Exploiting NOMA for Radio Resource Efficient Traffic Steering Use-case in O-RAN

Muhammad Waseem Akhtar[1], Aamir Mahmood[1], Sarder Fakhrul Abedin[1], Syed Ali Hassan[2], Mikael Gidlund[1]
[1] Department of Information Systems & Technology, Mid Sweden University, 851 70 Sundsvall, Sweden
[2] School of Electrical Engineering and Computer Science (SEECS)
National University of Sciences and Technology (NUST) H-12, Islamabad, Pakistan
Email: [1]{muhammadwaseem.akhtar, aamir.mahmood, sarder.fakhrulabedin, mikael.gidlund}@miun.se,
[2]ali.hassan@seecs.edu.pk

***Abstract*—In this work, we consider the design of a radio resource management (RRM) solution for traffic steering (TS) use-case in the open radio access network (O-RAN). The O-RAN TS deals with the quality-of-service (QoS)-aware steering of the traffic by connectivity management (e.g., device-to-cell association, radio spectrum, and power allocation) for emerging heterogeneous networks (HetNets) in 5G-and-beyond systems. However, TS in HetNets is a complex problem in terms of efficiently assigning/utilizing the radio resources while satisfying the diverse QoS requirements of especially the cell-edge users due to their poor signal-to-interference-plus-noise ratio (SINR). In this respect, we propose an intelligent non-orthogonal multiple access (NOMA)-based RRM technique for a small cell base station (SBS) within a macro gNB. A Q-learning-assisted algorithm is designed to allocate the transmit power and frequency sub-bands at the O-RAN control layer such that interference from macro gNB to SBS devices is minimized while ensuring the QoS of the maximum number of devices. The numerical results show that the proposed method enhances the overall spectral efficiency of the NOMA-based TS use case without adding to the system's complexity or cost compared to traditional HetNet topologies such as co-channel deployments and dedicated channel deployments.**

***Index Terms*—Internet-of-things (IoT), HetNets, Open RAN (O-RAN), non-orthogonal multiple access (NOMA), Q-learning, spectral efficiency, traffic steering.**

I. Introduction

Satisfying varying data rate requirements of emerging Internet-of-thing (IoT)-based use cases are challenging, especially for cell-edge users, with limited radio resources [1]–[3]. However, considering the accelerated deployment pace of IoT applications in 5G-and-beyond communication systems [4], intelligent solutions are required to achieve high area spectral efficiency, massive connectivity, and high data rates [5]–[8]. Such requirements have invoked the idea of non-orthogonal multiple access (NOMA) in future wireless networks [9], allowing multiple devices to utilize the same spectrum simultaneously. However, in high connection density scenarios, the performance of weak devices located far away from the next generation-NodeB (gNB) or in a low coverage area is still unsatisfactory. As a result, various techniques, including cooperative NOMA and HetNets, are under investigation in the literature to improve the performance of cell-edge devices. In particular, when compared to non-cooperative NOMA, integrating NOMA and cooperative communications can enable higher power- and spectral-efficiency while improving fairness for IoT networks with massive connectivity.

Cooperative relaying in NOMA systems is studied extensively in the literature [10]–[13]. For example, in [10], the authors proposed cooperative NOMA in which the strong devices can serve as relays for the weak devices (i.e., having bad channel conditions), increasing their diversity gain. However, this scheme increased the computational and power requirements by significantly increasing successive interference cancellation (SIC) operations at the device. Users may experience high energy costs because of many SIC in-network with limited energy, computation, and time resources. To reduce the number of SIC operations, a space-time block code (STBC)-based cooperative mechanism is proposed in [14], [15], which improves the throughput of the weak devices with a low number of SIC but at the cost of timing offsets due to distributed nature of wireless devices.

Meanwhile, the idea of using a small-cell base station (SBS) in the same spectral band in the low coverage area is known as heterogeneous network (HetNet) architecture, in which a high power gNB is used to cover the large area. In contrast, low-power gNBs are used to improve the coverage at the cell edge, or hot-spots [16], [17]. However, radio resource management (RRM) in HetNets is a complex problem in terms of efficiently assigning/utilizing the radio resources while satisfying the diverse quality-of-service (QoS) requirements of especially the cell-edge users for their poor SINR. In this context, the open RAN (O-RAN) traffic steering (TS) use case envisions machine learning (ML)-enabled TS–xApp at RAN intelligent controller (RIC) for QoS-aware steering of the traffic by connectivity management (e.g., device-to-cell association, radio spectrum, and power allocation) [1]–[3], [18]. xApps can be deployed by multiple sources (e.g., network operator, factory management) to include new features as 5G network control applications (e.g., industrial IoT applications [19]). Yet, the literature lacks a QoS-aware radio resource-efficient design of TS–xAPP in O-RAN control for HetNets.

