JOURNAL OF
Optical Communications
and Networking

Study on the impact of clustering for non-orthogonal multiple access based on multilevel code for radio-over-fiber fronthaul application

Yijie Tao,* Christina Lim, and Ampalavanapillai Nirmalathas

Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, VIC 3010, Australia
**Corresponding author: yijiet@student.unimelb.edu.au*

Received 17 August 2020; revised 5 November 2020; accepted 8 November 2020; published 8 January 2021 (Doc. ID 405506)

With the stringent requirements of fifth-generation and beyond wireless communications technology coupled with the potential bottleneck in the optical fronthaul, radio-over-fiber (RoF) technology offers an alternative solution to distribute the wireless signals while meeting these requirements. The inherent centralized architecture of RoF enables advanced coordination functionality to be performed within the fronthaul network. Advanced coordination functionality such as non-orthogonal multiple access (NOMA) has gained traction as a means to improve spectral efficiency for larger user-base networks. Among various NOMA implementations, the recently proposed multilevel code (MLC) was found to be superior to the commonly investigated superposition coding successive interference cancellation scheme in a two-user scenario. In this paper, we generalize the definition of the MLC to *n* user cases and discuss two power allocation schemes for the generalized MLC definition, which are based on the spectral efficiency (SE) criterion and symbol error rate (SER) criterion, respectively. Simulation-based performance analysis on the MLC–SE and the MLC–SER schemes was conducted over different cluster sizes and randomly generated user distributions. The results show, to the best of our knowledge, that the best performance in terms of user fairness and bit error rate can be obtained in a two-user MLC–SER scheme.

https://doi.org/10.1364/JOCN.405506

1. INTRODUCTION

To meet the stringent requirements of the fifth-generation (5G) and beyond mobile communication systems to provide unprecedented data rates while connecting to an increasing number of devices, there is a strong desire to shift the wireless operating frequency from the overcrowded lower microwave regions to millimeter-wave frequency bands. One of the consequences of this migration is an increased number of smaller antenna cell sites because of the higher propagation losses of wireless signals in this band [1]. As a result, very densely packed small cells would have to be deployed to maintain the desired geographical coverage. In addition, a significant amount of additional load will be created on the transport networks to deliver data to and from the expanding number of cells. To address this issue, a centralized/cloud radio access network (C-RAN) has been proposed that is capable of servicing a large number of remote radio heads (RRHs) via a centrally located baseband unit (BBU) pool connected by the fronthaul network, as shown in Fig. 1 [2]. In this architecture, the centralized BBU pool facilitates network management, a cooperative processing capability, and resource sharing.

Common Public Radio Interface (CPRI) is widely used in the fronthaul that uses an uncompressed digitization technique to digitize the wireless signals supporting up to 24 Gb/s [3]. Due to the stringent requirements of the 5G and beyond wireless specification coupled with an increased number of cell sites, the CPRI link does not scale well with the total optical bit rate that is expected to increase exponentially with the wireless channels and bandwidth [4]. One approach to overcome this bottleneck is to transport the wireless signals in an analog format using radio-over-fiber (RoF) technology over the fronthaul links. Analog-based RoF fronthaul also enables a simpler RRH design [5] while supporting a centralized control capability with minimal latency. The centralized control capability enables joint processing and advanced coordination functionalities such as coordinated multipoint (CoMP), non-orthogonal multiple access (NOMA), and multiple-input multiple-output (MIMO) to be performed within the fronthaul network [2,6].

NOMA is introduced to improve the spectral efficiency requirements of a larger user-base network. In conventional orthogonal multiple access (OMA) schemes, users are allocated orthogonal resources in the time/frequency/code domain so

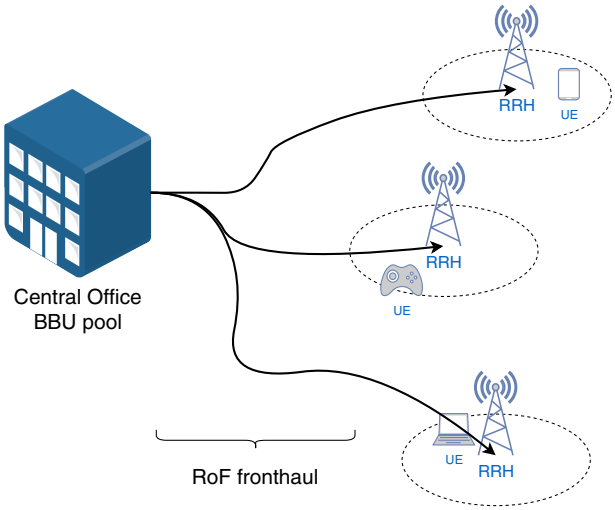

Fig. 1. Illustration of the fronthaul in C-RAN.

that the number of connected users is limited by the number of orthogonal resources [7–9]. To overcome this limitation, the NOMA scheme has been proposed and actively investigated with the main aim to improve the performance and spectral efficiency of a large user base, which may be limited by OMA [8]. In contrast to OMA, NOMA multiplexes multiple users in the power domain at the transmitter to improve the spectral efficiency (SE). The most commonly investigated NOMA scheme in the literature is based on a superposition code with a successive interference cancellation scheme (SPC–SIC) to support multiple users via interuser interference cancellation [10,11]. Despite its popularity, the performance of the SPC–SIC scheme is limited by error propagation and further compounded by the increased receiver complexities and latency.

