User pairing using laser chaos decision maker for NOMA systems

Zengchao Duan¹ᵃ), Aohan Li¹, Norihiro Okada¹, Yusuke Ito¹, Nicolas Chauvet², Makoto Naruse², and Mikio Hasegawa¹

¹ Department of Electrical Engineering, Tokyo University of Science
6-3-1 Niiyuku, Katsushika-ku, Tokyo 125-8585, Japan
² Department of Information Physics and Computing, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

ᵃ) 4320533@ed.tus.ac.jp

Received July 15, 2021; Revised August 31, 2021; Published January 1, 2022

Abstract: Non-Orthogonal Multiple Access is one of the most important technologies in 5G and Beyond 5G wireless communications, which improve system performance by power domain multiplexing. In realizing Non-Orthogonal Multiple Access, the pairing of multiple users is necessary where efficient principles are highly demanded in dynamically changing electromagnetic environments. In the meantime, ultrafast methods of solving multi-armed bandit problems have been developed using chaotic laser time series. In this paper, we consider the user pairing problem in Non-Orthogonal Multiple Access as a multi-armed bandit problem and propose an ultra-fast user pairing algorithm based on the laser chaos decision maker. We numerically demonstrate that the proposed scheme accomplishes higher throughputs compared with traditional user pairing algorithms, especially in cases with lower user density.

Key Words: non-orthogonal multiple access, laser chaos decision maker, multi-armed bandit problems, users pairing

1. Introduction
The ultrafast decision-maker has been demonstrated to resolve multi-armed-bandit (MAB) problems at the GHz order using chaotic laser time series [1, 2]. The MAB problem concerns how to accurately and quickly find the best slot machine in initially uncertain reward environments. In [1], the laser chaos time series is utilized in solving two-armed bandit problems at ultrahigh-speed, of which scalability is later accomplished by time-domain multiplexing [2]. As its applications, Takeuchi et al. adapted the laser chaos decision-maker to IEEE 802.11a-based wireless LAN (WLAN) system to realize dynamic channel selection [3]. Furthermore, Kanemasa et al. applied the laser chaos decision-maker for dynamic reconfiguration of channel bonding in WLAN [4].

In the meantime, to meet the demands of increasing user data rate and system capacity, the future wireless access aims to design and implement 5G and Beyond 5G mobile network [5–8]. In 5G, non-orthogonal multiple access (NOMA) is considered to be one of the most competitive candidates for enhancing system capacity because of its high spectral efficiency [9, 10]. In NOMA, the signals of...
multiple users are multiplexed in the power domain, meaning that the same frequency can be used by multiple users for information transmission. Multiplexed users are called user pairs. Pairing users with more channel gain differences can achieve higher frequency efficiency. In [11], a user pairing scheme based on user power demand is proposed. In previous studies, the solutions proposed in NOMA networks are all designed according to the distance between the user and the base station [12, 13]. If the users with small channel gain differences are paired with each other, the performance of successive interference cancellation (SIC) will be degraded, and accordingly, the capacity of these users will be reduced [13–15].

In this paper, we formulate the user pairing problem in NOMA systems as a MAB problem where each arm represents a combination of user pairing. We utilize the laser chaos decision-maker algorithm to quickly select the pairings that enhance the resultant total system throughput. The simulation results demonstrate that the proposed method realizes better throughput than the conventional ones. Furthermore, we demonstrate that the proposed method significantly outperforms conventional ones when the user density is small, indicating that the fast exploration by the bandit approach strongly affects the resultant performance.

The remainder of this paper is as follows. In Section 2, we describe the system model and the problem statement. Section 3 review the principle of the laser chaos decision-maker for solving MAB problems. A decision is determined by multiple comparisons of incoming laser chaos time series and the threshold levels. By updating the thresholds, a higher reward arm is recognized. Section 4 demonstrates our proposed user pairing algorithm for NOMA based on the laser chaos decision-maker. Section 5 examines the accuracy of the proposed method as well as performance comparisons to conventional methods. Finally, Section 6 summarizes the paper.