The properties of this cross-tier interference in NOMA-based HetNets differ from those of conventional orthogonal multiple access (OMA)-based macro-cellular network interference. Because of the limited association, macro IoT devices (MIDs) may be unable to connect to a SBS, even if it is the closest serving gNB, resulting in significant interference.

Cross-tier cellular architecture is also studied in [9], [20], [21]. The authors in [9] used the cognitive radio-based NOMA approach for cross-tier architecture. The author applied the greedy approach to maintain the QoS of the macro and small cell IoT devices (SIDs). However, this approach becomes more complex and limits the SID spectral efficiency when the number of devices increases. Similarly, classical approaches such as SIC and parallel interference cancellation (PIC) are not recommended due to their high computational requirements. An iterative joint resource allocation and device pairing in NOMA is proposed in [22], while [23] studied the Dinkelbach method to optimize the cross-tier interference in a heterogeneous NOMA network.

Most previous research has concentrated on cooperation mechanisms that enhance complexity and power consumption, especially for higher device densities. Also, the diverse data rate requirements are not considered, which ultimately does not guarantee the QoS of the devices with low data rate requirements. We design an intelligent interference control and efficient resource allocation for the traffic steering use case in O-RAN, termed as RRE-NOMA. In this scheme, the near-RT RIC handles cross-tier interference to SIDs so that the resulting interference can be efficiently mitigated, which means that the data rate requirements of the maximum number of devices are fulfilled by limiting interference to the low data rate devices. Also, the assignment of low power to the devices at overlapped spectral portion does not affect their data rate outage due to their low data rate requirements. Furthermore, MIDs with high data rate requirements are assigned to the non-overlapped spectral portion, limiting the interference from SBS. In this respect, our main contributions can be summarized as follows.

- Leveraging the O-RAN architecture and interfaces for reinforcement learning (RL)-based model training to enforce the long-term policy-based sub-band and power allocation to MIDs grouped in two sub-bands based on their data rate demands while applying NOMA in each sub-band independently. It leads to QoS-aware cross-tier and inter-devices (NOMA) interference reduction.
- Performing an extensive experimental analysis to demonstrate the efficacy of the proposed approach. The results indicate the overall network spectral and power-efficiency enhanced while achieving the QoS of the maximum number of devices.

The rest of the paper is organized as follows. In Section II, we describe the system model and explain the three HetNet deployment scenarios (i.e., co-channel, cross channel, and dedicated channel). Section III provides the proposed reinforcement learning approach, specially Q-learning, for interference management. In Section IV, we present the experimental results, and finally conclusions are given in Section V.

II. System Model

We consider a single antenna macro gNB and a single-antenna SBS, with the latter deployed at the cell edge to improve the coverage, as shown in Fig. 1. Macro gNB and SBS are assumed to be connected through a dedicated backhaul to

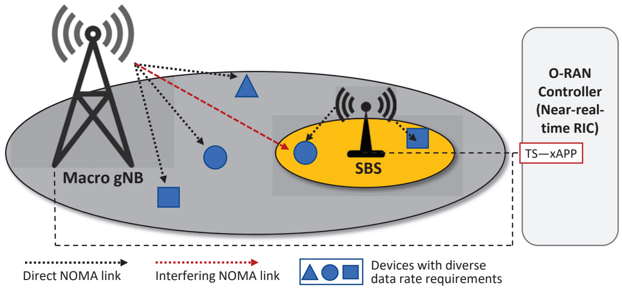

Fig. 1. A system model for NOMA-based traffic steering (TS) in O-RAN.

