Resource Allocation in Full-Duplex Uncoordinated Communication Systems with NOMA

Joseph Doumit, Marie-Josépha Youssef, Charbel Abdel Nour, Joumana Farah, Catherine Douillard

HAL Id: hal-02886510

https://hal-imt-atlantique.archives-ouvertes.fr/hal-02886510v3

Submitted on 6 Oct 2021

Resource Allocation in Full-Duplex Uncoordinated Communication Systems with NOMA

Joseph Doumit[(1)], Marie-Josépha Youssef[(1)], Charbel Abdel Nour[(1)], Joumana Farah[(2)], Catherine Douillard[(1)]

[(1)]IMT Atlantique, Lab-STICC, UMR CNRS 6285, F-29238 Brest, France
[(2)]Department of Electricity and Electronics, Faculty of Engineering,
Lebanese University, Roumieh, Lebanon

Abstract—The Internet of Things (IoT) is a major driving use case of future communication systems. 5G and beyond networks should be properly designed to meet the quality of service (QoS) requirements of IoT devices in terms of massive access, data rates, latency and reliability. In this paper, we study the uncoordinated spectrum and power allocation problems in an uplink IoT network. In the considered setting, IoT devices aim at organizing their transmissions, without any coordination, on the available frequency channels and power levels. To enhance performance, we propose the deployment of full-duplex relays as well as the use of non-orthogonal multiple access and queuing at the level of the relays. A new algorithm, based on the multi-player multi-armed bandits framework, is proposed with the aim of reducing the transmit power of IoT devices. Simulation results validate the convergence of the proposed algorithm. Results also confirm the superior performance of the proposed technique in terms of satisfying QoS requirements and reducing the needed user transmit power when compared to two baseline schemes.

Index Terms—Non-orthogonal multiple access, uncoordinated system, relays, multi-armed bandits, full duplex, power allocation.

I. INTRODUCTION

IN recent years, Internet of Things (IoT) devices have become a main component in our daily lives. Because of their dense deployment, machine-type devices (MTDs) or IoT devices produce a large amount of data to be transmitted. Moreover, the mobile traffic of IoT devices mostly consists of short packets concentrated in the uplink. Hence, to ensure successful transmissions and due to the heterogeneous nature of the underlying applications, a communication network with low-latency and low complexity features [1] is required.

Non-orthogonal multiple access (NOMA) has recently emerged as a promising multiple access technology. Power-domain NOMA consists of multiplexing the signals of multiple users on the same resource block by exploiting the power domain [2], [3], [4], [5], [6]. By doing so, NOMA enhances system spectral efficiency and increases the achieved data rates [7] with respect to orthogonal multiple access (OMA). At the receiver side, successive interference cancellation (SIC) is performed to retrieve the superimposed signals. Although most studies related to power-domain NOMA focus on the downlink setting, some work recently considered the application of NOMA in uplink scenarios. In [2] and [3], power allocation schemes are proposed. A backoff step between power levels is introduced in [2], then system outage probability and achievable sum rate are shown to outperform the OMA model. In [3], a

new power allocation scheme that aims at ensuring the user rate requirements is proposed. Other studies have also shown that NOMA helps achieve massive connectivity and low latency communications [8], making it a promising candidate to support IoT communications [9].

To reduce the signaling overhead in IoT networks, uncoordinated spectrum access has received significant interest in recent literature [10], [11]. That said, few works have considered the uncoordinated spectrum access in an uplink NOMA system. In fact, most of the available literature body on NOMA considers fully coordinated systems. Only a few recent works considered semi-grant free communication systems [12] and fully uncoordinated ones [13].

In addition to NOMA, the deployment of fixed or mobile relays has gained attention. In massive communication networks, adding relays could be a solution to alleviate network congestion. Relays can also increase the coverage area and the achieved throughput of wireless networks. In addition to that, they can restore communication links in damaged networks [14], [15]. In order to ensure low-latency and high data rates in massive networks, full-duplex capabilities at the relay level can be leveraged. By allowing the simultaneous transmission and reception of information on the same frequency resource, full-duplex links [15] reduce both the experienced latency and the need to store information at the relay.

The use of reinforcement learning and game theory for uncoordinated resource allocation has also recently garnered an increasing attention. Several studies [16], [17], [18] have considered reinforcement learning, and especially the multi-armed bandits (MAB) framework, to enable users to organize their transmissions on the available channels [16]. In the MAB framework, a set of players compete to find the actions, also called arms, that maximize their expected gain or welfare [19]. Since the arm rewards are previously unknown to the players, there is a need to achieve a trade-off between exploring the arms to learn their associated rewards, and exploiting the optimal arms to maximize the achieved gain. Multiple algorithms have been proposed to solve problems modeled using the MAB framework, such as the ϵ-greedy method, the ϵ-decreasing technique and the Upper Confidence Bound (UCB) algorithm [17]. These algorithms have been widely used to study, among others, the opportunistic spectrum access problem in cognitive radio systems [20], and the uncoordinated spectrum access

problem [16].