Recently, we proposed what we believe, to the best of our knowledge, is a new NOMA scheme based on the multilevel code (MLC) and experimentally demonstrated for 60 GHz mobile fronthaul to address the issues imposed by the SPC–SIC [12]. The results showed that the MLC scheme was able to eliminate the error propagation issue in the SPC–SIC and performed better than the SPC–SIC scheme.

However, the study in [12] was limited to only a two-user clustering scenario based on a RoF fronthaul, and the general benefits of the MLC over the SPC–SIC in clustering situations with less than two users remains uninvestigated. Learning how MLC behaves in such situations is vital to assess its performance in a practical NOMA-based RoF fronthaul application, because an operational NOMA is envisioned to have the advantage of dynamically allocating frequency resources and clustering users according to its channel state information (CSI) [13]. Hence, this motivates the study in this paper to further investigate the proposed MLC scheme to generalize the scheme to multiuser clustering scenarios, and study how the fronthaul performance varies with the number of users per cluster and how it compares to the SPC–SIC in such situations in a simulated environment.

The paper has four sections. In Section 2, the definition of the MLC for a generalized multiuser clustering will be provided, and the theoretical derivations for two optimal power allocation schemes will be discussed. Section 3 covers the simulation setup and the discussion of the results. Conclusions are summarized in Section 4.

2. THEORY

A. Principle of MLC for Multiuser Clustering

In this section, the principle of the MLC scheme will be presented and the developed algorithm to generate the constellation map for such a scheme will be provided.

The idea of the MLC is to integrate symbols for all users within a served cluster to one composite symbol on a conventional gray-coded quadrature amplitude modulation (QAM) constellation map and, as such, eliminate error propagation and receiver complexity by only using conventional QAM demodulation [12]. At the transmitter end, the symbols of each user are first bit interleaved to form one composite symbol on the QAM constellation map and then broadcast to all users within the cluster. At the receiver end, all users perform only one QAM demodulation to obtain the composite symbol, from which they then extract their designated symbol based on the position of the bits within the composite symbol. For this study, the scenario is limited to two bits per user per transmitted symbol.

The implementation was carried out in MATLAB, where the bits for each user are interleaved together to generate a composite symbol. To formulate this process, the following parameters are defined:

- User $n \subseteq \{1, \ldots, N\}$, where N is the total number of users, user 1 is the nearest user from the RRH, and user N is the furthest in a cluster.
- Transmitted bits for user n: $b_{1,n}b_{2,n}$, where $b_{1,n} \subseteq \{0, 1\}$, $b_{2,n} \subseteq \{0, 1\}$.
- Composite symbol S.
- Bit-interleaving function $f(\cdot)$.

Here the bits are treated as decimal integers, and the bit-interleaving function is defined as

$$S = f(b_{1,1}b_{2,1}, b_{1,2}b_{2,2}, \ldots, b_{1,N}b_{2,N}) = \sum_{m=1}^{N} 2^{2(N-m)+1} b_{1,m} + 2^{N-m} b_{2,m}. \quad (1)$$

The resulting composite symbol after interleaving in binary is given as

$$S = b_{1,1}b_{1,2} \ldots b_{1,N}b_{2,1}b_{2,2} \ldots b_{2,N}.$$

This symbol is then mapped onto a MLC constellation map for a corresponding complex signal. The mapping is achieved via a MLC algorithm. For this algorithm, we define the following parameters:

- Number of users/levels in the cluster N; e.g., the level for user 1 is the 1st level, the next level is for user 2, and so forth.

- Decimal integer representation i_n for the composite symbol for user n.
- Power for current level P_n.
- Current quadrant q.
- Constellation position c.
- Quadrant mapping matrix A, which is generated by extracting the gray-coding pattern with respect to the defined MLC levels. In matrix A shown below, the row number corresponds to the quadrant that symbol "00" is in, which in turns represents the symbol pattern in the current level. Each column number corresponds to a constellation symbol, and the index on that column indicates in which quadrant the corresponding symbol is located.

$$A = \begin{bmatrix} 1 & 4 & 2 & 3 \\ 2 & 3 & 1 & 4 \\ 3 & 2 & 4 & 1 \\ 4 & 1 & 3 & 2 \end{bmatrix}.$$

