOMA scheme in MBS only and HetNets environments. Then, we gauge the performance edge of H-NOMA in HetNets over traditional H-NOMA in MBS only. The main contributions of this work are listed below:

- We formulate an optimization problem incorporating UE admission in a cluster, UE association with a BS, power allocation, and H-NOMA, i.e., OMA and NOMA schemes in MBS only and HetNets environments.
- The problem formulated is a type of non-linear concave fractional programming (CFP) problem. The Charnes-Cooper transformation (CCT) is applied to the formulated non-linear CFP problem to convert it into a concave optimization.
- The MINLP problem is solved using an ϵ-optimal outer approximation algorithm (OAA). This is a two-phase algorithm, i.e., primal stage and master stage. In the primal phase, the MINLP problem is transformed into non-linear programming (NLP) problem to get the upper boundary of the optimal solution. In the master phase, the MINLP problem is transformed into a mixed-integer linear programming (MILP) problem to get a lower boundary of the optimal solution.
- The simulation results show that H-NOMA with HetNets outperforms H-NOMA with MBS only in terms of UE admission and association, throughput, and EE.

The presentation sequence of the paper is as follows: Section II includes the system model and problem formulation. Section III and IV propose algorithm and simulation results, respectively. The conclusion is given in section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. SPATIAL MODEL

Figure 1 shows the H-NOMA scheme based MBS-only and HetNets architecture. The perfect known CSI is assumed at MBS in the MBS-only network and additionally at SBS in HetNets. UEs are uniformly distributed and combined into various clusters based on the best channel gains. The UEs admitted in macro base station clusters (MBCs), i.e., clusters of MBS, and small base station clusters (SBCs), i.e., clusters of SBS, are served by an orthogonal SC rationed to each cluster that evades inter-cluster interference. Symbols used in problem formulation are defined in Table 2.

TABLE 2. Symbols and notations.

Symbol	Definition
N	Total UEs
K	Total clusters
$\mathcal{B}$	Set of BSs
$h^b_{n,k}$	Channel gain value
$\tilde{h}^b_{n,k}$	Rayleigh fading
G_o	Antenna gain
$d^b_{n,k}$	Distance between n^{th} UE admitted in k^{th} cluster with b^{th} BS
$i_{n,k}$	Binary admission variable
$j^b_{n,k}$	Binary association variable
$\lambda^b_{n,k}$	SINR of n^{th} UE admitted in k^{th} cluster associated with b^{th} BS
P^b	Total power at b^{th} BS
N_o	Noise power spectral density
p^b_k	Power assigned to k^{th} cluster associated with b^{th} BS
$p^b_{n,k}$	Received power by n^{th} UE admitted in k^{th} cluster associated with b^{th} BS
$r^b_{n,k}$	Throughput of n^{th} UE admitted in k^{th} cluster associated with b^{th} BS.
R^{min}_n	Minimum rate (Mbps) requirement
ξ	Zero mean gaussian random variable
σ	Standard deviation
α	Path loss exponent
t	Total admitted UEs in k^{th} cluster
d_o	Far-field reference distance

Let $\mathcal{N} = \{1, 2, \ldots, N\}$, where N is the number of UEs assumed in a network. $\mathcal{K} = \{1, 2, \ldots, K\}$ represents the number of clusters. $\mathcal{B} = \{m, s\}$ denotes the BSs, i.e., m = MBS and s = SBS. A downlink H-NOMA scheme is used in the clusters. More clearly, if a cluster is assigned an orthogonal SC that contains only one UE, then SC is defined as OMA-SC, and if a cluster is assigned an orthogonal SC that encompasses more than one UE, then SC is defined as NOMA-SC.

B. BINARY VARIABLES

A binary admission variable $i_{n,k}$ is defined to represent n^{th} UE's admission in k^{th} cluster, and is given below:

$$i_{n,k} = \begin{cases} 1, & \text{if } n^{th} \text{ UE is admitted in } k^{th} \text{ cluster} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

A binary association variable $j^b_{n,k}$ is defined to represent n^{th} UE's association with b^{th} BS, i.e., MBS in MBS-only network and HetNets, already admitted in k^{th} cluster, and is

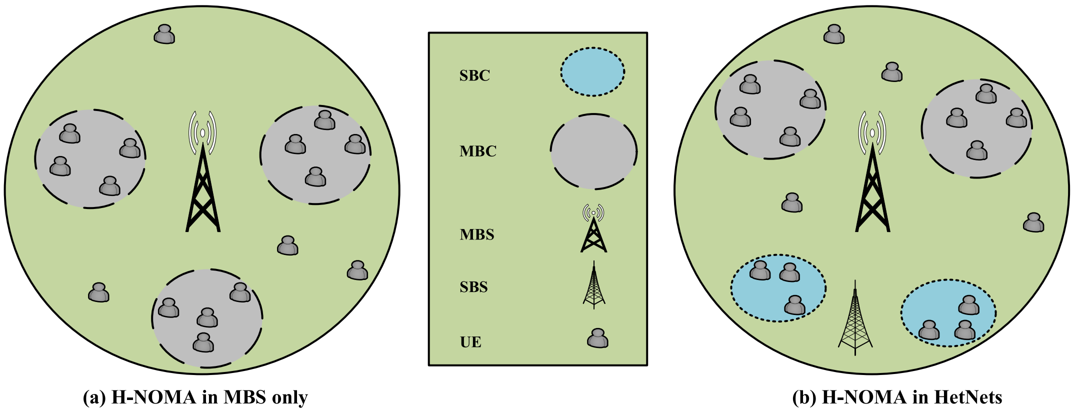

FIGURE 1. H-NOMA scheme in MBS-only and HetNets.

given below:

$$j^b_{n,k} = \begin{cases} 1, & \text{if } n^{th} \text{ UE is associated with } b^{th} \text{ BS} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

C. CHANNEL MODEL

Assume channel gain between n^{th} admitted UE in k^{th} cluster and associated with b^{th} BS is given below:

$$h^b_{n,k} = \tilde{h}^b_{n,k}\xi G_o(\frac{d_o}{d^b_{n,k}})^{\alpha}, \tag{3}$$

where rayleigh fading is indicated as $\tilde{h}^b_{n,k}$, antenna gain is G_o, α is path loss exponent, $d^b_{n,k}$ is the distance between n^{th} admitted UE in k^{th} cluster and associated with b^{th} BS, ξ is gaussian random variable with zero-mean and standard deviation σ in log-normal shadowing, and d_o is the reference distance [33].

D. UE-C WITH H-NOMA

According to the PD-NOMA principle, power domain multiplexing is used at the BS. UEs, decode and receive information by eliminating MUEI using SIC. The arrangement sequence of total t admitted UEs in k^{th} cluster associated with b^{th} BS is from lower channel gain UE to higher channel gain UE can be written as:

$$h^b_{1,k} \le h^b_{2,k} \ldots \le \ldots h^b_{t,k} \tag{4}$$

Accordingly, the received powers by total t UEs admitted in k^{th} cluster associated with b^{th} BS can be given as:

$$p^b_{1,k} \ge p^b_{2,k} \ldots \ge \ldots p^b_{t,k} \tag{5}$$

Assume P^b is the total power at b^{th} BS, (i.e., MBS or SBS) is divided among all the MBCs or SBCs, respectively. A weighting factor $w^b_{n,k}$ is adopted that allocates higher power to UE which has lower channel gain, and allocates lower power to UE which has higher channel gain in k^{th} cluster as given:

$$w^b_{n,k} = 1 - \frac{i_{n,k}j^b_{n,k}h^b_{n,k}}{\sum_{n' \in N} h^b_{n',k}}, \ \forall\ n', n \in \mathcal{N}, n' \ne n \tag{6}$$

where, $b \in \mathcal{B}$ and no UE is admitted and associated if $w^b_{n,k} = 1$. The value of weighting factor can vary between $1 \ge w^b_{n,k} > 0$. The received power of UE is computed as:

$$p^b_{n,k} = w^b_{n,k} * p^b_k \tag{7}$$

where p^b_k is the power allocated to k^{th} cluster in association with b^{th} BS, i.e., MBS or SBS.

