A Survey of Downlink Non-orthogonal Multiple Access for 5G Wireless Communication Networks

Wei, Z; Yuan, J; Ng, DWK; Elkashlan, M; Ding, Z

https://arxiv.org/abs/1609.01856

A Survey of Downlink Non-orthogonal Multiple Access for 5G Wireless Communication Networks

Zhiqiang Wei[*], Jinhong Yuan[*], Derrick Wing Kwan Ng[*],

Maged Elkashlan[†], and Zhiguo Ding[‡]

[*] The University of New South Wales, Sydney, Australia

[†]Queen Mary University of London, London, UK

[‡]Lancaster University, Lancaster, UK

Abstract

Non-orthogonal multiple access (NOMA) has been recognized as a promising multiple access technique for the next generation cellular communication networks. In this paper, we first discuss a simple NOMA model with two users served by a single-carrier simultaneously to illustrate its basic principles. Then, a more general model with multicarrier serving an arbitrary number of users on each subcarrier is also discussed. An overview of existing works on performance analysis, resource allocation, and multiple-input multiple-output NOMA are summarized and discussed. Furthermore, we discuss the key features of NOMA and its potential research challenges.

I. Introduction and Background

The fifth generation (5G) communication systems is on its way. It is widely believed that 5G is not just an incremental version of the fourth generation (4G) communication systems [1] due to not only the increasing demand of data traffic explosion, but also the expected new services and functionalities, such as internet-of-things (IoT), cloud-based architectural applications, etc [2]. These envisioned services pose challenging requirements for 5G wireless communication systems, such as much higher data rates ($100 \sim 1000\times$ of current 4G technology), lower latency (1 ms for a roundtrip latency), massive connectivity and support of diverse quality of service (QoS) (10^6 devices/km^2 with diverse QoS requirements) [1]. From a technical perspective, to meet the aforementioned challenges, some potential technologies, such as massive multiple-input multiple-output (MIMO) [3], [4], millimeter wave [5], [6], ultra densification and offloading
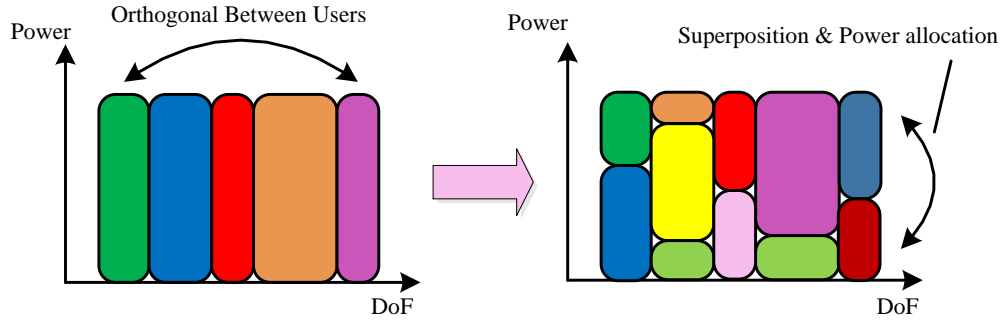
Fig. 1. From OMA to NOMA via power domain multiplexing.

[7]–[9], have been discussed extensively. Besides, it is expected to employ a future radio access technology for 5G, which is flexible, reliable [10], and efficient in terms of energy and spectrum [11], [12]. Radio access technologies for cellular communications are characterized by multiple access schemes, such as frequency-division multiple access (FDMA) for the first generation (1G), time-division multiple access (TDMA) for the second generation (2G), code-division multiple access (CDMA) used by both 2G and the third generation (3G), and orthogonal frequency-division multiple access (OFDMA) for 4G. All these conventional multiple access schemes are categorized as orthogonal multiple access (OMA) technologies, where different users are allocated to orthogonal resources in either time, frequency, or code domain in order to mitigate multiple access interference (MAI). However, OMA schemes are not sufficient to support the massive connectivity with diverse QoS requirements. In fact, due to the limited degrees of freedom (DoF), some users with better channel quality have a higher priority to be served while other users with poor channel quality have to wait to access, which leads to high unfairness and large latency. Besides, it is inefficient when allocating DoF to users with poor channel quality. In this survey, we focus on one promising technology, non-orthogonal multiple access (NOMA), which in our opinion will contribute to disruptive design changes on radio access and address the aforementioned challenges of 5G.

In contrast to conventional OMA, NOMA transmission techniques intend to share DoF among users via superposition and consequently need to employ multiple user detection (MUD) to separate interfered users sharing the same DoF, as illustrated in Fig. 1. NOMA is beneficial to enlarge the number of connections by introducing controllable symbol collision in the same DoF. Therefore, NOMA can support high overloading transmission and further improve the system capacity given limited resource (spectrum or antennas). In addition, multiple users with different types of traffic request can be multiplexed to transmit concurrently on the same DoF

to improve the latency and fairness. The comparison of OMA and NOMA is summarized in Table I. As a result, NOMA has been recognized as a promising multiple access technique for the 5G wireless networks due to its high spectral efficiency, massive connectivity, low latency, and high user fairness [13]. For example, multiuser superposition transmission (MUST) has been proposed for the third generation partnership project long-term evolution advanced (3GPP-LTE-A) networks [14]. Three kinds of non-orthogonal transmission schemes have been proposed and studied in the MUST study item. Through the system-level performance evaluation, it has been shown that the MUST can increase system capacity as well as improve user experience.