2. System model and problem formulation

2.1 System model

We consider a single-cell system with a central base station (BS) in which there are $M$ users where $M$ is an even number. Let $U_{1}, U_{2}, \cdots, U_{M}$ denotes each of $M$ users in the cell. $h_{i}$ is defined as the user’s channel gain from the BS to $i_{th}$ user, which satisfies the order of $|h_{1}|^{2} \geq |h_{2}|^{2} \geq \cdots \cdots \geq |h_{M}|^{2}$. Let $d_{i}$ be the distance from the BS to the $i^{th}$ user, which satisfies the order of $d_{1} \leq d_{2} \leq \cdots \leq d_{M}$. Figure 1 schematically illustrates a diagram of the downlink in a NOMA system. Conventionally, user pairing is intuitively considered as the following. We consider two users $UE_{a}$ and $UE_{b}$. Suppose that $UE_{a}$ is electromagnetically located in a bad channel condition or located far away from the BS, whereas $UE_{b}$ experiences a good channel condition or closely locates to the BS.

In such a case, by assigning more power for $UE_{a}$, $UE_{b}$ conducts SIC to detect and remove $UE_{a}$ in retrieving the data for $UE_{b}$. We call $UE_{a}$ and $UE_{b}$ by a far and a near user, respectively.

Thus, the transmitted signal $x$ at the BS can be written as follows:

$$x = \sqrt{\beta_{a}}P_{a}x_{a} + \sqrt{\beta_{b}}P_{b}x_{b}$$  \hspace{1cm} (1)

where $x_{a}$ and $x_{b}$ are the individuals information of $UE_{a}$ and $UE_{b}$, respectively. $P$ is the total transmission power of the BS. $\beta_{a}$ and $\beta_{b}$ represent the power coefficients for $UE_{a}$ and $UE_{b}$ ($\beta_{a} + \beta_{b} = 1$ and $\beta_{a} > \beta_{b}$).
The received signal $y_a$ and $y_b$ at the far user $UE_a$ and the near user $UE_b$ are respectively given by

$$y_a = h_a(\sqrt{\beta_a}P_x a + \sqrt{\beta_b}P_x b) + n_a \quad (2)$$

and

$$y_b = h_b(\sqrt{\beta_a}P_x a + \sqrt{\beta_b}P_x b) + n_b \quad (3)$$

where $h_a$ and $h_b$ represents the channel gain between the BS and $UE_a$ and $UE_b$. $n_a$ and $n_b$ represent Gaussian noise with a zero mean and a variance $\sigma^2$.

According to the principle of SIC, the signal of the user with a strong power is decoded first by treating the other signal as noise. In the subsequent step, SIC extracts the signal of the weak-power user. Therefore, the individual achievable data rate of each user of $i^{th}$ pair can be given by:

$$R_a = B \log_2 (1 + \frac{\beta_a\rho|h_a|^2}{\beta_a\rho|h_a|^2 + 1}) \quad (4)$$

and

$$R_b = B \log_2 (1 + \frac{\beta_b\rho|h_b|^2}{\beta_b\rho|h_b|^2}) \quad (5)$$

where $\rho$ is the transmit signal to noise ratio (SNR) per pair. $B$ is the bandwidth allocated to each pair. The total achievable data rate is given by

$$R_{sum} = \sum_{a=1}^{M/2} R_a + \sum_{b=1}^{M/2} R_b \quad (6)$$

2.2 Problem formulation

In this section, we address the optimization problem for maximizing the sum rate. The optimization problem formulated as

$$\max R_{sum} = \sum_{i=1}^{M/2} (R_{a,i} + R_{b,i}) \quad (7)$$

where $R_{a,i}$ and $R_{b,i}$ is $i^{th}$ paired user’s throughput. In this paper, we solve this problem through the laser chaos decision-maker.