E. SIGNAL TO INTERFERENCE PLUS NOISE RATIO (SINR) MODEL

In SIC, n^{th} UE first detects the message of UE n' and then removes the detected message from the desired information in a consecutive way. In an MBS-only network, the n^{th} UE's SINR admitted in the k^{th} cluster in association with b^{th} BS, (i.e., MBS) can be written as:

$$\lambda^b_{n,k} = \frac{i_{n,k}j^b_{n,k}p^b_{n,k}h^b_{n,k}}{h^b_{n,k}\sum_{n' \in N} p^b_{n',k} + N_o}, \quad \forall\ n', n \in \mathcal{N}, n' \ne n \tag{8}$$

where $b \in \mathcal{B}$ and $\mathcal{B} = \{m\}$, and $p^b_{n,k}$ is the received power of n^{th} UE, $p^b_{n',k}$ is the received power of UE n', N_o is the noise power spectral density. In any cluster, n^{th} UE can apply SIC if the received SINR of n^{th} UE is greater than or equal to the received SINR of UE n'.

More specifically, the following inequality is necessary to be satisfied for successful decoding and removing the UE n' signal from n^{th} UE's signal admitted in k^{th} cluster in association with b^{th} BS, (i.e., MBS or SBS) is given as:

$$\begin{aligned} \lambda^b_{n,k} &\ge \lambda^b_{n',k}, \ \forall\ n', n \in \mathcal{N}, n' \ne n \\ h^b_{n,k} &\le h^b_{n',k}, \ \forall\ n', n \in \mathcal{N}, n' \ne n \end{aligned} \tag{9}$$

In HetNets, in addition to the interference in the MBS-only network using UE-C-based downlink H-NOMA, UEs also experience interference between BSs (i.e., MBS and SBS). The n^{th} UE admitted in k^{th} cluster and association with b^{th} BS experience interference from b' BS. The SINR of n^{th} UE admitted in k^{th} cluster in association with b^{th} BS is written as:

$$\lambda_{n,k}^{b} = \frac{i_{n,k} j_{n,k}^{b} p_{n,k}^{b} h_{n,k}^{b}}{h_{n,k}^{b} \sum_{n' \in N} p_{n',k}^{b} + h_{n,k}^{b'} P^{b'} + N_o} \quad (10)$$

$$where\ b, b' \in \mathcal{B}, b' \neq b,\ n', n \in \mathcal{N}, n' \neq n$$

The factor $h_{n,k}^{b'} P^{b'}$ represents the added interference from other BS to the UE associated with one BS.

F. THROUGHPUT MODEL

The data rate achieved by UE is defined as throughput. The throughput of n^{th} UE admitted in k^{th} cluster and associated with b^{th} BS is given as:

$$r_{n,k}^{b} = \log_2(1 + \lambda_{n,k}^{b}) \quad (11)$$

where $b \in \mathcal{B} = \{m\}$ in MBS only network and $b \in \mathcal{B} = \{m, s\}$ in HetNets.

G. PROBLEM FORMULATION

In this subsection, We formulate an optimization problem incorporating UE admission in a cluster, UE association with a BS, and power allocation while considering H-NOMA, i.e., OMA and NOMA schemes in MBS only and HetNets environments. The objective function and constraints to model the optimization problem are defined below:

1) A utility function Δ for maximization of EE is defined using Eq. (1), (2), (7), and (11) as below:

$$\frac{\sum_{n \in N} \sum_{b \in B} \sum_{k \in K} i_{n,k} j_{n,k}^{b} r_{n,k}^{b}}{P_c + \sum_{n \in N} \sum_{b \in B} \sum_{k \in K} p_{n,k}^{b}}. \quad (12)$$

2) Using Eq. (1), the constraint to ensure that n^{th} UE admits in only one k^{th} cluster is given as:

$$\sum_{k \in K} i_{n,k} \leq 1, \ \ \forall\ n \in N. \quad (13)$$

3) Using Eq. (2), the constraint to ensure that n^{th} UE associates with only one b^{th} BS is given as:

$$\sum_{b \in B} \sum_{k \in K} j_{n,k}^{b} \leq 1, \ \ \forall\ n \in N. \quad (14)$$

4) Using Eq. (1) and (2), the constraint to ensure that n^{th} UE admitted in k^{th} cluster must be associated with b^{th} BS is given as:

$$i_{n,k} = j_{n,k}^{b}, \ \ \forall\ n \in N, k \in \mathcal{K}, b \in \mathcal{B}. \quad (15)$$

5) The constraint to ensure power allocation to each MBC or SBC is given as:

$$\sum_{k \in K} p_{k}^{b} \leq P^{b}, \ \ b \in \mathcal{B}. \quad (16)$$

6) Using Eq. (1), (2), and (11), the constraint to ensure QoS data rate for n^{th} UE admitted in k^{th} cluster, and association with b^{th} BS, i.e., MBS or SBS is given below:

$$r_{n,k}^{b} \geq i_{n,k} j_{n,k}^{b} R_{n}^{min}, \ \ \forall\ n \in N, b \in \mathcal{B}, k \in \mathcal{K}. \quad (17)$$

Maximization of EE is achieved when resource allocation results based on constraints defined above are fed to the objective function. Incorporating objective function and constraints defined above, the optimization problem based on H-NOMA, i.e., OMA and NOMA schemes for resource allocation in MBS only and HetNets can be formulated as:

$$\begin{aligned}
&\max_{i,j,p} \frac{\sum_{n \in N} \sum_{b \in B} \sum_{k \in K} i_{n,k} j_{n,k}^{b} r_{n,k}^{b}}{P_c + \sum_{n \in N} \sum_{b \in B} \sum_{k \in K} p_{n,k}^{b}} \\
&\text{s.t.} \\
&\quad C1: \sum_{k \in K} i_{n,k} \leq 1, \ \ \forall\ n \in N \\
&\quad C2: \sum_{b \in B} \sum_{k \in K} j_{n,k}^{b} \leq 1, \ \ \forall\ n \in N \\
&\quad C3: i_{n,k} = j_{n,k}^{b}, \ \ \forall\ n \in N, k \in \mathcal{K}, b \in \mathcal{B} \\
&\quad C4: \sum_{k \in K} p_{k}^{b} \leq P^{b}, \ \ b \in \mathcal{B} \\
&\quad C5: r_{n,k}^{b} \geq i_{n,k} j_{n,k}^{b} R_{n}^{min}, \ \ \forall\ n \in N, b \in \mathcal{B}, \\
&\qquad\quad and\ k \in \mathcal{K} \\
&\quad C6: i_{n,k} \in \{0, 1\}, j_{n,k}^{b} \in \{0, 1\} \\
&\quad C7: p_{k}^{b} \geq 0, p_{n,k}^{b} \geq 0
\end{aligned} \quad (18)$$

The objective of the function in (18) is the maximization of EE (Mbits/sec/watt) considering admission, association, and power allocation to UE subject to the constraints $C1$ to $C7$. The constraint $C1$ indicates that at a time, n^{th} UE can only admit in one k^{th} cluster. The constraint $C2$ indicates that at a time, n^{th} UE can only associate with one b^{th} BS (i.e., MBS in MBS only network and MBS/SBS in HetNets). The constraint $C3$ indicates that n^{th} UE associated with one b^{th} BS must be admitted to one k^{th} cluster. Constraint $C4$ allocates power to MBC or SBC such that entire power at MBS is allocated to MBCs and entire power at SBS is assigned to SBCs, respectively. The constraint $C5$ is related to the minimum rate requirement, i.e., QoS, of n^{th} UE admitted in k^{th} cluster in association with b^{th} BS.

