Data Throughput Maximization in a Novel Hybrid TDMA-NOMA System Based on Time Allocation

Jun Wang and Yang Liu

December 18, 2024

Data Throughput Maximization in a Novel Hybrid TDMA-NOMA System Based on Time Allocation

Jun Wang
Department of Electronic Engineering
Beijing University of Posts and Telecommunications
Beijing, China
junw@bupt.edu.cn

Yang Liu
Department of Electronic Engineering
Beijing University of Posts and Telecommunications
Beijing, China
ly@bupt.edu.cn

***Abstract*—By 2030, the number of mobile IoT connections in China will reach a scale of billions. Therefore, efficiently transmitting more data within limited time and resources has become an urgent problem to address. In this paper, we consider a novel hybrid non-orthogonal multiple access (NH-NOMA) system based on downlink transmission links with opportunistic time allocation. We develop a new hybrid non-orthogonal multiple access model based on time allocation, where the data transmitted by users in different time slots is rederived. To maximize the system's data transmission while meeting the users' data requirements, we transform the non-convex optimization problem into a convex optimization problem, obtaining the solution to the optimization problem. Finally, simulation results demonstrated that a time allocation based hybrid multiple access system achieves higher data transmission efficiency compared to traditional orthogonal multiple access systems and time division-based hybrid systems. This system effectively utilizes users with favorable channel conditions to maximize their data transmission while ensuring the basic communication needs of other users.**

***Index Terms*—Hybrid Non-Orthogonal Multiple Access, opportunistic time allocation, data throughput, non-convex optimization, priority service scenarios**

I. Introduction

In the development of 6G, Non-Orthogonal Multiple Access (NOMA) technology will play a crucial role in meeting the higher data rates, lower latency, and greater connection density required by 6G networks [1]. To further enhance system spectral efficiency and throughput, NOMA is often combined with other orthogonal multiple access (OMA) technologies to form hybrid NOMA schemes. Reference [2] indicates that Frequency Domain NOMA (FD-NOMA) not only significantly enhances spectral efficiency but also achieves higher data transmission rates and better user fairness through optimized spectrum resource allocation and power control. Time Domain NOMA (TD-NOMA), on the other hand, enables multi-user access within the same time interval [3]. In traditional TD-NOMA, each user transmits only in one time slot. However, literature [4] views NOMA as an additional component to OMA, allowing each user to be scheduled for data transmission similar to OMA, but unlike OMA, the time slot of the scheduled user can also be utilized by other users. Therefore, this paper adopts this user grouping method to achieve the coexistence of OMA and NOMA.

In NOMA systems, resource allocation and power management are among the key issues. Research indicates that by optimizing power allocation strategies and resource scheduling algorithms, it is possible to maximize system capacity and minimize energy consumption while meeting users' requirements of Quality of Service (QoS). For example, literature [5] proposed a dynamic power allocation method based on users' Channel State Information (CSI), which effectively reduces system power consumption. In [6], under a given total system power constraint, the maximum power P_k allocated to each user is limited to ensure fairness. The study addresses the problem of maximizing transmission rates in FD-NOMA systems using both centralized and distributed algorithms. In [7], under constraints on QoS and power, the study compares a deep reinforcement learning algorithm with traditional greedy and genetic algorithms, and the proposed deep reinforcement learning algorithm achieves higher throughput rates. In [8], considering joint power and time resource allocation, a Joint Generalized Energy Efficiency Maximization (GEE-Max) design is proposed for the downlink transmission of hybrid TDMA-NOMA systems. The study utilizes Sequential Convex Approximation (SCA) and novel Second-Order Cone (SOC) methods to address non-convex problems. It demonstrates that hybrid TDMA-NOMA systems with opportunistic time allocation outperform traditional resource allocation with equal time allocation in terms of minimal required transmit power and achieved total throughput. Incorporating an intelligent reflecting surface module in literature [9], the study focuses on reducing system energy consumption through optimized allocation of reflecting surfaces, user grouping, angle settings, and power distribution.