To illustrate the mapping procedure, an example is given for a three-user cluster, as shown in Fig. 2. Here, User 1 is located the furthest away from the RRH and will transmit "01"; User 2 is located 2nd furthest from the RRH and will transmit "10"; and User 3 is located closest to the RRH and will transmit "11." Based on the convention used in the mapping, User 1 is considered as Level 1, followed by User 2 at Level 2, and finally User 3 at Level 3. In this example, "00" is initialized to be in quadrant 3 ($q = 3$). The location of the composite symbol is mapped using the MLC constellation mapping algorithm together with matrix A. The row number of matrix A represents which quadrant "00" locates in on the current level and the symbol pattern on this level. The column number determines the location of a specific transmitting symbol on the current level given the current row number (the current value of q) and the symbol pattern going into the next level. For instance, User 1 transmitting bits of "01" determines that the composite symbol should be located within quadrant 2 of the constellation on Level 1 (obtained by using initial $q = 3$ (row 3 of matrix A) and 2nd column (symbol "01" is considered the 2nd symbol out of the predefined symbols sequence [0,1,10,11] via the equation in Algorithm 1, line 4) of row 3, which gives quadrant 2. This indicates that the composite symbol of the three-user cluster is located within quadrant 2 in Level 1 (as shown in Fig. 2) and q is updated to 2 as the algorithm goes into the next level. The bits transmitted by User 2 ("10") determine the composite symbol location on Level 2. To locate this, using the current quadrant q, which is already updated to 2, and the corresponding quadrant for Level 2 is obtained to be in quadrant 1 (row 2 that corresponds to $q = 2$, 3rd column that corresponds to the location of "10" in the predefined symbols sequence [0,1,10,11]). To find the final location of the composite symbol that will be determined by User 3, the current quadrant q is updated to 1, and the final location of the composite symbol is located in quadrant 3 (1st row, 4th column of matrix A as defined by "11," which is the fourth symbol within the predefined symbol sequence) for Level 3, as indicated in Fig. 2. As a result, we find the composite data "011101" located at the solid dotted constellation position. This procedure ensures gray coding while each level is associated with the transmitted data for the corresponding user. In addition, the constellation point position within each level is assigned a power, denoted P_1, P_2, and P_3, respectively. By adjusting these powers, each level is given a certain extent of robustness against noise. Thus, power can be allocated to achieve relatively even performance across users.

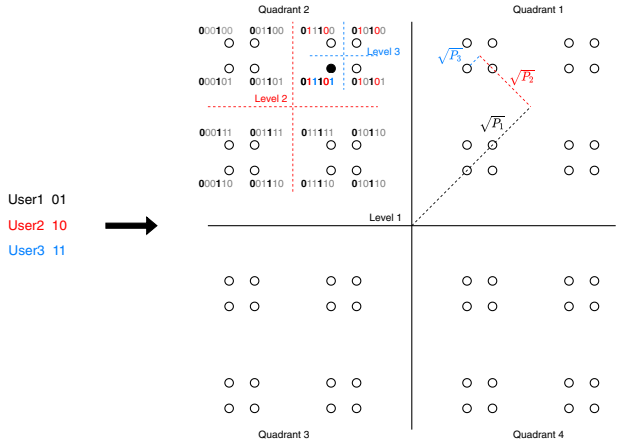

Fig. 2. Illustration of the MLC scheme for three-user clustering. The standard quadrant convention in the literature is followed where Quadrant 1 is located in the top right hand corner and 2, 3, and 4 are sequentially placed in the counterclockwise direction.

Algorithm 1. Algorithm for MLC Constellation Map Generation

1: $c = 0$.
2: $q = 3$ ▷q is chosen to be 3 for the constellation pattern shown in Fig. 2, where "01" is in the 2nd quadrant on level 1.
3: **For** n in N levels **do**.
4: $i_n = 2b_{1,n} + b_{2,n} + 1$.
5: $q = A(q, i_n)$.
6: $c = c + P_n e^{j(-\frac{\pi}{4} + \frac{q\pi}{2})}$.
7: **Return** c.

B. Optimal Power Allocation Schemes

As mentioned earlier, we need to determine the power for each user because it maximizes the user's performance. This can be explicitly established using an optimal power allocation scheme for the MLC. Here, we propose two power allocation schemes based on the SE criterion and symbol error rate (SER) criterion. We denote the two MLCs with their respective power allocating schemes as the MLC–SE and MLC–SER.

1. SE Criterion

A common scheme to maximize fairness for all users is the max–min fairness (MMF) scheme [14], which allocates power to a cluster of users such that the lowest spectral efficiency within the cluster is maximized. In other words, this scheme aims to improve the lower bound of a user cluster's performance. This is formulated as

$$\max \min_{n=1,\ldots,N} \{R_n\},$$

where R_n denotes the nth user's spectral efficiency, and N is the number of users within a cluster.

Assuming that we have a fixed total transmitted power P_t, and we are only concerned with the optimization within a cluster, the MMF is equivalent to finding a power ratio such that all users have the same transmission rate/spectral efficiency [14]. Since the signal is independently demodulated at each receiver, and the relevant bits for each user are independently extracted, the SE for each user is only a function of the transmitted power for that user, the channel gain, and the noise floor of that user's receiver (transmitter noise is ignored here). The SE for the nth user can then be expressed as

$$R_n = \log_2\left(1 + \frac{Gn P_n}{N_n}\right), \tag{2}$$

where G_n is the channel gain for the nth user, P_n is the transmitted power for the nth user, and N_n is the noise floor of the nth user's receiver [10]. Given a set of known channel gains of $\{G_{n=1,\_,N}\}$, we can readily write down the equation set for the MMF as

$$\log_2\left(1 + \frac{G_1 P_1}{N_1}\right) = \log_2\left(1 + \frac{G_2 P_2}{N_2}\right) = \cdots = \log_2\left(1 + \frac{G_N P_N}{N_N}\right). \tag{3}$$