Recently, several NOMA schemes have been proposed and received significant attention. According to the domain of multiplexing, the authors in [13] divided the existing NOMA techniques into two categories, i.e., code domain multiplexing (CDM) and power domain multiplexing (PDM). The CDM-NOMA techniques, including low-density spreading (LDS) [15]–[17], sparse code multiple access (SCMA) [18], pattern division multiple access (PDMA) [19], etc, introduce redundancy via coding/spreading to facilitate the users separation at the receiver. For instance, LDS-CDMA [15] intentionally arranges each user to spread its data over a small number of chips and then interleave uniquely, which makes optimal MUD affordable at receiver and exploits the intrinsic interference diversity. LDS-OFDM [16], [17], as shown in Fig. 2, can be interpreted as a system which applies LDS for multiple access and OFDM for multicarrier modulation. Besides, SCMA is a generalization of LDS methods where the modulator and LDS spreader are merged. On the other hand, PDM-NOMA exploits the power domain to serve multiple users in the same DoF, and performs successive interference cancellation (SIC) at users with better channel conditions.

In fact, the non-orthogonal feature can be introduced either in the power domain only or in the hybrid code and power domain. Although CDM-NOMA has the potential code gain to improve spectral efficiency, PDM-NOMA has a simpler implement since there is almost no big change in the physical layer procedures at the transmitter side compared to current 4G technologies. In addition, PDM-NOMA paves the way for flexible resource allocation via relaxing the orthogonality requirement to improve the performance of NOMA, such as spectral efficiency [20], [21], energy efficiency [22], and user fairness [23]. Therefore, this paper will focus on the PDM-NOMA, including its basic concepts, key features, existing works, and future research challenges.

TABLE I

COMPARISON OF OMA AND NOMA.

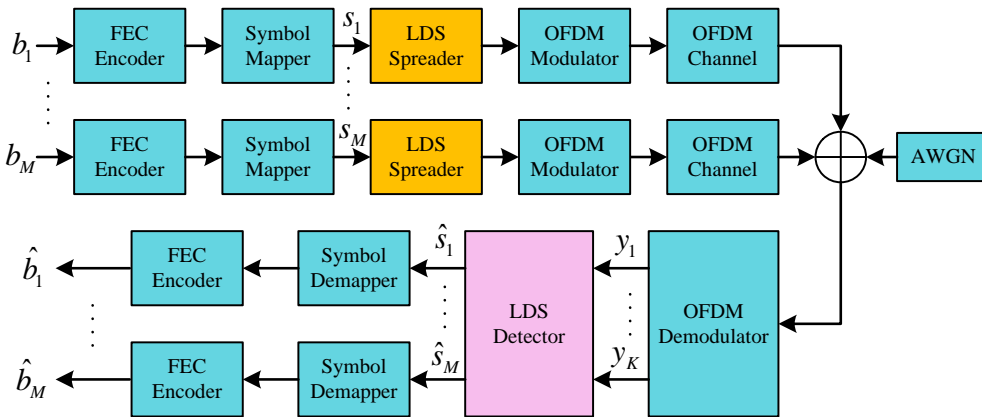|	Advantages	Disadvantages
OMA	• Simpler receiver detection	• Lower spectral efficiency
		• Limited number of users
		• Unfairness for users
NOMA	• Higher spectral efficiency	• Increased complexity of receivers
	• Higher connection density	• Higher sensitivity to channel uncertainty
	• Enhanced user fairness	
	• Lower latency	
	• Supporting diverse QoS	

Fig. 2. Block diagram of an uplink LDS-OFDM system.

II. FUNDAMENTALS OF NOMA

This section presents the basic model and concepts of single-antenna downlink NOMA. The first subsection presents a simple downlink single-carrier NOMA (SC-NOMA) system serving two users simultaneously, while the second subsection presents a more general multi-carrier NOMA (MC-NOMA) model for serving an arbitrary number of users in each subcarrier.

1) Two-user SC-NOMA: [1] Benjebbour and Saito et al. [24], [25] proposed the system model of downlink NOMA with superposition transmission at the base station (BS) and successive interference cancellation (SIC) at the user terminals, which can be illustrated in Fig. 3 in case of one BS and two users.

The BS transmits the messages of both user 1 and user 2, i.e., s_1 and s_2, with different transmit

[1]In this paper, a two-user NOMA systems means that two users are multiplexed on each subcarrier simultaneously. Similarly, a multiuser NOMA systems means that an arbitrary number of users are multiplexed on each subcarrier simultaneously.
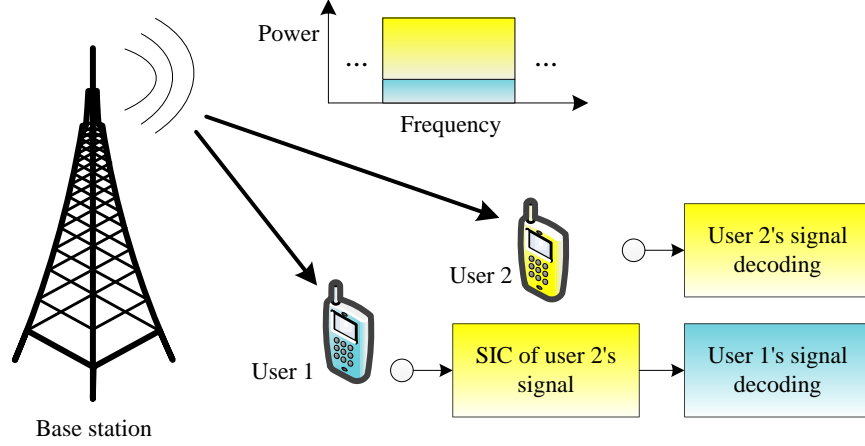
Fig. 3. A downlink NOMA model with one base station and two users.