3. Laser chaos decision-maker

The oscillation from semiconductor lasers becomes unstable, leading to chaotic oscillation when a part of the output light is feedback to the laser cavity, called time-delayed feedback [16]. The delayed optical feedback of the reflector to the semiconductor laser generates high-frequency laser chaos. The resultant ultrafast chaotic time series have been utilized for random number generations [17], secure optical communications [18], photonic reservoir computing [19], among others. One of the latest development of chaotic lasers is reinforcement learning, in particular, multi-armed bandit problems, which we call laser chaos decision-makers [1, 2].

3.1 Two-armed laser chaos decision-maker

Figure 2 shows a schematic diagram of Tug-of-War (ToW) [20] based decision making using the laser chaos waveforms given by semiconductor lasers. The ToW is a model that utilizes the “efficiency of movement of objects” that satisfies the law of conservation of volume.

The decision-making process of a two-armed bandit problem based on laser chaos is summarized as follows. Here the two slot machines are called slot machines A and B. First, set the initial value of the threshold. A decision is made by comparing the sampled laser chaos time series data with the threshold value. If the sampled time series data is larger than the threshold, the decision is to select slot machine A. Otherwise, the slot machine B is selected. The threshold is adjusted according to whether the selected slot machine is rewarded or not, so that the slot machine with a higher reward can be selected with a high probability.

More precisely, the threshold level is at cycle $t$ is given by
where $TA(t)$ is the threshold adjuster value at cycle $t$, $a$ is a constant that controls the size range of the threshold. $[x]$ means to rounding a value $x$ to an integer toward zero. We assume that the $T(t)$ takes a value of $-Z, \cdots, -1, 0, 1, \cdots, Z$ where $Z$ is a natural number. Therefore, the number of threshold levels is $2Z + 1$. The range of $TA(t)$ is limited between $-kZ$ and $kZ$.

If the selected slot machine A receives a reward, the threshold adjuster is updated according to the following formula:

$$TA(t+1) = -\delta + \alpha TA(t),$$

(9)

whereas when the selected slot machine B receives a reward,

$$TA(t+1) = \delta + \alpha TA(t)$$

(10)

where $\alpha$ represents the forgetting parameter. The initial $TA$ value was zero and $\delta$ is a constant increment, which is considered to be unity hereafter ($\delta = 1$).

In the meantime, when the selected slot machine A does not obtain a reward, the threshold adjuster is updated according to the following:

$$TA(t+1) = \omega + \alpha TA(t),$$

(11)

whereas when the selected slot machine B does not obtain a reward,

$$TA(t+1) = -\omega + \alpha TA(t).$$

(12)

Here, $\omega$ is determined by considering the betting history. Suppose that the number of times to select slot machine $j$ by the cycle $t$ is $F_j$ and the number of times to obtain a reward by selecting the slot machine $j$ by the cycle $t$ is $L_j$, where $j$ is A or B. The estimated reward probability of slot machines A and B are given by

$$P_A = \frac{L_A}{F_A}, P_B = \frac{L_B}{F_B}.$$ 

(13)

Based on these estimated probabilities, $\omega$ is given by

$$\omega = \frac{P_A + P_B}{2 - (P_A + P_B)}.$$ 

(14)

### 3.2 Scalable laser chaos decision-maker

The aforementioned bandit method has been extended to more than two-arm bandit problems model as schematically shown in Fig. 3 where the number of arms is given by $2^K$ where $K$ is a natural number. Let the identity of the slot machine be presented by a binary code; the decision is determined from the most significant bit (MSB) toward the least significant bit (LSB).

First, the signal level of the laser chaos time series measured at $t = t_1$ is compared with the threshold value of $TH_1$. If the signal level is smaller than or equal to the threshold, the MSB of the slot machine to be chosen is determined by 0, which we denote as $D_1 = 0$; otherwise, $D_1 = 1$. 