TS–xAPP at the O-RAN controller (i.e., near-real-time—*near-RT RIC*) for applying the intelligent resource allocation policy guidance received from orchestration and automation O-RAN layer. Meanwhile, we assume 3GPP RAN functional split 7.2 for the control unit (CU), distributed unit (DU), and radio unit (RU) [18]. Further, there are M MID devices randomly distributed in the coverage area of macro gNB and F SID devices under the coverage area of SBS. Both macro gNB and SBS employ the NOMA transmission scheme in the downlink. The MIDs and SIDs are arranged as per descending order of their channel conditions. Let $\mathbb{U} \in \{U_1, U_2, \ldots, U_m, \ldots, U_M\}$ be the set of all MIDs and $\mathbb{K} \in \{K_1, K_2, \ldots, K_f, \ldots, K_F\}$ be the set of all SIDs. Therefore, U_1 and K_1 are, respectively, the strongest macro and SIDs, while U_M and K_F are the weakest macro and SIDs. The power assigned by the macro gNB to the MID m is p_m, whereas the power assigned by the SBS to the SID f is p_f. We also consider all the channels as independent, identically distributed (i.i.d.) Rayleigh model and with block fading channel conditions. Using this model, we describe the possible HetNet deployment scenarios in the following subsections.

A. Deployment Scenarios

There are three possible deployment scenarios for spectrum allocation for the considered cross-tier O-RAN architecture.

1) Co-channel deployment: In this scenario, complete bandwidth of B Hz is concurrently assigned to both macro gNB and SBS, and consequently, underlayed SIDs face a strong interference from the overlayed macro gNB. The SINR at a SID f is given as

$$\gamma_f^{co} = \frac{|h_f|^2 p_f}{I_{macro} + \sum_{j=1}^{f-1} |h_f|^2 p_j + \sigma^2}, \tag{1}$$

where h_f is the channel between f-th SID and the SBS, and $I_{macro} = \sum_{i=1}^{M} |h_f|^2 p_i$ is the interference received from the macro gNB to the SIDs. Using (1), the capacity of SID is

$$C_f^{co} = B \log_2 \left(1 + \gamma_f^{co}\right), \tag{2}$$

Similarly, sum data rate of small cell (R_f^{co}) and macro (R_m^{co}) devices (R_m^{co}) can be, respectively, determined as

$$R_f^{co} = \sum\nolimits_{f=1}^{F} C_f^{co}, \tag{3}$$

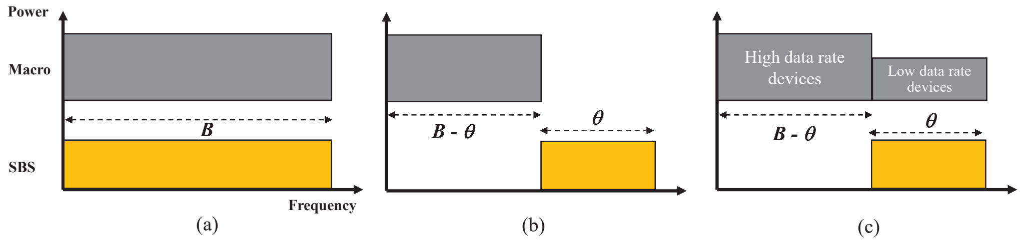

Fig. 2. Spectrum management techniques for Macro-SBS RAN architecture using NOMA: (a) co-channel deployment (b) dedicated channel deployment (c) proposed radio resource management technique.

$$R_m^{co} = \sum\nolimits_{m=1}^{M} B \log_2 \left(1 + \gamma_m^{co}\right), \tag{4}$$

2) Dedicated channel deployment: In this scenario, orthogonal bandwidth is allocated to both macro gNB and SBS, such that macro gNB is assigned $B - \theta$ and SBS is assigned θ of the bandwidth, where $\theta \leq B/2$, as shown in Fig. 2(b). In the dedicated channel deployment, SIDs do not face interference from the macro gNB, however, at the cost of less bandwidth. For the dedicated channel deployment, we have $I_{macro} = 0$, and therefore, the SINR and achievable data rate of the f-th SID is given as

$$\gamma_f^{ded} = \frac{|h_f|^2 p_f}{\sum_{j=1}^{f-1} |h_f|^2 p_j + \sigma^2}, \tag{5}$$

and

$$C_f^{ded} = \theta \log_2 \left(1 + \gamma_f^{ded}\right), \tag{6}$$

respectively.