To prevent data loss in case of system congestion, relays can be equipped with queuing capabilities. These capabilities were mostly considered in coordinated systems where users have information about the buffer status [21]. However, no previous work has considered the use of full-duplex relays in an uncoordinated uplink NOMA system with transmit power limits and queuing capabilities at the relays. In this paper, we consider the uplink of an uncoordinated spectrum access system where IoT devices aim at organizing their transmissions with the deployed relays using MAB frameworks without communicating with each other. Each IoT device has a power budget and is allowed to connect to one relay only. In addition to that, the devices adopt continuous transmissions. Leveraging full-duplex communications, the deployed relays forward the received messages to a base station (BS). Also, queuing is applied, when needed, at the relay level.

The rest of this paper is organized as follows. In section II, the system model is described and the problem is modeled using the MAB framework. Section III presents the proposed solution for the joint distributed spectrum and power allocation, while section IV considers queuing at the level of the relays. Simulation results are presented in section V and conclusions in section VI.

II. SYSTEM MODEL

Consider a hexagonal cell with one BS located at the cell center, R relays and N IoT devices. Let $\mathcal{D} = \{D_1, D_2, \ldots, D_N\}$ and $\mathcal{R} = \{r_1, r_2, \ldots, r_R\}$ be the sets of devices and relays respectively. The IoT devices are randomly distributed in the cell. Moreover, the relays are distributed uniformly at a mid-distance between the BS and the cell edges.

We assume an uncoordinated or grant-free uplink communication system where no communication occurs between the MTDs. A two-hop communication network, where relays forward the received signals from the IoT devices to the BS, is considered. We assume that IoT devices cannot communicate directly with the BS due to large distances between them and the BS. As shown in Figure 1, each IoT device is allowed to communicate with one relay at a time using one channel. The relays operate in full-duplex, i.e., they receive signals from the MTDs and transmit them to the BS by simultaneously using the same resources. Thereby, relays may suffer from self-interference, possibly impeding the successful decoding of the received signals. Each relay has access to a separate set of subbands which will be used simultaneously in the access and the backhaul links. Hence, relays do not suffer from inter-relay interference. Also, clustering the MTDs over the channels using NOMA allows their clustering between the relays. For multiple access, NOMA is considered between the devices and the relays, while OMA is applied between the relays and the BS.

A. Power Allocation Scheme

To enable SIC decoding at the level of the relays, devices transmitting on the same channel should have their signals
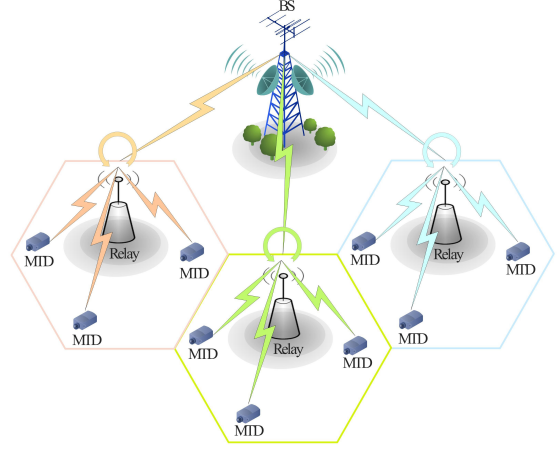

Figure 1: System model with full-duplex relaying.

received with different power levels. To this aim, we generalize the uplink NOMA power allocation scheme proposed in [3] to the case of full-duplex transmissions.

We consider L available received power levels per channel at the relay side and let Γ_{req} be the signal-to-interference-plus-noise ratio (SINR) requirement for each MTD. Inspired by [3], it can be shown that the lth received power level $v_l, l = 1, ..., L$ ($v_1 > v_2 > ... > v_L$), at the relay side, to guarantee the target SINR, is given by:

$$v_l = \left(\sigma^2 + C_{SI} \times P_r^{max}\right) \times \Gamma_{req}(\Gamma_{req} + 1)^{L-l}. \tag{1}$$

σ^2 is the additive Gaussian noise power, C_{SI} the residual self-interference factor, and P_r^{max} the maximum transmit power of relay r on each of its channels. Hence, $C_{SI} \times P_r^{max}$ is the maximum self-interference power experienced at the relay.

In this work, MTDs are assumed to perform a perfect channel estimation so as to estimate the channel gains of their links with the relays. Moreover, these channel gains are supposed to depend only on the experienced path loss and shadowing, i.e., only long-term variations of the channel gains are considered. To reach a received power level v_l at the rth relay level, each device D_i calculates its needed transmit power according to:

$$P_{D_i,l}^t = \frac{v_l}{h_{D_i,r}^2}, \tag{2}$$

where $h_{D_i,r}$ is the channel gain between user D_i and relay r.

As previously said, to ensure successful SIC, the signals of devices transmitting on the same channel must be received with different power levels at the relay side. Therefore, if two or more devices, associated with the same relay and transmitting on the same channel, choose the same power level v_l, a collision occurs, leading to a failed signal decoding. In uplink NOMA, SIC is performed in decreasing order of channel gains. Therefore, signals of devices having chosen weaker received power levels than the collided signals are also undecodable. The collision can even cause upper-level signals to be undecodable if their target SINR is not reached.