H. ALTERNATIVE PROBLEM FORMULATION

Problem (18) contains concave function in numerator and convex function in denominator, hence categorized as CFP problem, where real-valued functions defined on R^n are $r_{n,k}^{b}$ and $p_{n,k}^{b}$. We have used CCT to convert CFP problem of (18) in to concave optimization problem by substituting $p_{n,k}^{b} = (\frac{u_{n,k}^{b}}{v})$, $p_{k}^{b} = (\frac{u_{k}^{b}}{v})$. Appendix provides CCT to transform CFP problem into a concave optimization problem. The equivalent concave optimization problem is given below:

$$\max_{i,j,v} v \sum_{n \in N} \sum_{b \in B} \sum_{k \in K} i_{n,k} j_{n,k}^{b}$$

$$\log_2(1 + \frac{i_{n,k} j_{n,k}^{b} u_{n,k}^{b} h_{n,k}^{b}}{h_{n,k}^{b} \sum_{n' \in N} u_{n',k}^{b} + h_{n,k}^{b'} P^{b'} v + v N_o})$$

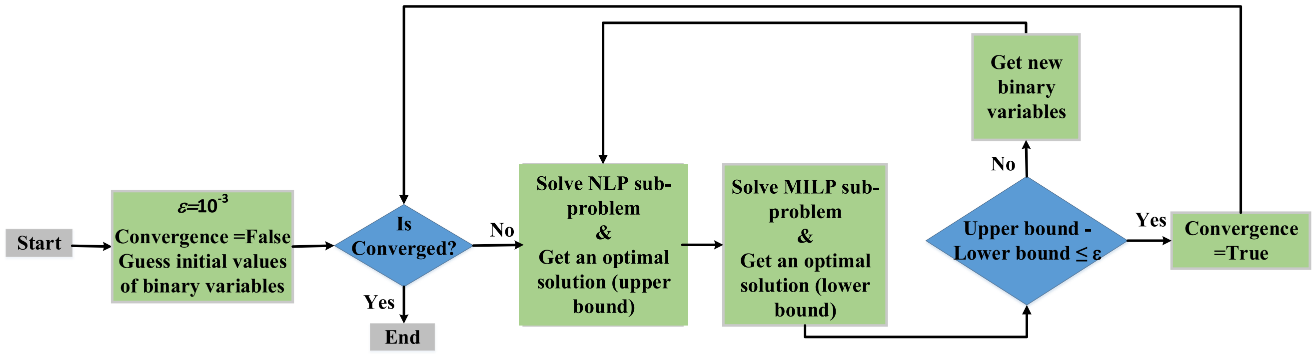

FIGURE 2. Outer approximation algorithm.

s.t.

$$
\begin{aligned}
&C1: \sum_{k\in K} i_{n,k} \le 1, \ \forall\ n \in N\\
&C2: \sum_{b\in B}\sum_{k\in K} j^{b}_{n,k} \le 1, \ \forall\ n \in N\\
&C3: i_{n,k} = j^{b}_{n,k}, \ \forall\ n \in N, b \in \mathcal{B}, k \in \mathcal{K}\\
&C4: \sum_{k\in K} u^{b}_{k} \le P^{b} v, \ \ b \in \mathcal{B}\\
&C5: \log_2(1 + \frac{i_{n,k} j^{b}_{n,k} u^{b}_{n,k} h^{b}_{n,k}}{h^{b}_{n,k}\sum_{n'\in N} u^{b}_{n',k} + h^{b'}_{n,k} P^{b'} v + v N_o})\\
&\qquad \ge i_{n,k} j^{b}_{n,k} R^{min}_{n}, \forall\ n', n \in \mathcal{N}, n' \ne n,\\
&\qquad and\ b, b' \in \mathcal{B}, b' \ne b,\\
&C6: i_{n,k} \in \{0,1\}, j^{b}_{n,k} \in \{0,1\}\\
&C7: u^{b}_{k} \ge 0, u^{b}_{n,k} \ge 0\\
&C8: P_c v + \sum_{n\in N}\sum_{b\in B}\sum_{k\in K} u^{b}_{n,k} = 1
\end{aligned} \tag{19}
$$

The problem in (19) is non-deterministic polynomial-time (NP)-hard in nature and categorized as MINLP. The intensification of discrete variables in polynomial time imposes restrictions on finding an optimal solution with any algorithm. This problem contains binary variables $i_{n,k}, j^{b}_{n,k}$ as well as continuous variables $u^{b}_{n,k}$. There will be an exponential increase in the search space of (19) if the total number of UEs (N) will increase. An optimal solution can be obtained using an exhaustive search algorithm (ESA) on binary variables. But its complexity is high because we will have to solve $2^{|N|}$ optimization problems if the search space for binary variables is $2^{|N|}$. That's why we have proposed OAA to reach a near-optimal solution to guarantee convergence [33]. A detailed implementation of OAA [34] for the MINLP problem in (19) is given in the next section.

III. OUTER APPROXIMATION ALGORITHM

The MINLP problem in (19) contains binary, integer, and continuous variables. The binary variables are UEs admission and UEs association. The integer variables are the numbers of UE. Continuous variables are related to the received power by UE. The medley of all mentioned variables and their non-linear behavior shapes the problem in (19) significantly challenging and complex. Nevertheless, the special design of the problem directed us to employ OAA to solve this problem. OAA delivers ϵ-optimal values of u^b and v, employing these values in (18), we obtain ϵ-optimal values for EE. Figure 2 presents flow diagram of OAA.

A. DESCRIPTION OF ALGORITHM

Assume objective function as $\mho$ for C1 to C8, set of constraints as ψ_{C1-C8}, $\pi = \{p^{b}_{n,k}\}$ where $b \in \mathcal{B}$, and $\chi = \aleph \cup \pi$ in (19). The problem in (19) satisfies the following propositions.

1) π is convex, compact, and non-empty. Objective function $\mho$ and the constraints ψ_{C1-C8} are convex in π, if the values of χ are fixed.
2) Objective function $\mho$ and constraints ψ_{C1-C8} are continually differentiable once.
3) Fixing χ solves each nonlinear continuous subproblem for satisfying the constraints.
4) Fixing χ makes the possibility for the absolute solution of the NLP problem.

All four propositions for the problem in (19) are satisfied as to the sequences of non-decreasing lower bounds and non-increasing upper bounds for MINLP and are computed iteratively with OAA. OAA is converged in an infinite number of iterations with convergence capability ϵ [35].