Previous studies on novel hybrid NOMA (NH-NOMA) [4] have primarily focused on minimizing system energy consumption while meeting user demands. However, this approach does not directly reflect the specific data transmission volumes for each user. Additionally, past research has derived channel capacities under equal time allocation, which limits the ability to leverage advantages for users with better channel conditions in priority service scenarios, thereby potentially reducing overall system performance. The main contribution of this paper is to consider time allocation and rederive the achievable data transmission volumes for each user in a NH-NOMA system. Resource allocation aims to achieve maximum data throughput while meeting each user's minimum data requirement. By introducing slack variables or Taylor expansions, the non-convex optimization problem transforms into a convex problem. Finally, the optimization results are discussed and analyzed based on a set of benchmark scenarios.

The rest of the paper is organized as follows: Section II introduces the NH-NOMA model and provides a detailed derivation of the user's data transmission under this model. Section III describes the problem and its solutions. Simulation results are presented in Section IV. Section V concludes the paper.

II. System Model

Consider a downlink transmission of a traditional Single-Input Single-Output (SISO) TDMA network consisting of K users. In this system, both the base station and all users are equipped with a single antenna, with the users assumed to be uniformly distributed on one side of the base station. Let h_m denote the channel gain of user U_m. Assuming the base station can obtain CSI from users. Specifically, it broadcasts pilot signals to all users, users receive these pilot signals, estimate their channel gains based on the received signals, and then feed back the estimated channel gains to the base station. After receiving CSI feedback from all users, the base station sorts the users based on their channel quality, specifically $|h_1|^2 > |h_2|^2 > \ldots > |h_K|^2$, the users are correspondingly sorted as $U_1, U_2, ..., U_K$.To enable each user to achieve OMA transmission, the available transmission time T is divided into K slots, which are used for orthogonal transmission for each of the K users, with the length of each slot denoted as t_i for $i = 1, 2, ..., K$, t_i are optimization variables and may not necessarily be equal. Allowing users with better channel conditions to transmit first, while also enabling weaker users to cluster together in each time slot, as illustrated in Figure 1. Considering each time slot, using NOMA technology to superimpose the information of users with poorer channel conditions onto the current time slot. This allows all users to achieve orthogonal transmission between time slots while enabling non-orthogonal transmission within each time slot. Therefore, the signal transmitted by the base station in the i-th time slot is:

$$X_i = \sum_{j=i}^{K} \sqrt{p_{i,j}} x_{i,j} \tag{1}$$

where $u_{i,j}$ represents the U_j in the i-th time slot ($i \leq j$), $p_{i,j}$ denotes the power allocated to $u_{i,j}$, and $x_{i,j}$ represents the signal sent to $u_{i,j}$.

Assuming the channel follows quasi-static flat Rayleigh fading, which means the channel coefficients remain constant within each transmission block but independently vary between different blocks.

To ensure each user can decode without signal interference in their corresponding time slot as OMA transmission, the decoding process in each time slot begins with decoding the

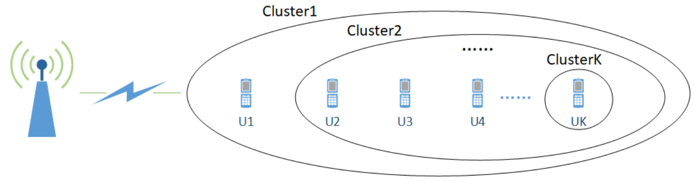


Fig. 1: novel hybrid NOMA transmission system.