![Laser chaos decision-maker](image-url)
Next, we decide the second most significant bit. Assume that the MSB has been determined by $D_1 = 0$. In this case, we examine the signal level of the chaotic laser time series measured at $t = t_2$ with the threshold value of $TH_{2,0}$. The first number 2 in the suffix indicates that the threshold value is related to the second MSB, and the second number 0 in the suffix indicates that the previous decision was to be 0, $D_1 = 0$. If the time series is less than or equal to the threshold $TH_{2,0}$, the second MSB is to be 0 ($D_2 = 0$), otherwise $D_2 = 1$.

Similarly, the remaining bits are determined. For example, in the case when both MSB and the second MSB have been determined by zero. Then, the third bit will be determined by the comparison of the laser chaos time series sampled at $t = t_3$ and the threshold value $TH_{3,0,0}$. The number 3 indicates that the threshold value is related to the third MSB, and the second number 0 in the suffix indicates that MSB was determined $D_1 = 0$, and the third number 0 indicates that the second MSB was determined $D_2 = 0$. Such threshold comparisons continue until all bits of information that specify the slot machine have been determined.

If the selected slot machine obtains a reward, the threshold values are updated in the following manner.

\begin{align}
TH_{K,S_1,S_2,\ldots,S_{K-1}}(t + 1) &= \delta + \alpha TH_{K,S_1,S_2,\ldots,S_{K-1}}(t)(S_K = 0) \\
TH_{K,S_1,S_2,\ldots,S_{K-1}}(t + 1) &= -\delta + \alpha TH_{K,S_1,S_2,\ldots,S_{K-1}}(t)(S_K = 1)
\end{align}

where $\alpha$ and $\delta$ means a forgetting parameter and a constant increment, respectively, which is the with the two-armed bandit case examined in the former subsection. Hereafter, we assume that $\delta = 1$.

When the selected slot machine does not receive a reward, the threshold values are updated by the following:

\begin{align}
TH_{K,S_1,S_2,\ldots,S_{K-1}}(t + 1) &= \omega_{K,S_1,S_2,\ldots,S_{K-1}} + \alpha TH_{K,S_1,S_2,\ldots,S_{K-1}}(t) \\
TH_{K,S_1,S_2,\ldots,S_{K-1}}(t + 1) &= -\omega_{K,S_1,S_2,\ldots,S_{K-1}} + \alpha TH_{K,S_1,S_2,\ldots,S_{K-1}}(t),
\end{align}

where $\omega_{K,S_1,S_2,\ldots,S_{K-1}}$ is the increment defined as follows.

Let the number of times of the $K_{th}$ bits are selected be denoted by $F_{S_1,S_2,\ldots,S_K}$. Let the number of times of wins when the machines was selected is given by $L_{S_1,S_2,\ldots,S_K}$. The estimated reward probability of slot machines whose first $K$ bit are specified by $S_1,S_2,\ldots,S_K$ is given by

\begin{equation}
P_{S_1,S_2,\ldots,S_K} = \frac{L_{S_1,S_2,\ldots,S_K}}{F_{S_1,S_2,\ldots,S_K}},
\end{equation}

which leads to $\omega_{K,S_1,S_2,\ldots,S_{K-1}}$ given by

\begin{equation}
\omega_{K,S_1,S_2,\ldots,S_{K-1}} = \frac{P_{S_1,S_2,\ldots,S_K}}{2 - P_{S_1,S_2,\ldots,S_K}}.
\end{equation}
4. Laser chaos decision-maker for user pairing

As schematically shown in Fig. 4, we transform the user pairing problem in NOMA systems into a MAB problem. That is, a user pairing is associated with an arm, and the problem of finding the best user pairs combination is solved by laser chaos decision-maker. Then reward in the MAB problem is evaluated by the throughput corresponding to the selected user combination.