powers p_1 and p_2, on the same subcarrier, respectively. The corresponding transmitted signal is represented by

$$x = \sqrt{p_1}s_1 + \sqrt{p_2}s_2, \tag{1}$$

where transmit power is constrained by $p_1 + p_2 = 1$. The received signal at user i is given by

$$y_i = h_i x + v_i, \tag{2}$$

where h_i denotes the complex channel coefficient including the joint effect of large scale fading and small scale fading. Variable v_i denotes the additive white Gaussian noise (AWGN), and $v_i \sim \mathcal{CN}(0, \sigma_i^2)$, where $\mathcal{CN}(0, \sigma_i^2)$ denotes the circularly symmetric complex Gaussian distribution with mean zero and variance σ_i^2. Assume that user 1 is the cell-center user with a better channel quality (*strong user*), while user 2 is the cell-edge user with a worse channel quality (*weak user*), i.e., $\frac{|h_1|^2}{\sigma_1^2} \geq \frac{|h_2|^2}{\sigma_2^2}$. According to the NOMA protocol [26], the BS will allocate more power to the weak user to provide fairness and facilitate the SIC process, i.e., $p_1 \leq p_2$.

In downlink SC-NOMA, the SIC process is implemented at the receiver side. The optimal SIC decoding order is in the descending order of channel gains normalized by noise. It means that user 1 will decode s_2 first and remove the inter-user interference of user 2 by subtracting s_2 from the received signal y_1 before decoding its own message s_1. On the other hand, user 2 does not perform interference cancellation and directly decodes its own message s_2 with interference from user 1. Fortunately, the power allocated to user 2 is larger than that of user 1 in the aggregate received signal y_2, which will not introduce much performance degradation compared to allocating user 2 on this subcarrier exclusively. The rate region of SC-NOMA is illustrated in Fig. 4 in comparison with that of OMA, where it has been proved that NOMA schemes
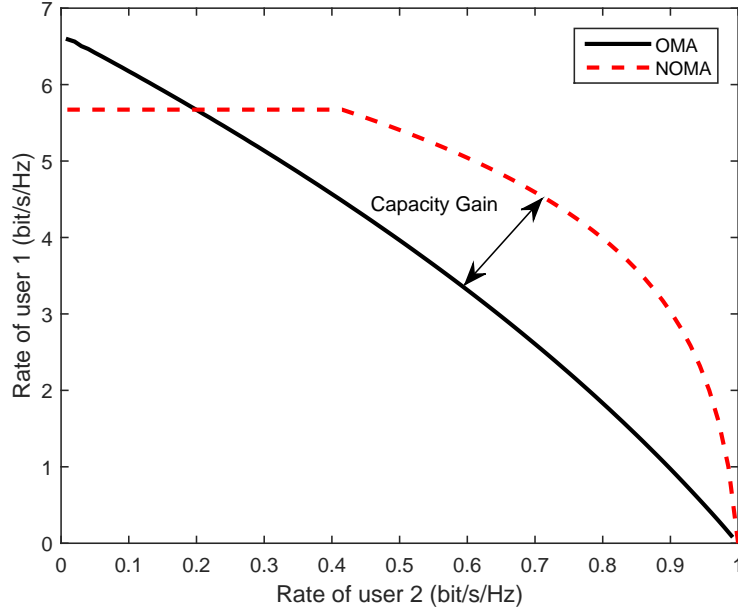
Fig. 4. The rate region of two-user SC-NOMA in comparison with that of OMA. User 1 is strong user with $\frac{|h_1|^2}{\sigma_1^2} = 100$, while user 2 is weak user with $\frac{|h_2|^2}{\sigma_2^2} = 1$.

are very likely to outperform OMA schemes in [27]. Note that the rate region of NOMA only covers a part of the capacity region of broadcast channel with SIC receiver [28] due to the power constraint $p_1 \leq p_2$.

2) Multiuser MC-NOMA: For a downlink MC-NOMA system with one BS serving an arbitrary number of users, such as K, the available bandwidth is divided into a set of N subcarriers, where $N < K$, i.e., an overloading scenario that OFDMA cannot afford. The channel between user k and the BS on subcarrier n is denoted by $h_{k,n}$, and is assumed to be perfectly known at both transmitter and receiver side. The BS schedules all users across all subcarriers by ξ_k and ζ_n, where ξ_k denotes a user set allocated on subcarrier k and ζ_n denotes a subcarrier set occupied by user n. Without loss of generality, the channel gains of all users allocated on subcarrier k are sorted as $\left|h_{k,b(1)}\right|^2 \geq \left|h_{k,b(2)}\right|^2 \geq \cdots \geq \left|h_{k,b(|\xi_k|)}\right|^2$, where $|\xi_k|$ denotes the cardinality of the user set ξ_k and $b(\cdot)$ indicates the mapping between the sorted channel gain order and the original one. For instance, for subcarrier k occupied by three users $\xi_k = \{1, 2, 3\}$ and $|h_{k,2}|^2 \geq |h_{k,3}|^2 \geq |h_{k,1}|^2$, we will have $b(1) = 2$, $b(2) = 3$, and $b(3) = 1$, respectively. Note that the mapping functions are different on different subcarriers due to the different frequency selective fading patterns of different users.

According to NOMA protocol [26], all users in ξ_k share subcarrier k by different transmission

power $p_{k,b(l)}$ based on the given channel gain, where $l = 1, 2, \cdots, |\xi_k|$ and $p_{k,b(1)} \leq p_{k,b(2)} \leq \cdots \leq p_{k,b(|\xi_k|)}$. The sharing strategy saves the subcarriers those might be wasted by only transmitting the messages of the weak users and accommodates more users with diverse QoS requirements, which is favorable to massive connectivity and IoT in 5G networks.