last user U_K first, proceeding in reverse order of user indices from the highest to the lowest, where channel gains are correspondingly from smallest to largest, using SIC as illustrated in Figure 2. In the i-th time slot, the Signal-to-Interference-plus-Noise Ratio (SINR) for weak user $U_m (m \geq i)$ decoding its own signal is:$SINR_{i,m}^{m} = \frac{|h_m|^2 p_{i,m}}{|h_m|^2 \sum_{j=i}^{m-1} p_{i,j} + \sigma^2}$, The SINR with which strong user $U_n (i \leq n < m)$ decodes the signal of weak user U_m is:$SINR_{i,m}^{n} = \frac{|h_n|^2 p_{i,m}}{|h_n|^2 \sum_{j=i}^{m-1} p_{i,j} + \sigma^2}$. According to the SIC order, the SINR for the weak user U_m must satisfy $SINR = min\{SINR_{m,i}^{m}, SINR_{m,i}^{n}\} = SINR_{m,i}^{m}$. Therefore, if U_m can successfully decode its own signal, U_n will certainly be able to decode and remove $U_m^{\prime} s$ signal with effective SIC. Then the SINR of the U_j in the i-th time slot is:

$$SINR_{i,j} = \frac{|h_j|^2 p_{i,j}}{|h_j|^2 \sum_{k=i}^{j-1} p_{i,k} + \sigma^2} \tag{2}$$

Where $|h_j|^2$ represents the squared magnitude of the channel gain for U_j, and σ^2 denotes the noise power at U_j in the i-th time slot. Assuming the noise power is normalized, i.e., $\sigma^2 = 1$. When $j = i$, $SINR_{i,i} = \frac{|h_i|^2 p_{i,i}}{\sigma^2}$, at which point the user's interference is reduced to only noise.

Then, The total amount of data that the system can transmit in the i-th time slot is given by:

$$Sum(i) = t_i \times \sum_{j=i}^{K} \log(1 + SINR_{i,j}) \tag{3}$$

Meanwhile, the data of U_k can be transmitted before the k-th time slot, then the total amount of data that U_k can transmit in one time frame is:

$$N_k = \sum_{i=1}^{k} t_i \times \log(1 + SINR_{i,k}) \tag{4}$$

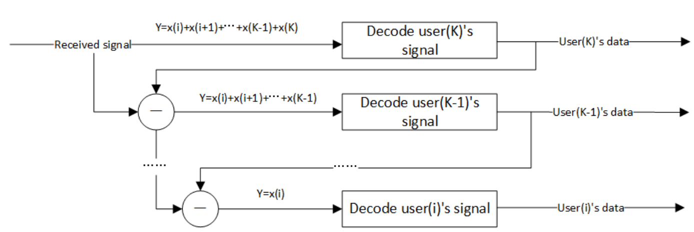


Fig. 2: Successive Interference Cancellation (SIC) Technique Employed in the i-th Time Slot.

III. Problem Formulation And Solution

A. Problem Formulation

In this section, we have developed a data transmission maximization optimization design for the NH-NOMA system. In this optimization design, the slot length is also treated as an optimization variable, jointly optimized with power allocation. It requires that the total length of segmented slots equals the length of one time frame T, and the total power allocated to users across all time slots does not exceed the total power P_{total} that the base station can provide. To ensure the optimization problem is realistic, we require that the power allocated to users must be greater than or equal to 0, furthermore, to ensure each user operates in an orthogonal transmission mode within their corresponding slot, the power allocation for U_i in the i-th slot must not be zero, i.e., $P_{i,i} > 0$. In addition, we require that under the given total power, minimum data rate requirement N_{min} must be satisfied to ensure QoS for every user.