We denote $C_n$ as the $n^{th}$ combination.

![Fig. 4. The user pairing issue in NOMA is considered as a MAB problem in this study.](image)

Assuming that there are in total $M$ users, and hence in total $M/2$ user pairs, the throughput $R_{\text{sum}}^n$ regarding the combination $C_n$ at time $t$ is calculated as follows:

$$R_{\text{sum}}^n(t) = \frac{M}{2} \sum_{i=1}^{\frac{M}{2}} R_{a,i} + \frac{M}{2} \sum_{i=1}^{\frac{M}{2}} R_{b,i},$$  \hspace{1cm} (21)

where $R_{a,i}$ and $R_{b,i}$ respectively denote the throughput of paired users $a$ and $b$ at time $t$ when the combination is given by $C_n$. We consider the average throughput $R_{\text{aver}}^n$, which is defined by

$$R_{\text{aver}}^n(t) = \frac{1}{\tau} \sum_{j=-\tau+1}^{t} R_{\text{sum}}^n(j),$$  \hspace{1cm} (22)

which means the average of the throughput by the combination $C_n$ over the past $\tau$ rounds. We define the reward if the observed throughput is greater than the average throughput otherwise the reward is not dispensed.

Figure 5 illustrates the case of a six-user NOMA system. In this case, the total number of user pair combinations are 15. Hence, four bits are needed to specify which arm is the best by the laser chaos decision-maker. We introduce how the principle works in the following. We determined the redundant arm’s throughput is 0, when the redundant arm is selected, the throughput output is 0, according to our strategy, this arm is not received a reward, and the threshold will be updated according to Eq. (18).

First, the signal level of the laser chaos measured at $t = t_1$ is compared with the threshold $TH_1$. When the laser chaos signal level is larger than the threshold value $TH_1$, the first bit is determined to be 0. Next, the time series measured at $t = t_2$ is compared with the threshold of $TH_{2,0}$. When the time series is larger than $TH_{2,0}$, the 2nd bit is determined to be 0. Similarly, the laser chaos time series measured at $t = t_3$ is compared with the threshold value $TH_{3,0,0}$. If the time series is larger than the threshold value, the third bit is determined as 0. Then, the time series measured at $t = t_4$ is compared with the threshold $TH_{4,0,0,0}$, which results in 0 if the time series is larger than $TH_{4,0,0,0}$. Finally, the arm that should be selected is determined by the binary sequence of 0000, which corresponds to the combination $C_0$. 

77
Fig. 5. Laser chaos decision-maker for solving six users NOMA system. In this case, there in total 15 user pair combinations ($C_0, C_1, \cdots, C_{14}$). That is, the problem is to solve a 15-armed bandit problem.

We measure the throughput of the user pairing corresponding to the $C_0$ combination. Consequently, we compare the measured throughput with the average throughput. If the measured throughput is greater than the average throughput, a reward is dispensed. The associated thresholds are updated by

$$
\begin{align*}
TH_{4,0,0,0}(t+1) &= -\delta + \alpha TH_{4,0,0,0}(t) \\
TH_{3,0,0}(t+1) &= -\delta + \alpha TH_{3,0,0}(t) \\
TH_{2,0}(t+1) &= -\delta + \alpha TH_{2,0}(t) \\
TH_1(t+1) &= -\delta + \alpha TH_1(t).
\end{align*}
$$

