

High-Speed Resource Allocation Algorithm Using a Coherent Ising Machine for NOMA Systems

Tepei Otsuka , Aohan Li , *Member, IEEE*, Hiroki Takesue , Kensuke Inaba , Kazuyuki Aihara ,
and Mikio Hasegawa , *Member, IEEE*

Abstract—Non-orthogonal multiple access (NOMA) technique is important for achieving a high data rate in next-generation wireless communications. A key challenge to fully utilizing the effectiveness of the NOMA technique is the optimization of the resource allocation (RA), e.g., channel and power. However, this RA optimization problem is NP-hard, and obtaining a good approximation of a solution with a low computational complexity algorithm is not easy. To overcome this problem, we propose the coherent Ising machine (CIM) based optimization method for channel allocation in NOMA systems. The CIM is an Ising system that can deliver fair approximate solutions to combinatorial optimization problems at high speed (millisecond order) by operating optimization algorithms based on mutually connected photonic neural networks. The performance of our proposed method was evaluated using a simulation model of the CIM. We compared the performance of our proposed method to simulated annealing, a conventional-NOMA pairing scheme, deep Q learning based scheme, and an exhaustive search scheme. Simulation results indicate that our proposed method is superior in terms of speed and the attained optimal solutions.

Index Terms—Non-orthogonal multiple access, resource allocation, coherent Ising machine, mutually connected neural network.

I. INTRODUCTION

WITH the rapid increase in the number of mobile devices, there is an urgent need to increase the spectrum efficiency of mobile communications [1]. A non-orthogonal multiple access (NOMA) technique has been proposed as one of the candidate technologies to meet the requirements of the next

Manuscript received 10 August 2022; revised 2 December 2022, 11 March 2023, and 17 July 2023; accepted 21 July 2023. Date of publication 21 August 2023; date of current version 17 January 2024. This work was supported in part by JST Moonshot R & D under Grant JPMJMS2021, in part by AMED under Grant JP23dm0307009, and in part by AI Center of UTokyo, Institute of AI and Beyond of UTokyo, JSPS Kakenhi Grants JP20H05921 and JP22K14263. The review of this article was coordinated by Dr. Zhiguo Ding. (*Corresponding author: Aohan Li.*)

Tepei Otsuka and Mikio Hasegawa are with the Department of Electrical Engineering, Tokyo University of Science, Tokyo 1258585, Japan (e-mail: t-otsuka@haselab.ee.kagu.tus.ac.jp; hasegawa@ee.kagu.tus.ac.jp).

Aohan Li is with the Department of Electrical Engineering, Tokyo University of Science, Tokyo 1258585, Japan, also with the Graduate School of Informatics and Engineering, The University of Electro-Communications, Tokyo 1828585, Japan, and also with the Meta-Networking Research Center, The University of Electro-Communications, Tokyo 1828585, Japan (e-mail: aohanli@ieee.org).

Hiroki Takesue and Kensuke Inaba are with the NTT Basic Research Laboratories, NTT Corporation, Atsugi 2430198, Japan (e-mail: hiroki.takesue.km@hco.ntt.co.jp; kensuke.inaba.yg@hco.ntt.co.jp).

Kazuyuki Aihara is with the International Research Center for Neurointelligence, The University of Tokyo, Tokyo, Japan (e-mail: kaihara@g.ecc.u-tokyo.ac.jp).

Digital Object Identifier 10.1109/TVT.2023.3300920

generation of wireless networks [2], [3]. Unlike conventional orthogonal multiple access (OMA) schemes, NOMA schemes can effectively increase the spectrum efficiency by introducing extra-power-domains, which enables the data of multiple different users to be multiplexed on the same channel [3]. More specifically, by applying superposition coding in the transmitter and successive interference cancellation (SIC) technology in the receiver, signals from multiple users that are multiplexed can be distinguished.

In a NOMA system, resources such as channels and power must be appropriately allocated to improve the system performance, such as the data rate, and avoid system outages such as an SIC error [4], [5]. Previous studies have shown that optimal resource allocation (RA) can increase the total achievable data rate of a system in comparison with random RA [6]. Specifically, channel allocation methods that consider the differences in channel gain between users, and power allocation methods that balance system data rates and user fairness, are key aspects in a NOMA system design. In addition, it is important to allocate resources to avoid the failure conditions of SIC and maximize the performance, including the total data rate, quality of service (QoS), or fairness of the NOMA system. However, joint channel and power allocation problems have been proven to be NP-hard [7], and their high complexity hinders an efficient and reliable RA, resulting in a low performance of NOMA systems.

To solve the NP-hard RA optimization problems in NOMA systems, many approaches such as machine learning (ML) and heuristic approaches have been proposed [5], [8], [9], [10], [11], [12], [13]. Optimization using ML has the advantage of a rapid RA in large-scale problems by using trained models that have been learned in advance. However, a trained model may no longer be usable and must be retrained when the system environment changes. This has a disadvantage that the RA may not be sufficiently fast when the communication environments change. By contrast, optimization using heuristic methods is faster owing to a lower computational complexity; however, a fundamental trade-off emerges between the solution accuracy and computational complexity. By contrast, hardware-based search algorithms, such as quantum computers and Ising machines, have been proposed and are expected to solve various optimization problems at high speed [14], [15], [16], [17], [18]. As described in [19], quantum technologies are emerging technologies that are expected to be deployed in 6G networks around 2030, although it is still unclear when a practical quantum computer will be available. D-wave [14] is a quantum annealing machine that

has already been commercialized and is available for general users. The coherent Ising machine (CIM) [15], [16], [17], [18] is an optimization machine based on laser networks, which can deliver fair approximate solutions to combinatorial optimization problems at high speed on the order of milliseconds. Many NP-complete and NP-hard problems can be transformed into ground-state search problems of the Ising model [20]. Here, the CIM and D-Wave are Ising machines that can rapidly obtain a state near the ground state of the Ising Hamiltonian which represents a good solution to the optimization problem. As the main difference between the D-Wave and CIM, D-Wave reproduces a coupling of magnetic spins with a chimera topology having sparse connections, whereas in CIM, the Ising network of optical pulses generated by a laser oscillator is fully coupled [15], [16].

A previous study [17] demonstrated that the CIM is more effective than D-Wave for large-scale large-density problems, and it will be more suitable for the next-generation large-scale wireless communications. In [18], the authors showed that 100,000 fully connected spins can be implemented using the CIM. In previous studies [21], [22], [23], [24], various wireless communication systems are optimized using CIMs. A CIM-based optimization approach has been applied to solve the NP-hard optimization problems in wireless communications. In [25], it is shown that the optimal solutions obtained can be empirically scaled as $O(1)$ using the CIM. It is therefore possible to achieve a real-time optimization of RA in a large-scale NOMA system using the CIM. In this article, we propose an RA method for NOMA systems using the CIM. The main contributions of this article can be summarized as follows.

- We propose a high-speed RA method using the CIM for NOMA systems, especially for the channel assignment. The proposed method is a hardware-based method with superior speed (milliseconds-order) compared to other existing methods.
- To the best of our knowledge, previous works related to RA using the CIM in wireless communication assumed that only one user could access each channel simultaneously. In contrast to previous studies, this study focuses on the RA optimization problem using the CIM for NOMA systems where multiple users can be multiplexed in the same channel. Hence, novel solutions are necessary to solve such problems which are different from those in our previous work. For instance, dummy users need to be introduced and the constraints are different, which leads to different strength of the interaction between spins and that of the external magnetic field on each spin. In this study, we formulated the RA combinatorial optimization problem carefully to meet the characteristics of the RA in the NOMA system and the feasibility of solving it using the CIM.
- To solve the RA optimization problem in NOMA systems using the CIM, it is necessary to derive the parameters of the Ising model and introduce them into the CIM. Therefore, we first transformed the formulated RA combinatorial problem into a ground state search problem of the Ising Hamiltonian. Then, we derived the parameters of the Ising model that needed to be introduced into the CIM to obtain

the solutions, i.e., the interaction between spins and the external magnetic field.

- To verify the effectiveness of the proposed method, we first evaluated the convergence of the proposed method. Simulation results demonstrated the convergence of the proposed method. Then, we evaluated the performance of the proposed method in the total data rate and compared it with that of other methods including simulated annealing (SA), exhaustive search (ES), the conventional-NOMA (C-NOMA), deep Q learning (DQN), and random methods. The simulation results indicate that our proposed approach can achieve higher data rates than other methods. Moreover, the computation time of the proposed, SA, DQN and ES methods was evaluated, which demonstrates that the proposed method can obtain better performance at faster speed.

The rest of this article is organized as follows. Section II reviews previous related studies. Section III introduces the system model and problem formulation. Section IV introduces the principle of the CIM in solving the optimization problems. Section V presents the proposed CIM-based RA method for NOMA systems. Section VI demonstrates the simulation results. Section VII presents the discussion and future work. Section VIII provides some concluding remarks regarding this research. The abbreviations used in this article are listed in Table I.

II. RELATED WORK

In this section, we review the related studies on RA. We first review the heuristic RA approaches and ML-based RA approaches for NOMA systems. We then review the hardware-based RA, particularly for the existing studies on CIM-based RA in wireless communications that are not limited to NOMA RA problems.

A. Heuristic RA Approaches

The high complexity of the RA problem in NOMA systems can be reduced using heuristic methods by dividing the problem into sub-problems [5], [8], [9], [10], [11]. In [5], the authors propose a low complexity and efficient method for RA in a downlink NOMA system. In this study, user fairness, data rate, and energy efficiency are optimized. The RA problem is divided into the sub-problems of power allocation and channel allocation. The power allocation solution is derived from a theoretical analysis. The channel allocation is carried out based on the derived power allocation solution. In [8], the RA in full duplex-NOMA (FD-NOMA) systems is studied. In this study, the RA problem is successfully solved with low computational complexity by dividing the channel and power problems based on the block coordinate descent method. In [9], an algorithm for downlink NOMA systems that analytically examines the optimal power allocation (OPA) and ideal pairing search algorithm (IPSA) in terms of fairness is proposed. Simulation results indicate that the computational complexity is significantly reduced in comparison to the ES. In [10], NOMA-based device-to-device (D2D) communications are considered. It is shown that communication performance can be improved to a significant extent using the

TABLE I
ABBREVIATIONS USED IN THIS PAPER

Abbreviation	Description	Abbreviation	Description
ANN	Attention-based Neural Network	FD-NOMA	Full Duplex-NOMA
BS	Base Station	IPSA	Ideal Pairing Search Algorithm
BPNN	Backpropagation Neural Network	ML	Machine Learning
BHD	Balanced Homodyne Detection	NOMA	Non-Orthogonal Multiple Access
CNR	Carrier-to-Noise Ratio	OPA	Optimal Power Allocation
CIM	Coherent Ising Machine	OMA	Orthogonal Multiple Access
C-NOMA	Conventional-NOMA	PSA	Phase Sensitive Amplifier
DQN	Deep Q Learning	QoS	Quality of Service
DRL	Deep Reinforcement Learning	RL	Reinforcement Learning
DOPO	Degenerate Optical Parametric Oscillator	RA	Resource Allocation
D2D	Device-to-Device	SA	Simulated Annealing
ES	Exhaustive Search	SIC	Successive Interference Cancellation
FPGA	Field Programmable Gate Array	WLAN	Wireless LAN

NOMA technique. The RA solution is obtained with low computational complexity by applying an optimal power allocation after a channel allocation algorithm that assigns sub-carriers to the D2D groups. In [11], the RA in a NOMA system is optimized using the SA, which is a meta-heuristic algorithm. In this study, a joint optimization of the channel allocation, sub-channel power allocation, and inter-user power allocation is considered. Simulation results indicate that a sufficiently reliable solution in terms of system throughput can be obtained with low computational complexity in comparison to an ES.