B. Uncoordinated Joint Channel and Power Allocation

To allow devices to organize their transmissions in an uncoordinated manner, the problem of joint channel and power allocation is modeled using the multi-player MAB framework with zero-reward on collision [19]. The set of players is the set of IoT devices $\mathcal{D}$ and the set of arms is the set of available channel-power level pairs. Having a transmit power budget $P_{D_i}^{max}$, each device D_i builds its action profile $\mathcal{A}_{D_i}$. The latter consists of channel and power level pairs, or arms, on which device D_i can transmit without violating its power budget. Let $\mathcal{A}_{D_i} = \{a_1, a_2, \dots, a_{K_{D_i}}\}$, with K_{D_i} representing the number of available arms for user D_i.

Upon choosing an action, each device D_i receives a reward or a utility from the relay. In this work, the reward is taken as the device achieved rate. It is given by:

$$Rate_{i,a_k} = B_c \log_2(1 + \Gamma_{i,a_k}), \tag{3}$$

where B_c is the channel bandwidth and Γ_{i,a_k} is the achieved SINR of device D_i, given by:

$$\Gamma_{i,a_k} = \frac{v_l}{\sum_{j=l+1}^{L} v_j + C_{SI} P_r^t + \sigma^2}, \tag{4}$$

with $\sum_{j=l+1}^{L} v_j$ being the interference of the NOMA signals not canceled with SIC, and P_r^t the transmit power of relay r per channel. In (4), v_l is the power level chosen in action a_k.

If two or more devices choose the same action, a collision occurs, resulting in a null utility for some users. In this case, the relay returns a zero reward to the concerned devices. Let η_i represent the collision indicator for device D_i on arm a_k:

$$\eta_{i,a_k} = \begin{cases} 0, & \text{in the case of a collision,} \\ 1, & \text{otherwise.} \end{cases} \tag{5}$$

Therefore, the utility U_{i,a_k} received by device D_i when picking arm a_k can be expressed as:

$$U_{i,a_k} = \eta_{i,a_k} \times Rate_{i,a_k}. \tag{6}$$

In the considered MAB framework, no communication is allowed between players. Hence, each player can only observe its own chosen actions and achieved rewards. The objective of the following study is to maximize the welfare of all devices, which is expressed as:

$$w = \sum_{i=1}^{n} U_{i,a_k}. \tag{7}$$

C. Relay-BS Communication

In full-duplex communications, the relay receives signals from the IoT devices and transmits them to the BS simultaneously. That said, at timeslot t, the relay transmits the signals received at timeslot $t - 1$ while receiving new signals from the devices. At timeslot $t = 1$, the relay has not received any messages beforehand, hence does not communicate with the BS. Starting at $t = 2$, the relay transmits the signals received at $t = 1$ to the BS, while the devices transmit new messages to the relay, and so on.

To guarantee the users QoS requirements when forwarding the device messages to the BS, the transmit power of the relay should be optimized. Let X_r^t denote the number of correctly decoded signals at the level of relay r at timeslot t. The minimum rate required for the relay to be able to transmit all successfully decoded signals is equal to:

$$R_r^{req} = X_r^t \times Rate^{req}, \tag{8}$$

where $Rate^{req}$ stands for the required rate per device and is related to Γ_{req} through (3).

Each relay is assumed to transmit its data over C orthogonal channels in an OMA manner. Hence, the rate achieved by relay r on channel c is given by:

$$R_{r,c} = B_c \times \log_2\left(1 + \frac{h_{r,BS}^2 P_r^t}{\sigma^2}\right). \tag{9}$$

$h_{r,BS}$ is the channel gain between relay r and the BS. It is common to all C channels since only large-scale channel variations are considered. Assuming an equal rate repartition among the C channels, to reach its rate requirement expressed in (8), the transmit power of relay r over a channel c should satisfy:

$$P_r^t = min\left\{\frac{\sigma^2}{h_{r,BS}^2}\left(2^{\frac{X_r^t \times Rate^{req}}{C \times B_c}} - 1\right), P_r^{max}\right\}. \tag{10}$$

III. PROPOSED ALLOCATION TECHNIQUE

Classical MAB algorithms for uncoordinated resource allocation such as the ϵ-greedy, ϵ-decreasing, and UCB techniques [17] generally lead to system convergence. However, the solutions proposed in the literature do not take into consideration the transmit power limit of devices. In MTD applications, there is a need to save power in order to maximize the battery life of devices. To fulfill this objective, we propose to modify the utility function of the UCB algorithm to account for the devices power budget.

Each IoT device aims at achieving a rate larger than or equal to its requested rate, while consuming the least amount of transmit power. Thus, two parameters should be taken into consideration in the design: the transmit power of device D_i, $P_{D_i,k}^t$, and its achieved rate, $Rate_{i,a_k}$. As shown in (2), $P_{D_i,k}^t$ depends on both the received power level chosen by D_i and the channel gain between the relay and D_i. An increase in the channel gain reduces the amount of required transmit power. Hence, the minimum transmit power value is reached when the device chooses its best radio channel, i.e., the one with the largest gain. However, having each device choose its best arm in terms of channel gain might not always lead to system convergence. Constrained by its power budget, user D_i might not be able to transmit on all arms associated with channel c. In such a case, the choice of all users must be adapted in a way that system welfare is maximized.