This optimization problem can be formulated as follows:

$$\text{(P1)}: \max_{p_{i,j},t_i} \quad \sum_{i=1}^{K} Sum(i) \tag{5}$$

$$\text{s.t.} \quad \sum_{i=1}^{K} t_i = T, \tag{6}$$

$$\sum_{i=1}^{K}\sum_{j=i}^{K} p_{i,j} \leqslant P_{total}, \tag{7}$$

$$p_{i,j} \geqslant 0, \forall j > i, i \in \{1, ..., K\}, \tag{8}$$

$$p_{i,i} > 0, i \in \{1, ..., K\}, \tag{9}$$

$$N_k \geqslant N_{min}, k \in \{1, ..., K\} \tag{10}$$

However, there are some challenges in solving P1. Firstly, unlike the equal time allocation considered in the work [4], the joint allocation of time and power resources introduces additional complexity, making the resolution of the problem and the evaluation of design parameters more challenging. Secondly, the optimization objective turns to the total data transmitted by the system from a convex function of system energy consumption. It can be seen from P1 that not only is the objective function non-convex, but also the constraints (10) form a non-convex set. The non-convex nature increases the complexity of solving the optimization problem. Finally, due to the constraints in (7) and (10), if available power budget cannot meet the minimum quality of service requirements for users, the optimization problem P1 may become infeasible. Therefore, these factors must be comprehensively considered when solving P1.

B. Problem Solution

In this optimization scheme, we introduce slack variables and the Taylor series expansion to transform the problem into a convex optimization problem for solving. On one hand, if the non-convex problem can be successfully transformed into a convex problem, it benefits from stronger theoretical guarantees. On the other hand, using the Taylor series for approximation can ensure the convergence and stability of iterative methods.

We will first address the objective function. Due to the involvement of high-dimensional variables and complex functions in the objective function, using the Taylor series expansion may become impractical. Therefore, we introduce new slack variables γ to relax the original non-convex parts and transform the objective function into a convex form. The objective function can be expressed as:

$$\max_{p_{i,j},t_i,\gamma} \gamma \tag{11}$$

a new constraint needs to be added:

$$\sum_{i=1}^{K}\sum_{j=i}^{K} t_i \log(1 + SINR_{i,j}) \geqslant \gamma \tag{12}$$

Notice that the objective function (11) becomes convex, but a new non-convex constraint (12) is also introduced. Now, we handle the non-convexity of (12) by introducing new slack variables $\alpha_{i,j}$ and $\beta_{i,j}$ as follows:

$$(1 + SINR_{i,j}) \geq \alpha_{i,j} \tag{13a}$$

$$\log(1 + SINR_{i,j}) \geq \beta_{i,j} \tag{13b}$$

$$\alpha_{i,j} \geq 2^{\beta_{i,j}} \tag{13c}$$

$$\sum_{i=1}^{K}\sum_{j=i}^{K} t_i \beta_{i,j} \geq \gamma \tag{13d}$$

Where $i \in \{1, ..., K\}; j \in \{i, ..., K\}$; (13a) is intended to separate and convexly approximate the variables in $SINR_{i,j}$ further, while (13b) to (13d) are aimed at constraining the variables in (13a) to satisfy equation (12).

Clearly, the constraint in expression (13c) forms a convex set. To overcome the non-convexity issue of (13a), new slack variables $\phi_{i,j}$ are introduced, ensuring that

$$\frac{|h_j|^2 p_{i,j}}{|h_j|^2 \sum_{k=i}^{j-1} p_{i,k} + \sigma^2} \geq \frac{(\alpha_{i,j} - 1)\phi_{i,j}}{\phi_{i,j}} \tag{14}$$

Accordingly, the constraints in (14) can be decomposed into the following two constraints:

$$|h_j|^2 p_{i,j} \geq (\alpha_{i,j} - 1)\phi_{i,j} \tag{15a}$$

$$|h_j|^2 \sum_{k=i}^{j-1} p_{i,k} + \sigma^2 \leq \phi_{i,j} \tag{15b}$$

Clearly, (15b) represents a convex set since both sides of the inequality are linear expressions. However, (15a) remains non-convex, where the left side is a linear expression and the right side is a simple function composed of two variables. Therefore, we consider an approximation using a first-order Taylor series expansion, expanding the simple function on the right side into a linear expression of two variables. Then, it transforms into the following constraint:

$$\begin{aligned} |h_j|^2 p_{i,j} \geq (\alpha_{i,j}^{(0)} - 1)\phi_{i,j}^{(0)} + \phi_{i,j}^{(0)}(\alpha_{i,j} - \alpha_{i,j}^{(0)}) \\ + (\alpha_{i,j}^{(0)} - 1)(\phi_{i,j} - \phi_{i,j}^{(0)}) \end{aligned} \tag{16}$$