As discussed above, we can see that the heuristic RA approaches can obtain a solution with low computational complexity. However, there is a fundamental trade-off between the performance of the solution and the computational complexity, and there have been no previous studies that can quickly optimize on the order of milliseconds.

B. ML Based RA Approaches

With the development of ML applications in wireless communications, ML solutions for RA optimization problems in NOMA systems have been proposed [4], [12], [13]. In [4], deep reinforcement learning (DRL) based RA scheme is proposed for NOMA systems. Here, the framework of the DRL method is designed based on an attention-based neural network (ANN). Specifically, the joint optimization problem of channel and power allocation is formulated, and the channel allocation is learned according to the policy learned from the ANN. Simulation results indicate that the proposed method can achieve the same performance for six users with lower computational complexity in comparison to the ES. In [12], a DQN is adopted to achieve a joint RA for a large number of users, where reinforcement learning (RL) frames are used for the joint optimization of clustering, power allocation, and beamforming. Specifically, the clustering is adjusted using a DQN, and a backpropagation neural network (BPNN) is designed to learn the power coefficients for each cluster. Simulation results indicate that the performance of the proposed method is close to that of the ES. In [13], a method based on DRL is proposed to assign users to time-varying channels. Recently, DRL has been expected to be applied as a solution for such real-time dynamic decision-making problems in practical environments

where the conventional RA algorithms are difficult to be applied. A complexity analysis and simulation results show that the proposed approach provides a superior performance with lower computational complexity in comparison to conventional optimization algorithms.

Therefore, ML-based RA methods can obtain near-optimal solutions with lower computational complexity than the ES. However, to obtain near-optimal solutions, a proper training time is required.

C. Hardware-Based RA

Optimization problems such as channel and power allocation are important not only in NOMA systems but also in a variety of other wireless communication systems. Optimization problems in wireless communication systems have recently become increasingly complex, and the demand for efficient solutions to meet the requirements of next-generation wireless communications has increased [26]. Hardware-based approaches, such as quantum computers and Ising machines, have been expected to be potential methods for solving optimization problems at high speed [14], [15], [16], [17]. In previous studies, several RA optimization problems in wireless communications have been studied using the CIM [21], [22], [23], [24]. In [22], the CIM is applied to RA problems in wireless communication systems such as IEEE802.11 wireless LAN (WLAN) systems and D2D communications. In addition, it has been shown that the CIM can approximately solve these optimization problems within the order of milliseconds. In [23], the CIM is applied to the optimization of scheduling problems for distributed antenna systems. In [24], the CIM is applied to the optimization problem of channel assignment in dense WLAN systems. In this study, the maximization of the system throughput in high-density WLAN systems using the CIM is considered. The performance of the proposed CIM-based channel assignment approach in comparison to other optimization approaches such as the SA and greedy algorithms is evaluated. Simulation results indicate that the CIM can obtain the optimal solution well with constant computational time, whereas other optimization algorithms increase the computational time and decrease the accuracy of the obtained solution as the number of users in the system increases.

In these studies, the CIM is assumed to work either as a base station (BS) or as a controller in the centralized systems. The

ability of the CIM to operate at room temperature [27] and its simple components make it suitable for integration into many wireless transmission systems. Hence, the CIM is significantly more suitable for large-scale wireless networks if the size of the problem is up to the number of spins that can be achieved with the CIM.

In summary, the CIM may obtain optimal solutions to combinatorial optimization problems without the need for complex computer calculations. The principle of searching for optimal solutions using the CIM is based on the spontaneous energy reduction in physical phenomena. Existing research has demonstrated the higher speed of a CIM-based optimization. On the contrary, DRL or other ML algorithms require training samples to obtain the optimal solution, and trial and error to increase performance. In a real communication environment, such as a vehicle network, the environment changes frequently, and real-time optimization may not be possible for the ML algorithms because retraining is needed when the environment changes. In [4], [13], the complexity due to training in the ML algorithm is studied. In contrast to the complexity of training presented in these papers, the proposed CIM-based RA method can always solve the optimization problem at constant high speed, i.e., millisecond order [25]. Therefore, real-time RA can be achieved using our proposed method, which may improve the spectrum efficiency under dynamic changing environments.

III. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a downlink NOMA system where the BS transmits data to multiple users, as shown in Fig. 1.

Both the BS and users have a single antenna. At the BS, the transmitter sends signals with multiplexing the data of different users on the same channel. Assume that the signal to each user is multiplexed in one channel at the same time. At the receiver side, the SIC technology is used to restore the signal of each user. The BS is equipped with the CIM. Once the required communication information is gathered in the BS side, the channel will be allocated by the CIM. Let the set of users in the NOMA system be $\mathbf{U} = \{1, 2, \dots, u, \dots, N_u\}$ and the set of sub-channels be $\mathbf{C} = \{1, 2, \dots, c, \dots, N_c\}$, where N_u and N_c are the numbers of the users and channels in the NOMA system. Note that, $N_u > 2$. Assume that the total bandwidth is B while the bandwidth is divided into N_c sub-channels equally. Hence, the bandwidth of the c th channel is $B_c = B/N_c$. The multiplexed signal on the c th channel can be expressed as follows:

$$x_c = \sum_{i=1}^{M_c} \sqrt{P_i^c} b_i, \quad (1)$$

where M_c is the number of users multiplexed on the c th channel, P_i^c is the power allocated to the i th user transmitted on the c th channel, and b_i is the transmission symbols of the i th user. At the receiver side, the multiplexed signal of the u th user in the c th channel can be expressed as follows:

$$y_u^c = \sqrt{P_u^c} h_u^c b_u + \sum_{\substack{i=1 \\ i \neq u}}^{M_c} \sqrt{P_i^c} h_i^c b_i + z_u^c, \quad (2)$$

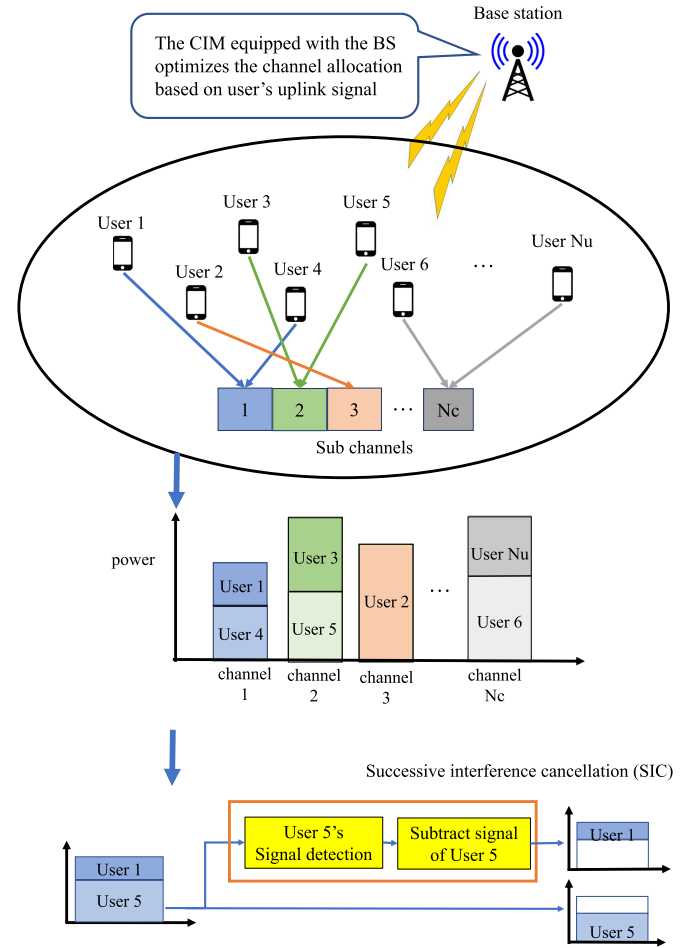


Fig. 1. System model.

where z_u^c is additive white Gaussian noise with zero mean and variance $\sigma_{z_c}^2$. h_u^c is the channel response between the BS and the u th user, which can be calculated as follows:

$$|h_u^c|^2 = (g_u^c)^2 d_u^{-\alpha}, \quad (3)$$

where g_u^c is a coefficient that follows a Rayleigh distribution, d_u is the distance between the BS and the u th user, and α is the path loss coefficient. The carrier-to-noise ratio (CNR) in the c th channel of the u th user is defined as follows:

$$\Gamma_u^c = \frac{|h_u^c|^2}{\sigma_{z_c}^2}. \quad (4)$$

Herein, we assume that the CNR of each user assigned to a channel c is $\Gamma_1^c > \dots > \Gamma_u^c > \dots > \Gamma_{M_c}^c$. In the NOMA protocol, higher power is allocated to the users with lower CNR [5], [28], i.e., $P_1^c < \dots < P_u^c < \dots < P_{M_c}^c$. In addition, SIC is performed at the receiver side, which successively removes the signal of the user with the higher power. Therefore, the signal of the u th user can be decoded by removing the signal of the i th user when $P_i^c > P_u^c$. The signal of the u th user is considered as the interference of the i th user at that time. Based on the principle discussed above, when the SIC is fully working in the power domain NOMA, i.e., when there is a sufficient difference in gain between users that are multiplexed in the same channel, the signal-to-interference-plus-noise ratio of the u th user can be

calculated as follows:

$$\gamma_u^c = \frac{P_u^c \Gamma_u^c}{1 + \sum_{i=1}^{u-1} P_i^c \Gamma_u^c}. \quad (5)$$

The data rate achievable by the u th user in the c th channel can then be calculated as follows:

$$R_u^c = B_c \log_2 \left(1 + \frac{P_u^c \Gamma_u^c}{1 + \sum_{i=1}^{u-1} P_i^c \Gamma_u^c} \right). \quad (6)$$