IoT devices should choose their channels and corresponding received power levels in a way to strike a trade-off between the achieved rate and the consumed transmit power. Therefore,

instead of choosing the arm that maximizes the achieved rate, each device D_i chooses arm $a_i^* \in \mathcal{A}_{D_i}$ that satisfies:

$$a_i^* = \underset{a_k \in \mathcal{A}_{D_i}}{\operatorname{argmax}} \left(\alpha \frac{Rate_{i,a_k}}{Rate_{max}} - \beta \frac{P_{D_i,k}^t}{P_{D_i}^{max}} \right). \qquad (11)$$

In (11), α and β are positive weight parameters relative to the rate and power respectively, with $\alpha + \beta = 1$. $Rate_{max} = B_c \log_2(1 + \frac{v_1}{\sigma^2})$ is the maximum achievable rate by all users. The proposed technique to solve the uncoordinated joint channel and power allocation problem is summarized in Algorithm 1, where T is the time horizon and $n_i(a_k,t)$ the number of times arm a_k is played by device i until timeslot t.

Algorithm 1: UCB Power Minimization

Initialization: Each device i generates a random permutation of $\mathcal{A}_{D_i}$. During the first $|\mathcal{A}_{D_i}|$ iterations, i plays the arms iteratively and updates the utilities accordingly.

for $t=1:T$ **do**
  for $i=1:N$ **do**
    // Best arm identification:
    $a_i^* = \underset{a_k \in \mathcal{A}_{D_i}}{\operatorname{argmax}} \left(\alpha \frac{Q(i,a_k)+\sqrt{\frac{2\log(t)}{n_i(a_k,t)}}}{Rate_{max}} - \beta \frac{P_{D_i,k}^t}{P_{D_i}^{max}} \right)$
    // Parameters update:
    $s_i(a_i^*, t+1) = s_i(a_i^*, t) + U_{i,a_i^*}$
    $n_i(a_i^*, t+1) = n_i(a_i^*, t) + 1$
    $Q(i, a_i^*) = \frac{s_i(a_i^*,t+1)}{n_i(a_i^*,t+1)}$
  end
end

IV. QUEUING AT THE LEVEL OF THE RELAYS

When the relay successfully decodes the signals received from the MTDs, it forwards them to the BS. Occasionally, in the case of congestion or bad channel states between the relay and the BS, the relay may not be able to transmit all the received data to the BS within a unique timeslot. To avoid data loss in such a case, buffering the unsent signals at the relay can be used. In fact, this is necessary when the required power to transmit the received signals from the devices, P_r^t, exceeds the relay power budget per channel, i.e., when:

$$P_r^t \geq P_r^{max}. \qquad (12)$$

This can be translated as the number of correctly decoded signals at the relay level, X_r^t, being greater than some threshold value $X_{0,r}$. The latter is deduced from (10) after setting the transmit power value equal to the relay power budget. $X_{0,r}$ relates to the maximum number of received IoT signals that a relay can forward to the BS per timeslot without resorting to buffering. It is expressed as:

$$X_{0,r} = \frac{B_c \times C}{Rate^{req}} \log_2 \left(1 + \frac{P_r^{max} h_{r,BS}^2}{\sigma^2} \right). \qquad (13)$$

When buffering occurs, a certain number of MTDs should be incited to connect to another, less congested, relay. This is necessary to enable each relay to forward all received messages to the BS during one timeslot, thereby satisfying the QoS requirements of connected devices. The system being uncoordinated, very little information is shared between the relay and the connected devices. Thus, MTDs must learn on their own the need to connect to another relay. To do so, the relay having reached its overflow threshold, i.e., a full buffer possibly entailing information loss, returns a zero reward to all its connected devices for a certain period. This period is referred to as *Timeout* period which is determined in practice depending on the system parameters.

During the *Timeout* period of relay r, devices associated with r are encouraged to perform a new exploration phase. In other words, those devices may try to access arms associated with another relay, if their power budgets allow. By exploring new arms associated with less congested relays, IoT devices aim at reaching a new stable state leading to system convergence. Note that this solution could lead to some devices remaining silent, i.e., refraining from transmitting, for multiple timeslots if they are not able to connect to another relay.

To represent the zero utility in the case of a buffering overflow, the collision indicator is modified as follows: $\eta_{i,a_k} = 0$ in case of collision or buffer overflow, and 1 otherwise.