3) Proposed spectrum management technique: In this case, the sub-band θ is assigned to the SBS, whereas the whole bandwidth B is assigned to macro gNB. The macro gNB further divides bandwidth B into two sub-bands(i.e., θ and $B - \theta$), where θ is a portion of the spectrum overlapping with SBS, and $B - \theta$ is the portion of the spectrum orthogonal to SBS. We consider that the sub-band θ contains the set of low data rate MIDs whereas the sub-band $B - \theta$ contains the set of high data rate MIDs. Now let b the generalized term for each sub-band, such that $b \in \{\theta, B - \theta\}$, then the set of MIDs in each sub-band is $\Psi_{b,l} = \{l_{b,1}, l_{b,2}, l_{b,3}, \ldots, l_{b,N}\}$, where N the total number of MIDs in sub-band b, such that $\sum_{b=1}^{2} N = M$. The devices in each sub-band follow the NOMA principle.

Our objective is to reduce the interference from macro gNB to the SIDs, which ensures the QoS demands of a large number of SIDs, as well as the MIDs, as shown in Fig. 2(c). If the MIDs in the non-overlapped spectrum is allocated with P (W), and $P - \eta$ (W) is the total power allocated to MIDs in the overlapped spectrum, where $\eta < P$, then the interference received at the SIDs from macro gNB is given as

$$I_{macro}^{prop} = \sum_{i=1}^{N} |h_{\theta,l}|^2 p_{\theta,i}, \tag{7}$$

where $h_{\theta,l}$ is the channel between macro gNB and and the device l in the sub-band θ, $h_{\theta,l}$ is the channel between macro gNB and device l, $p_{\theta,l}$ is the power allocated to the device l in the sub-band θ. Therefore, the SINR and achievable data rate of SIDs in this case is

$$\gamma_f^{prop} = \frac{|h_f|^2 p_f}{I_{macro}^{prop} + \sum_{j=1}^{f-1} |h_f|^2 p_j + \sigma^2}, \tag{8}$$

and

$$C_f^{prop} = \theta \log_2 \left(1 + \gamma_f^{prop}\right), \tag{9}$$

respectively. The sum desired data rate (Φ_{sum}) of SIDs is calculated as follows

$$\Phi_{sum} = \sum_{f=1}^{F} \Phi_f, \tag{10}$$

where Φ_f is the f-th SID's desired data rate. Similarly, the sum achievable data rate (Υ_{sum}) of SIDs is

$$\Upsilon_{sum} = \sum_{f=1}^{F} C_f^{prop}. \tag{11}$$

The macro gNB in the proposed NOMA system has to allocate sub-band and power to the devices to maximize the sum data rate of the overall heterogeneous RAN. However, a stochastic transition model for the macro gNB between states as a function of interfering power cannot be defined due to the intrinsic dynamics of a wireless channel. Therefore, we propose a radio resource-efficient (RRE) NOMA system using a reinforcement learning model technique, specifically the Q-learning method, to reduce interference. The control strategy for the interference mitigation problem described in the next section is a self-managing process that belongs to the domain of self-organizing future networks.

III. Proposed Methodology

We propose an RL technique, especially Q-learning, for interference management of the SIDs. We apply RL to this problem due to its suitability for handling such issues that involve deciding on multiple options. Macro gNB gets the information about SIDs from the backhaul that connects the macro gNB and SBS and then decides about the power assignment to the SIDs that are operated on the overlapped spectral portion of the small cell. Hence, with the help of a reinforcement learning algorithm, macro gNB can enhance the QoS of the SIDs located at the cell edge.

We apply the Q-learning algorithm due to its proven ability to perform in a highly dynamic environment. The main components of Q-learning algorithms are i) Agent with an environment, ii) state, iii) action, iv) Cost (a reward or penalty), and v) Q-table that stores the cost-values for all possible actions in any state. In the following, we provide a detailed discussion on the Q-learning model that includes the design of the state and action spaces, and cost/reward function.