Assuming that the total transmission power from the BS is P_T , the relation between the power coefficient P_u^c and total transmission power P_T can then be expressed as follows:

$$\sum_{u=1}^{N_u} \sum_{c=1}^{N_c} P_u^c \leq P_T. \quad (7)$$

As the number of users assigned to the same channel increases, the complexity of SIC decoding at the receiver side increases, which leads to high complexity of the power allocation scheme to efficiently avoid SIC errors [5], [13]. SIC and power allocation are needed to be carefully considered if we assume that more than 2 NOMA access the same channel simultaneously, which is beyond the scope of this article. The contribution of this article is to achieve fast RA using the CIM for NOMA systems. Therefore, we focus on introducing how to use the CIM to solve the RA problem in NOMA. In other words, we consider the scenario where $M_c = 2$ for $\forall c \in \mathbf{C}$. Thus, the data rate that can be achieved by any user in any channel can be expressed

$$R_{NOMA}^{ijk} = \begin{cases} B_c \log_2 \left(1 + P_{ik}^j \Gamma_i^j \right), & \text{if } \Gamma_i^j > \Gamma_k^j, \\ B_c \log_2 \left(1 + \frac{P_{ik}^j \Gamma_i^j}{1 + P_{ki}^j \Gamma_i^j} \right), & \text{otherwise,} \end{cases} \quad (8)$$

where for $i \in \mathbf{U}$, $k \in \mathbf{U}$, and $j \in \mathbf{C}$, R_{NOMA}^{ijk} is the data rate achieved by the i th user when multiplexed with the k th user in the j th channel, and P_{ik}^j is the transmission power allocated to the i th user when the i th user is multiplexed into the j th channel with the k th user. In addition, when there is only one user i is assigned to channel j , the data rate of the i th user can be expressed as follows:

$$R_{OMA}^i = B_c \log_2 \left(1 + P_i^j \Gamma_i^j \right), \quad (9)$$

where P_i^j is the transmit power allocated to the i th user on the j th channel. Here, the expression of the relationship between the number of users and that of channels below must be stratified to transform the formulated problem into an Ising problem. That is, $N_u = 2N_c$. However, we assume that up to 2 users can be multiplexed to the same channel, which means that one or no users may be assigned to the channel. The transmissions will be OMA when only one user is assigned to the channel. Therefore, in addition to the user set \mathbf{U} , we define the dummy user set as $\mathbf{D} = \{1, \dots, N_d\}$. If the user from set \mathbf{U} is multiplexed with the user from set \mathbf{D} , OMA transmission will be carried out by the user from set \mathbf{U} , which means that the user in the set \mathbf{D} is not a really existing user in the real world. As described above, we can understand that the total number of users and that of the dummy users is twice the number of channels i.e., $N_d + N_u = 2N_c$. Hence, the throughput achieved by users in the NOMA system

can be redefined as follows:

$$R_{ik}^j = \begin{cases} R_{NOMA}^{ijk}, & \text{if } i, k \in \mathbf{U}, \\ R_{OMA}^i, & \text{if } i \in \mathbf{U}, k \in \mathbf{D}, \\ 0, & \text{if } i \in \mathbf{D}, \forall k. \end{cases} \quad (10)$$

In this article, our goal is to maximize the total data rate of the NOMA system. The objective function of this article is expressed as follows:

$$\max_x \sum_{i=1}^{N_u+N_d} \sum_{j=1}^{N_c} \sum_{\substack{k=1 \\ k \neq i}}^{N_u+N_d} \left(R_{ik}^j + R_{ki}^j \right) x_{ij} x_{kj} \quad (11)$$

$$s.t. \quad R_{ik}^j \geq (R_{ij})_{\min}, \text{ for } \forall k, \quad (11.a)$$

$$\sum_{i=1}^{N_u+N_d} \sum_{j=1}^{N_c} \left(P_{ik}^j + P_{ki}^j \right) x_{ij} x_{kj} \leq P_T, \quad (11.b)$$

$$\sum_{j=1}^{N_c} x_{ij} = 1, \text{ for } \forall i, \quad (11.c)$$

$$\sum_{i=1}^{N_u+N_d} x_{ij} = 2, \text{ for } \forall j, \quad (11.d)$$

where x_{ij} represents the channel assignment variable, which can be expressed:

$$x_{ij} = \begin{cases} 1, & \text{if the } i\text{th user uses the } j\text{th channel,} \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

Constraint (11.a) indicates the minimum data rate that the user should achieve. In addition, $(R_{ij})_{\min}$ in constraint (11.a) is the QoS constraint, which represents the minimum data rate that the i th user using the j th channel must achieve. Constraint (11.b) indicates the power constraint, i.e., the total power allocated to the users should not exceed the total transmission power from the BS. Constraints (11.c) and (11.d) indicate that one user, and up to two users, can be multiplexed into a single channel, respectively.

In this article, power allocation is optimized using the method in [5], whereas channel allocation is optimized using the CIM. Specifically, we maximize (11) while satisfying constraints (11.c) and (11.d) using the CIM, and the power factor considering constraints (11.a) and (11.b) is then optimized using the method described in [5]. The method of power allocation in [5] can maximize the total data rate while satisfying the QoS constraints and total power constraint, i.e., constraints (11.a) and (11.b) can be satisfied by allocating power using the method described in [5]. Here, we assume that $\Gamma_i^j > \Gamma_k^j$ and $i, k \in \mathbf{U}$. Then, the optimal P_{ik}^j and P_{ki}^j in (11) can be derived as

$$P_{ik}^j = \frac{\Gamma_k^j q_j - A_k^j + 1}{A_k^j \Gamma_k^j},$$

$$P_{ki}^j = q_j - P_{ik}^j, \quad (13)$$

where $A_k^j = 2^{\frac{(R_{kj})_{\min}}{B_c}} > 2$, and q_j is the transmission power from the BS allocated to the j th channel. By using this equation, the signals can be successfully decoded as avoiding SIC failure conditions is considered. Note that if user $i \in \mathbf{U}$ is multiplexed

with dummy user $k \in \mathbf{D}$ for OMA communication, P_{ik}^j is set as a constant q_j because the power allocated to the dummy user is zero. The value of q_j maximizing the total data rate can be obtained using the following equations:

$$q_j = \left[\frac{B_c}{\Lambda} - \frac{A_k^j}{\Gamma_i^j} + \frac{A_k^j}{\Gamma_k^j} - \frac{1}{\Gamma_k^j} \right]_{\gamma_j}^{\infty},$$

$$\gamma_j = \frac{A_k^j (A_i^j - 1)}{\Gamma_i^j} + \frac{A_k^j - 1}{\Gamma_k^j}, \quad (14)$$

where Λ needs to be selected to satisfy $\sum_{j=1}^{N_c} q_j = P_T$. $[\bullet]$ is a well-known water-filling form [5]. It means the power allocation solution for the formulated combinatorial optimization problem, i.e., (11) is uniquely determined if q_j is in the range $[\gamma_j, +\infty)$. In the following, we introduce the optimal channel allocation using the CIM.

IV. COHERENT ISING MACHINE

In this section, we first introduce the Ising Hamiltonian. Then, we present the measurement feedback CIM, which is a measurement-feedback type CIM that has been proposed to easily increase the number of couplings between spins to solve larger-scale problems. Finally, we investigate the operation process of the measurement feedback CIM.

A. Ising Hamiltonian

The CIM is an Ising-type machine that can artificially reproduce an Ising model, a physical model of magnetic spins. The Ising model consists of two states of magnetic spins: upward and downward. Spins are mutually coupled and subject to interactions from other spins and effects from external magnetic fields. We consider Ising spins with an $N \times M$ two-dimensional structure. Here, the energy function of the Ising model, i.e., the Ising Hamiltonian is expressed as follows:

$$E(\sigma) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{k=1}^N \sum_{l=1}^M J_{ijkl} \sigma_{ij} \sigma_{kl} + \sum_{i=1}^N \sum_{j=1}^M \lambda_{ij} \sigma_{ij}, \quad (15)$$

where $\sigma_{ij} \in \{-1, +1\}$ is the spin direction of the (i, j) th spin, J_{ijkl} is the strength of the interaction between the (i, j) th and (k, l) th spins, and λ_{ij} is the strength of the external magnetic field on the (i, j) th spin. Here, the CIM is the machine that can obtain the ground state of the Ising Hamiltonian at high speed. In other words, by setting J_{ijkl} and λ_{ij} to correspond to the optimal solution of the optimization problem to the minimum of (15), we can approximately obtain a fast optimal solution using the CIM.

B. Measurement Feedback Coherent Ising Machine

Initially, the CIM was a laser network consisting of one master laser and multiple slave lasers [29]. Currently, a measurement-feedback type CIM has been proposed to easily increase the number of couplings between spins to solve larger scale problems [15], [16]. A schematic diagram of the measurement feedback-type CIM is shown in Fig. 2. As shown in Fig. 2,

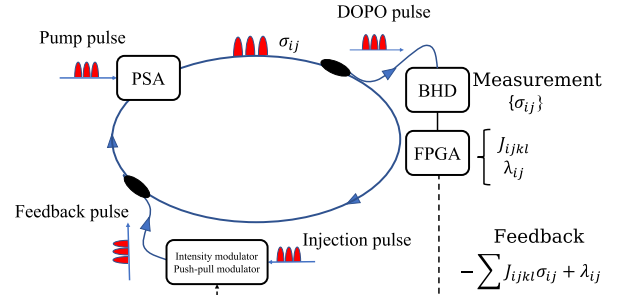


Fig. 2. Coherent Ising Machine [15], [16].

parameters of the Ising model J_{ijkl} and the λ_{ij} , are pre-configured in a field programmable gate array (FPGA) module. The FPGA module calculates the feedback value using the pre-configured parameters and feeds it back to the original pulse. Therefore, interactions between spins can be programmatically reproduced, which makes it possible to fully connect spins. In the measurement-feedback CIM, the phase of the degenerate optical parametric oscillator (DOPO) pulse circulating on an optical fiber represents Ising spins σ_{ij} . In [18], it is shown that 100,000 DOPO pulses can be realized over 5 km of polarization-maintaining optical fiber and reach the reference score of the MAX-CUT problem for 100,000 nodes within a time of 593 μ s.

C. Operation Process of the Measurement Feedback CIM

The CIM is an Ising machine that can rapidly obtain the ground state of the Ising Hamiltonian by getting the combinations of the Ising spin directions. The set of the Ising spin directions that minimize the Ising Hamiltonian corresponds to the solution of the combinatorial optimization problem. The operation process of the CIM can be summarized as follows. As shown in Fig. 2, the initial pulses are first generated by a phase sensitive amplifier (PSA) on the optical fiber. The initial pulses are repeatedly amplified by PSA with a pump pulse each cycle and become a DOPO pulse. The phase of the DOPO pulse circulating on the fiber during oscillation is used to reproduce the Ising spin on the optical fiber in the CIM. Here, gradually increasing the pump pulse amplitude until the threshold is exceeded; the DOPO pulse will oscillate and take on a π or 0 phase representing the orientation of the Ising spin σ_{ij} , where $\sigma_{ij} \in \{-1, +1\}$. 0 phase represents the up spin, i.e., $\sigma_{ij} = 1$. Meanwhile, the π phase represents the down spin, i.e., $\sigma_{ij} = -1$ while the DOPO pulse is being amplified, a part of the DOPO pulse on the fiber is separated by a coupler, and the in-phase amplitude is measured through a balanced homodyne detection (BHD). The measured in-phase amplitude σ_{ij} is input into the FPGA module. A feedback signal $-\sum J_{ijkl} \sigma_{ij} + \lambda_{ij}$ is calculated using the mutual coupling between Ising spins J_{ijkl} and external magnetic field λ_{ij} . Finally, the feedback signal is injected into the original DOPO pulse as a feedback pulse through the intensity modulator and push-pull modulator. By repeating the operation process of the CIM in this manner, we can eventually obtain a combination of σ_{ij} that minimizes (15), which corresponds to the optimal solution to the optimization problem.