Conversely, if the measured throughput is smaller than the average throughput, a reward is not provided, leading to the threshold update given by

$$
\begin{align*}
TH_{4,0,0,0}(t+1) &= \omega_{4,0,0,0} + \alpha TH_{4,0,0,0}(t) \\
TH_{3,0,0}(t+1) &= \omega_{3,0,0} + \alpha TH_{3,0,0}(t) \\
TH_{2,0}(t+1) &= \omega_{2,0} + \alpha TH_{2,0}(t) \\
TH_1(t+1) &= \omega_1 + \alpha TH_1(t),
\end{align*}
$$

where $\omega$ are configured by

$$
\begin{align*}
\omega_{4,0,0,0} &= P_{4,0,0,0} + P_{4,0,0,1} - P_{4,0,0,0} - P_{4,0,0,1} \\
\omega_{3,0,0} &= P_{3,0,0} + P_{3,0,1} - P_{3,0,0} - P_{3,0,1} \\
\omega_{2,0} &= P_{2,0} + P_{2,1} - P_{2,0} - P_{2,1} \\
\omega_1 &= P_1 - P_1
\end{align*}
$$

where $P$ is defined in Eq. (19).
5. Simulation result

We evaluate the proposed method by comparing the system performance obtained with the Conventional-NOMA pairing scheme (C-NOMA) and the Unified Channel Gain Difference pairing scheme (UCGD-NOMA) in terms of the number of paired users and the total achieved throughput in the network. In the simulation, we consider a cell with a radius of 1000 m where users are randomly distributed. The forgetting parameter $\alpha$ is given by 1. We quoted the research conclusions in [1], using a chaotic signal with a sampling interval of 50 ps. We performed simulations under the conditions of bandwidth of 20 MHz, and the transmit signal-to-noise ratio is specified by $SNR = 20$ dB. The number of users is set as 10, which means that the total number of user pair combinations is 945. The combination $C$ calculation formula is given by $C = \frac{M!}{\left(\frac{M}{2}\right)!\left(\frac{M}{2}\right)!}$. Therefore, a total of 10 bits of decision levels are utilized in the laser chaos decision-maker. The length of duration in considering the reward is given by $\tau = 5$. The power allocation scheme is the same as that in [21]. The minimum

![Fig. 6. Time evolution of the selected combination of user pairing by the laser chaos decision maker algorithm. The panels (a), (b), (c), (d), (e), and (f) examine the cases with different number of threshold levels specified by $N$ of 2, 4, 8, 16, 32, and 64, respectively.](image)
transmission rate for all users is 1 bit/sec/Hz. The channel gain is assumed to following a Rayleigh distribution. Without loss of generality, we assume that the channel gain obeys the standard Rayleigh distribution at the node 100 m away from the BS.

5.1 Algorithm characterization

Figure 6 summarizes the result of the user pairing combination determined by the laser chaos decision-maker algorithm where the horizontal axis indicates the time cycle ranging from 0 to 1000 while the vertical axis represents the index of combinations ranging from 0 to 944. Figures 6(a), (b), (c), (d), (e), and (f) concern different threshold levels specified by $N$ taking the values of 2, 4, 6, 16, 32, and 64, respectively. The actual number of threshold levels is given by $2N + 1$. When the value of $N$ is too small, the throughput may converge when it has not reached the optimum. For example, when $N = 2$, ground-truth has occurred, but it will not converge. As the value of $N$ increases, the probability of

![Fig. 7. Time evolution of the total throughput by the laser chaos decision-maker algorithm. The panels (a), (b), (c), (d), (e), and (f) examine the cases with different number of threshold levels specified by $N$ of 2, 4, 8, 16, 32, and 64, respectively.](image-url)
throughput converging to the optimum will increase. For example, ground-truth has occurred when \( N = 32 \) and \( N = 64 \), but we can know from Fig. 8 that when \( N = 64 \), the throughput obtained by convergence is greater. So, we set \( N = 64 \) for the simulation of the comparison object.

From Fig. 6, especially Fig. 6(d), the selected combination converges after a certain initial exploration cycle, indicating that the pairing combination with the larger throughput is selected. Actually, in this simulation, the combination number of 578 is the ground truth that maximizes the total throughput. In fact, the ground-truth state has been successfully discovered by the cases when \( N \) is given by 2, 32, and 64.