Further assuming an equal noise floor for all users, the above equation set can be reduced to

$$G_1 P_1 = G_2 P_2 = \cdots = G_N P_N. \tag{4}$$

Using this equation set in conjunction with the total fixed transmitted power, we can determine the power for all users based on the SE criterion. If we denote the power ratio between two adjacent user as $r_n = \frac{P_n}{P_{n+1}}$, we have

$$\begin{cases} r_1 = \frac{P_1}{P_2} = \frac{G_2}{G_1} \\ \cdots \\ r_{N-1} = \frac{P_{N-1}}{P_N} = \frac{G_N}{G_{N-1}} \\ \sum_{n=1}^{N} P_n = P_t \end{cases}. \tag{5}$$

Then the power for each user P_n can be obtained as

$$\begin{aligned} P_n &= P_t \frac{P_n}{P_t} \\ &= P_t \frac{r_n r_{n+1} \ldots r_{N-1} P_N}{P_t} \\ &= P_t \frac{\left(\prod_{m=n}^{N-1} r_m\right) P_N}{\left(1 + \sum_{l=1}^{N-1} \prod_{o=l}^{N-1} r_o\right) P_N} \\ &= P_t \frac{\prod_{m=n}^{N-1} r_m}{1 + \sum_{l=1}^{N-1} \prod_{o=l}^{N-1} r_o}. \end{aligned} \tag{6}$$

2. SER Criterion

The second proposed power allocation scheme is based on maximizing user fairness. This scheme equalizes performance across all users by equalizing each user's SER. Specifically,

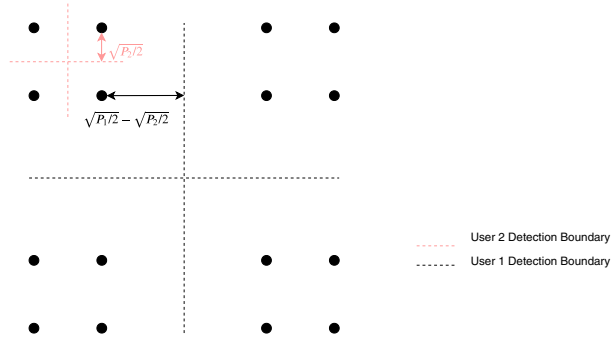

Fig. 3. Illustration of the MLC–SER criterion for two users.

for an additive white Gaussian noise (AWGN) maximum-likelihood (ML) detector, the symbol error rate is determined by the distance between symbols and its detection boundary [15]. Therefore, the idea for this scheme is to equalize the distance between the corresponding user symbols and the ML detection boundary after taking into account the relevant user channel gain. For illustration purposes, the two-user case is shown in Fig. 3.

From basic trigonometry, the distance of a near user symbol to its detection boundary is $\sqrt{\frac{P_2}{2}}$, and the distance of the far user symbol to its detection boundary is $\sqrt{\frac{P_1}{2}} - \sqrt{\frac{P_2}{2}}$. The proposed SER criterion scheme asserts that, to equalize the SER for both users, $\sqrt{G_2}\sqrt{\frac{P_2}{2}} = \sqrt{G_1}\left(\sqrt{\frac{P_1}{2}} - \sqrt{\frac{P_2}{2}}\right)$, with G_1 and G_2 being the far user and near user channel gain, respectively.

Generalizing this to N users, we have

$$\begin{aligned} &\sqrt{G_1}\left(\sqrt{P_1} - \sum_{n=2}^{N}\sqrt{P_n}\right) \\ &= \sqrt{G_2}\left(\sqrt{P_2} - \sum_{m=3}^{N}\sqrt{P_m}\right) \\ &= \cdots \\ &= \sqrt{G_i}\left(\sqrt{P_i} - \sum_{l=i+1}^{N}\sqrt{P_l}\right) \\ &= \cdots \\ &= \sqrt{G_N}\sqrt{P_N}. \end{aligned} \tag{7}$$

With a fixed total power $\sum_{n=1}^{N} P_n = P_t$, we can solve Eq. (7) for the power for each user within the cluster. Denoting the square root of the power ratio between adjacent users as $R_n = \frac{\sqrt{P_n}}{\sqrt{P_{n+1}}}$,

$$R_n = \begin{cases} \sqrt{\frac{G_N}{G_{N-1}}}, & \text{for } n = N-1 \\ 1 + \sqrt{\frac{G_{n+1}}{G_n}} + \left(1 - \sqrt{\frac{G_{n+1}}{G_n}}\right)\sum_{m=n+1}^{N-1}\prod_{l=n+1}^{m}\frac{1}{R_l}, & \text{otherwise} \end{cases}. \tag{8}$$

The power for user n can then be obtained as

$$P_n = P_t\left(\frac{\prod_{m=n}^{N-1} R_m}{1+\sum_{l=1}^{N-1}\prod_{o=l}^{N-1} R_o}\right)^2. \quad (9)$$

3. SIMULATION

A. Assumptions

The following assumptions are made for the simulation.

For the transmission in the fronthaul:

- It is assumed that the losses associated with the optical-to-electrical conversions, optical components, and fiber dispersion were constant. Furthermore, as the received powers at the users were normalized for the analysis, as shown in the simulation parameters in Table 1, these losses did not play a significant role in the investigations. As such, only the transmission from the RRH to the user equipment (UE) is modeled in the simulation.