Where $\alpha_{i,j}^{(0)}$ and $\phi_{i,j}^{(0)}$ represent the initial values of $\alpha_{i,j}$ and $\phi_{i,j}$, respectively. Using these approximations, the constraints in (13a) can be rewritten as convex constraints in (15b) and (16).

Now, we address the non-convexity of the constraint in (13d). Similar to the previous approximation, the non-convex constraint in (13d) is rewrite by using a new slack variable $z_{i,j}$ as

$$t_i\beta_{i,j} \geq z_{i,j} \tag{17a}$$

$$\sum_{i=1}^{K}\sum_{j=i}^{K} z_{i,j} \geq \gamma \tag{17b}$$

Then, the right side of (17a) is approximated using an upper bound convex approximation through the first-order Taylor series expansion. Therefore, (17a) can be reformulated as:

$$t_i^{(0)}\beta_{i,j}^{(0)} + \beta_{i,j}^{(0)}(t_i - t_i^{(0)}) + t_i^{(0)}(\beta_{i,j} - \beta_{i,j}^{(0)}) \geq z_{i,j} \tag{18}$$

Where $t_i^{(0)}$ and $\beta_{i,j}^{(0)}$ represent the initial values of t_i and $\beta_{i,j}$, respectively. The constraint in (13d) can be rewritten as convex constraints in (17b) and (18).

In summary, the non-convex constraint of (12) can be equivalently represented by the set of convex constraints: (13c), (15b), (16), (17b), and (18). Therefore, the constraint on the minimum data transmission requirement N_k for user k can be redefined as the following convex constraints:

$$\sum_{i=1}^{k} z_{i,k} \geq N_{min} \tag{19}$$

Through the aforementioned relaxation, the original non-convex optimization problem P1 can be equivalently written as the following approximate convex optimization problem:

$$\text{(P2):} \quad \max_{\Gamma} \gamma \tag{20}$$

$$\text{s.t.} \quad (6) \sim (9), (13c), (15b), (16), (17b), (18), (19) \tag{21}$$

Where Γ consists of all optimization variables, i.e., $\Gamma = \{p_{i,j}, t_i, \gamma, \alpha_{i,j}, \beta_{i,j}, \phi_{i,j}, z_{i,j}\}$ for any $i \in \{1, \ldots, K\}$ and $j \in \{i, i+1, ..., K\}$.

Specifically, by solving the approximate convex optimization problem P2, we obtain the solution to the original non-convex optimization problem P1. However, when using Taylor series expansion to transform a non-convex optimization problem into a convex one, the choice of initial point is crucial. That's because Taylor series expansion is a local approximation method, and the choice of the expansion center (i.e., the initial point) directly affects the accuracy of the approximation and the quality of the optimization results. Therefore, based on previous experience, the time slots are initially set to equal lengths, i.e.,$t_i = T/K(\forall i \in \{1, ..., K\})$. The optimal power allocation from the P-min problem is used as the initial power allocation for this optimization problem. These initial values are then used to determine the corresponding slack variables by substituting them into equations (13a), (13b), (14), (17a), and (17b). Since P2 is a convex problem, the optimal solution can be quickly obtained using established optimization tools.