To avoid returning a zero reward for a long period by the congested relay, i.e., to help devices react faster in the *Timeout* period, they may take into consideration the total transmit power that has been spent on each specific arm a_k, reflected by $P_{D_i,k}^t \times n_i(a_k,t)$. Since the congested relay is replying with zero rewards, the decision function on this relay depends mostly on the consumed power, leading to free some arms and to try new ones. Therefore, to choose their actions, instead of using (11), IoT devices will adopt the following metric:

$$a_i^* = \underset{a_k \in \mathcal{A}_{D_i}}{\operatorname{argmax}} \left(\alpha \frac{Rate_{i,a_k}}{Rate_{max}} - \beta \frac{P_{D_i,k}^t}{P_{D_i}^{max}} \times \frac{n_i(a_k,t)}{t} \right). \qquad (14)$$

To summarize, Algorithm 2 is proposed to account for buffering needs at the relays.

Algorithm 2: UCB Power Minimization and Queuing in an Uncoordinated System

while $X_r^t \leq X_{0,r}$ **do**
  Run Algorithm 1
end
while $X_r^t \geq X_{0,r}$ *and* $t \leq$ *Timeout* **do**
  for *devices connecting to r* **do**
    $U_{i,a_k} = 0 \forall i$
    Rerun Algorithm 1 with the decision function in
    (14) to change arms if possible
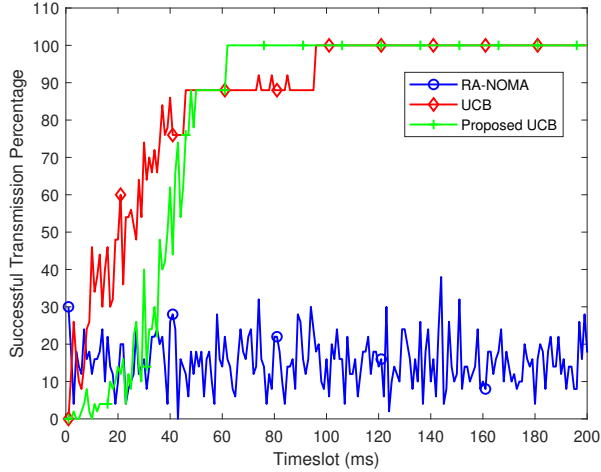 **end**
  $t = t + 1$
end

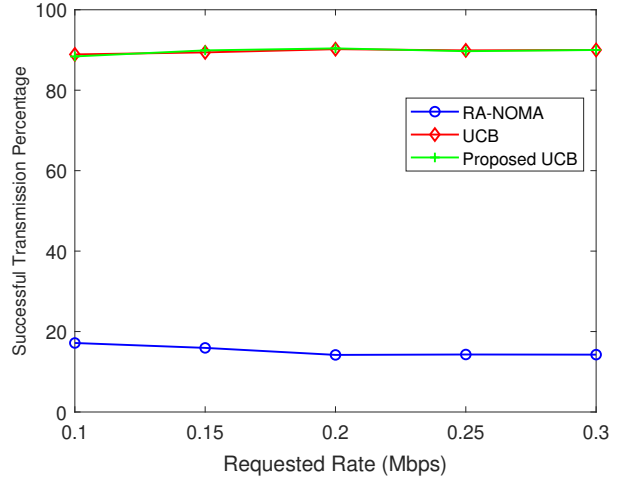Figure 2: Instantaneous success probability.

Figure 3: Average success percentage for a varying requested rate.

V. NUMERICAL RESULTS

Extensive simulations of the proposed allocation technique were conducted to validate its performance. First, we considered an uplink communication system with one BS at the cell center, $N = 50$ transmitting IoT devices, and one relay. The cell radius is 250 m. A total of $C = 10$ channels and $L = 5$ power levels is considered, leading to a maximum of $C \times L = 50$ arms. The MTDs are scattered around the relay according to a Poisson distribution [22]. The channel propagation model corresponds to a path loss exponent of 3.76, and an 8 dB zero mean lognormal shadowing. The noise power spectral density is $N_0 = 4 \times 10^{-21}$ W/Hz. Other design parameters are presented in Table I.

Table I: Design parameters values

Requested rate	0.3 Mbps	$P_{D_i}^{max}$	23 dBm
$C \times P_r^{max}$	5W	C_{SI}	-120 dB
B_c	156 kHz	Timeout period	50 ms
α	0.5	β	0.5

The performance of the proposed method is analyzed over $T = 200$ timeslots and compared to that of two baseline schemes. The first is the UCB method in [17] that does not take into consideration the power budget. The second baseline method is referred to as Random Access NOMA (RA-NOMA), where devices choose their channels and power levels uniformly at random.