- **State:** A state is the current required data rate of macro IoT device, channel conditions, and sub-band allocated to that IoT device and is given as

$$s \in \mathbb{S} \triangleq \{\mathcal{R}, CQI, b\}, \tag{12}$$

  where $\mathcal{R}$ is the device required data rate, CQI is the channel quality indicator.
- **Action:** Action is the assignment of power levels to the devices that is given by

$$a \triangleq \{a_p\}_{p \in \mathbb{A}}, \tag{13}$$

  where a is the action taken in the current state s, $p \in \mathbb{A}$ is a set of power levels with $\mathbb{A}$ as the maximum number of power levels. We use three discrete power levels as actions depending upon the data rate demands of the devices.
- **Cost or Reward:** In our optimization problem, the goal of any device is to achieve the minimum required data rate. If the devices meet the required data rate, they are rewarded in terms of better spectral efficiency; otherwise, they receive penalty. The reward function is given as

$$r(s,a) = \frac{\Phi_{sum}}{\Upsilon_{sum}}\Big(C_f^{prop} - \Phi_f\Big), \tag{14}$$

  where Φ_{sum}, Υ_{sum}, and C_f^{prop} are given in (10), (11), and (9), respectively. where Φ_f is the data rate threshold for f-th SID. Our goal is to maximize the reward so that the SID can achieve the minimum data rate requirement and small cell, on the other hand, tries to reduce interference which can ultimately enhance the spectral efficiency of the overall network.

The proposed Q-table for RRE-NOMA is updated according to the following rule

$$\begin{aligned} Q(s,a) \leftarrow\ & Q(s,a) + \alpha \cdot (r(s,a) \\ & + \vartheta \cdot \max_{a'} Q(s',a') - Q(s,a)), \end{aligned} \tag{15}$$

where $0 \leq \alpha \leq 1$ is the learning rate and $\vartheta \in [0,1]$ is the discount factor. The terms s and a refer to the current state and current action, whereas s' and a' denote the next state and the next action, respectively. The action-selection policy causes the system to shift into a new state. This policy is developed during the exploration phase and implemented during the exploitation phase. The policy is known as the ε-greedy action selection policy. It involves observing a reward for all potential actions for a given state, then updating the state-action pair with the highest reward in the Q-table until all state-action pairs have been updated.

Algorithm 1 Q-learning algorithm for O-RAN traffic steering (TS)–xApp

Input: Information about small cell and MIDs, B (Hz), θ (Hz), and P_{noma}.
Output: Total number of MIDs on the overlapped spectral portion. Power and bandwidth assignment to the SIDs on the overlapped spectral portion.
Initialization :
Organize devices into two groups based on their data rate requirements (i.e., low and high data rate devices) w.r.t a certain data rate requirement threshold.
Assign the overlapped spectral part of the bandwidth to low data rate devices and the non-overlapped spectral part to high data rate devices and organize the devices in each sub-band using the NOMA principle.
For overlapped spectral portion θ, start Q-value $Q(0,0) = 0$.
for each episode,
select state $s \in \mathbb{S}$ randomly **do**
  for for each step of episode **do**
    if (rand$(\cdot) < \varepsilon$) **then**
• Observe the current state $s \in \mathbb{S}$ in the overlapped spectral portion.
• Assign fixed power as an action a to the devices.
• Take the immediate reward as in (3).
• Update the Q-values in the Q-table as per (15).
• Move to the new state $s^{'} \in \mathbb{S}$.
• Change the action, i.e., assign more or less power to the devices based on the new state.
    end if
• Repeat the process until the algorithm determines the best power level for each device.
  end for
end for

The goal of the agent is to find the optimal action selection/transition policy $\pi^*(s)$ for each state [4], which maximizes the reward with time, which is given by

$$W^{\pi}(s) = W^{\pi^*}(s) = \max_{a'}\Big(r(s,a) + \vartheta \sum_{s' \in \mathbb{S}} P_{s,s'} W^*(s')\Big), \tag{16}$$

where $r(s,a)$ is the reward for performing action a in state s, $a' \in \mathbb{A}$ is the next feasible action that maximizes the reward, and Ps, s' is the transition probability from state $s \to s'$ as

$$P_{s,s'} = \frac{e^{Q(s,a)}}{\sum_{a \neq a'} e^{Q(s,a')}}. \tag{17}$$

We describe the strategy for the interference management for future NOMA-based heterogeneous RAN as in Algorithm 1. In the next section, we analyze the outage probability of SID and compare its performance for different deployment schemes.