Problem (19) is divided into the master and primal problems for finding the lower and upper bound sequences, respectively. The primal problem is described by fixing χ and χ^{y} is integer variable's value at the y^{th} iteration for problem in (19) is given as:

$$
\begin{aligned}
&\min_{\pi} - \mho\left(\chi^{y}, \pi\right)\\
&\text{subject to}\\
&\qquad\qquad \psi_{C1-C8}\left(\chi^{y}, \pi\right) \le 0
\end{aligned} \tag{20}
$$

The solution to the primal problem in (20) is provided by exercising outer approximation (successive linearization) with fixed values of binary variables. The problem in (20) is solved to find the values of π^{y} used in the master problem. The solution of the primal problem provides upper bounds and the solution of the master problem provides lower bounds. Primal solution π^y helps in deriving the master problems. The master problem is derived around the primal solution π^y by using OAA to make the constraint functions ψ_{C1-C8} and the objective function $\mho$ linear [36], [37].

The master problem is solved to find the integer variables that are applied in the next iteration χ^{y+1}. The difference between the upper and lower bounds becomes minimum if the algorithm continues. When the difference between these two values becomes less than ϵ, then the algorithm quits. The problem in (19) is modified as below:

$$\begin{aligned} &\min_{\chi}\min_{\pi} -\mho\left(\chi^y,\pi\right)\\ &\text{subject to}\\ &\qquad\qquad \psi_{C1-C8}\left(\chi^y,\pi\right)\leq 0 \end{aligned} \tag{21}$$

The problems in (21) is revised as below:

$$\begin{aligned} &\min_{\chi} -\mu\left(\chi\right)\\ &\textit{such that}: \mu\left(\chi\right)=\min_{\pi} -\mho\left(\chi^y,\pi\right)\\ &\text{subject to } \psi_{C1-C8}\left(\chi^y,\pi\right)\leq 0 \end{aligned} \tag{22}$$

The projection of (19) on χ space is provided in the problem (22). All constraints are contained in primal problem (20) for all χ^y. Therefore, the solution of the problem (22) is written as below:

$$\begin{aligned} &\min_{\psi}\min_{\pi} -\mho\left(\chi^y,\pi^y\right)-\nabla\mho\left(\chi^y,\pi^y\right)\begin{pmatrix}\pi-\pi^y\\ \chi-\chi^y\end{pmatrix}\\ &\text{subject to}\\ &\qquad \psi_{C1-C8}\left(\chi^y,\pi^y\right)\\ &\qquad -\nabla\psi_{C1-C8}\left(\chi^y,\pi^y\right)\begin{pmatrix}\pi-\pi^y\\ \chi-\chi^y\end{pmatrix}\leq 0 \end{aligned} \tag{23}$$

The equivalent minimization problem with the introduction of a new variable κ can be given as below:

$$\begin{aligned} &\min_{\chi,\pi,\kappa}\kappa\\ &\text{subject to}\\ &\qquad \kappa\geq -\mho\left(\chi^y,\pi^y\right)-\nabla\mho\left(\chi^y,\pi^y\right)\begin{pmatrix}\pi-\pi^y\\ \chi-\chi^y\end{pmatrix}\\ &\qquad \psi_{C1-C8}\left(\chi^y,\pi^y\right)\\ &\qquad -\nabla\psi_{C1-C8}\left(\chi^y,\pi^y\right)\begin{pmatrix}\pi-\pi^y\\ \chi-\chi^y\end{pmatrix}\leq 0 \end{aligned} \tag{24}$$

The master problem in (24) results in lower bound values. Upon satisfaction of the corresponding propositions 1, 2 and 3, the master problem in (24) becomes equivalent to (19). The branch and bound algorithm is used to find the solution to the MILP problem, i.e., master problem in (24) [38].

B. ALGORITHM CONVERGENCE AND OPTIMALITY

OAA converges linearly [35], [36]. OAA optimally solves concave optimization problems having convex constraints and objective function by fixing values of χ. The branch and bound architecture is used to provide an optimal solution within $\epsilon = 10^{-3}$ by fixing values of χ in OAA. If an optimal solution is found within $\epsilon = 10^{-3}$ after satisfying all four propositions optimally for fixed values of χ then it results in termination of OAA in a finite number of steps. If κ is greater than $\mho\left(\chi^y,\pi^y\right)$ for any feasible point in (24) then it shows the optimality of π in (24). For master problem, feasible solution does not exist for given value of χ when κ is less than $\mho\left(\chi^y,\pi^y\right)$. If there exists no feasible solution for any value of χ^y in (24), it will not be included for subsequent master problems. Hence, the algorithm's convergence is achieved.

C. COMPLEXITY OF ϵ-OPTIMAL ALGORITHM

The complexity of the ϵ-optimal algorithm is determined in this section. The complexity is determined by counting F flops. A flop is a real floating-point operation. 1 flop is added for each addition, multiplication, or division operation. 2 flops are added for complex addition and 4 flops are added for complex multiplication. Moreover, the addition or removal of an element from the set adds to 1 flop. 2abc flops are added for the multiplication of matrix a × b with matrix b × c. Likewise, each assignment operator and logical operator adds to 1 flop [39].

Therefore, the initialization of OAA adds 5 flops. Solution of the NLP problem adds $2AE$ flops. The upper bound of the optimal solution adds $4AE\Im$ flops. Solution of MILP problem adds $4AE\Im$ flops. The lower bound of the optimal solution adds $2AE\Im$ flops. 2 flops are added for comparing the upper and lower bounds. 4 flops are added for the initialization of new binary variables. The sum of F_{OAA} flops is given as:

$$\begin{aligned} F_{OAA} &= 5+2AE+4AE\Im+4AE\Im+2AE\Im+4\\ F_{OAA} &= 9+2AE+10AE\Im\\ F_{OAA} &\approx 2AE+10AE\Im \end{aligned} \tag{25}$$

The complexity of OAA is given in (25). Likewise, the complexity of OAA is determined using Big O notation as O(A × E) + O(A × E × $\Im$). The numbers of UE are A, the numbers of BS are E, and the total constraints are $\Im$.

IV. SIMULATION RESULTS

The key performance indicators (KPIs) of the proposed scenario in HetNets and the MBS-only network are evaluated using ϵ-optimal OAA. Basic open-source non-linear mixed integer (BONMIN) programming [40] is implemented in OAA. Admission of UE in a cluster, the association of UE with BS, network throughput, and EE are the considered KPIs. The simulation parameters used are given in Table 3.

It is assumed that the minimum numbers of UE are 5 with an increment of 10 UEs up to the maximum allowed 85 UEs. The maximum coverage area of MBS and SBS is 1000 m and 300 m, respectively. The minimum required rates for a UE (i.e., R_n^{min}) are set to {0.5, 1, 2, 3} Mbps. The antenna gain (i.e., G_o) is 50. The far-field reference distance (i.e., d_o) is 10 m, path loss exponent (i.e., α) is 2, and Gaussian random variable in log-normal shadowing (i.e., ξ) is 10 dB. The total

TABLE 3. Simulation parameters.