For the transmission from the RRH to the UE:

- All transmission channels from the RRH to the user equipment are AWGN channels. The noise has zero mean and variance of σ^2. A zero-mean AWGN channel is commonly used to model the noise in communication channels [16].
- The signal operates in the millimeter-wave band. Therefore, the path loss (PL) can be modeled by the general Friis free space path loss (FSPL) model with obstruction loss neglected [17], so

$$\mathrm{PL}(f,r)[\mathrm{dB}] = 20\log_{10}\left(\frac{4\pi f}{c}\right) + (10l)\log_{10}(r),$$

where f is the transmission frequency, c is the speed of light, l is the path loss exponent, and r is the distance between the RRH and the served user. And the channel gain is given by

$$G = \frac{G_{TX}G_{RX}}{PL},$$

where G_{TX} and G_{RX} are the transmitter and receiver antenna gains, respectively.

- In the research literature, various receiver techniques with satisfying results have been proposed for the millimeter-wave band [18,19]. Therefore, an assumption is made that the up- and down-conversion process can be carried out perfectly at the transmitter and receiver. Thus, only a baseband channel is modeled.

Table 1. Table of Parameters Used in Simulation

Power	Per user power budget P	1 W
	Noise power at user equipment σ^2	0.002 W
	Transmitter antenna gains G_{TX}	1000
	Receiver antenna gains G_{RX}	1000
FSPL	Speed of light c	3e8 m/s
	Transmission frequency f	60 GHz
	Path loss exponent l	2.1

- All receivers experience the same noise floor, which is valid under the assumption that the interference from neighboring cells is negligible.
- The signal power is recovered to its transmitted power level at the receiver. This is to simplify the computation steps as explained below and can be done by tuning the receiver gain.

B. Simulation Setup

The simulation setup is illustrated in Fig. 4. A total of 12 user locations are randomly generated in uniform distribution between 0.5 m to 4.5 m away from the RRH transmitter, while the RRH has a range of 10 m in diameter. These users are sorted according to their distance in ascending orders, and then scenarios are simulated where they are clustered in two-, three-, four-, or six-user groups. For clustering of N users, the users that are $\frac{12}{N}-1$ position(s) away on the sorted order will be clustered into one group; e.g., in six-user clustering, users 1, 3, 5, 7, 9, and 11 will be clustered together, while users 2, 4, 6, 8, 10, and 12 will be clustered into the other group, as illustrated in Fig. 5. The users are clustered this way so that the users under the same cluster are always a considerable distance apart. This prevents a situation where one cluster's users may be overcrowded compared to another cluster, which may introduce extra variation to different clusters' performances.

The user data are first used to generate the composite symbols as defined by the MLC-specific bit interleaving scheme, then mapped onto the MLC constellation map via a custom Simulink block. The MLC constellation map has the optimized power allocation incorporated with either the SE criterion or the SER criterion. The channel gain information is considered known in real time and is converted from the user

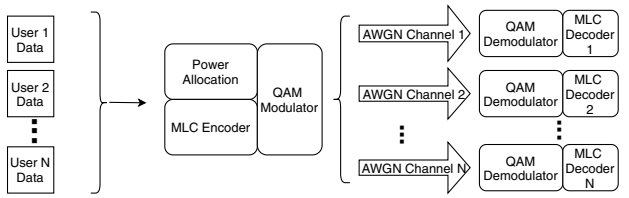

Fig. 4. Illustration of the simulation setup. Data is randomly generated for each user, then sent through the power allocation and MLC algorithm to be combined as a composite symbol. This symbol is then QAM-modulated and transmitted through N parallel AWGN channels, each channel with a noise level that reflects the receiver SNR after the signal experiences the respective path loss.

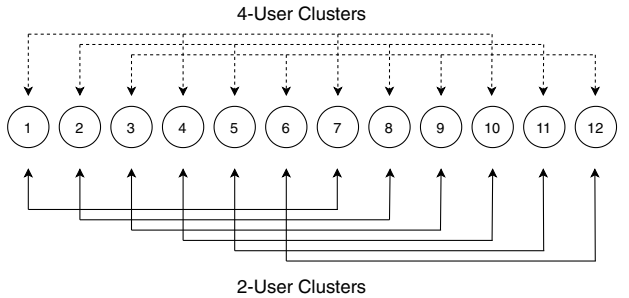

Fig. 5. Illustration of the clustering format of two-user clusters and four-user clusters.

distance via Friis' FSPL equation. The corresponding signal is then generated as a QAM signal. In the AWGN channel, instead of attenuating the transmitted signal, the noise is scaled up to match the targetted signal-to-noise ratio (SNR). This is based on the assumption that the transmitted power level is perfectly recovered at the receiver.

At the receiver, the signal is demodulated as QAM using the same MLC constellation map, and the relevant bits are extracted for each user and compared to the sent bits for the bit error rate (BER) calculation.

For the transmitter power, the power budget toward a cluster is proportional to the user number in the cluster; that is, for an N user cluster, the transmitted power budget is NP, where P is the per-user power budget. Therefore, the SNR before demodulation at the nth user is $\frac{G_n NP}{\sigma^2}$.

At the receiver, because the power is restored to the transmitted level according to the assumption above and all UEs experience the same noise level, the noise is scaled up inverse proportionally to the channel loss. Thus, to obtain the correct SNR at the receiver, the simulation puts the signal through an AWGN lossless channel, with the noise level inverse proportional to the calculated channel loss based on the Friis model above.

C. Results and Discussion

The parameters used in the simulation are presented in Table 1.