A. Convergence and Performance Assessment of the Proposed Solution

In Figure 2, the instantaneous successful transmission probability is shown. While both UCB-based algorithms converge to a success probability of 1, our proposed method reaches convergence slightly faster than the baseline UCB technique. Both methods greatly outperform RA-NOMA.
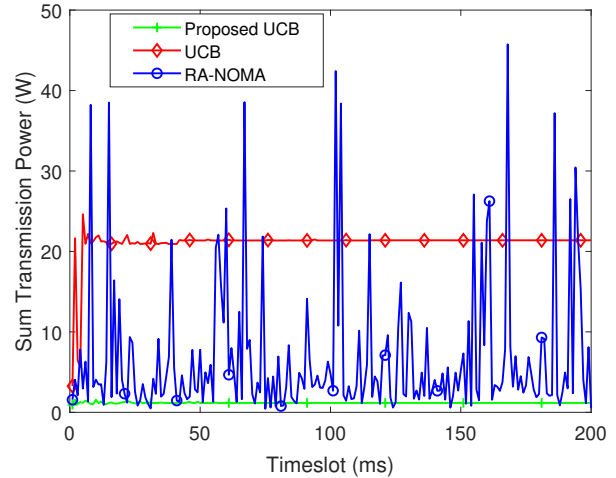

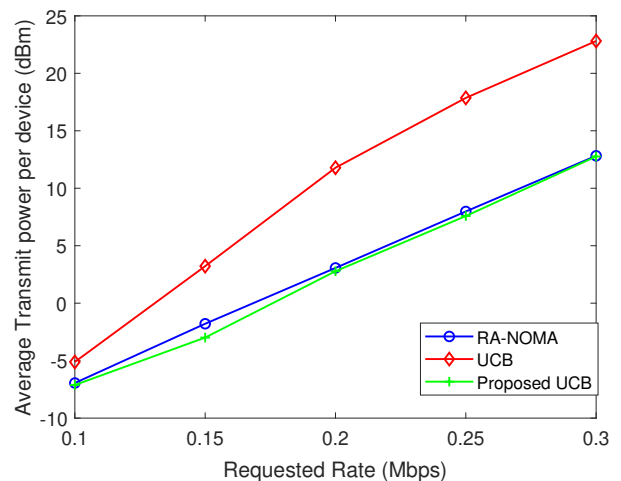Figure 4: Instantaneous transmit power.

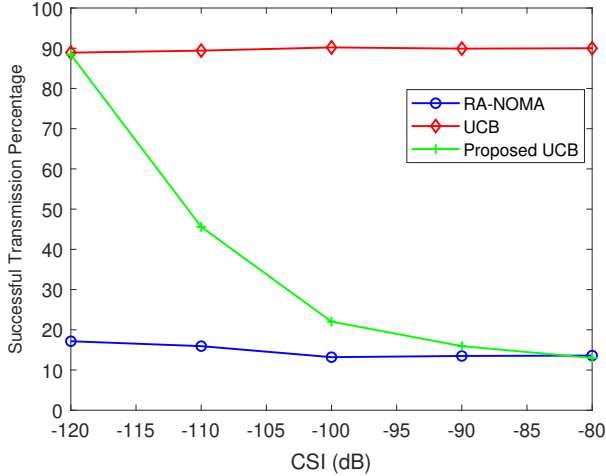Figure 5: Average transmit power for a varying requested rate.

Figure 6: Success rate for a varying C_{SI}.

Figure 3 shows the successful transmission percentage for a varying IoT requested data rate, averaged over 2000 simulations. For the UCB-based methods, the successful transmission percentage is evaluated during both the exploration and the exploitation phases. Figure 3 shows that both UCB-based algorithms converge to an almost 90% successful transmission probability for all considered values of the IoT requested rate. Moreover, the UCB-based methods greatly outperform RA-NOMA.

In Figure 4, the instantaneous sum transmission power of all IoT devices is shown. The proposed method greatly decreases the needed transmit power when compared to the baseline UCB technique that does not account for power minimization and outperforms RA-NOMA regarding Peak to Average Power Ratio (PAPR).

Figure 5 shows the average transmit power per device for different values of the required rate. Here also, the proposed method outperforms the baseline UCB technique by reducing the average needed transmit power, with a performance gap that increases with the required rate. To conclude, using the proposed algorithm, the devices can limit their transmit power while reaching their rate requirements.

B. Performance Assessment for a Varying Residual Self-Interference Factor

In the following, the effect of self-interference on system performance is evaluated by varying the value of the residual self-interference factor C_{SI}. A system with $R = 4$ full-duplex relays and a total of $N = 40$ transmitting devices is now considered. $C = 2$ channels and $L = 5$ power levels are set for each relay, leading to a maximum of $C \times L = 10$ arms per relay. Having four relays, the MTDs are now uniformly distributed over the cell.

Figure 6 shows the successful transmission percentage with a varying C_{SI} factor. By keeping constant the power budget and increasing C_{SI}, the performance of the proposed method with

full-duplex relays decreases when compared to the baseline UCB technique. In fact, when C_{SI} increases, the effect of the self-interference becomes dominant, thereby lowering system performance. It should be noted that the performance of the baseline UCB technique is not affected by an increasing C_{SI} value since it does not account for the devices power budgets. That said, the baseline UCB technique requires a much larger amount of power when C_{SI} increases (around a 450% increase in the power consumption of the baseline UCB compared to the proposed UCB, for $C_{SI} = -80$ dB). This may lead to unacceptable power levels for such type of devices.

A suitable alternative solution for high C_{SI} values could be reached by introducing additional channels to the system.

Table II: Success percentage of the proposed technique for a varying number of channels per relay, with $L = 5$ power levels.