Parameters	Values
Min UEs	5
UE Increment	10
Max UEs	85
UE Distribution	Uniform
MBS Coverage	1000 m
SBS Coverage	300 m
R_n^{min}	$\{0.5, 1, 2, 3\}$ Mbps
G_o	50
d_o	10 m
α	2
ξ	10 dB
P^m	43 dBm
P^s	33 dBm
P_c	-30 dBm

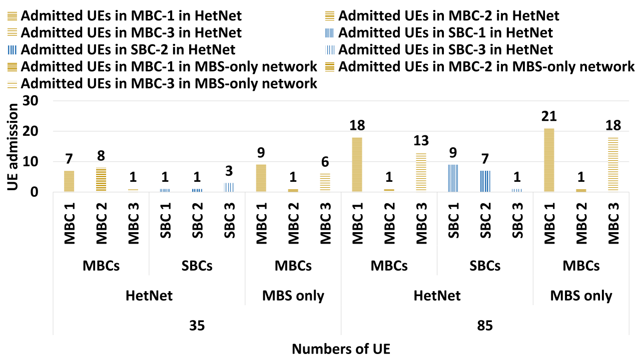

FIGURE 3. UE admission vs number of UEs for required rate $R_n^{min} =$ 2 Mbps.

power for MBS (i.e., P^m) and SBS (i.e., P^s) is 43 dBm and 33 dBm, respectively. The circuit power (i.e., P_c) is -30 dBm.

Figure 3 presents the UE admission in the clusters concerning the number of UEs in an MBS-only network and HetNets. Figure 3 shows that the H-NOMA scheme employs OMA when only one UE is admitted in MBC or SBC. However, the H-NOMA scheme employs NOMA when more than one UE is admitted in MBC or SBC. It is evident in Figure 3 that UE admission increases with the increase in the number of UEs in both the networks. But in the HetNets, the admission of the UE in clusters is higher than in the MBS-only network. This performance edge of the UEs admission in the HetNets over MBS only is because a UE based on the best SINR can be admitted either in MBC or SBC in the HetNets. The proposed scheme offers UEs to be admitted in UE-C using OMA and NOMA in the MBC and SBC in the HetNets. Here, SBC of SBS offers services to the UEs in the dead zones of MBC of MBS. However, H-NOMA in the MBS only could not reach the UEs in the dead zones of MBS in the MBS only. Therefore, the H-NOMA scheme in HetNets outperforms the H-NOMA scheme in MBS only in terms of UE admission in the clusters.

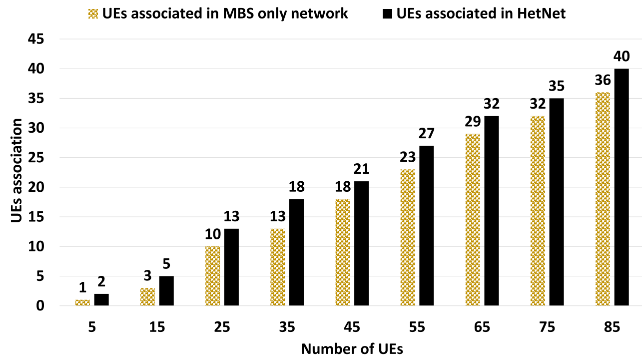

FIGURE 4. UE association vs number of UEs.

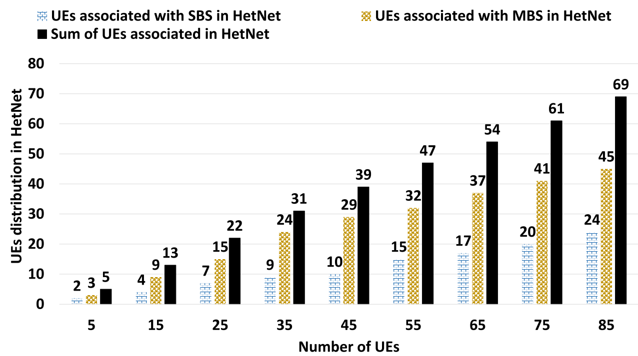

FIGURE 5. UE distribution/association in HetNets for required rate $R_n^{min} = 0.5$ Mbps.

Figure 4 shows performance in terms of energy-efficient UE association employing UE-C based downlink H-NOMA for $\mathcal{N} = \{5, 15, 25, 35, 45, 55, 65, 75, 85\}$ UEs and required rate $R_n^{min} = 3$ Mbps in MBS only network and HetNets. It can be seen that if we increase the number of UEs, then the numbers of associated UEs also increase in both networks. But UE association (i.e., based on best SINR) is more in HetNets than in the MBS-only network due to the availability of SBS in addition to MBS. Availability of diverse transmit power BSs, i.e., MBS and SBS, etc coupled with the H-NOMA scheme in HetNets offer multiple options to UEs for association (i.e., based on best SINR) using OMA and NOMA access schemes. This results in more UEs association in the HetNets as compared to the MBS only.

Figure 5 shows UEs distribution/association in the HetNets. The HetNets offers diverse transmit power BSs, i.e., MBS and SBS, etc coupled with UE-C-based NOMA and OMA access schemes for UEs association in the network. In the HetNets, MBS extends seamless coverage and SBS covers dead zones not covered by the MBS in the network. Here, MBS overrides SBS in terms of transmit power, therefore, more UEs are associated with MBS and fewer UEs are associated with SBS in the HetNets. This results in more UEs association in the HetNets as compared to the MBS only.

Figure 6 shows the plot of associated UEs versus required rate $R_n^{min} = \{0.5, 1, 2, 3\}$ Mbps and numbers of UEs for $\mathcal{N} = \{5, 15, 25, 35, 45, 55, 65, 75, 85\}$ in MBS-only network and HetNets. It is clear in the UEs association plot that the

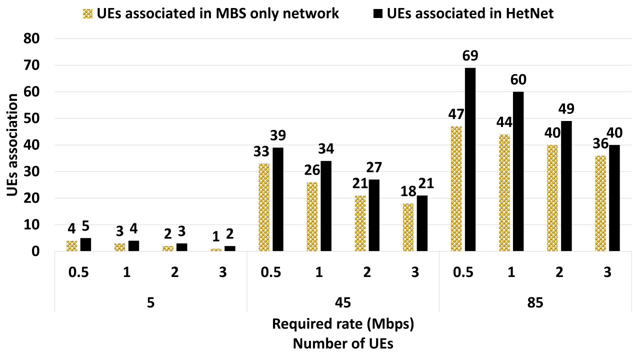

FIGURE 6. Associated UEs with respect to numbers of UE and required rate.

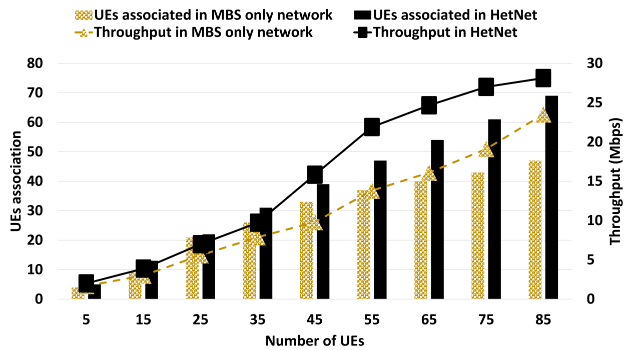

FIGURE 7. UE association and Throughput vs number of UEs.