The wireless carrier frequency is chosen to be 60 GHz to represent the signals in the millimeter-wave band for the RoF fronthaul system. The path loss exponent l is 2.1, which is empirically obtained for a 60 GHz millimeter-wave signal in [20]. The per-user power budget P is normalized to 1 and the noise level σ^2 is adjusted to 0.002 W with respect to the normalized power budget and antenna gains G_{TX} and G_{RX} to produce an appropriate range of the BER data for demonstration.

The result is produced by running the above simulation setup with the proposed SE criterion and the SER criterion, respectively, over 122 simulation runs with 1e8 samples generated for each run. As a reference to the performance of the MLC, the signal in each run is also transmitted and received using the SPC–SIC with the max–min optimal power allocation scheme proposed in [14]. For each simulation run, the 12 user locations are uniformly redistributed over the specified range. The lined curve represents the mean value at each user over these 122 runs and the vertical error bar represents the variance at each user over these 122 runs. These observations can be made from the results in Fig. 6:

- When comparing across different clustering schemes, despite being assigned the same amount of power, smaller clusters perform better under the same SNR; hence, they are more robust against noise and are sufficient for zero-error transmissions such as the FEC limit [21]. This performance improvement in smaller clusters is observed for all cases of the SPC, MLC–SE, and MLC–SER.
- When comparing the MLC–SER and MLC–SE schemes, the SER exhibits higher consistency in terms of a tighter upper and lower bound for the BER within one clustering scenario, as well as a smaller variance across all users, as seen in Table 2. This translates to better user fairness. Thus, the SER scheme is considered the better scheme to adopt for the MLC.
- Between the MLC–SER and SPC, the MLC–SER scheme in general achieves a better BER level across all users, as shown by the mean values in Table 2. The disadvantage of SPC becomes particularly noticeable in a six-user clustering format.
- A periodical "wavy" pattern is observed in the MLC–SER. This is because the users in one cluster can achieve relatively similar BER, but the further away a cluster is from the RRH, the higher the BER gets for the cluster. Consider a three-user MLC–SER for an example: Users 1, 5, and 9 are in the closest cluster to the RRH, with their BERs around –7, while the BER for the second closest cluster (users 2, 7, and 10) increases to around –5. This behavior agrees with how the clustering format is set up and the design of the MLC–SER to achieve better user fairness.

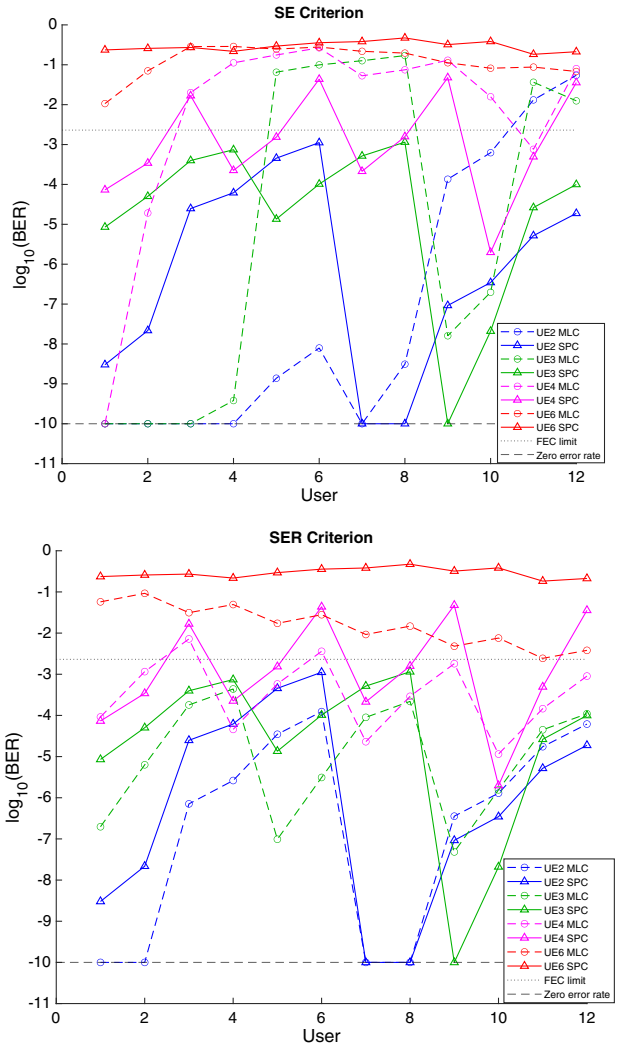

Fig. 6. Plot of $\log_{10}(\text{BER})$ per user for SE criterion (top) and SER criterion (bottom). The line at $\log_{10}(\text{BER}) = -10$ represents the zero error rate in this simulation. Results under the SPC scheme are also plotted for comparison. The line at $\log_{10}(\text{BER}) = \log_{10}(2.3e-3)$ represents the forward error correction (FEC) limit with 7% overhead.