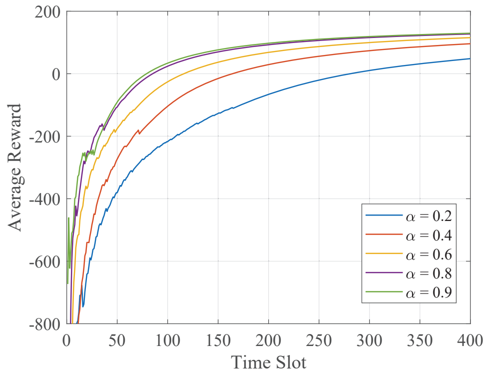

Fig. 3. Impact of learning rate on the convergence of the algorithm, with $B = 5$ MHz, $\theta = 2$ MHz, and $F = 4$.

IV. Performance Evaluation

This section provides and discusses the results for our proposed system design and compares these with the baseline schemes. We consider a macrocell with a radius of 1 km, and M devices are uniformly distributed in its area. We further assume that these devices' data rates are uniformly distributed in the range [5, 8000] kbps unless stated otherwise. There is one SBS located at the edge of the macrocell, with a radius of 200 m. The transmit power of the macro gNB is 45 dBm (unless it is decided to change to 35 dBm or 25 dBm using Q-learning), whereas the transmit power of each SBS is 20 dBm. Thermal noise density is considered as -173 dBm/Hz, and path loss exponent as 4.

Fig. 3 shows the learning rate's influence on the algorithm's convergence in the proposed scheme. It can be observed that the algorithm converges quickly for higher learning rate values (i.e., as $\alpha \to 1$) and vice versa. Also, the convergence time is almost the same for the learning rate values of ≥ 0.4. Therefore, we choose the learning rate of 0.4 in the following simulation results.

Fig 4 shows the impact on the spectral efficiency of the weakest SID by using different power levels (i.e., 45 dBm, 35 dBm, and 25 dBm) from the macro gNB in the sub-band θ. Higher transmit power at the macro gNB will generate higher interference at the SIDs and vice versa. Spectral efficiency (SE) of the weakest SID is simulated, corresponding to the signal-to-noise ratio (SNR) on the dB scale and interference-to-signal ratio (β) on the linear scale. We can observe that for the values of $\beta = 1$, and SNR = 15 dB, it can achieve a SE of 5 bps/Hz with the macro gNB transmit power of 25 dBm, whereas its SE is degraded to 3 bps/Hz, and 1 bps/Hz for the transmit power of 35 dBm, and 45 dBm, respectively.

Fig. 5 shows the average capacity of the weakest SID for an increasing number of SIDs in different NOMA deployment topologies, such as co-channel, dedicated channel, and the proposed RRE-NOMA. It is observed that the average capacity of the SID decreases with the increase in the total number of SIDs. Since the inter-devices, NOMA interference is increased with the increase in SIDs. For $F = 2$, the weakest SID achieves 7.9 bps/Hz, 5.5 bps/Hz, and 3.3 bps/Hz for proposed