All messages of users in ξ_k are superimposed on subcarrier k, where the transmitted signal is given by

$$x_k = \sum_{l=1}^{|\xi_k|} \sqrt{p_{k,b(l)}} s_{k,b(l)}, \tag{3}$$

where $s_{k,b(l)}$ and $p_{k,b(l)}$ denote the message and allocated power of user $n(l)$ on subcarrier k, respectively.

Assuming the independent and identically distributed (IID) AWGN over all subcarriers and all users for simplicity, the user scheduling, power allocation, and the SIC decoding order only depends on the channel gain order. At the receiver side, the received signal of user $b(l)$ on subcarrier k can be represented by

$$y_{k,b(l)} = h_{k,b(l)} x_k + v \tag{4}$$

$$= h_{k,b(l)} \sum_{l'=1}^{|\xi_k|} \sqrt{p_{k,b(l')}} s_{k,b(l')} + v, \ \forall l \in \{1, 2, \cdots, |\xi_k|\}, \tag{5}$$

where v denotes the AWGN, i.e., $v \sim CN(0, \sigma^2)$, and σ^2 denotes the noise power.

On subcarrier k, the scheduled users in ξ_k will perform SIC to eliminate inter-user interference. Similar to the case of two-user NOMA, the optimal SIC decoding order is in the descending channel gain order, i.e., $\{b(1), b(2), \cdots, b(|\xi_k|)\}$. It means that the user $b(l)$ first decodes and subtracts the message $s_{k,b(l')}$, $\forall l' > l$, in descending order from $|\xi_k|$ to $l + 1$, and then decodes its own message $s_{k,n(l)}$ by treating $s_{k,n(l')}$, $\forall l' < l$, as interference.

III. PERFORMANCE AND KEY FEATURES OF NOMA

In this section, we first present the performance characteristics of NOMA in existing works, and then discuss the pros and cons of NOMA schemes.

A. Performance of NOMA

It has been shown that NOMA offers considerable performance gain over OMA in terms of spectral efficiency and outage probability [25]–[27], [29]–[31]. Initially, the performance of NOMA was evaluated through simulations given perfect CSI by utilizing the proportional fairness scheduler [25], [29], fractional transmission power allocation (FTPA) [25], and tree-search based

transmission power allocation (TTPA) [30]. These works showed that the overall cell throughput, cell-edge user throughput, and the degrees of proportional fairness achieved by NOMA are all superior to those of OMA. In [27], the author analyzed a two-user SC-NOMA system under statistical CSI from an information theoretic perspective, where it proved that NOMA outperforms native TDMA with high probability in terms of both the sum rate and individual rates. In [26], for a fixed power allocation, the performance of a multiuser SC-NOMA system in terms of outage probability and ergodic sum rates under statistical CSI was investigated in a cellular downlink scenario with randomly deployed users. With the proposed asymptotic analysis, it showed that user n experiences a diversity gain of n and NOMA is asymptotically equivalent to the opportunistic multiple access technique. Furthermore, the authors in [32] analyzed the performance degradation of a multiuser SC-NOMA system on outage probability and average sum rates due to partial CSI. It showed that NOMA based on second order statistical CSI always achieves a better performance than that of NOMA based on imperfect CSI, while it can achieve similar performance to the NOMA with perfect CSI in the low SNR region.

In summary, most of the existing works on performance analysis of NOMA focused on a SC-NOMA system since the user scheduling in MC-NOMA complicates the analysis due to its combinatorial nature. A remarkable work in [31] characterized the impact of user pairing on the performance of a two-user SC-NOMA system with fixed power allocation and cognitive radio inspired power allocation, respectively. The authors proved that, for fixed power allocation, the performance gain of NOMA over OMA increases when the difference in channel gains between the paired users becomes larger. However, further exploration on performance analysis of MC-NOMA system should be carried out in the future since user scheduling is critical for performance of NOMA.

B. Pros

1) Higher spectral efficiency: By exploiting the power domain for user multiplexing, NOMA systems are able to accommodate more users to cope with system overload. In contrast to allocate a subcarrier exclusively to a single user in OMA scheme, NOMA can utilize the spectrum more efficiently by admitting strong users into the subcarriers occupied by weak users without compromising much their performance via utilizing appropriate power allocation and SIC techniques.

2) Better utilization of heterogeneity of channel conditions: As we mentioned before, NOMA schemes intentionally multiplex strong users with weak users to exploit the heterogeneity

of channel condition. Therefore, the performance gain of NOMA over OMA is larger when channel gains of the multiplexed users become more distinctive [31].

3) Enhanced user fairness: By relaxing the orthogonal constraint of OMA, NOMA enables a more flexible management of radio resources and offers an efficient way to enhance user fairness via appropriate resource allocation [23].

4) Applicability to diverse QoS requirements: NOMA is able to accommodate more users with different types of QoS requests on the same subcarrier. Therefore, NOMA is a good candidate to support IoT which connects a great number of devices and sensors requiring distinctive targeted rates.

C. Cons

1) The BS needs to know the perfect channel state information (CSI) to arrange the SIC decoding order, which increases the CSI feedback overhead.

2) The SIC process introduces a higher computational complexity and delay at the receiver side, especially for multicarrier and multiuser systems.

3) The strong users have to know the power allocation of the weaker users in order to perform SIC, which also increases the system signalling overhead.

4) Allocating more power to the weak users, who are generally in the cell-edge, will introduce more inter-cell interferences into the whole system.

IV. DESIGN OF NOMA SCHEMES

Due to the remarkable performance gain of NOMA over conventional OMA, a lot of works on design of NOMA schemes have been proposed in literatures. In this section, we present the existing works on resource allocation of NOMA and MIMO-NOMA, and then briefly introduce other works associated with NOMA.