Table 2. Means and Variance in Linear Scales Across All Users Under Each Clustering Format and Each Scheme

Cluster Format	Mean			Variance		
	MLC–SER	MLC–SE	SPC	MLC–SER	MLC–SE	SPC
UE2	2.02e-5	5.70e-3	1.40e-4	1.97e-9	3.69e-4	1.32e-7
UE3	9.11e-5	4.25e-2	2.59e-4	2.72e-8	4.08e-3	1.62e-7
UE4	1.30e-3	7.73e-2	1.23e-2	7.44e-6	7.85e-3	7.56e-4
UE6	2.65-2	1.62-1	2.99e-1	7.52e-4	1.17e-2	9.30e-3

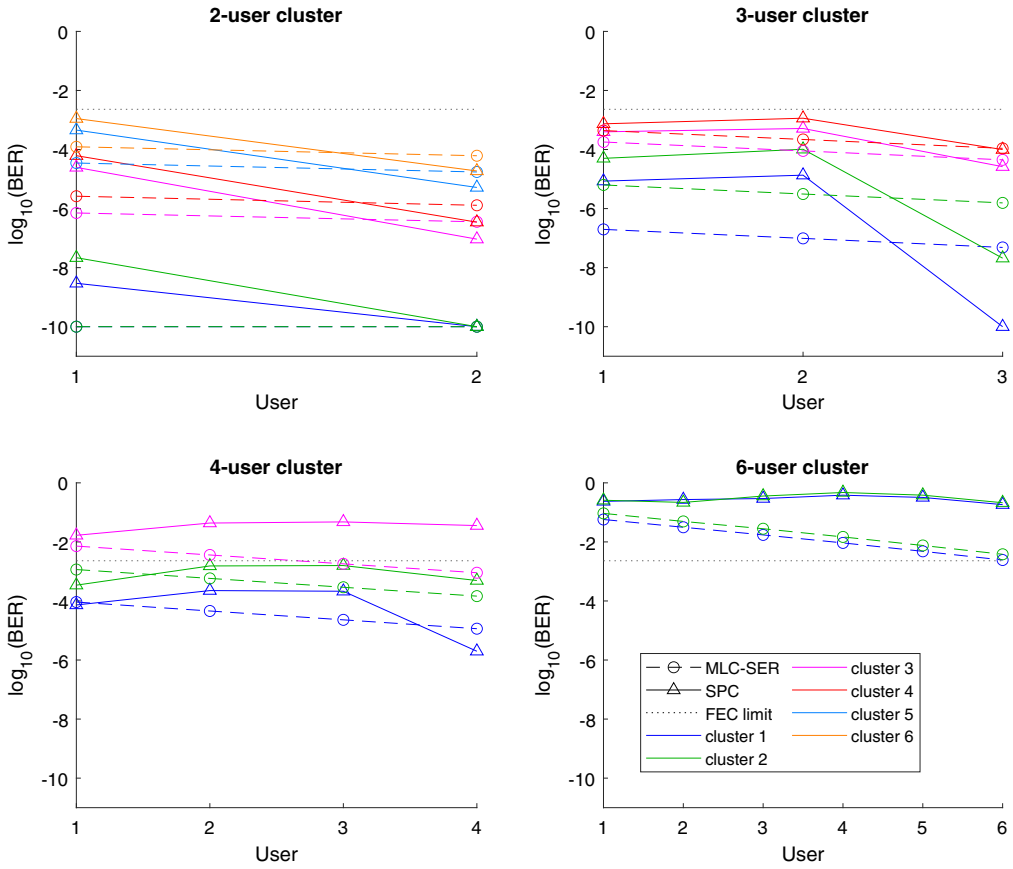

Fig. 7. Zoomed-in comparison on the performances of the MLC–SER and SPC. The users are overlapped along the x-axis by clusters. User 1 to N corresponds to the near to far users in each N-user cluster. Cluster 1 is the closest to the RRH while Cluster 6 is the furthest away from the RRH.

To further analyze the results for the MLC–SER and SPC, we have plotted the results of different user clusters in Fig. 7. The calculated BER curves for each cluster within the N cluster are overlapped onto each other. From these results, it can be clearly seen that the clusters located further away from the RRH have a poorer performance for both the MLC–SER and SPC for all the cases investigated. Also clearly indicated is that the far user within each cluster for the SPC with $N < 4$ always has a better performance compared to the rest of the users in that cluster, which highlights SPC error propagation characteristics that degrade the performance of near users. For the MLC–SER, the BER performance is relatively constant for fewer user clusters; however, a slight BER degradation is observed for users closer to the RRH for larger user clusters. This is because the power allocation scheme for the MLC–SER only considered the demodulation decision boundary within the respective quadrant where the symbol sits. If, however, the noise-imposed symbol is shifted outside the respective quadrant, the demodulation would fail. In other words, near users have less tolerance for noise with the current MLC–SER power allocation design.

Overall, we have concluded that the best performance is achieved from the MLC with SER criterion among all investigated, based on observations that the MLC–SER has the minimum mean BER across all users compared to the MLC–SE and SPC. The results also indicated that the MLC–SER is able to provide better user fairness across the users within the cluster. Based on the results obtained, it is clear that to achieve good NOMA performance, fewer user clusters is preferred, with $N < 4$ in this investigation.

Although the investigations carried out in this study targeted at quantifying the performance as a function of cluster size

with appropriate assumptions made, there are still many areas to be further explored, such as the improvement in the current MLC–SER power allocation design to alleviate the bias toward far users, a more accurate fronthaul model taking into account the possible fiber impairment effects including power loss and dispersion, the effect of a larger user pool that expands further toward the cell edge, and the possible adaptation of dynamic modulation and channel coding, which is beyond the scope of this paper.