From the above, to obtain the optimal solution to the optimization problem using the CIM, we need to transform the objective function into the form of an Ising Hamiltonian, and then derive the corresponding J_{ijkl} and λ_{ij} of the objective function. Then, by pre-configuring the derived these parameters into the FPGA module of the CIM, we can search for the ground state of the Ising Hamiltonian, which corresponds to the optimal solution to the optimization problem. In the next section, we introduce the problem transformation and the derivation of the corresponding J_{ijkl} and λ_{ij} for our formulated objective function in (11).

V. APPLYING THE CIM TO THE RA PROBLEM IN NOMA

To solve the combinatorial optimization problem in the CIM, we must set J_{ijkl} and λ_{ij} in the FPGA module of the CIM. However, it is difficult to transform the formulated objective function, i.e., (11), directly into an Ising Hamiltonian using the $\{-1, +1\}$ variables corresponding to the value of the Ising spin of the CIM. We therefore first express (11) using binary variables of $\{0, 1\}$, and then transform it into the form of an Ising Hamiltonian using the $\{-1, +1\}$ variables. In this study, we use a mutually connected neural network to represent (15) using binary variables $\{0, 1\}$. Mutually connected neural networks have been proposed as algorithms for solving various minimization problems. In practice, the traveling salesman problem is solved using this type of neural network in [30], and can also be applied to the problem defined in (11). The mutually connected neural network and the Ising Hamiltonian have similar energy structures; therefore, after formulating (11) in the mutually connected neural network, we can derive the parameters of the Ising Hamiltonian, J_{ijkl} and λ_{ij} , by converting the output from $\{0, 1\}$ into $\{-1, +1\}$. In the following, we first introduce mutually connected neural networks. We then present how to solve the RA in a NOMA system using a mutually connected neural network. Next, we present the RA in a NOMA system using the CIM. Finally, the time and computational complexity of the proposed method are discussed.

A. Mutually Connected Neural Network

Let us consider a neural network with a two-dimensional structure of $N \times M$; the mutually connected neural network has the following energy structure:

$$E(x) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{k=1}^N \sum_{l=1}^M w_{ijkl} x_{ij} x_{kl} + \sum_{i=1}^N \sum_{j=1}^M \theta_{ij} x_{ij}, \quad (16)$$

where $x_{ij} \in \{0, 1\}$ is the state of the (i, j) th neuron, w_{ijkl} is the coupling weight between the (i, j) th and (k, l) th neurons, and θ_{ij} is the firing threshold of the (i, j) th neuron. From the similarity of the energy structure of the Ising Hamiltonian and that of the mutually connected neural network, i.e., (15) and (16), the parameters of the CIM, J_{ijkl} and λ_{ij} , can be derived using the parameters (i.e., w_{ijkl} and θ_{ij}) of the mutually coupled neural network. Here, the state of a neuron in a network is defined by

the following equation:

$$x_{ij}(t+1) = u \left[\sum_{k=1}^N \sum_{l=1}^M w_{ijkl} x_{kl} - \theta_{ij} \right], \quad (17)$$

where u is the Heaviside step function, namely, $u[y] = 0$ for $y \leq 0$, and $u[y] = 1$ for $y > 0$. The energy function of the mutually connected neural network, i.e., (16) always decreases and converges to a locally minimum value by updating the neurons iteratively according to (17) if the following conditions are satisfied. 1). The self-connections of all neurons are zero, i.e., $w_{ijij} = 0$. 2). The mutual connections are symmetric, i.e., $w_{ijkl} = w_{klij}$.

To solve the optimization problem using a mutually connected neural network, we must formulate the optimization problem as an objective function using a neuron x_{ij} . Then, the formulated objective function is compared with (16) to derive w_{ijkl} and θ_{ij} . By iteratively updating the neurons according to (17) using the derived parameters w_{ijkl} and θ_{ij} , (16) converges to a local minimum value, and the states of the neurons at that time correspond to the approximate optimal solution to the optimization problem. In the next subsection, we define neurons x_{ij} and derive w_{ijkl} and θ_{ij} for optimizing the RA in a NOMA system.

B. Mutually Connected Neural Network to Solve RA in NOMA System

To optimize the RA problem in a NOMA system using a mutually connected neural network, we define the binary variables in (12) as neurons. Then, our formulated object function (11) can be transformed into the energy function of the mutually connected neural network, which is expressed as follows:

$$\begin{aligned} E_1 &= \sum_{i=1}^{N_u+N_d} \sum_{j=1}^{N_c} \sum_{\substack{k=1 \\ k \neq i}}^{N_u+N_d} - \left(R_{ik}^j + R_{ki}^j \right) x_{ij} x_{kj} \\ &= \sum_{i=1}^{N_u+N_d} \sum_{j=1}^{N_c} \sum_{k=1}^{N_u+N_d} \sum_{l=1}^{N_c} -\delta_{jl} (1 - \delta_{ik}) \left(R_{ik}^j + R_{ki}^l \right) x_{ij} x_{kl}, \end{aligned} \quad (18)$$

where δ_{ij} is Kronecker's delta, namely, $\delta_{ij} = 1$ when $i = j$, and $\delta_{ij} = 0$, otherwise. In addition, the constraints (11.c) and (11.d) can be formulated as constraint terms using neurons, which can be expressed as follows:

$$\begin{aligned} E_2 &= \sum_{i=1}^{N_u+N_d} \left(\sum_{j=1}^{N_c} x_{ij} - 1 \right)^2 \\ &= \sum_{i=1}^{N_u+N_d} \sum_{j=1}^{N_c} \sum_{k=1}^{N_u+N_d} \sum_{l=1}^{N_c} \delta_{ik} (1 - \delta_{jl}) x_{ij} x_{kl} \\ &\quad - \sum_{i=1}^{N_u+N_d} \sum_{j=1}^{N_c} x_{ij}, \end{aligned} \quad (19)$$

$$\begin{aligned}
E_3 &= \sum_{j=1}^{N_c} \left(\sum_{i=1}^{N_u+N_d} x_{ij} - 2 \right)^2 \\
&= \sum_{i=1}^{N_u+N_d} \sum_{j=1}^{N_c} \sum_{k=1}^{N_u+N_d} \sum_{l=1}^{N_c} \delta_{jl} (1 - \delta_{ik}) x_{ij} x_{kl} \\
&\quad - 3 \sum_{i=1}^{N_u+N_d} \sum_{j=1}^{N_c} x_{ij}. \tag{20}
\end{aligned}$$

Note that the above constraints term can be fully satisfied even if the number of users in the cell N_u is less than $2N_c$, since we define the number of the dummy users N_d to always satisfy $N_u + N_d = 2N_c$. Using E_1 , E_2 , and E_3 , the energy function of the RA optimization problem in a NOMA system can be obtained as follows:

$$E_{NOMA} = \epsilon E_1 + \zeta E_2 + \eta E_3, \tag{21}$$

where ϵ , ζ , and η are parameters that adjust the scaling of each term, which are used to adjust the objective and constraint terms, causing the energy structure to more likely reach the optimal solution. By comparing (21) with (16), the coupling weight w_{ijkl} and firing threshold θ_{ij} for solving the RA optimization problem in a NOMA system using mutually connected neural networks can be obtained as follows:

$$w_{ijkl} = 2\{\epsilon\delta_{jl}(1 - \delta_{ik})(R_{ijk} + R_{kli}) - \zeta\delta_{ik}(1 - \delta_{jl}) - \eta\delta_{jl}(1 - \delta_{ik})\}, \tag{22}$$

$$\theta_{ij} = -(\zeta + 3\eta). \tag{23}$$

Using w_{ijkl} and θ_{ij} obtained in (22) and (23), the neurons were updated according to (17). Thereby, the combination of neurons minimizes the energy function in (16), i.e., the optimal solution to the NOMA RA problem can be obtained. However, owing to the monotonically decreasing nature of the energy function, this neural network is likely to fall into a local minimum of the network state. Hence, the optimal solutions may not be achieved using the mutually connected neural network. Further, it has been shown that the Ising Hamiltonian near the optimal solution can be reached with high probability by amplifying the DOPO pulse to near the oscillation threshold in CIM [31]. Therefore, in the next subsection, we will use the derived parameters w_{ijkl} and θ_{ij} to derive J_{ijkl} and λ_{ij} which are the parameters used to solve the optimization problem in the CIM.

C. CIM Used to Solve RA in NOMA System

As previously noted, the Ising Hamiltonian of the CIM and the energy function of the mutually connected neural network have similar energy structures. As a difference, the neuron $x_{ij} \in \{0, 1\}$ in the mutually connected neural network has states of 0 and 1, whereas the Ising-spin $\sigma_{ij} \in \{-1, +1\}$ in the Ising model has states of -1 and $+1$. To solve the RA problem using the CIM, we redefine the output of the neurons in a mutually connected neural network as -1 and $+1$. Using the neuron $\tilde{x}_{ij} \in \{-1, +1\}$ with a redefined output, the update equation for the neuron is

expressed as follows:

$$\tilde{x}_{ij}(t+1) = r \left[\sum_{k=1}^N \sum_{l=1}^M \tilde{w}_{ijkl} \tilde{x}_{kl} - \tilde{\theta}_{ij} \right], \tag{24}$$

where $r[y] = -1$ for $y \leq 0$, and $r[y] = +1$ for $y > 0$, and \tilde{w}_{ijkl} and $\tilde{\theta}_{ij}$ are the coupling weight and firing threshold of the neural network with output $\{-1, +1\}$, respectively. We then can obtain the following equation by transforming the internal state of (17) using $\tilde{x}_{ij} = 2x_{ij} - 1$:

$$\begin{aligned}
\sum_{k=1}^N \sum_{l=1}^M w_{ijkl} x_{kl} - \theta_{ij} &= \sum_{k=1}^N \sum_{l=1}^M \frac{w_{ijkl}}{2} \tilde{x}_{kl} \\
&\quad - \left(\theta_{ij} - \sum_{k=1}^N \sum_{l=1}^M \frac{w_{ijkl}}{2} \right). \tag{25}
\end{aligned}$$

By comparing (16) and (25), we can obtain the following equations for \tilde{w}_{ijkl} and $\tilde{\theta}_{ij}$, respectively.

$$\tilde{w}_{ijkl} = J_{ijkl} = \frac{w_{ijkl}}{2}, \tag{26}$$

$$\tilde{\theta}_{ij} = \lambda_{ij} = \theta_{ij} - \sum_{k=1}^N \sum_{l=1}^M \frac{w_{ijkl}}{2}. \tag{27}$$

By setting these derived parameters in the FPGA module, a fast RA optimization of the NOMA system using the CIM can be performed.

D. RA Optimization Procedure of the NOMA System Using the CIM

In this subsection, we present the RA optimization procedure using the CIM in NOMA systems. First, we give an overview of the RA optimization procedure using the CIM. Then, we introduce the proposed solution construction algorithm for channel assignment in detail.