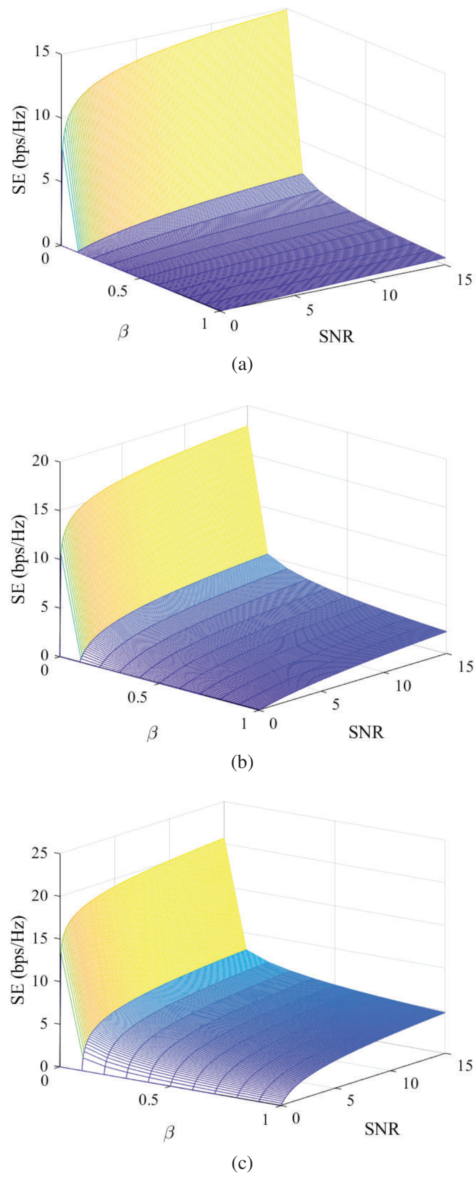

Fig. 4. Link-level simulation for the proposed RAN architecture. β is the interference-to-signal ratio (ISR) on linear scale, and SINR is optimized using Q-learning to achieve desired QoS requirement of the SID when F = 2, and the macro gNB transmit power is: (a) 45 dBm, (b) 35 dBm, and (c) 25 dBm.

deployment, dedicated channel deployment, and co-channel deployment, respectively. However, the values of average capacity are decreased to 5.8 bps/Hz, 3 bps/Hz, and 0.6 bps/Hz with $F = 12$. Also, these results show that the weakest SID in the proposed RRE-NOMA achieves significantly higher capacity for any number of SIDs.

Fig. 6 depicts the average number of SIDs in an outage for an increase in the total number of SIDs with θ = 6 MHz, and the target data rate requirements range [5, 800] kbps. Our proposed RRE-NOMA outperforms the conventional NOMA and orthogonal multiple access (OMA) schemes for all cases of the total number of SIDs. However, we can also observe

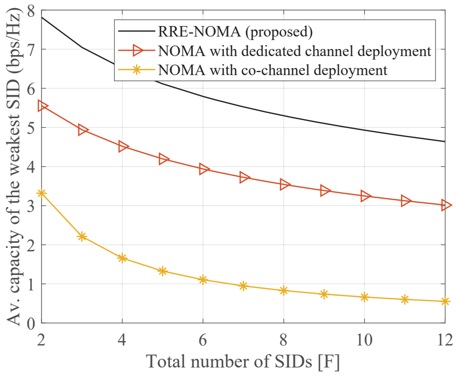

Fig. 5. Average capacity of the weakest SID with respect to the total number of SIDs served.

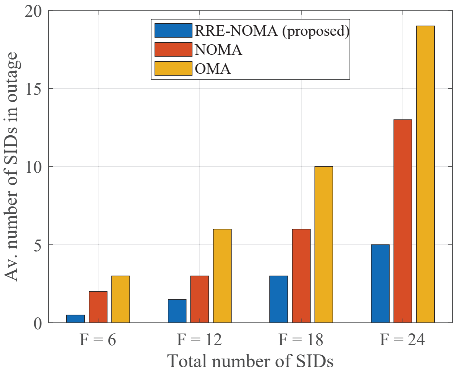

Fig. 6. Average number of SIDs in outage with respect to increase in the total number of SIDs with data rate range [5, 800] kbps for OMA, NOMA, and the proposed Q-learning-based NOMA with θ = 6 MHz, and 20 dBm of transmit power at SBS in the proposed deployment.

that the performance gap between our proposed scheme and that of NOMA and OMA also increases with the increase in the total number of SIDs. This is because, with the increase in F, OMA and NOMA schemes can not manage the SIDs' resources (i.e., power and bandwidth). On the other hand, the proposed RRE-NOMA scheme manages the resources intelligently by assigning the power and sub-bands such that interference is minimized at the SIDs, ultimately maintaining the QoS requirements of maximum SIDs and enhancing the overall SE of the network.

V. Conclusions

This paper proposed a novel Q-learning-based interference control and radio resource management technique for NOMA-enabled macro-SBS HetNet architecture. Instead of using the conventional interference cancellation techniques, we intelligently minimize the interference at the SIDs without adding extra cost or complexity to the network. The overlapped spectral portion is allocated to the MIDs with low data rate requirements, whereas the MIDs with high required data rates are allocated to the non-overlapped spectral portion. Macro gNB assigns the power, as per the NOMA principle, to the SIDs located at the overlapped portion so that their QoS is not affected. As a result, the interference received at the SIDs from macro gNB is significantly reduced, which ultimately enhances the spectral efficiency of the weakest SID, and the overall QoS and sum data rate of the network is improved. In the future, we will study the user association, NOMA user grouping, and resource efficiency as a joint problem in the O-RAN traffic steering use case.