4. CONCLUSION

In summary, we have further investigated the application of a MLC NOMA scheme for a RoF fronthaul system to scenarios of multiuser clustering. Two power allocation schemes were investigated for the MLC scheme based on the SE and SER, respectively. An AWGN channel simulation was conducted where a group of 12 users underwent different clustering scenarios. The results show that a MLC with a SER criterion scheme offers better user fairness across all the clusters and higher consistency in terms of variations in the transmission distance. A SE criterion scheme, on the other hand, achieves a lower minimum BER, but provides poorer fairness. In addition, a larger cluster with more users is more prone to noise and more easily fails the FEC error-free limit, justifying the typically used two-user clustering scheme. We believe this study provides valuable insights into the behavior of the MLC scheme in dynamic user clustering scenarios.

REFERENCES

1. Y. Tong, C.-W. Chow, G.-H. Chen, C.-W. Peng, C.-H. Yeh, and H. K. Tsang, "Integrated silicon photonics remote radio frontend (RRF) for single-sideband (SSB) millimeter-wave radio-over-fiber (ROF) systems," IEEE Photon. J. **11**, 1–8 (2019).
2. C. Ranaweera, E. Wong, A. Nirmalathas, C. Jayasundara, and C. Lim, "5G C-RAN with optical fronthaul: an analysis from a deployment perspective," J. Lightwave Technol. **36**, 2059–2068 (2017).
3. CPRI Specification v7.0 (2015).
4. T. Pfeiffer, "Next generation mobile fronthaul and midhaul architectures," J. Opt. Commun. Netw. **7**, B38–B45 (2015).
5. D. Perez-Galacho, D. Sartiano, and S. Sales, "Analog radio over fiber links for future 5G radio access networks," in *21st International Conference on Transparent Optical Networks (ICTON)* (IEEE, 2019).
6. R. Singh, H. Zhu, and J. Wang, "Performance of non-orthogonal multiple access (NOMA) in a C-RAN system," in *IEEE 86th Vehicular Technology Conference (VTC-Fall)* (IEEE, 2017).
7. M. Morelli, C.-C. J. Kuo, and M.-O. Pun, "Synchronization techniques for orthogonal frequency division multiple access (OFDMA): a tutorial review," Proc. IEEE **95**, 1394–1427 (2007).
8. L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," IEEE Commun. Surv. Tutorials **20**, 2294–2323 (2018).
9. M. Vaezi, Z. Ding, and H. V. Poor, *Multiple Access Techniques for 5G Wireless Networks and Beyond* (Springer, 2019).
10. A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *International Symposium on Intelligent Signal Processing and Communication Systems* (IEEE, 2013), pp. 770–774.
11. S. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: potentials and challenges," IEEE Commun. Surv. Tutorials. **19**, 721–742 (2016).
12. Y. Tian, K.-L. Lee, C. Lim, and A. Nirmalathas, "Demonstration of non-orthogonal multiple access scheme using multilevel coding without successive interference cancellation with 60 GHz radio-over-fiber fronthaul," in *Optical Fiber Communication Conference and Exposition (OFC)* (2018), paper Tu3J.4.
13. M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems," IEEE Access **4**, 6325–6343 (2016).
14. S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," IEEE Signal Process Lett. **22**, 1647–1651 (2015).
15. J. S. Yeom, H. S. Jang, K. S. Ko, and B. C. Jung, "BER performance of uplink NOMA with joint maximum-likelihood detector," IEEE Trans. Veh. Technol. **68**, 10295–10300 (2019).
16. D. Tse and P. Viswanath, *Fundamentals of Wireless Communication* (Cambridge University, 2005).
17. A. I. Sulyman, A. Alwarafy, G. R. MacCartney, T. S. Rappaport, and A. Alsanie, "Directional radio propagation path loss models for millimeter-wave wireless networks in the 28-, 60-, and 73-GHz bands," IEEE Trans. Wireless Commun. **15**, 6939–6947 (2016).
18. A.-L. Yi, L.-S. Yan, C. Liu, M. Zhu, J. Wang, L. Zhang, C.-H. Ye, and G.-K. Chang, "Frequency offset compensation and carrier phase recovery for differentially encoded 16-QAM vector signal in a 60-GHz ROF system," IEEE Photon. J. **5**, 7200807 (2013).
19. S. An, Z. S. He, J. Chen, H. Han, J. An, and H. Zirath, "A synchronous baseband receiver for high-data-rate millimeter-wave communication systems," IEEE Microwave Wireless Compon. Lett. **29**, 412–414 (2019).
20. W. Keusgen, R. J. Weiler, M. Peter, M. Wisotzki, and B. Göktepe, "Propagation measurements and simulations for millimeter-wave mobile access in a busy urban environment," in *39th International Conference on Infrared, Millimeter, and Terahertz waves (IRMMW-THz)* (IEEE, 2014).
21. S. Okabe, T. Ikeda, F. Suginoshita, K. Shogen, A. Hirata, M. Yaita, N. Kukutsu, and Y. Kado, "10-Gbps forward error correction system for 120-GHz-band wireless transmission," in *IEEE Radio and Wireless Symposium (RWS)* (IEEE, 2010), pp. 472–475.