Figure 7 demonstrates the time evolution of the observed total throughput of all pairs by the laser chaos decision-maker. After initial exploration duration, the total throughput increases and exhibits approximately 68 Mbps on average. In Fig. 7, we did not consider the arrangement of combinations. When \( N \) is small, the value of convergence is the local optimal solution instead of the global optimal solution.

Figure 8 compares the average of the total throughput as a function \( N \), which determines the number of the threshold levels. We observe that \( N = 64 \) yields the maximum throughput of about 68.767 Mbps. We consider that, since there are many arms in the present problem, too few threshold levels result in insufficient exploration that inhibits finding a better user pairing. When \( N = 64 \), the number of times the convergent pair needs to converge is about 200 times, that is, 10\( \text{ns} \). In Fig. 8, when \( N \) gradually increases, theoretically the probability of obtaining ground-truth will increase, but at the same time, since the number of threshold updates to be performed will also increase, the convergence speed will decrease. We choose \( N = 64 \) for comparison and simulation under the premise of considering the convergence speed.

5.2 Performance comparison
In Fig. 9, we compared the total throughputs among the proposed method, C-NOMA, and UCGD-NOMA under different user densities. While maintaining the number of users being 10, the maximum distance of the user from the BS is configured differently from 1000 to 2000. We can clearly observe that the proposed method, denoted by red markers, provides greater throughputs compared with C-NOMA and UCGD-NOMA, which are marked by green and blue markers, respectively. Indeed, the proposed method accomplishes almost optimal throughput represented by the yellow markers.

Furthermore, the significance of the proposed method over C-NOMA and UCGD-NOMA increases as the distance increases or the user density decreases. We consider that since C-NOMA and UCGD-
NOMA deterministically decide the combination based only on the distance information of the users and the base station, they lead to user pairing with unavoidable SIC demodulation errors when the user density is low. On the contrary, the proposed method explores a variety of combinations using the bandit strategy based on the observed total throughputs; hence it can find a better combination, which is actually almost the ground truth.

6. Conclusion

This paper examines the user pairing issue in 5G and Beyond 5G NOMA systems as a multi-armed bandit problem in dynamically changing electromagnetic environments. We propose and demonstrate a user pairing principle in NOMA systems based on the chaotic laser-based ultrafast decision-maker. The accuracy of the proposed method is characterized. Furthermore, compared with conventional methods of C-NOMA and UCGD-NOMA, the proposed algorithm achieves superior total throughputs, especially in cases when the user density is small, indicating the impact of the fast exploration ability of laser chaos decision-makers. As future works, we consider that channel allocation is an associated important issue related to the current research. That is, a simultaneous realization of user pairing and channel assignment will be of our important next research direction. At the same time, the case of high-density users is our next work, and we will explore whether this method is beneficial in the case of high-density users in the future.

References

[1] M. Naruse, Y. Terashima, A. Uchida, and S.J. Kim, “Ultrafast photonic reinforcement learning based on laser chaos,” Scientific Reports, vol. 7, 8772, 2017.
[2] M. Naruse, T. Mihana, H. Hori, H. Saigo, K. Okamura, M. Hasegawa, and A. Uchida, “Scalable photonic reinforcement learning by time-division multiplexing of laser chaos,” Scientific Reports, vol. 8, 10890, July 2018.
[3] S. Takeuchi, M. Hasegawa, K. Kanno, A. Uchida, N. Chauvet, and M. Naruse, “Dynamic channel selection in wireless communications via a multi-armed bandit algorithm using laser chaos time series,” Scientific Reports, vol. 10, 1574, 2020.
[4] H. Kanemasu, A. Li, M. Naruse, N. Chauvet, and M. Hasegawa, “Dynamic channel bonding using laser chaos decision maker in WLANs,” International Conference on Artificial Intelligence in Information and Communication (ICAIIC), April 2021
[5] Q. Li, H. Niu, A. Papathanassiou, and G. Wu, “5G network capacity: key elements and technologies,” IEEE Vehicular Technology Magazine, vol. 9, no. 1, pp. 71–78, January 2014.