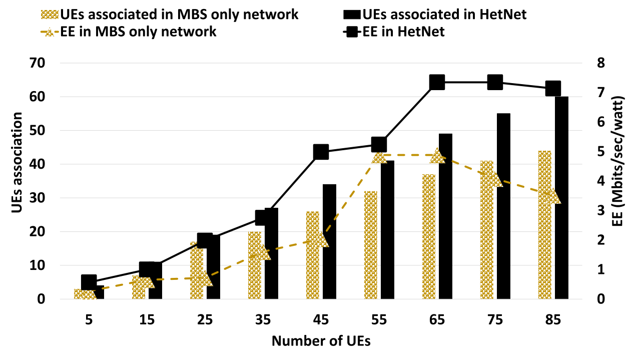

FIGURE 8. EE and UE association vs numbers of UE.

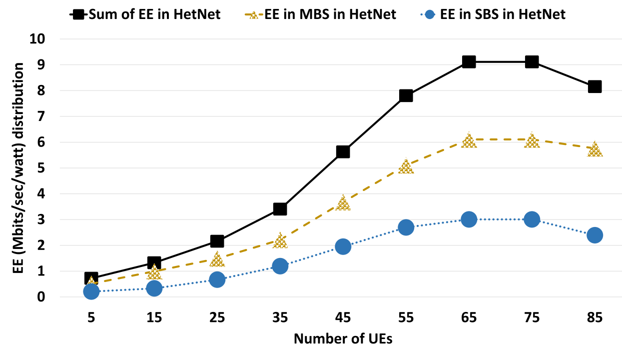

FIGURE 9. EE distribution in HetNets vs number of UEs.

numbers of associated UEs increase with the increase of the number of UEs in the network. However, the number of associated UEs decrease with the increase in the QoS required data rate. As the QoS required data rate is increased, fewer UEs qualify to maintain the higher QoS required data rates. Since, the HetNets offers diverse transmit power BSs, i.e., MBS and SBS, etc coupled with UE-C-based NOMA and OMA access schemes for UEs association in the network. Therefore, the HetNets outperforms the MBS-only network at higher QoS required data rates.

Figure 7 shows the performance in terms of UEs association versus throughput in UE-C based downlink H-NOMA scheme in MBS only and HetNets. Simulation is run for a minimum of 5 UEs and a maximum of 85 UEs with an increment of 10 UEs in each iteration. The simulation results of throughput and UEs association for the QoS required data rate $R_n^{min} = 0.5$ Mbps in both HetNets and the MBS-only network are shown in Figure 7. UEs association viz-a-viz throughput simulation result in the HetNets outperforms the simulation result in MBS only network. Since an optimal solution is provided with UE-C based downlink H-NOMA technique for effective resource allocation. Consequently, the UE-C-based downlink H-NOMA strategy in HetNets outperforms the UE-C-based downlink H-NOMA strategy in the MBS-only network in terms of throughput and UE association in the network.

Figure 8 shows EE versus UE association when the numbers of UEs are $\mathcal{N} = \{5, 15, 25, 35, 45, 55, 65, 75, 85\}$ and the QoS required data rate is $R_n^{min} = 1$ Mbps. The result shows that EE and UE association both increase as the number of UEs increases up to a certain limit in the network. If we further raise the number of UEs, EE decreases because the system's maximum capacity is achieved. Hence, the performance regarding UE association and EE versus the number of UEs in HetNets is better than in the MBS-only network.

Figure 9 shows the performance of UE-C based downlink H-NOMA strategy in HetNets for the required rate $R_n^{min} = 0.5$ Mbps, and the numbers of UEs $\mathcal{N} = \{5, 15, 25, 35, 45, 55, 65, 75, 85\}$ in terms of EE distribution versus the number of UEs. It depicts that if the number of UEs rises, then EE also rises to a certain limit, and after that if the number of UEs further increases then EE decreases due to the system capacity limit achieved. But MBS performs better compared to SBS in HetNets.

Figure 10 shows the plot of throughput and EE versus the required rate $R_n^{min} = \{0.5, 1, 2, 3\}$ Mbps for 85 UEs to evaluate the performance of UE-C based downlink H-NOMA in MBS-only network and HetNets. This plot shows the effect of an increase in QoS required data rate on average throughput and EE when we employ H-NOMA in MBS only and HetNets. It was found in figure 6 that UEs association starts dropping when QoS required data rate is increased. Effect of UEs association dropping at higher QoS required data rates also have a negative impact on throughput and EE when QoS required data rates are increased in the simulations as evident in the figure 10. However, this negative effect has a major effect in the MBS case and a minor effect in

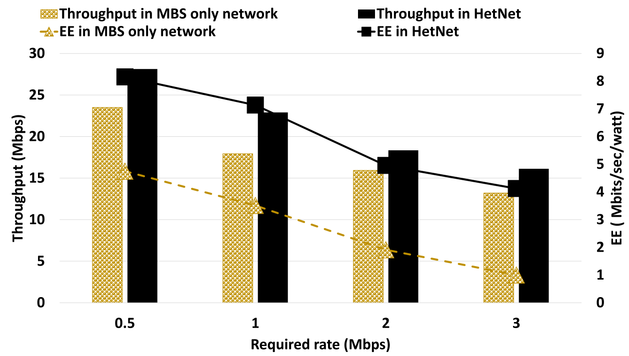

FIGURE 10. Throughput and EE vs required rate.

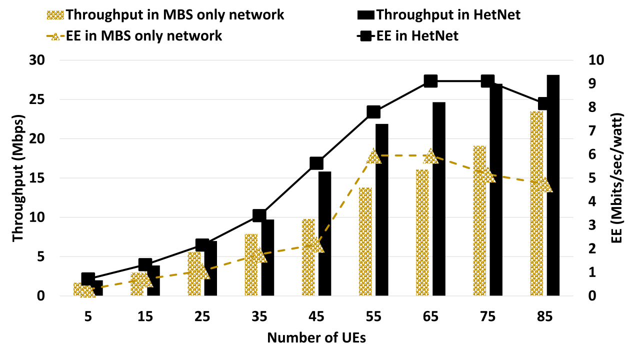

FIGURE 11. Throughput (Mbps) and EE (Mbits/sec/watt) vs numbers of UE for required rate $R_n^{min} = 0.5$ Mbps.

the HetNets case. Thus, the H-NOMA scheme in HetNets outperforms the H-NOMA scheme in MBS only.

Figure 11 shows the the performance of UE-C based downlink H-NOMA in HetNets and MBS-only network is analyzed for throughput and EE versus the numbers of UEs. The result shows that the throughput and EE in both HetNets and the MBS-only network increase when we increase the number of UEs up to a certain number. But if we further increase the numbers of UEs in the network then a slight decrease in EE appears but throughput increases. This figure also shows the better performance of HetNets over the MBS-only network regarding throughput and EE.

V. CONCLUSION

This research work considers an optimization problem incorporating UE admission in a cluster, UE association with a BS, power allocation and H-NOMA, i.e., OMA and NOMA schemes in MBS only and HetNets environments. The HetNets offers to diversify coverage and transmit power by BSs as compared to MBS only where coverage and transmit power is fixed. In the HetNet, the SBS offers services to the UEs in the dead zones of the MBS. However, the services of the H-NOMA in the MBS only could not reach the UEs in the dead zones of MBS in the MBS only. The proposed H-NOMA scheme employs OMA when only one UE is admitted in MBC or SBC. However, the H-NOMA scheme employs NOMA when more than one UE are admitted to MBC or SBC. Therefore, the H-NOMA scheme in HetNets outperforms the H-NOMA scheme in MBS only in terms of UEs admission in the clusters, UEs association with the BSs, throughput and EE. Spectrum resource is becoming scarce in terms of demand and supply in B5G wireless communication. This research work will be extended in Space-Air-Ground Integrated Vehicular Networks (SAGVN) to address this challenging issue in B5G cellular networks.