A. Resource Allocation

Resource allocation has received significant attention since it is critical to improve the performance of NOMA. However, optimal resource allocation is very challenging for MC-NOMA systems, since user scheduling and power allocation couple with each other severely. Some initial works on resource allocation in [25], [29], [30] have been reported, but they are far from optimal. In [33], [34], the authors studied a two-user MC-NOMA system by minimizing the number of subcarriers assigned under the constraints of maximum allowed transmit power and requested data rates, and further introduced a hybrid orthogonal-nonorthogonal scheme.

Furthermore, the authors in [21] studied a joint power and subcarrier allocation problem for a two-user MC-NOMA system. They proposed an optimal scheme and a suboptimal scheme with close-to-optimal performance based on monotonic optimization and difference of convex function programming, respectively.

Besides, there are also several works on resource allocation for multiuser MC-NOMA systems. In [35], the authors formulated the resource allocation problem to maximize the sum rate, which is a non-convex optimization problem due to the binary constraint and the existence of the interference term in the objective function. Interestingly, they proposed a suboptimal solution by employing matching theory and water-filling power allocation. In [36], the authors presented a systematic approach for NOMA resource allocation from a mathematical optimization point of view. They formulated the joint power and channel allocation problem of a downlink multiuser MC-NOMA system, and proved its NP-hardness based on [37] via defining a special user. Furthermore, they proposed a competitive suboptimal algorithm based on Lagrangian duality and dynamic programming, which significantly outperforms OFDMA as well as NOMA with FTPA.

Most of works aforementioned focus on the optimal resource allocation for maximizing the sum rate. However, fairness is another objective to optimize for resource allocation of NOMA. Proportional fairness (PF) has been adopted as a metric to balance the transmission efficiency and user fairness in many works [38], [39]. In [40], the authors proposed a user pairing and power allocation scheme for downlink two-user MC-NOMA based on the PF objective. A prerequisite for user pairing was given and a closed-form optimal solution for power allocation was derived. Apart from PF, max-min or min-max methods are usually adopted to achieve user fairness. Given a preset user group, the authors in [23] studied the power allocation problem from a fairness standpoint by maximizing the minimum achievable user rate with instantaneous CSI and minimizing the maximum outage probability with average CSI. Although the resulting problems are non-convex, simple low-complexity algorithms were developed to provide close-to-optimal solutions. Similarly, another paper [41] studied the outage balancing problem of a downlink multiuser MC-NOMA system to maximize the minimum weighted success probability with and without user grouping. Joint power allocation and decoding order selection solutions were given, and the inter-group power and resource allocation solutions were also provided.

In summary, many existing works focus on the resource allocation for NOMA systems under perfect CSI at the transmitter side. However, there are only few works on the joint user scheduling and power allocation problem for MC-NOMA systems under imperfect CSI, not to mention the
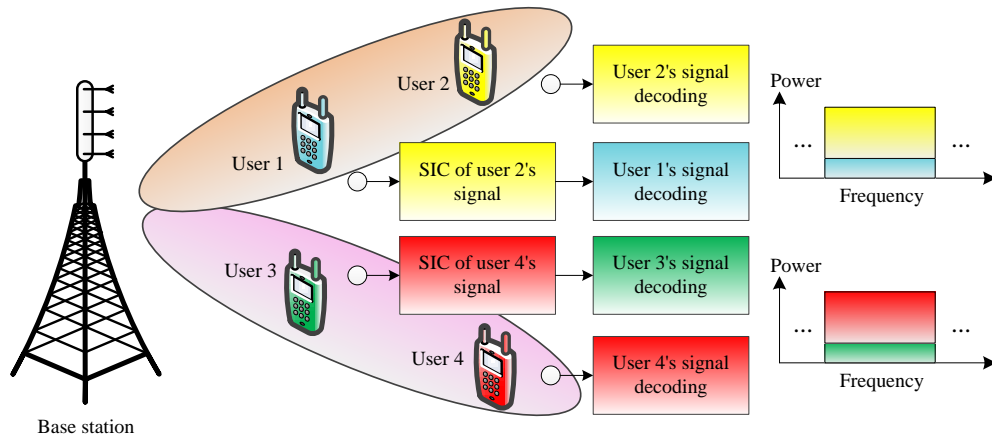
Fig. 5. A downlink MIMO-NOMA model with one base station and four users.

SIC decoding order selection problem. In fact, under imperfect CSI, the SIC decoding order cannot be determined by channel gain order, and some other metrics, such as distance, priority, and target rates, are potential criteria to decide the SIC decoding order.

B. MIMO-NOMA

The application of MIMO techniques to NOMA systems is important to enhance the performance gains of NOMA. Therefore, MIMO-NOMA is another hot topic that has been researched, where the BS and users are equipped with multiple antennas, and multiple users in the same beam are multiplexed on power domain. Fig. 5 illustrates a simple MIMO-NOMA system with one base station and four users. Initially, the concept of MIMO-NOMA was proposed in [30], [42], [43], which demonstrated that MIMO-NOMA outperforms conventional MIMO OMA. The authors in [44] proposed a two-user MIMO-NOMA scheme with a clustering and power allocation algorithm, where the correlation and channel gain difference were taken into consideration to reduce intra-beam interference and inter-beam interference simultaneously. In [45], the authors proposed a minimum power multicast beamforming scheme and applied to two-user NOMA systems for multi-resolution broadcasting. The proposed two-stage beamforming method outperforms the zero-forcing beamforming scheme in [44].