1) *Overview of the RA Optimization Procedure Using the CIM:* In this study, the BS performs channel allocation using the CIM for the NOMA system to achieve the optimal total reachable data rate. A schematic of the optimization procedure of the NOMA system is shown in Fig. 3. For simplicity, we present a NOMA system with 4 users and 2 channels as an example. As shown in Fig. 3, each user transmits the data to the BS. Then, the channel assignment is determined by the CIM equipped in the BS. There are multiple DOPO pulses on the optical fiber of the CIM, each of which behaves as an Ising spin. Because this problem involves 4 users and 2 channels, the RA problem can be solved using the CIM implemented with $2 \times 4 = 8$ DOPO pulses on the fiber. When the CIM is operated, the states of the in-phase components shown on the left side of Fig. 3 are obtained. Here, the cycle number represents the number of times the DOPO pulse travels around the optical fiber of the CIM. As the cycle count increases, the in-phase component of the amplitude of each DOPO pulse increases and eventually splits into positive or negative values, i.e., phase 0 or phase π . In the CIM, the spin orientation is determined by the in-phase components of the amplitude. Here, when user i uses channel j , the spin is $+1$, i.e., an upward spin. Thus, the obtained spin combination corresponds to the channel assignment. Finally,

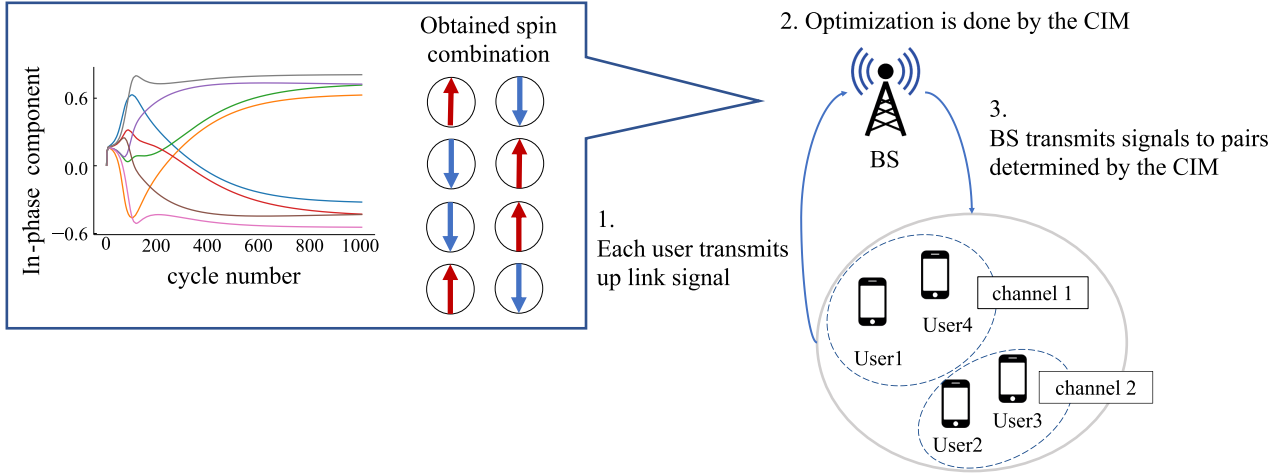


Fig. 3. Overview of our proposed CIM-based RA method.

Algorithm 1: Optimization of RA Using the CIM in NOMA System.

Input: Distance between user i and the BS d_i , and the channel coefficient of user i in channel j g_i^j

Output: channel allocation σ_{ij} , power allocation q_j , P_{ik}^j .

1: **Initialize:**

2: Initialize the channel response h_i^j and the CNR Γ_i^j according to (3), (4)

3: Set the power allocation P_{ik}^j according to (13) with $q_j = 1.0 \forall j$.

4: Calculate data rate for users according to (10).

5: **Channel allocation by the CIM:**

6: Calculate the J_{ijkl} and λ_{ij} according to (26), (27) and setting these parameters to FPGA.

7: Obtaining combination of in-phase amplitude c_{ij} by the CIM as described in Section IV.

8: Getting σ_{ij} indicating that ground state of the Ising Hamiltonian by applying solution construction algorithm.

9: **Optimal power allocation:**

10: q_j is obtained by water filling algorithms [5] and re-calculate P_{ik}^j .

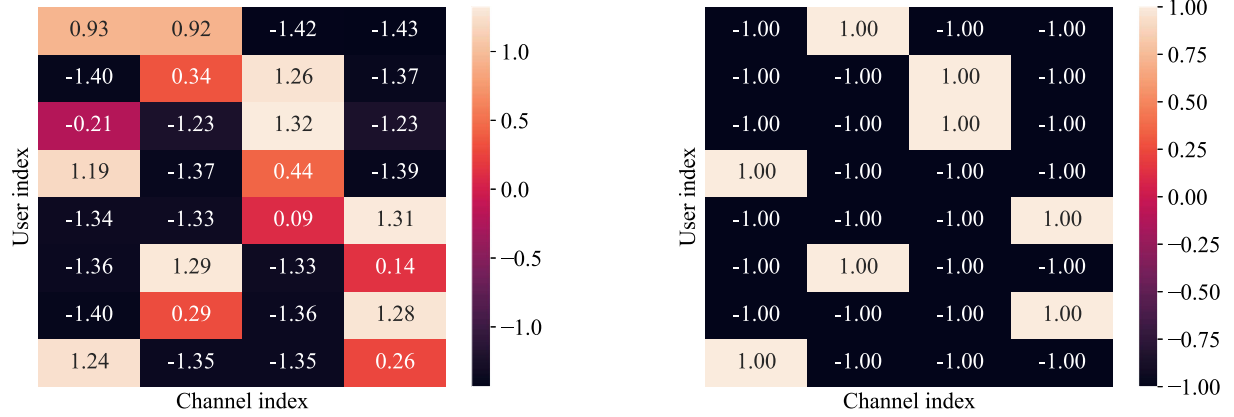
based on this channel assignment, the BS sends a signal to the users after assigning the optimal power using (13). The pseudo-code for the algorithm is presented as Algorithm 1.

2) *Solution Construction Algorithm for Channel Assignment:* In the CIM-based RA method, the solution of the RA problem in the NOMA systems is represented by positive and negative values of the in-phase amplitudes of the DOPO pulses. However, since the CIM is hardware that generates in-phase amplitude based on the physical phenomenon, which is unstable because of the fluctuation of experimental parameters, it is difficult to guarantee that solutions can always satisfy the constraints in the formulated objective function all the time. Specifically, the Ising model may converge to a spin group that could not satisfy (11.c) nor (11.d), and this spin group cannot be interpreted as the

solution of our formulated RA optimization problem to calculate the data rate. To guarantee the constraints are satisfied, we propose a solution construction algorithm. By this, the obtained combinations of in-phase amplitudes can be interpreted as spin groups that always satisfy the constraints. The main idea of the solution construction algorithm is that spins are constructed according to the ordinal order of magnitude of the amplitude values. Specifically, the in-phase amplitude values are obtained from the CIM operation first. Then, for each user i , the spin σ_{ij} corresponding to the largest in-phase amplitude c_{ij} is constructed as +1. That is, user i is allocated to channel j . The above process continues until two users, possibly including dummy users, are allocated to channel j . After two users are allocated to channel j , the spins σ_{il} for all l but j are constructed as -1. Fig. 4 shows an example of applying the solution construction algorithm for channel assignment in a NOMA system with 8 users, possibly including a dummy user, and 4 channels. Fig. 4(a) is the in-phase amplitude values obtained from the CIM operation before applying the solution construct algorithm, which could not satisfy the constraints (11.c) and (11.d), i.e., each user can only access one channel while at most two users can access the same channel. Fig. 4(b) shows the constructed spin groups obtained by applying the solution construct algorithm, which can satisfy the constraints (11.c) and (11.d). As described above, by using the solution construction algorithm, the combination of assignment variable, i.e., the values of the spins that satisfy the constraints (11.c) and (11.d) (shown in Fig. 4(b)), can be constructed based on the in-phase amplitudes obtained as result of the CIM operation (shown in Fig. 4(a)). The solution construction algorithm is summarized in Algorithm 2.

E. Complexity Analysis of the CIM

In this subsection, we discuss the time and computational complexity of the proposed method. In [25], it is demonstrated that the CIM takes a constant amount of time to obtain the ground state as long as the fiber length is constant. That is, the time complexity of the CIM is $\mathcal{O}(1)$ experimentally. The reason is that the computation time of one cycle for the CIM is related



(a) In-phase amplitude values before applying the solution construction algorithm

(b) Spin groups constructed by applying the solution construction algorithm

Fig. 4. Examples of applying solution construction algorithm for channel assignment. (a) In-phase amplitude values before applying the solution construction algorithm. (b) Spin groups constructed by applying the solution construction algorithm.

Algorithm 2: Solution Construction Algorithm.

Input: In-phase amplitude c_{ij} obtained as a result of the CIM operation, where $i \in \mathbf{U} \cup \mathbf{D}$ and $j \in \mathbf{C}$.

Output: Combination of assignment variable σ_{ij} that satisfies the constraints (11.c), (11.d).

1: **Initialize:**

2: Variable that counts up the number of users multiplexed to the j th channel: $count_j = 0, \forall j$

3: **Solution construction:**

4: **While** $count_j < 2, \forall j$ **do**

5: **for** user index i **do**

6: Find the largest c_{ij} and get the indices (i, j)

7: Update $\sigma_{ij} = +1$ and $count_j = count_j + 1$

8: Set $c_{kj} = -1.0$ and $\sigma_{kj} = -1$ for all k but i

9: **if** $count_j == 2$ **then**

10: Set $c_{il} = -1.0$ and $\sigma_{il} = -1$ for all l but j

to the length of a fiber. For the CIM with 2000 spins and 1 km fiber, 1000 cycles running of the CIM can be achieved within 5 ms while the spins can converge to the ground state within 1000 cycles by amplifying pump pulse amplitude appropriately. In our article, the problems can be solved using the CIM with 2000 spins. Hence, the computation time is within 5 ms.

Then, we discuss the computation complexity of the main computation part of the CIM, i.e., FPGA. Ref. [16] shows that the memory resources required for matrix computation and power consumption in FPGAs are scaled by $\mathcal{O}(N^2)$. A typical CPU runs 10–100Gflop per second, so its power efficiency is often below 1 Gop/J. Similarly, GPUs also suffer from high power consumption. Hence, CPUs and GPUs are difficult to apply to applications that require low power consumption [32]. In addition, although parallel processing is possible for modern CPUs with multi-core, unnecessary parts also need to be parallelized since the entire core is used. In the case of FPGA, only necessary processing needs to be programmed. Hence, parallel processing can be performed without wasting computation resources,

increasing the computation efficiency. Moreover, owing to the merit of the parallel processing of the FPGA, the FPGA uses less circuit area when performing the same processing compared to the CPU/GPU and can perform processing without resource waste. It is natural with high power efficiency. In summary, our proposed method may be superior to the resource allocation method computed using the traditional computer in computation complexity.

VI. PERFORMANCE EVALUATION

In this section, we evaluate the performance in the convergence, data rate, and computation time of our proposed CIM-based RA for a NOMA system and compare it with the SA (a meta-heuristic optimization method), a conventional-NOMA pairing scheme (C-NOMA) [6], a DRL and the ES methods. Specifically, the RA is applied as follows. First, the channels are allocated using each method. The optimal power allocation is then applied using (13) and (14). The following of this section is organized as follows. First, the simulation settings are introduced. Then, the comparison methods are presented. Next, the performance in convergence, data rate, computation time of our proposed method and the comparison with the other methods are evaluated.

A. Simulation Settings

In the simulation, a circular cell of a NOMA system with a radius of 500 m is considered. A BS is placed at the center of the cell. Users are randomly placed in the cell. Table II shows the detailed parameter setting used in our simulation.