APPENDIX.
FRACTIONAL PROGRAMING AND CHARNES COOPER TRANSFORMATION

FP contains objective function as a ratio of two nonlinear functions generally. A FP can be described mathematically as in (26):

$$\max_{t \in S} \frac{x(t)}{y(t)}$$
subject to
$$C1 : h_n(t) \leq 0 \tag{26}$$

where $S \subset R^n$ is a set of real values, described by $x(t)$, $y(t)$ and $h_n(t)$ where $\mathcal{N}=\{1, 2, \ldots, N\}$.

In (26), assuming S is convex set, $x(t)$ is $+ive$ and concave, $y(t)$ is $+ive$ and convex makes FP as CFP. CCT reduce a CFP to a concave programme [41], as given below in (27), (28):

$$u = \frac{t}{y(t)} \tag{27}$$

$$v = \frac{1}{y(t)} \tag{28}$$

The concave problem for (26) equivalently can be described as below in (29):

$$\max_{\frac{u}{v} \in S} vx_o\frac{u}{v}$$
subject to
$$C1 : vy(\frac{u}{v}) = 1,$$
$$C2 : vh_n(\frac{u}{v}) \leq 0, \forall\ n = 1, 2, 3, \ldots, N. \tag{29}$$

Only if the problem in (29) have optimal solution, then the problem in (26) can have optimal solution.

REFERENCES

[1] P. Karmakar, R. V. Rajakumar, and R. Roy, "A survey on energy efficient cellular mobile communication," *Wireless Pers. Commun.*, vol. 120, no. 2, pp. 1475–1500, Sep. 2021.

[2] L. Williams, B. K. Sovacool, and T. J. Foxon, "The energy use implications of 5G: Reviewing whole network operational energy, embodied energy, and indirect effects," *Renew. Sustain. Energy Rev.*, vol. 157, Apr. 2022, Art. no. 112033.

[3] X. Cheng, Y. Hu, and L. Varga, "5G network deployment and the associated energy consumption in the U.K.: A complex systems' exploration," *Technol. Forecasting Social Change*, vol. 180, Jul. 2022, Art. no. 121672.

[4] H. Z. Khan, M. Ali, I. Rashid, A. Ghafoor, and M. Naeem, "Cell association for energy efficient resource allocation in decoupled 5G heterogeneous networks," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC-Spring)*, May 2020, pp. 1–5.

[5] H. Z. Khan, M. Ali, I. Rashid, A. Ghafoor, M. Naeem, A. A. Khan, and A. M. Saddiqui, "Resource allocation for energy efficiency optimization in uplink–downlink decoupled 5G heterogeneous networks," *Int. J. Commun. Syst.*, vol. 34, no. 14, p. e4925, Sep. 2021.

[6] H. Z. Khan, M. Ali, M. Naeem, I. Rashid, S. Mumtaz, A. A. Khan, and A. N. Akhtar, "Secure resource management in beyond 5G heterogeneous networks with decoupled access," *Ad Hoc Netw.*, vol. 125, Feb. 2022, Art. no. 102737.

[7] U. Ghafoor, M. Ali, H. Z. Khan, A. M. Siddiqui, and M. Naeem, "Throughput maximization in hybrid NOMA assisted beyond 5G heterogeneous networks," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Jun. 2021, pp. 991–996.

[8] H. Z. Khan, M. Ali, M. Naeem, I. Rashid, A. N. Akhtar, and F. Akram, "Joint DL/UL decouple user association in microwave and mmWave enabled beyond 5G heterogeneous networks," *IEEE Access*, vol. 9, pp. 134703–134715, 2021.

[9] X. Sun and S. Wang, "Resource allocation scheme for energy saving in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4407–4416, Aug. 2015.

[10] A. S. Marcano and H. L. Christiansen, "A novel method for improving the capacity in 5G mobile networks combining NOMA and OMA," in *Proc. IEEE 85th Veh. Technol. Conf. (VTC Spring)*, Jun. 2017, pp. 1–5.

[11] W. Yu, L. Musavian, and Q. Ni, "Link-layer capacity of NOMA under statistical delay QoS guarantees," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4907–4922, Oct. 2018.

[12] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[13] K. Deka and S. Sharma, "Hybrid noma for future radio access: Design, potentials and limitations," *Wireless Pers. Commun.*, vol. 123, no. 4, pp. 1–16, 2021.

[14] D. Zhang, Q. Wu, M. Cui, G. Zhang, and D. Niyato, "Throughput maximization for IRS-assisted wireless powered hybrid NOMA and TDMA," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1944–1948, Sep. 2021.

[15] F. Fang, H. Zhang, J. Cheng, and V. C. Leung, "Energy efficiency of resource scheduling for non-orthogonal multiple access (NOMA) wireless network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–5.

[16] Y. Zhang, Q. Yang, T.-X. Zheng, H.-M. Wang, Y. Ju, and Y. Meng, "Energy efficiency optimization in cognitive radio inspired non-orthogonal multiple access," in *Proc. IEEE 27th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2016, pp. 1–6.

[17] Z. Liu, G. Kang, L. Lei, N. Zhang, and S. Zhang, "Power allocation for energy efficiency maximization in downlink CoMP systems with NOMA," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–6.

[18] H. Zhang, B. Wang, C. Jiang, K. Long, A. Nallanathan, V. C. Leung, and H. V. Poor, "Energy efficient dynamic resource optimization in NOMA system," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 5671–5683, Sep. 2018.

[19] Y. A. Sambo, M. Z. Shakir, K. A. Qaraqe, E. Serpedin, M. A. Imran, and B. Ahmed, "Energy efficiency improvements in hetnets by exploiting device-to-device communications," in *Proc. 22nd Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2014, pp. 151–155.

[20] W. Nie, Y. Zhong, F.-C. Zheng, and W. Zhang, "Local delay and energy efficiency analysis in HetNets with random DTX scheme," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 1857–1862.

[21] J. Yu, G. Yu, Y. Wang, and Q. Zhang, "Energy-efficient radio resource management in a HetNet downlink system with JT CoMP," in *Proc. IEEE/CIC Int. Conf. Commun. China-Workshops (CIC/ICCC)*, Nov. 2015, pp. 142–147.

[22] S. Chege and T. Walingo, "Energy efficient resource allocation for uplink hybrid power domain sparse code nonorthogonal multiple access heterogeneous networks with statistical channel estimation," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 1, p. e4185, Jan. 2021.

[23] H. Zhang, M. Feng, K. Long, G. K. Karagiannidis, V. C. M. Leung, and H. V. Poor, "Energy efficient resource management in SWIPT enabled heterogeneous networks with NOMA," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 835–845, Feb. 2019.

[24] H. Zhang, F. Fang, J. Cheng, K. Long, W. Wang, and V. C. M. Leung, "Energy-efficient resource allocation in NOMA heterogeneous networks," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 48–53, Apr. 2018.

[25] L. Lei, E. Lagunas, S. Maleki, Q. He, S. Chatzinotas, and B. Ottersten, "Energy optimization for full-duplex self-backhauled HetNet with non-orthogonal multiple access," in *Proc. IEEE 18th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2017, pp. 1–5.