The design of precoding and detection algorithms also received considerable attention since they are the key to eliminate or reduce the inter-cluster interferences. The authors in [46] studied the ergodic sum capacity maximization problem of a two-user MIMO-NOMA system under statistical CSI with the total power constraint and minimum rate constraint for the weak user. This paper derived the optimal input covariance matrix, and proposed the optimal power allocation

scheme as well as a low complexity suboptimal solution. Furthermore, in [47], the authors studied the sum rate optimization problem of two-user MIMO-NOMA under perfect CSI with the same constraints, while different precoders were assigned to different users. The optimal precode covariance matrix was derived by utilizing the duality between uplink and downlink, and a low complexity suboptimal solution based on singular value decomposition (SVD) was also provided. In [48], the authors proposed a new design of precoding and detection matrices for a downlink multiuser MIMO-NOMA system, then analyzed the impact of user pairing as well as power allocation on the sum rate and outage probability of MIMO-NOMA system. Furthermore, in [49], a transmission framework based on signal alignment was proposed for downlink and uplink two-user MIMO-NOMA systems. The authors in [20] studied the sum rate maximization problem of a downlink multiuser multiple-input single-output (MISO) NOMA system. They showed that MISO NOMA transmission outperforms conventional OMA schemes, particularly when the transmit SNR is low, and the number of users is greater than the number of BS antennas. Recently, a multiuser MIMO-NOMA scheme based on limited feedback was proposed and analyzed in [50].

In summary, most of the existing works on MIMO-NOMA focused on design of precoding and detection algorithms, and their performance analyses. However, user scheduling and power allocation were rarely discussed in the spatial domain, which play important roles to improve the spatial efficiency of MIMO-NOMA.

C. Other Works on NOMA

In addition to the above two aspects, there are many other works associated NOMA. We will not discuss further in detail due to the limited space. Compared to downlink NOMA, uplink NOMA was also studied in several works [51]–[57]. Moreover, asynchronous NOMA has also been investigated in uplink scenarios [58], [59]. Cooperative NOMA, where strong users serve as relays for weak users, was studied in [60], [61]. In addition, several works on NOMA combined with other techniques were also reported, such as energy harvesting [62], [63], cognitive radio networks [64], visible light communication [65], and physical layer security [66].

V. RESEARCH CHALLENGES

As discussed above, NOMA can be employed to improve the spectral efficiency, user fairness, as well as to support massive connections with diverse QoS requirements. Based on our overview of existing works on NOMA and its potential applications in practical systems, we present the research challenges of NOMA in the following three aspects.

A. Resource Allocation under Imperfect CSI

Most of existing works on resource allocation of NOMA are based on the assumption of perfect CSI at the transmitter side, which is difficult to obtain in practice due to either the estimation error or the feedback delay. Therefore, it is nature to investigate how CSI error affects the performance of NOMA and to consider robust resource allocation under imperfect CSI [67]. Since NOMA is expected to offer lower latency in order to support delay-sensitive applications in 5G, one promising solution is the outage-based robust approach for designing the resource allocation of NOMA. In this direction, the SIC decoding order under imperfect CSI is still an open problem. Furthermore, it is important to study the joint optimization of power allocation, user scheduling and SIC decoding order selection of NOMA under imperfect CSI.

B. Cooperative NOMA

A key feature of NOMA is that the strong users have prior information of the weak users, which has not been fully exploited in existing works. In cooperative NOMA, the strong users can serve as relays for the weak users, which has the potential to utilize the spatial DoF even for users with a single antenna. Some preliminary works showed that cooperative NOMA can achieve the maximum diversity gain for all the users [60], [61]. It is important to study the optimal resource allocation for cooperative NOMA. Besides, distributed beamforming can be employed in cooperative NOMA to harvest the spatial DoF without much signalling overhead. Considering that cooperative NOMA will introduce more complexity and extra delay into systems, it is important to investigate the tradeoffs among the system performance, complexity, and delay.

C. QoS-based NOMA

As we mentioned before, NOMA has great potential to support diverse QoS requirements. The heterogeneity of QoS requirements might in turn facilitate the power allocation and user scheduling of NOMA, which is also an interesting topic to explore in the future. For example, users in NOMA systems can be categorized according to their QoS requirements, instead of their channel conditions, which offers two following benefits. One is that the SIC decoding order, power allocation, and user scheduling can be designed more appropriately to meet the users QoS requests. The other is to make NOMA communications more general, e.g., applicable to scenarios in which users channel conditions are the same.

VI. Conclusion

In this article, a promising multiple access technology for 5G networks, NOMA, is discussed. A two-user SC-NOMA scheme and a multiuser MC-NOMA scheme were presented and discussed to illustrate the basic concepts and principles of NOMA. A literature review about performance analyses of NOMA, resource allocation for NOMA, and MIMO-NOMA were discussed. Furthermore, we presented the key features and potential research challenges of NOMA.

References

[1] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[2] G. Wunder, P. Jung, M. Kasparick, T. Wild, F. Schaich, Y. Chen, S. Brink, I. Gaspar, N. Michailow, A. Festag, L. Mendes, N. Cassiau, D. Ktenas, M. Dryjanski, S. Pietrzyk, B. Eged, P. Vago, and F. Wiedmann, "5GNOW: non-orthogonal, asynchronous waveforms for future mobile applications," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 97–105, Feb. 2014.

[3] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[4] J. Zhu, D. W. K. Ng, N. Wang, R. Schober, and V. K. Bhargava, "Analysis and design of secure massive MIMO systems in the presence of hardware impairments," *arXiv preprint arXiv:1602.08534*, 2016.

[5] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, June 2011.

[6] T. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. Wong, J. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[7] J. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. Reed, "Femtocells: Past, Present, and Future," *IEEE J. Select. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.