[6] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, “Nonorthogonal multiple access for 5G and beyond,” Proceedings of the IEEE, vol. 105, no. 12, pp. 2347–2381, December 2017.

[7] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, “Concept and practical considerations of nonorthogonal multiple access (NOMA) for future radio access,” Proc. Int. Symp. Intell. Signal Process. Commun. Syst., Naha, Japan, pp. 770–774, November 2013.

[8] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X.G. Xia, “Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications,” IEEE Trans. Wireless Commun., vol. 17, no. 5, pp. 2961–2974, May 2018.

[9] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, “Non-orthogonal multiple access (NOMA) for cellular future radio access,” 2013 IEEE 77th Vehicular Technology Conference (VTC Spring), Dresden, Germany, pp.1–5, 2013.

[10] 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. Sel. Areas Commun., vol. 35, no. 10, pp. 2181–2195, October 2017.

[11] Z. Al-Abbasi and D.K. So, “User-pairing based nonorthogonal multiple access (NOMA) system,” Proceedings of IEEE VTC Spring, Nanjing, China, July 2016.

[12] L. Zhu, J. Zhang, Z. Xiao, X. Cao, and D. Wu, “Optimal user pairing for downlink non-orthogonal multiple access,” IEEE Wireless Commun. Lett. vol. 8, no. 2, pp. 328–331, April 2019.

[13] M.B. Shahab, M. Irfan, M.F. Kader, and Y.S. Soo, “User pairing schemes for capacity maximization in non-orthogonal multiple access systems,” Wireless Commun. Mobile Comput. vol. 16, no. 17, pp. 2884–2894, 2016.

[14] K. Saito, A. Benjebbour, Y. Kishiyama, Y. Okumura, and T. Nakamura, “Performance and design of SIC receiver for downlink NOMA with open-loop SU-MIMO,” 2015 IEEE International Conference on Communication Workshop (ICCW), London, UK, pp. 1161–1165, 2015.

[15] S. Mounchili and S. Hamouda, “Pairing distance resolution and power control for massive connectivity improvement in NOMA systems,” IEEE Trans. Veh. Technol. vol. 69, no. 4, pp. 4093–4103, April 2020.

[16] A. Uchida, “Optical communication with chaotic lasers: applications of nonlinear dynamics and synchronization,” John Wiley Sons, 2021

[17] A. Uchida, K. Amano, M. Inoue, H. Hirano, S. Naito, H. Someya, I. Owada, T. Kurashige, M. Shiki, and S. Yoshimori, “Fast physical random bit generation with chaotic semiconductor lasers,” Nature Photonics, pp. 728–732, November 2008.

[18] A. Argyris, D. Syvridis, L. Larger, V. Annovazzi-Lodi, P. Colet, I. Fischer, J. Garcia-Ojalvo, C.R. Mirasso, L. Pesquera, and K.A. Shore, “Chaos-based communications at high bit rates using commercial fibre-optic links,” Nature, vol. 437, 7066, pp. 343–346, November 2005.

[19] L. Appeltant, M.C. Soriano, S.G. Van, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C.R. Mirasso, and I. Fischer, “Information processing using a single dynamical node as complex system,” Nature Communications, vol. 2, 468, September 2011.

[20] S. Kim, M. Aono, and M. Hara, “Tug-of-war model for the two-bandit problem: Nonlocally correlated parallel exploration via resource conservation,” BioSystems, vol. 101, pp. 29–36, 2010.

[21] M. Zamani, M. Eslami, and M. Khorramizadeh, “Optimal Sum-Rate Maximization in a NOMA System with Channel Estimation Error,” Proc. Iranian Conference on Electrical Engineering (ICEE), May 2018.