References

[1] L. Bonati, S. D'Oro, *et al.*, "Intelligence and learning in O-RAN for data-driven NextG cellular networks," *IEEE Communications Magazine*, vol. 59, no. 10, pp. 21–27, 2021.

[2] I. Chih-Lin, S. Kuklinskí, and T. Chen, "A perspective of O-RAN integration with MEC, SON, and network slicing in the 5G era," *IEEE Network*, vol. 34, no. 6, pp. 3–4, 2020.

[3] S. K. Singh *et al.*, "The evolution of radio access network towards open-RAN: Challenges and opportunities," in *IEEE WCNC Workshops*, 2020, pp. 1–6.

[4] M. W. Akhtar *et al.*, "Q2A-NOMA: A Q-learning-based QoS-aware NOMA system design for diverse data rate requirements," *IEEE Trans. Ind. Informat.*, 2022.

[5] H. Munir *et al.*, "Resource optimization in multi-tier HetNets exploiting multi-slope path loss model," *IEEE Access*, vol. 5, pp. 8714–8726, 2017.

[6] A. Mahmood *et al.*, "Industrial IoT in 5G-and-beyond networks: Vision, architecture, and design trends," *IEEE Trans. Ind. Informat.*, vol. 18, no. 6, pp. 4122–4137, 2022.

[7] J. Kim *et al.*, "A novel resource allocation scheme for NOMA-V2X-Femtocell with channel aggregation," in *IEEE GLOBECOM*, 2020, pp. 1–6.

[8] Z. Liu *et al.*, "Robust resource allocation in two-tier NOMA heterogeneous networks toward 5G," *Computer Networks*, vol. 176, p. 107299, 2020.

[9] I. Budhiraja *et al.*, "Cross layer NOMA interference mitigation for femtocell users in 5G environment," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4721–4733, 2019.

[10] Z. Ding *et al.*, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, 2015.

[11] Z. Yang *et al.*, "Novel relay selection strategies for cooperative noma," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10 114–10 123, 2017.

[12] W. Mei and R. Zhang, "Uplink cooperative noma for cellular-connected uav," *IEEE J. Sel. Topics Signal Process*, vol. 13, no. 3, pp. 644–656, 2019.

[13] P. Xu *et al.*, "Optimal relay selection schemes for cooperative NOMA," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7851–7855, 2018.

[14] M. W. Akhtar *et al.*, "STBC-aided cooperative NOMA with timing offsets, imperfect successive interference cancellation, and imperfect channel state information," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11 712–11 727, 2020.

[15] ——, "On the symbol error probability of STBC-NOMA with timing offsets and imperfect successive interference cancellation," *Electronics*, vol. 10, no. 12, p. 1386, 2021.

[16] A. J. Muhammed *et al.*, "Energy-efficient resource allocation for NOMA based small cell networks with wireless backhauls," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3766–3781, 2020.

[17] Q.-T. Trinh *et al.*, "Performance analysis of NOMA for wireless downlink in multi-tiers heterogeneous network," in *International Conference on System Science and Engineering (ICSSE)*, 2019, pp. 329–334.

[18] O-RAN Alliance, "O-RAN use cases and deployment scenarios," White Paper, Feb. 2020.

[19] S. F. Abedin *et al.*, "Elastic O-RAN slicing for industrial monitoring and control: A distributed matching game and deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, pp. 1–15, 2022.

[20] I. Budhiraja *et al.*, "An energy-efficient resource allocation scheme for SWIPT-NOMA based femtocells users with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 69, no. 7, pp. 7790–7805, 2020.

[21] F. Zhou *et al.*, "Energy-efficient NOMA enabled heterogeneous cloud radio access networks," *IEEE Network*, vol. 32, no. 2, pp. 152–160, 2018.

[22] I. Budhiraja *et al.*, "An energy-efficient resource allocation scheme for swipt-noma based femtocells users with imperfect csi," *IEEE Trans. Veh. Technol.*, vol. 69, no. 7, pp. 7790–7805, 2020.

[23] H. F. Saeed *et al.*, "Joint pairing and resource allocation for backhaul of small cells using NOMA," *Journal of Computational Science*, vol. 45, p. 101197, 2020.