Number of channels	2	3	4	5	6	7
Success Percentage	15.98	17.54	25.05	69.77	79.91	89.54

Table II shows the effect of an increasing number of channels per relay on the performance of our proposed method for $C_{SI} = -80$ dB. The results show that, to reach a success percentage of 90%, the number of channels should be increased to 7 per relay, resulting in a total of $35 \times 4 = 140$ available arms per device.

C. Performance Assessment with Queuing Needs at the Relays

In the following, we consider a system with $R = 4$ relays and $N = 160$ uniformly distributed MTDs. $C = 10$ channels and $L = 5$ power levels are considered per relay, leading to a maximum of $R \times C \times L = 200$ arms per IoT device. The purpose of this setting is to validate the role of queuing at the relays. As explained in Section IV, to avoid the need for queuing, relay r can support a maximum of $X_{0,r}$ connected devices. The value of $X_{0,r}$, estimated using (13), for one simulation of the considered setting is shown in Table III.

Table III: Maximum number of supported IoT users per relay, $X_{0,r}$.

Relay	1	2	3	4
Arms	50	50	50	50
$X_{0,r}$	43	87	36	75

Table III shows that there is no need for queuing at the level of relays 2 and 4. In fact, the number of users that can connect to relays 2 and 4 per timeslot, i.e., the number of users able to choose unique arms, is lower than the maximum number of supported users. This is not the case for relays 1 and 3 where a queue may be needed if the number of connected devices exceeds the maximum supported number $X_{0,r}$.

In Figure 7, the instantaneous number of successfully connected devices per relay is shown for this setup. The devices
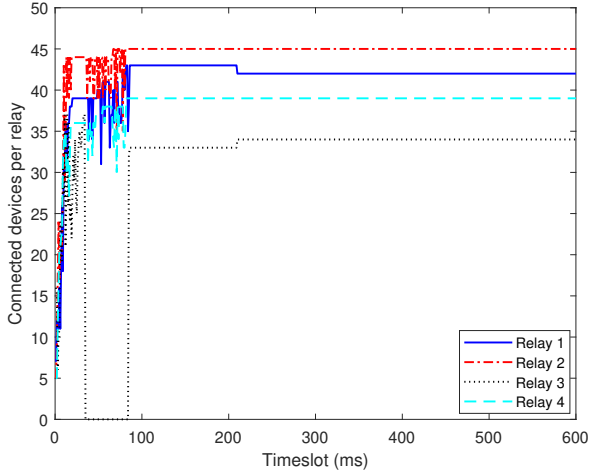
Figure 7: Number of connected devices per relay.

start by exploring the available arms until reaching convergence. Relay 3 reaches its overflow requirements since the number of connected devices to this relay when reaching the first convergence state is larger than its maximum supported number $X_{0,r}$. Hence, relay 3 enters a timeout period between timeslots 40 and 90, in which a zero reward is returned to all devices trying to associate with it. During this period, users previously associated with relay 3 attempt to connect to other less congested relays. Having succeeded to do so, a first convergence is reached within 100 timeslots and then full convergence is reached within 200 timeslots, as shown in Figure 7. Upon reaching convergence, the rate requirements of all relays and IoT users are satisfied.

VI. Conclusion

In this paper, uncoordinated joint channel and power allocation in a NOMA uplink system equipped with full-duplex relays was studied. The considered problem was modeled using the multi-armed bandits framework with zero-reward on collision. A solution based on the UCB algorithm that does not require any communication between devices and that satisfies the QoS requirements of all devices was proposed. Simulation results confirm the convergence of the proposed algorithm. Moreover, they show that, compared to two baseline schemes, the proposed allocation technique increases the number of satisfied IoT users whilst decreasing the needed transmit power. The proposed algorithm was then extended to account for buffering requirements in case of congestion at the relay level. In the future, we will extend this study to the case of heterogeneous systems and multiple-input multiple-output relays.

References

[1] E. Balevi, F. T. A. Rabee, and R. D. Gitlin, "ALOHA-NOMA for Massive Machine-to-Machine IoT Communication," in *2018 IEEE Int. Conf. on Commun. (ICC)*, Kansas City, MO, USA, May 2018, pp. 1–5.

[2] N. Zhang, J. Wang, G. Kang, and Y. Liu, "Uplink Non-orthogonal Multiple Access in 5G Systems," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 458–461, Mar. 2016.

[3] J. Choi, "NOMA-Based Random Access With Multichannel ALOHA," *IEEE J. on Sel. Areas in Commun.*, vol. 35, no. 12, pp. 2736–2743, Oct. 2017.

[4] A. Kilzi, J. Farah, C. Abdel Nour, and C. Douillard, "Mutual Successive Interference Cancellation Strategies in NOMA for Enhancing the Spectral Efficiency of CoMP Systems," *IEEE Trans. on Commun.*, vol. 68, no. 2, pp. 1213–1226, Feb. 2020.

[5] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access," in *2013 IEEE 77th Veh. Technol. Conf. (VTC-Spring)*, Dresden, Germany, June 2013, pp. 1–5.