[8] J. Andrews, "Seven ways that HetNets are a cellular paradigm shift," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, Mar. 2013.

[9] D. Ramasamy, R. Ganti, and U. Madhow, "On the capacity of picocellular networks," in *Proc. IEEE Intern. Sympos. on Inf. Theory*, July 2013, pp. 241–245.

[10] D. W. K. Ng, E. S. Lo, and R. Schober, "Robust beamforming for secure communication in systems with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4599–4615, Aug. 2014.

[11] ——, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3292–3304, Sept. 2012.

[12] ——, "Wireless information and power transfer: Energy efficiency optimization in OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 6352–6370, Dec. 2013.

[13] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sept. 2015.

[14] "Study on downlink multiuser supersition transmission (MUST) for LTE (Release 13)," 3GPP TR 36.859, Tech. Rep., Dec. 2015.

[15] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1616–1626, Apr. 2008.

[16] R. Hoshyar, R. Razavi, and M. Al-Imari, "LDS-OFDM an efficient multiple access technique," in *Proc. IEEE Veh. Techn. Conf.*, May 2010, pp. 1–5.

[17] R. Razavi, M. Al-Imari, M. A. Imran, R. Hoshyar, and D. Chen, "On receiver design for uplink low density signature OFDM (LDS-OFDM)," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3499–3508, Nov. 2012.

[18] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE Personal, Indoor and Mobile Radio Commun. Sympos.*, Sept. 2013, pp. 332–336.

[19] X. Dai, S. Chen, S. Sun, S. Kang, Y. Wang, Z. Shen, and J. Xu, "Successive interference cancelation amenable multiple access (SAMA) for future wireless communications," in *Proc. IEEE Intern. Commun. Conf.*, Nov. 2014, pp. 222–226.

[20] M. F. Hanif, Z. Ding, T. Ratnarajah, and G. K. Karagiannidis, "A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 76–88, Jan. 2016.

[21] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for MC-NOMA systems," *arXiv preprint arXiv:1603.08132*, 2016.

[22] Q. Sun, S. Han, C.-L. I, and Z. Pan, "Energy efficiency optimization for fading MIMO non-orthogonal multiple access systems," in *Proc. IEEE Intern. Commun. Conf.*, June 2015, pp. 2668–2673.

[23] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1647–1651, Oct. 2015.

[24] A. Benjebbour, A. Li, Y. Saito, Y. Kishiyama, A. Harada, and T. Nakamura, "System-level performance of downlink NOMA for future LTE enhancements," in *Proc. IEEE Global Commun. Conf.*, Dec. 2013, pp. 66–70.

[25] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE Personal, Indoor and Mobile Radio Commun. Sympos.*, Sept. 2013, pp. 611–615.

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

[27] P. Xu, Z. Ding, X. Dai, and H. V. Poor, "A new evaluation criterion for non-orthogonal multiple access in 5G software defined networks," *IEEE Access*, vol. 3, pp. 1633–1639, Sept. 2015.

[28] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[29] N. Otao, Y. Kishiyama, and K. Higuchi, "Performance of non-orthogonal access with SIC in cellular downlink using proportional fair-based resource allocation," in *Proc. IEEE Intern. Sympos. on Wireless Commun. Systems*, Aug. 2012, pp. 476–480.

[30] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE Veh. Techn. Conf.*, June 2013, pp. 1–5.

[31] Z. Ding, P. Fan, and V. Poor, "Impact of user pairing on 5G non-orthogonal multiple access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1–1, 2015.

[32] Z. Yang, Z. Ding, P. Fan, and G. K. Karagiannidis, "On the performance of non-orthogonal multiple access systems with partial channel information," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 654–667, Feb. 2016.

[33] M. R. Hojeij, J. Farah, C. Nour, and C. Douillard, "Resource allocation in downlink non-orthogonal multiple access (NOMA) for future radio access," in *Proc. IEEE Veh. Techn. Conf.*, May 2015, pp. 1–6.

[34] ——, "New optimal and suboptimal resource allocation techniques for downlink non-orthogonal multiple access," *Wireless Personal Commun.*, vol. 87, no. 3, pp. 837–867, May 2015.

[35] B. Di, S. Bayat, L. Song, , and Y. Li, "Radio resource allocation for downlink non-orthogonal multiple access (NOMA) networks using matching theory," in *Proc. IEEE Global Commun. Conf.*, Dec. 2015, pp. 1–6.

[36] L. Lei, D. Yuan, C. K. Ho, and S. Sun, "Joint optimization of power and channel allocation with non-orthogonal multiple access for 5G cellular systems," in *Proc. IEEE Global Commun. Conf.*, Dec. 2015, pp. 1–6.

[37] Y.-F. Liu and Y.-H. Dai, "On the complexity of joint subcarrier and power allocation for multi-user OFDMA systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 583–596, Feb. 2014.

[38] H. Kim, K. Kim, Y. Han, and S. Yun, "A proportional fair scheduling for multicarrier transmission systems," in *Proc. IEEE Veh. Techn. Conf.*, vol. 1, Sept 2004, pp. 409–413.

[39] C. Wengerter, J. Ohlhorst, and A. G. E. von Elbwart, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA," in *Proc. IEEE Veh. Techn. Conf.*, vol. 3, May 2005, pp. 1903–1907.

[40] F. Liu, P. Mahonen, and M. Petrova, "Proportional fairness-based user pairing and power allocation for non-orthogonal multiple access," in *Proc. IEEE Personal, Indoor and Mobile Radio Commun. Sympos.*, Aug. 2015, pp. 1127–1131.