[26] L. Xiang and H. Chen, "Energy-efficient and fair power allocation approach for NOMA in ultra-dense heterogeneous networks," in *Proc. Int. Conf. Cyber-Enabled Distrib. Comput. Knowl. Discovery (CyberC)*, Oct. 2017, pp. 89–94.

[27] M. Zeng, A. Yadav, O. A. Dobre, and H. V. Poor, "Energy-efficient joint user-RB association and power allocation for uplink hybrid NOMA-OMA," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5119–5131, Jun. 2019.

[28] S. Sheikhzadeh, M. R. Javan, and N. Mokari, "Cooperative multiple access cognitive radio transmission with renewable energy sources," *Phys. Commun.*, vol. 40, Jun. 2020, Art. no. 101049.

[29] J. Shi, W. Yu, Q. Ni, W. Liang, Z. Li, and P. Xiao, "Energy efficient resource allocation in hybrid non-orthogonal multiple access systems," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3496–3511, May 2019.

[30] F. Tanaka, H. Suganuma, and F. Maehara, "Hybrid multiple access scheme using NOMA and OMA simultaneously considering user request," in *Proc. 24th Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Dec. 2021, pp. 1–5.

[31] A. Kumar and K. Kumar, "A game theory based hybrid NOMA for efficient resource optimization in cognitive radio networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 4, pp. 3501–3514, Oct. 2021.

[32] X. Wei, H. Al-Obiedollah, K. Cumanan, Z. Ding, and O. A. Dobre, "Energy efficiency maximization for hybrid TDMA-NOMA system with opportunistic time assignment," *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 8561–8573, Aug. 2022.

[33] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[34] R. Fletcher and S. Leyffer, "Solving mixed integer nonlinear programs by outer approximation," *Math. Program.*, vol. 66, no. 1, pp. 327–349, Aug. 1994.

[35] M. A. Duran and I. E. Grossmann, "An outer-approximation algorithm for a class of mixed-integer nonlinear programs," *Math. Program.*, vol. 36, no. 3, pp. 307–339, Oct. 1986.

[36] C. A. Floudas and P. M. Pardalos, *Encyclopedia of Optimization*. USA: Springer, 2008.

[37] C. A. Floudas and E. N. Pistikopoulos, "Non-linear and mixed-integer optimization. fundamentals and applications," *J. Global Optim.*, vol. 12, no. 1, p. 108, 1998.

[38] A. H. Land and A. G. Doig, "An automatic method for solving discrete programming problems," in *50 Years of Integer Programming 1958–2008*. Berlin, Germany: Springer, 2010, pp. 105–132.

[39] G. Golub and L. Van, "CF (1983). Matrix computations," 1st ed., Johns Hopkins Univ. Press, Baltimore, MD, USA, Tech. Rep., 1983.

[40] P. Bonami, P. Belotti, J. Forrest, L. Ladanyi, and C. Laird, "Basic open-source nonlinear mixed integer programming," Aug. 2019. [Online]. Available: http://www.coin-or.org/Bonmin/

[41] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals," *Naval Res. Logistics Quart.*, vol. 9, nos. 3–4, pp. 181–186, 1962.

UMAR GHAFOOR received the M.S. degree in electrical engineering from the National University of Sciences and Technology, Pakistan. He has been associated with the profession of teaching and research, since 2013. His research interests include wireless communications, resource allocation, optimization theory, heterogeneous networks, and cognitive radio networks.

HUMAYUN ZUBAIR KHAN (Senior Member, IEEE) received the B.E. degree in electrical (telecommunication) engineering from the National University of Sciences and Technology (NUST), Pakistan, in 2006, the Master of Business Administration degree with majors in finance from the Virtual University of Pakistan, in 2011, the M.S. degree in electrical (telecommunication) engineering from the NUST, in 2013, and the Ph.D. degree in electrical engineering from the Military College of Signals (MCS), NUST, in 2020. His research interests include resource allocation, interference reduction, and spectrum and energy efficiency in 5G heterogeneous networks. He was a recipient of the Indigenous Ph.D. Fellowship Program Scholarship.

MUDASSAR ALI received the B.S. degree in computer engineering and the M.S. degree in telecom engineering with a major in wireless communication from the University of Engineering and Technology, Taxila, Pakistan, in 2006 and 2010, respectively, and the Ph.D. degree from the School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Pakistan, in 2017. From 2006 to 2007, he worked as a Network Performance Engineer with Mobilink (An Orascom Telecom Company). From 2008 to 2012, he worked as a Senior Engineer Radio Access Network Optimization with Zong (A China Mobile Company). Since 2012, he has been an Assistant Professor at the Telecom Engineering Department, University of Engineering and Technology, Taxila, Pakistan. His research interests include 5G wireless systems, heterogeneous networks, interference coordination, and energy efficiency in 5G green heterogeneous networks.

ADIL MASOOD SIDDIQUI received the bachelor's degree in telecom engineering from the Military College of Signals, Rawalpindi, Pakistan, in 1994, and the master's degree in electronics and telecommunication and the Ph.D. degree in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2005 and 2009, respectively. He has been the Faculty of the Military College of Signals, National University of Sciences and Technology, since 2009. He has number of research publications at his credit. His research interests include image registration, de-noising, image enhancement, and defogging.

MUHAMMAD NAEEM received the B.S. and M.S. degrees in electrical engineering from the University of Engineering and Technology, Taxila, Pakistan, in 2000 and 2005, respectively, and the Ph.D. degree from Simon Fraser University, Burnaby, BC, Canada, in 2011. From 2000 to 2005, he was a Senior Design Engineer with Comcept (Pvt) Ltd., Islamabad, Pakistan, where he participated in the design and development of smart card-based GSM and CDMA pay phones with the Department of Design. From 2012 to 2013, he was a Postdoctoral Research Associate with the Wireless Networks and Communications Research (WINCORE) Laboratory, Ryerson University, Toronto, ON, Canada. From 2013 to 2016, he was an Assistant Professor, and since 2016, he has been an Associate Professor with the Department of Electrical Engineering, COMSATS University Islamabad, Wah Campus, Wah, Pakistan. Since 2013, he has been a Research Associate with the WINCORE Laboratory. He is also a Microsoft Certified Solution Developer. His research interests include optimization of wireless communication systems, nonconvex optimization, resource allocation in cognitive radio networks, and approximation algorithms for mixed integer programming in communication systems. He was a recipient of the NSERC CGS Scholarship.

IMRAN RASHID received the B.E. degree in electrical (telecomm) engineering from the National University of Sciences and Technology, Pakistan, in 1999, the M.Sc. degree in telecomm engineering (optical communication) from the DTU, Denmark, in 2004, and the Ph.D. degree in mobile communication from The University of Manchester, U.K., in 2011. He has been qualified for four EC-Council certifications, i.e., Certified Ethical Hacker, Computer Hacking Forensic Investigator, EC-Council Certified Security Analyst, and EC-Council Certified Incident Handler. He is also a Certified EC-Council Instructor and has conducted numerous trainings. He is currently the Chief Instructor at Engineering Wing, MCS, National University of Sciences and Technology. His research interests include mobile and wireless communication, MIMO systems, compressed sensing for MIMO OFDM systems, massive MIMO systems, M2M for mobile systems, cognitive radio networks, cyber security, and information assurance.