IV. Simulation Result

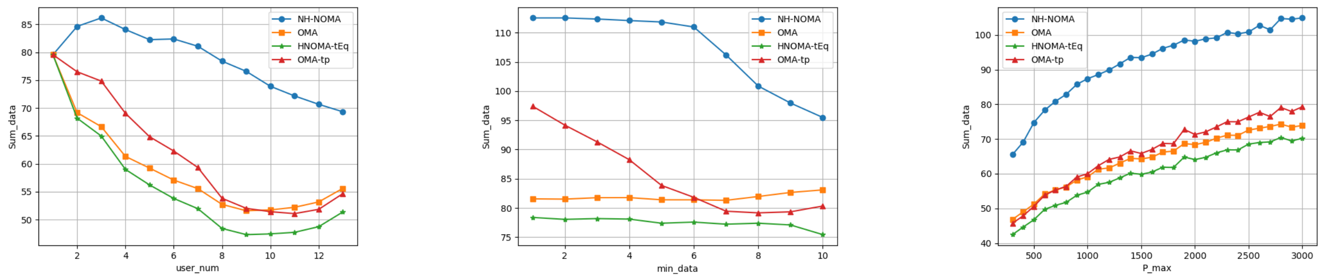
In this section, simulation results are provided to demonstrate that within a single time frame, the proposed novel hybrid NOMA (NH-NOMA) based on joint time and power allocation achieves higher data transmission, OMA with equal time and power allocation, OMA with joint time and power optimization (OMA-tp), and HNOMA with equal time allocation (HNOMA-tEq) are used as benchmark schemes. All schemes are assumed to use the same minimum quality of service requirements. Table I presents the time allocation scheme for NH-NOMA. In the simulation, it is assumed that the users' channels are independently and identically distributed (i.i.d.) complex Gaussian random variables, with the noise power at all users being $\sigma^2 = 1$. The length of a time frame is set to 10 seconds.

It can be observed from Figure 3a that the proposed NH-NOMA achieves the highest data transmission. Specifically, NH-NOMA significantly outperforms OMA-tp, which also uses joint time and power allocation. HNOMA-tEq exhibits the worst performance, this is because equal time allocation means that users with poor channel conditions still receive the same transmission time, leading to inefficiency. Additionally, as the number of users increases, the performance of all four transmission schemes declines. This is due to the relative reduction in power allocated to each user and the increase in interference, leading to a decrease in each user's signal quality, and formulating an optimal power allocation strategy becomes more complex and challenging, potentially decreasing resource utilization efficiency.

Figure 3b shows that as the minimum data transmission requirement for users increases, the total data throughput achieved by NH-NOMA decreases. This is because, on one hand, the stronger users' signals increase interference for weaker users. On the other hand, the increased signal requirements of weak users mean that the system needs to allocate more time and power to "accommodate" users with poor channel conditions, thus consuming a significant amount of resources and reducing overall performance.

As shown in Figure 3c, as the maximum system power increases, the total data transmission capacity of all transmission schemes rises, with the proposed NH-NOMA scheme exhibiting the most significant increase. This is because with increased available power, better power scheduling can be achieved, resulting in a higher signal-to-interference ratio. The NH-NOMA scheme optimally allocates time and power based on user channel conditions, further enhancing resource utilization efficiency.

Figure 4 illustrates that under the same minimum data requirement constraint, NH-NOMA based on time allocation can meet the data demands of most users, users with better channel conditions achieve data rates far exceeding other users. It can be seen from table I that the time slot for the user with the best channel conditions is allocated the majority of the time. This is because, allocating the same resources to the user who can transmit data in the form of OMA and has the best



(a) Total data throughput vs. user_num. (b) Total data throughput vs. min_data. (c) Total data throughput vs. P_max

Fig. 3: The relationship between data throughput and some metrics for different transmission schemes.

TABLE I: the time allocation of NH-NOMA

	Time allocation in NH-NOMA									
Channels	$t_1(s)$	$t_2(s)$	$t_3(s)$	$t_4(s)$	$t_5(s)$	$t_6(s)$	$t_7(s)$	$t_8(s)$	$t_9(s)$	$t_{10}(s)$
Channel 1	4.757	0.951	0.512	0.514	0.507	0.510	0.544	0.562	0.586	0.555
Channel 2	5.574	0.825	0.458	0.460	0.446	0.437	0.431	0.461	0.448	0.459
Channel 3	5.570	0.569	0.484	0.488	0.491	0.494	0.483	0.490	0.488	0.443

channel yields the highest return. This system ensures that key users can still enjoy high-speed, low-latency communication even under heavy overall system load. It is suitable for priority service scenarios in resource-limited environments, such as prioritizing critical tasks in enterprise networks or meeting high service quality demands for specific user groups.