We use the following simulation model of the CIM to evaluate our proposed method:

$$\frac{dc_{ij}}{dt} = (-1 + p - c_{ij}^2 - s_{ij}^2) c_{ij} + \sum_{k=1}^N \sum_{l=1}^M J_{ijkl} c_{kl} - \lambda_{ij}, \quad (28)$$

TABLE II
SIMULATION PARAMETERS

Parameters	Values
Total Bandwidth B	5.0 MHz
Minimal data rate of each users R_{min}	2 bps/Hz
Cell radius	500 m
Minimal distance between user and BS	50 m
Maximum number of multiplexes users	2 users
Noise power spectral density	-170 dBm/Hz

$$\frac{ds_{ij}}{dt} = (-1 - p - c_{ij}^2 - s_{ij}^2) s_{ij} + \sum_{k=1}^N \sum_{l=1}^M J_{ijkl} s_{kl} - \lambda_{ij}, \quad (29)$$

where c_{ij} and s_{ij} are the in-phase and quadrature-phase amplitude of the (i, j) th optical pulses, respectively, and p is the pump pulse, which is used to amplify c_{ij} . By running the simulation with a gradual increase in p , the in-phase and quadrature components of the DOPO pulse can be obtained. As described in Section IV, the measurement feedback of the CIM reproduces the Ising spins using the amplitudes of DOPO pulses cycling over long fibers. In other words, the simulation model can be used to reproduce the behavior of the CIM and obtain the combination of spins that minimizes the Hamiltonian. Ising spin σ_{ij} is implemented by c_{ij} : If $c_{ij} < 0$, then $\sigma_{ij} = -1$, whereas if $c_{ij} > 0$, then $\sigma_{ij} = +1$.

B. Compared Methods

1) *SA-based RA*: In SA-based RA, the Boltzmann machine model [33] is used as a computational model. The Boltzmann machine is a type of mutually connected neural network that incorporates the statistical behavior. Specifically, the output function of the mutually connected neural network is a sigmoid function with temperature T introduced, and the probability that the value of neuron $X \in \{0, 1\}$ changes with the temperature T . In our simulation, the initial temperature of the SA was set as $T_{ini} = 5.0$, and is cooled according to the following equation for each iteration t as described in [34]:

$$T(t) = \frac{T_{ini}}{\log(1+t)}, \quad (30)$$

where $T(t)$ is the temperature of the SA at iteration t . The number of SA iterations is measured as 10,000.

2) *C-NOMA*: In the C-NOMA, users are divided into a set of near-users that are located near the center of the cell and far-users that are located far from the center of the cell based on criteria such as the distance or channel gain. Thereafter, channels are allocated by pairing one user from the near-user set and one user from the far-user set, and assigning them to the same channel.

3) *DRL-Based RA Method*: As a DRL-based RA method for comparison, we designed a DQN method to solve our formulated RA optimization problem shown in (11). In the designed DQN method, channel assignment is based on the following information at the decision time t . CSI matrix information $\mathbf{CSI}_t \in \mathbb{R}^{N_u \times N_c}$, user assignment variables $\mathbf{x}_{t-1} \in \mathbb{R}^{N_u \times N_c}$, user to be assigned to the channel in the next step $\mathbf{A}_t \in \mathbb{R}^{N_u}$, the achieved data rate of each user in the last step $\mathbf{R}_{t-1} \in \mathbb{R}^{N_u}$.

That is, the state at the decision time t , which is denoted as S_t is defined as $S_t = [\mathbf{CSI}_t, \mathbf{x}_{t-1}, \mathbf{A}_t, \mathbf{R}_{t-1}]$. Here, the action is defined as the channel allocation, i.e., the channel allocated to which user. The policy used in our designed method is ϵ -greedy method while the reward is defined as the data rate [13]. The Q network and target network of DQN consist of 3 layers, and the size of each layer is $|S_t|$, 256, and N_c , respectively. RMSProp is used to update the parameters of neural network, and the learning rate is set to 10^{-4} .

4) *ES*: In the ES, the channel allocation is the optimal solution by exhaustive searching.

C. Convergence of the Proposed Method

First, we evaluate the convergence of the proposed method. Specifically, we evaluated the convergence of the proposed method in the NOMA system with 12 users and 6 subchannels, i.e., $N_u = 12$ and $N_c = 6$. Since the trend of the DOPO pulse over time depends on the initial values of the spin in-phase amplitude and scaling parameters, we show two sets of the results under different initial values of the spin in-phase amplitude and scaling parameters, i.e., (a) Example 1 and (b) Example 2 in Fig. 5. Specifically, the left, the center, and the right subfigures of each set of results are the trend of the in-phase amplitude of the DOPO pulse over time, the corresponding energy of the Ising Hamiltonian, and the total data rate, respectively. Here, for the optimization time, we used the time required for 1000 cycles on an actual CIM machine as shown in [15]. Each line of the subfigures on the left side represents the in-phase amplitude of the number of Ising spins of $N_u \times N_c$, respectively. In addition, to verify whether the proposed CIM-based RA method can achieve the optimal solutions or not while showing the superiority of the NOMA methods compared with OMA, we also plot the total data rate achieved by the optimal solution and OMA after calculating in the rightmost subfigure of Fig. 5.

From Fig. 5, we can see that when the in-phase amplitude of the spins converges, the minimum Ising Hamiltonian reaches, while the total data rate of the CIM-based RA achieves the optimal value. This means that in the process of the CIM operations, the CIM spontaneously selected a combination of spins that exhibited the ground state of the Ising Hamiltonian, which corresponds to the solutions to the formulated optimization problem. Hence, the convergence of our proposed method has been verified. Moreover, from the rightmost subfigures of Fig. 5, we can get that the proposed CIM-based RA method can achieve the optimal solutions while improving the performance to a great extent in terms of the total data rate compared with OMA. Furthermore, from example 1 in Fig. 5, it can be confirmed that the Ising Hamiltonian and total data rate converge even though the in-phase amplitudes do not converge around 500 cycles. This may be because of the application of the solution construction algorithm described in Section V.D. The combination of σ_{ij} indicating channel assignment is determined by the ordinal order of in-phase amplitude values. In the simulation scenario for example 1 in Fig. 5, the ordinal order is fixed at around 500 cycles.

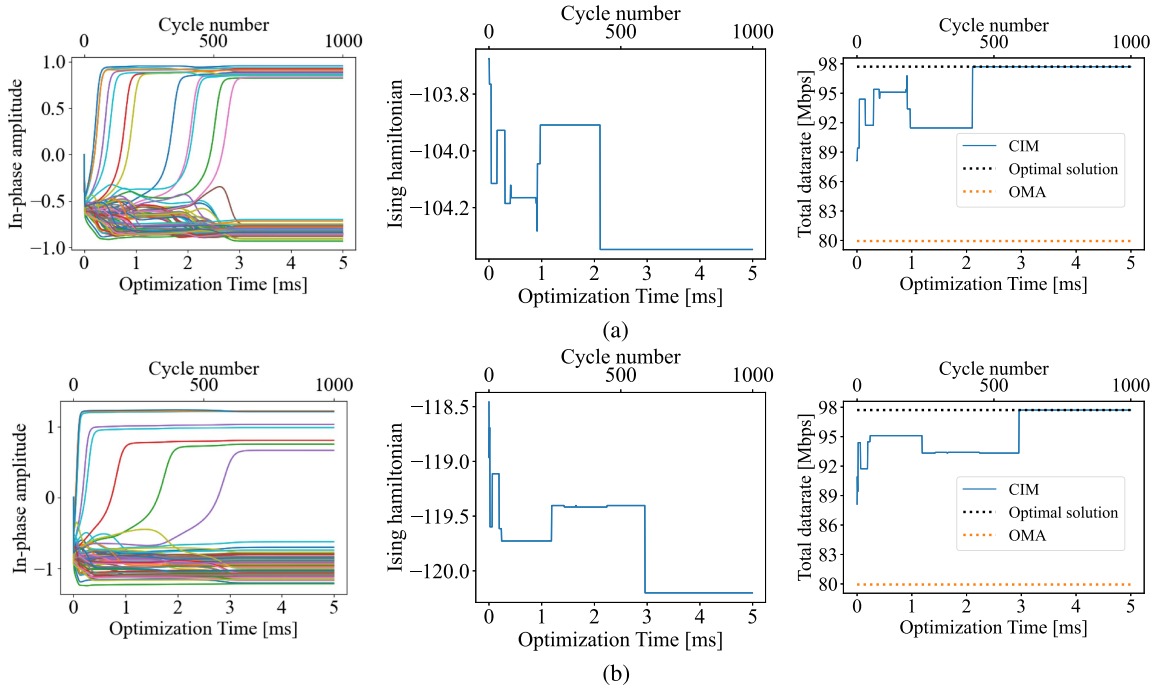


Fig. 5. Convergence of in-phase components of DOPO pulses (left) and the corresponding Ising Hamiltonian at that time (center) and comparison of the total data rate obtained by the CIM-based RA with the optimal and the OMA methods at that time (right) in the NOMA system with $N_u = 12$ users and $N_c = 6$ sub-channels. (a) Example 1. (b) Example 2.

In addition, it is shown in [18] that for certain MAX-CUT problems, the real CIM machines can provide a stable state approximately 1000 times faster than SA using state-of-the-art CPUs. Thus, our proposed method has a significant advantage in convergence speed. Moreover, in [35], it has been demonstrated that the phase of the DOPO pulse during oscillation converges to the ground state within 5 ms by amplifying the DOPO pulse with an appropriate pump pulse amplitude. The setting of the pump pulse amplitude in real CIM is out of the scope of this manuscript, while considering the space of the manuscript, we omit the detailed analytical proof process, which can be found in [18], [31], [35]. In this article, we set the pump pulse amplitude, i.e., p in equations (28) and (29), to an appropriate value by parameter search during the simulations. Thus, the in-phase amplitude of the DOPO pulse can converge within 5 ms, as shown in Fig. 5.

In the following subsections, the comparison of the proposed CIM-based RA methods with the benchmarks in terms of the computation time and total data rate under various kinds of settings will be performed.