[6] J. Farah, A. Kilzi, C. Abdel Nour, and C. Douillard, "Power Minimization in Distributed Antenna Systems Using Non-Orthogonal Multiple Access and Mutual Successive Interference Cancellation," *IEEE Trans. on Veh. Technol.*, vol. 67, no. 12, pp. 11 873–11 885, Dec. 2018.

[7] A. Benjebbour, K. Saito, A. Li, Y. Kishiyama, and T. Nakamura, "Non-orthogonal multiple access (NOMA): Concept, performance evaluation and experimental trials," in *2015 Int. Conf. on Wireless Netw. and Mobile Commun. (WINCOM)*, Marrakech, Morocco, Oct. 2015, pp. 1–6.

[8] C. Xiao, J. Zeng, W. Ni, X. Su, R. P. Liu, T. Lv, and J. Wang, "Downlink MIMO-NOMA for Ultra-Reliable Low-Latency Communications," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 780–794, Feb. 2019.

[9] A. Mazin, M. Elkourdi, and R. D. Gitlin, "Comparison of Slotted Aloha-NOMA and CSMA/CA for M2M Communications in IoT Networks," in *2018 IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Chicago, IL, USA, USA, Aug. 2018, pp. 1–5.

[10] X. Li and J. Hwu, "A Frequency Hopping Spread Spectrum Transmission Scheme For Uncoordinated Cognitive Radios," in *2009 IEEE Int. Conf. on Acous., Speech and Signal Proc. (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 2345–2348.

[11] M. Bande and V. V. Veeravalli, "Adversarial Multi-user Bandits for Uncoordinated Spectrum Access," in *2019 IEEE Int. Conf. on Acous., Speech and Signal Proc. (ICASSP)*, Brighton, United Kingdom, May 2019, pp. 4514–4518.

[12] Z. Ding, R. Schober, P. Fan, and H. V. Poor, "Simple Semi-Grant-Free Transmission Strategies Assisted by Non-Orthogonal Multiple Access," *IEEE Trans. on Commun.*, vol. 67, no. 6, pp. 4464–4478, Mar. 2019.

[13] H. Jiang, Q. Cui, Y. Gu, X. Qin, X. Zhang, and X. Tao, "Distributed Layered Grant-Free Non-Orthogonal Multiple Access for Massive MTC," in *2018 IEEE 29th Annu. Int. Symp. on Pers., Indoor and Mobile Radio Commun. (PIMRC)*, Bologna, Italy, Sept. 2018, pp. 1–7.

[14] M. J. Youssef, C. Abdel Nour, J. Farah, and C. Douillard, "Backhaul-Constrained Resource Allocation and 3D Placement for UAV-Enabled Networks," in *2019 IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Honolulu, HI, USA, USA, Sept. 2019, pp. 1–7.

[15] M. J. Youssef, J. Farah, C. Abdel Nour, and C. Douillard, "Full-Duplex and Backhaul-Constrained UAV-Enabled Networks Using NOMA," *IEEE Trans. on Veh. Technol.*, vol. 69, no. 9, pp. 9667–9681, Sept. 2020.

[16] M. J. Youssef, V. V. Veeravalli, J. Farah, and C. Abdel Nour, "Stochastic Multi-Player Multi-Armed Bandits with Multiple Plays for Uncoordinated Spectrum Access," in *2020 IEEE 31st Annu. Int. Symp. on Pers., Indoor and Mobile Radio Commun. (PIMRC)*, London, United Kingdom, United Kingdom, Sept. 2020, pp. 1–7.

[17] M. A. Adjif, O. Habachi, and J. P. Cances, "Joint Channel Selection and Power Control for NOMA: A Multi-Armed Bandit Approach," in *2019 IEEE Wireless Commun. and Netw. Conf. Workshop (WCNC)*, Marrakech, Morocco, Morocco, Apr. 2019, pp. 1–6.

[18] D. P. Zhou and C. J. Tomlin, "Budget-Constrained Multi-Armed Bandits with Multiple Plays," *CoRR*, vol. abs/1711.05928, 2017.

[19] I. Bistritz and A. Leshem, "Distributed Multi-Player Bandits - a Game of Thrones Approach," in *Advances in Neural Inf. Process. Syst.*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., vol. 31. Curran Associates, Inc., 2018, pp. 7222–7232.

[20] K. Liu and Q. Zhao, "Distributed Learning in Multi-Armed Bandit With Multiple Players," *IEEE Trans. on Signal Process.*, vol. 58, no. 11, pp. 5667–5681, Nov. 2010.

[21] J. Li, Y. Q. Zhao, F. R. Yu, and X. Huang, "Queuing Analysis of Two-Hop Relay Technology in LTE/LTE-A Networks With Unsaturated and Asymmetric Traffic," *IEEE Internet of Things J.*, vol. 3, no. 3, pp. 378–385, June 2016.

[22] M. Afshang and H. S. Dhillon, "Poisson Cluster Process Based Analysis of HetNets With Correlated User and Base Station Locations," *IEEE Trans. on Wireless Commun.*, vol. 17, no. 4, pp. 2417–2431, Jan. 2018.