[41] S. Shi, L. Yang, and H. Zhu, "Outage balancing in downlink non-orthogonal multiple access with statistical channel state information," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, Mar. 2016.

[42] Y. Lan, A. Benjebboiu, X. Chen, A. Li, and H. Jiang, "Considerations on downlink non-orthogonal multiple access (NOMA) combined with closed-loop SU-MIMO," in *Proc. IEEE Intern. Conf. on Signal Process. and Commun. Syst.*, Dec. 2014, pp. 1–5.

[43] X. Chen, A. Benjebbour, Y. Lan, A. Li, and H. Jiang, "Impact of rank optimization on downlink non-orthogonal multiple access (NOMA) with SU-MIMO," in *Proc. IEEE Intern. Conf. on Commun. Systems*, Nov. 2014, pp. 233–237.

[44] B. Kim, S. Lim, H. Kim, S. Suh, J. Kwun, S. Choi, C. Lee, S. Lee, and D. Hong, "Non-orthogonal multiple access in a downlink multiuser beamforming system," in *Proc. IEEE Mil. Commun. Conf.*, Nov. 2013, pp. 1278–1283.

[45] J. Choi, "Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 791–800, Mar. 2015.

[46] Q. Sun, S. Han, C.-L. I, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 405–408, Aug. 2015.

[47] Q. Sun, S. Han, Z. Xu, S. Wang, I. Chih-Lin, and Z. Pan, "Sum rate optimization for MIMO non-orthogonal multiple access systems," in *Proc. IEEE Wireless Commun. and Networking Conf.*, Mar. 2015, pp. 747–752.

[48] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 537–552, Jan. 2016.

[49] Z. Ding, R. Schober, and H. V. Poor, "A general MIMO framework for NOMA downlink and uplink transmission based on signal alignment," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4438–4454, June 2016.

[50] Z. Ding and H. V. Poor, "Design of massive-MIMO-NOMA with limited feedback," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 629–633, May 2016.

[51] T. Takeda and K. Higuchi, "Enhanced user fairness using non-orthogonal access with SIC in cellular uplink," in *Proc. IEEE Veh. Techn. Conf.*, Sept. 2011, pp. 1–5.

[52] M. Al-Imari, P. Xiao, M. A. Imran, and R. Tafazolli, "Uplink non-orthogonal multiple access for 5G wireless networks," in *Proc. IEEE Intern. Sympos. on Wireless Commun. Systems*, Aug. 2014, pp. 781–785.

[53] S. Chen, K. Peng, and H. Jin, "A suboptimal scheme for uplink NOMA in 5G systems," in *Proc. IEEE Intern. Wireless Commun. and Mobile Comput. Conf.*, Aug. 2015, pp. 1429–1434.

[54] Y. Chen, J. Schaepperle, and T. Wild, "Comparing IDMA and NOMA with superimposed pilots based channel estimation in uplink," in *Proc. IEEE Personal, Indoor and Mobile Radio Commun. Sympos.*, Aug. 2015, pp. 89–94.

[55] M. Al-Imari, P. Xiao, and M. A. Imran, "Receiver and resource allocation optimization for uplink NOMA in 5G wireless networks," in *Proc. IEEE Intern. Sympos. on Wireless Commun. Systems*, Aug. 2015.

[56] B. Kim, W. Chung, S. Lim, S. Suh, J. Kwun, S. Choi, and D. Hong, "Uplink NOMA with multi-antenna," in *Proc. IEEE Veh. Techn. Conf.*, May 2015, pp. 1–5.

[57] 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.

[58] H. Haci, H. Zhu, and J. Wang, "A novel interference cancellation technique for non-orthogonal multiple access (NOMA)," in *Proc. IEEE Global Commun. Conf.*, Dec. 2015, pp. 1–6.

[59] H. Haci, "Non-orthogonal multiple access (NOMA) with asynchronous interference cancellation," Ph.D. dissertation, Dept. Electron. Eng., Univ. of Kent, Kent, England, Mar. 2015.

[60] Z. Ding, M. Peng, and H. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, Aug. 2015.

[61] Z. Ding, H. Dai, and H. V. Poor, "Relay selection for cooperative NOMA," *IEEE Wireless Commun. Lett.*, vol. PP, no. 99, pp. 1–1, June 2016.

[62] Y. Liu, Z. Ding, M. Eikashlan, and H. Poor, "Cooperative non-orthogonal multiple access in 5G systems with SWIPT," in *Proc. European Signal Process. Conf.*, Aug. 2015, pp. 1999–2003.

[63] P. D. Diamantoulakis, K. N. Pappi, Z. Ding, and G. K. Karagiannidis, "Optimal design of non-orthogonal multiple access with wireless power transfer," *arXiv preprint arXiv:1511.01291*, 2015.

[64] Y. Liu, Z. Ding, M. Elkashlan, and J. Yuan, "Non-orthogonal multiple access in large-scale underlay cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1–1, Feb. 2016.

[65] H. Marshoud, V. M. Kapinas, G. K. Karagiannidis, and S. Muhaidat, "Non-orthogonal multiple access for visible light communications," *IEEE Photonics Technol. Lett.*, vol. 28, no. 1, pp. 51–54, Jan. 2016.

[66] Y. Zhang, H. M. Wang, Q. Yang, and Z. Ding, "Secrecy sum rate maximization in non-orthogonal multiple access," *IEEE Commun. Lett.*, vol. PP, no. 99, pp. 1–1, Mar. 2016.

[67] Z. Wei, D. W. K. Ng, and J. Yuan, "Power-efficient resource allocation for MC-NOMA with statistical channel state information," *arXiv preprint arXiv:1607.01116*, 2016.