D. Performance Evaluation of the Data Rate Under Static Environment

1) *Data Rate vs the Number of Users:* Then, we evaluated the total data rate when the number of users N_u and that of subchannels N_c are varied. Note that we consider that up to 2 users are multiplexed in each subchannel. Therefore, as N_u increases, N_c also increases as $N_c = \lceil N_u/2 \rceil$. In other words, N_u is varied from 12 to 30 in this simulation, while at the same time N_c is varied from 6 to 15. In addition, the transmission

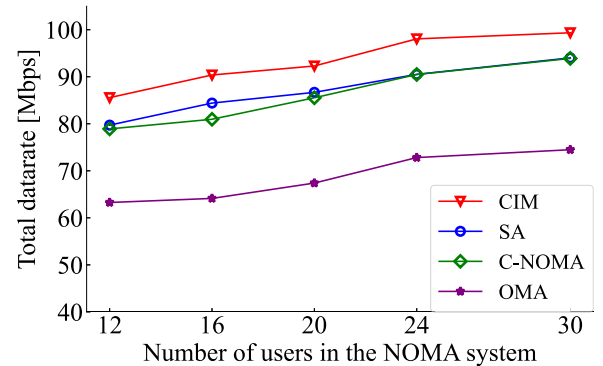


Fig. 6. Total data rate versus number of users N_u in the NOMA system with total transmit power $P_T = 12$ from the BS.

power P_T is set to 12. Fig. 6 shows the simulation results. From Fig. 6, we can see that our proposed method obtained the highest total data rate among the compared methods. This indicates that the most suitable joint channel and power allocation is possible using the proposed method. In addition, we can observe that the NOMA methods achieve a higher total data rate than the existing OMA method. During this simulation, we set up to 30 users in the NOMA system. As N_u is increased, the number of combinations of user-channel pair increases, making it difficult to obtain an optimal solution using the ES. By contrast, $N_c \times N_u = 450$ spins are required when solving this RA problem with 30 users and 15 channels using the CIM. Because the actual CIM can currently solve problems with up to 100,000 spins [18], this RA problem can be solved using the CIM. In other words, using the CIM, it is possible to obtain an effective solution to the RA

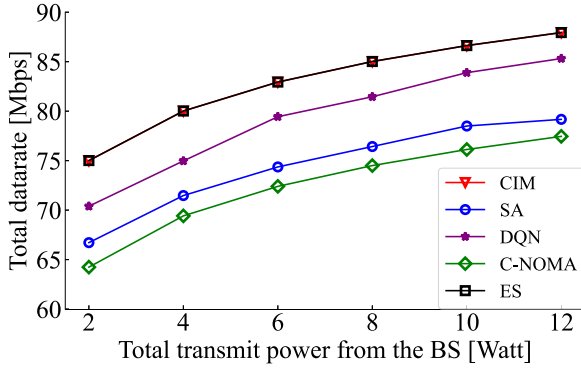


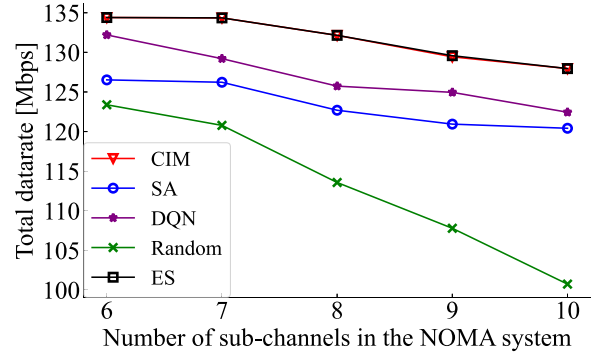
Fig. 7. Total data rate versus total transmit power P_T from the BS with number of users $N_u = 12$ and number of sub-channels $N_c = 6$ in the NOMA system.

optimization problem in a large-scale NOMA system that cannot be solved using the ES at a speed on the order of milliseconds.

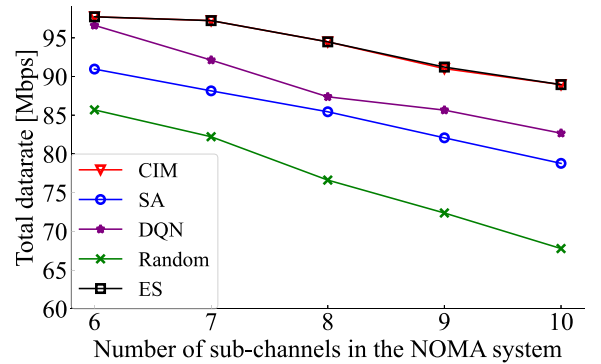
2) *Data Rate vs Transmit Power*: Next, we evaluated the total data rate of the NOMA system when varying the total transmission power from the BS, i.e., P_T . In this simulation, the numbers of users and sub-channels in the NOMA system are set to $N_u = 12$ and $N_c = 6$, respectively. Here, P_T is set to 2, 4, 6, 8, 10, and 12 Watts. Fig. 7 shows the simulation results. From Fig. 7, we can see that our proposed CIM-based method can obtain a reachable optimal solution regardless of the transmission power from the BS P_T .

3) *Data Rate vs the Number of Channels*: Next, we evaluated the total data rate for varying number of channel N_c . In this simulation, the transmission power from the BS is set to $P_T = 12$. The number of users is set to $N_u = 12$. N_c is varied from 6 to 10. In this study, the total bandwidth $B = 5.0$ MHz is constant, and thus the sub-carrier width is $5.0/N_c$ MHz. Note that this means that the sub-carrier width changes as the number of sub-channels changes. Fig. 8 shows the simulation results when the path loss coefficient α is set to 3 and 4. From Fig. 8, we can see that the proposed method can achieve the same total data rate as the ES even when N_c increases. Hence, we can say that the performance in total data rate of the proposed method is comparable to that of the ES. However, when N_c increases, the number of channel-user combinations highly increases. At that time, the total data rate of the random allocation method decreases rapidly. This indicates that the optimization using the CIM can result in an effective channel allocation. In addition, the total data rate decreases as N_c increases. One of the reasons for this is that the power allocated per channel decreases as N_c increases because the power from the BS is constant. Another reason to be considered can be summarized as follows. When N_c increases, the bandwidth for each sub-channel becomes small, whereas the OMA is applied in more channels. By contrast, when N_c decreases, the bandwidth for each sub-channel increases while the NOMA is applied in more channels. Because the NOMA technique is more efficient than the OMA technique, the total data rate decreases with the increase in the number of sub-channels.

4) *Data Rate vs the Number of Users Under a Fixed Number of Channels*: Then, we evaluated the total data rate for the



(a)



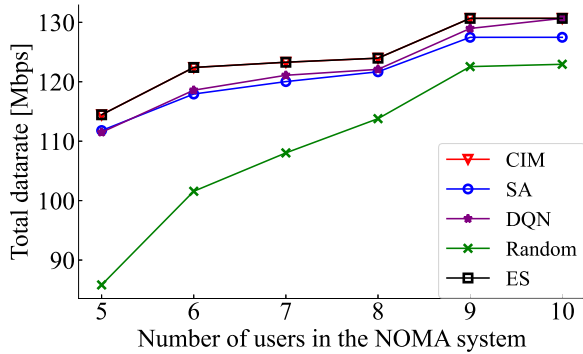
(b)

Fig. 8. Total data rate versus number of sub-channels N_c with number of users $N_u = 12$ in the NOMA system. Note that the sub-carrier width changes as the number of sub-channels changes. (a) $\alpha = 3.0$. (b) $\alpha = 4.0$.

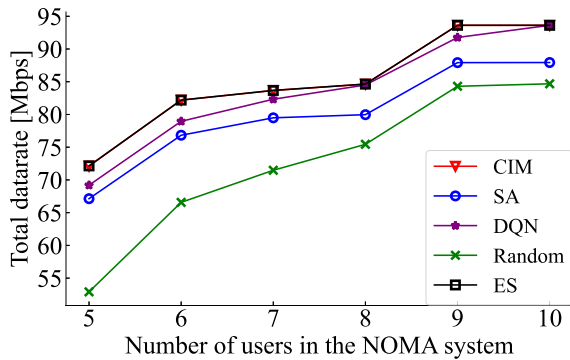
varying number of users N_u . In the simulation, N_u is varied from 5 to 10. The number of sub-channels is set to $N_c = 5$. Under this setting, the OMA and NOMA coexist when N_u is less than 10. The transmission power from the BS is set to $P_T = 12$. Fig. 9 shows the simulation results when the path loss coefficient α is set to 3 and 4. From Fig. 9, we can see that the proposed method can achieve the same total data rate as the ES even when N_u increases. This indicates that the CIM can achieve the optimal channel selection even when the NOMA and OMA coexist. In addition, the total data rate of the system increases as N_u increases. As the reason for this, as N_u increases, the number of channels to be multiplexed increases. In other words, the advantages of the NOMA system were fully utilized.

E. Comparison of the Computation Time

Then, we evaluated the computation time of the SA, the ES, the DQN-based, and the CIM-based RA method under the varying number of users. Here, the computation time of the SA and the ES is evaluated using an Intel Xeon Gold 5222 CPU @ 3.80 GHz and C language, and the DQN-based RA is evaluated using the same CPU and Python with PyTorch, respectively. The number of users in the NOMA system N_u is set as 6, 8, 10, 12, 14, 16, and 18 in the performance evaluation. The numbers of SA iterations under corresponding different N_u were set as 10, 50, 100, 500, 1000, 5000, and 10000 times, respectively. In addition, the number of training episodes for the DQN-based RA under



(a)



(b)

Fig. 9. Total data rate versus number of users N_u with number of sub-channels $N_c = 5$ in the NOMA system. (a) $\alpha = 3.0$. (b) $\alpha = 4.0$.

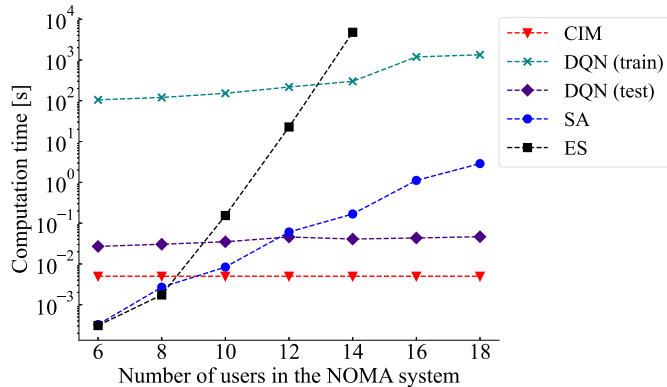


Fig. 10. Comparison of computation time.

different N_u is set as 100, 250, 400, 550, 700, 850, and 1000 times, respectively. Here, each episode consists of 120 steps. The BS collects information and trains the neural network for each step. The information consists of the CSI information, the channel assignment information, and the achieved data rate of the previous step of each user. The reason for the settings of the SA and DQN-based RA is that the larger the scale of the problem is, the larger number of iterations and training episodes required for SA and DQN-based RA methods. Here, owing to the enormous amount of computation time involved for the ES with the growth of the number of users, the method was evaluated for up to 14 users. Fig. 10 shows the evaluation results. The computation time of the CIM in Fig. 10 is the time cost for RA using the CIM in the real world. In [16], it has been shown that the Ising problem

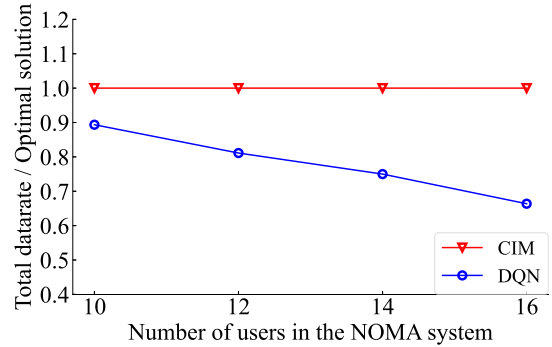


Fig. 11. Ratio of the data rate to the optimal solution in the NOMA system with mobile users.

can be solved using the CIM in 5 ms regardless of the problem size when the length of the fiber is 1-km-long. The computation time of DQN (train) and DQN (test) represents that for training the neural network and testing (making decisions) using the trained neural network, respectively. As shown in Fig. 10, the CIM-based and DQN-based RA for testing methods demonstrate almost constant computation time even as the number of users in the system increases. In contrast, other algorithms require a much longer time as the problem size increases. However, the DQN-based RA method requires training to obtain a good solution when the communication environment changes. The training time increases with the number of users in the NOMA system. Thus, as described above, the CIM is superior in speed compared to other methods.