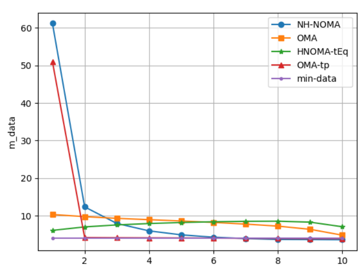


Fig. 4: The amount of data transmitted per user under different transmission schemes under identical conditions.

V. Conclusion

In this paper, we propose a novel hybrid NOMA system (NH-NOMA) based on time allocation for downlink transmission. Specifically, we consider allowing users to access multiple time slots while ensuring that each user can achieve OMA transmission in their respective slots, and jointly allocate time and power to maximize the system's data transmission capacity. Due to the non-convex nature of the optimization problem, we employ slack variables and Taylor expansion to transform the problem into a convex one for resolution. The results demonstrate that under the same minimum required transmission data, NH-NOMA with time allocation consistently achieves higher total data transmission compared to equal time allocation hybrid NOMA and traditional TDMA. In addition, the results demonstrate that NH-NOMA exhibits favorable properties in priority service scenarios, where users with favorable channel conditions can maintain high service quality even under heavy system loads, while other users receive basic service.

References

[1] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends," IEEE J. Select. Areas Commun., vol. 35, no. 10, pp. 2181–2195, Oct. 2017.

[2] S. M. R. Islam, M. Zeng, O. A. Dobre, and K.-S. Kwak, "Resource Allocation for Downlink NOMA Systems: Key Techniques and Open Issues," IEEE Wireless Commun., vol. 25, no. 2, pp. 40–47, Apr. 2018.

[3] Y. Zhu, X. Yuan, Y. Hu, T. Wang, M. C. Gursoy, and A. Schmeink, "Low-Latency Hybrid NOMA-TDMA: QoS-Driven Design Framework," IEEE Trans. Wireless Commun., vol. 22, no. 5, pp. 3006–3021, May 2023.

[4] Zhiguo Ding, Robert Schober, and H. Vincent Poor, "Design of Downlink Hybrid NOMA Transmission." unpublished.

[5] X. Li, C. Li, and Y. Jin, "Dynamic Resource Allocation for Transmit Power Minimization in OFDM-Based NOMA Systems," IEEE Commun. Lett., vol. 20, no. 12, pp. 2558–2561, Dec. 2016.

[6] Y. Fu, L. Salaun, C. W. Sung, and C. S. Chen, "Subcarrier and Power Allocation for the Downlink of Multicarrier NOMA Systems," IEEE Trans. Veh. Technol., vol. 67, no. 12, pp. 11833–11847, Dec. 2018.

[7] Cirine Chaieb, Fatma Abdelkefi, and Wessam Ajib, "Deep Reinforcement Learning for Resource Allocation in Multi-Band and Hybrid OMA-NOMA Wireless Networks," IEEE Trans. Commun., vol. 71, no. 1, pp. 187–198, Jan. 2023.

[8] Xinchen Wei, Haitham Al-Obiedollah, Kanapathippillai Cumanan, Zhiguo Ding, and Octavia 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.

[9] Haitham Al-Obiedollah, Haythem A. Bany Salameh, Kanapathippillai Cumanan, Zhiguo Ding, and Octavia A. Dobre, "Self-Sustainable Multi-IRS-Aided Wireless Powered Hybrid TDMA-NOMA System," IEEE Access, vol. 11, pp. 57428–57436, 2023.