F. Performance Evaluation of the Data Rate Under Dynamic Environment

Finally, we evaluate the performance of the proposed method when more realistic user activation is considered. In the simulation, the user activity follows the setup for the urban case in 3GPP TR 36.885 [36], [37]. In this scenario, the initial locations of the users are randomly generated according to the spatial Poisson process. Each user drives on the road at a constant speed. The speed of each user is generated randomly between 36 km/h to 54 km/h. The number of lanes is 2 in each direction and 4 in total in each street. The road grid size by the distance between intersections is 433 m \times 250 m. The simulation area size is 1299 m \times 750 m. When the user moves to the intersection, the direction is changed with a certain probability that follows the uniform distribution. In the performance evaluation, the number of users is set to 10, 12, 14, and 16. The time interval between each decision making is set to 20 ms [13]. In this scenario, the wireless environment changes faster than the resource allocation using the ES, SA, and DQN (train) methods due to their long computation time, as shown in Fig. 10. These methods may no longer apply to mobile NOMA systems. Hence, we only evaluate the performance of the CIM-based and the DQN (test)-based RA methods under these settings. The evaluation results in terms of the ratio of the total data rate and the optimal solution are shown in Fig. 11. From Fig. 11, we can see that the CIM-based RA can

achieve a higher data rate than the DQN (test)-based RA regardless of the number of users in the NOMA system. The reason can be summarized as follows. For the DQN (test)-based RA method, the time consumption for resource allocation is longer than the time interval between the decision-making. However, the user location and channel information are changed after the time interval. Hence, the information used by the DQN (test)-based RA method is the previous information that is different from the current information when allocating resources. On the other hand, the time consumption for the CIM-based RA method is shorter than the time interval between the decision-making. The CIM-based RA method can allocate resources using the current information. Hence, the optimal solution of the RA can almost be achieved by the CIM-based RA method. In summary, the proposed CIM-based RA method is applicable to the NOMA system with mobile users.

VII. DISCUSSION AND FUTURE WORK

A. Multiple Users Multiplexing in the NOMA System

In this subsection, we discuss multiplexing more users (more than 2) in the same channel using the proposed CIM-based RA method for the NOMA system. This article discusses NOMA systems where up to 2 users are multiplexed into the same channel. That is: in (11.d), we set a constraint term to limit the maximum number of users that are multiplexing in the same channel to up to 2. We believe that fast channel allocation in NOMA systems with more multiplexed users can be achieved by extending the constraint term in (11.d) as an arbitrary number of users n . Specifically, given the constraint, the energy function in (20) can be calculated. Finally, by recomputing the Ising model parameters, J_{ijkl} and λ_{ij} , RA in the NOMA system with n users multiplexed can be realized. Note that the number of dummy users N_d to must be set as $N_u + N_d = nN_c$ to satisfy the above constraint. In addition, the number of channel assignment variables x_{ij} , x_{kj} in (11) needs to be increased to correspond to the number of users multiplexed into the same channel.

B. Scalability of the Proposed Method

In this subsection, we discuss the scalability of the proposed CIM-based RA methods in the practical settings with 2 or more users in the NOMA systems. As previously described, the CIM is the machine that can obtain the ground state of the Ising Hamiltonian at high speed by getting the optimal combination of the spin directions. The ground state of a quantum-mechanical system is its stationary state of lowest energy, which corresponds to the optimal solution of combinatorial optimization problems. In this article, the combinatorial optimization problem is the RA problem in the NOMA system. The RA problem in the NOMA system can be solved by transforming the formulated RA combinatorial optimization problem into the form of the Ising Hamiltonian. The spin direction of the spin in the Ising Hamiltonian corresponds to the channel assignment variable of the RA problem in the NOMA system, i.e., whether the channel is allocated to the user or not. Each channel-user pair corresponds to one spin. The value of each spin direction is +1 or -1. If the

spin direction is +1, the channel is allocated to the corresponding user. Otherwise, the channel is not allocated to the corresponding user. For example, channel i -user j pair corresponding to the spin direction σ_{ij} . If σ_{ij} is +1, the channel i is allocated to user j . Otherwise, the channel i is not allocated to user j . Hence, to solve the RA optimization problem in the NOMA system, $N_u * N_c$ spins are necessary, where N_u and N_c are the numbers of the users and channels, respectively. The related work [18] shows that the CIM can realize problems up to 100000 spins. This means that up to 446 users can be realized using the present CIM in our scenario, where up to 2 users are allocated to one channel simultaneously. In general, however, clustering with an arbitrary number of n users of multiplexing is more attractive. Here, to achieve multiplexing at an arbitrary user n using CIM, we need to change the definition of the assignment variable as described in the last subsection. Specifically, the number of indexes should be increased by considering which user i is multiplexed with which user on channel j in (12). Therefore, as n increases, the number of spins required to solve the problem also increases. For example, when $n = 3$, the required number of spins is $N_u^3/3$, and RA for up to 67 users can be realized using the present CIM with 100000 spins. With the development of the CIM, much larger-scale optimization problems are expected to be solved by increasing the number of spins in the future.

C. Unequal Bandwidth Allocation

In this study, the RA in the NOMA system is under the assumption of equal bandwidth allocation. According to [38], the performance of NOMA systems can be improved by performing unequal bandwidth and allocating higher power and wider bandwidth to users with stronger channel conditions. Therefore, we will consider unequal bandwidth allocation in our future work. We believe that our proposed CIM-based RA method can be easily extended to the NOMA system with unequal bandwidth allocation by assigning more than one resource block to one use, specifically, by considering extending the constraint in (11.c) to an arbitrary number n . Using these constraints, the energy function (19) can be computed. Finally, by recomputing J_{ijkl} and λ_{ij} , more than one bandwidth can be allocated to one user using the CIM. Here, by adjusting the value of n , the upper limit of the bandwidth allocated to users with stronger channel conditions can be set. Thus, better system performance may be achieved with unequal bandwidth allocation compared to that with equal bandwidth allocation.

D. Optimal Power Allocation Using the CIM

In this study, the channel allocation is optimized by the CIM, and then the power is assigned based on the optimizing methods in [5] using conventional computers. However, in large-scale NOMA systems, the time consumption for allocating power using conventional computers may reduce the speed of the optimization, which may bring a bottleneck to CIM-based optimization. Therefore, channel and power optimization will be jointly considered using the CIM in our future work.

E. Diversity Schemes in NOMA System

In this study, we assumed that both of the BS and users have a single antenna, and the signal to each user is multiplexed in one channel at the same time. To improve the performance of the system, we will consider diversity schemes such as receiver diversity and antenna diversity in our future work. Specifically, the receiver diversity can improve the communication performance, such as bit error rate, by combining the same data in the different channels to achieve a diversity gain [39]. Antenna diversity can be used to improve the outage performance of downlink NOMA systems combined with signal user multiple-input multiple-output [40].

F. Implementation of the CIM

Although quantum technologies as emerging technologies are expected to be deployed in 6G networks around 2030, it is still unclear when a practical quantum computer will be available. Hence, we will work with the communications and quantum computing industries in the future to try to achieve a practical implementation of the CIM in the next generation of wireless communication.

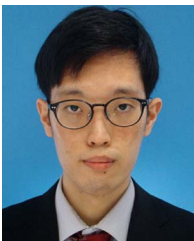
VIII. CONCLUSION

In this study, we proposed a CIM-based RA optimization method for NOMA systems. Initially, we formulated the RA problem as a combinatorial optimization problem. To apply RA optimization for NOMA systems in the CIM, we derived the interaction and external magnetic field by mapping the Ising Hamiltonian to the formulated objective function. Herein, the interaction and the external magnetic field are the parameters of the Ising model, which are used to search the optimal solutions. For correspondence between the Ising Hamiltonian and the formulated objective function, we transformed our formulated objective function into a similar form with an Ising Hamiltonian by the aid of the mutually connected neural network. The parameters of the coupling weights and firing threshold corresponding to the energy function of the mutually connected neural network were then derived. Subsequently, these parameters were used to derive the Ising model interaction between the spin and external magnetic field used for optimizing the RA in the CIM. To evaluate the performance of the proposed method, we conducted simulations using the simulation model of the CIM. Moreover, we compared our proposed method with other optimization methods and pairing algorithms. The simulation results indicated that the CIM is not only faster but also superior in searching an optimal solution.

REFERENCES

- [1] Z. Shi, X. Xie, H. Lu, H. Yang, J. Cai, and Z. Ding, "Deep reinforcement learning-based multidimensional resource management for energy harvesting cognitive NOMA communications," *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3110–3125, May 2022.
- [2] P. S. Bouzinis, P. D. Diamantoulakis, and K. Karagiannidis, "Wireless federated learning (WFL) for 6G networks—Part II: The compute-then-transmit NOMA paradigm," *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 8–12, Jan. 2022.
- [3] M. Nikooroo and Z. Becvar, "Optimal positioning of flying base stations and transmission power allocation in NOMA networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1319–1334, Feb. 2022.
- [4] C. He, Y. Hu, Y. Chen, and B. Zeng, "Joint power allocation and channel assignment for NOMA with deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 10, pp. 2200–2210, Oct. 2019.
- [5] J. Zhu, J. Wang, Y. Huang, S. He, X. You, and L. Yang, "On optimal power allocation for downlink non-orthogonal multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2744–2757, Dec. 2017.
- [6] L. Zhu, J. Zhang, X. Cao, and D. O. Wu, "Optimal user pairing for downlink non-orthogonal multiple access (NOMA)," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 328–331, Apr. 2019.
- [7] L. Salaün, C. S. Chen, and M. Coupechoux, "Optimal joint subcarrier and power allocation in NOMA is strongly NP-hard," in *Proc. IEEE Int. Conf. Commun.*, 2018, pp. 1–7.
- [8] A. Abrardo, M. Moretti, and F. Saggese, "Power and subcarrier allocation in 5G NOMA-FD systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8246–8260, Dec. 2020.
- [9] E. C. Cejudo, H. Zhu, and J. Wang, "Resource allocation in downlink multicarrier NOMA under a fairness constraint," in *Proc. IEEE 92nd Veh. Technol. Conf.*, 2020, pp. 1–5.
- [10] C. Kai, Y. Wu, M. Peng, and W. Huang, "Joint uplink and downlink resource allocation for NOMA-enabled D2D communications," *IEEE Wireless Commun. Lett.*, vol. 10, no. 6, pp. 1247–1251, Jun. 2021.
- [11] O. Abuajwa, M. B. Roslee, and Z. B. Yusoff, "Simulated annealing for resource allocation in downlink NOMA systems in 5G networks," *Appl. Sci.*, vol. 11, no. 10, May 2021, Art. no. 4592.
- [12] Y. Cao, G. Zhang, G. Li, and J. Zhang, "A deep Q-network based-resource allocation scheme for massive MIMO-NOMA," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1544–1548, May 2021.
- [13] J. Zheng, X. Tang, X. Wei, H. Shen, and L. Zhao, "Channel assignment for hybrid NOMA systems with deep reinforcement learning," *IEEE Wireless Commun. Lett.*, vol. 10, no. 7, pp. 1370–1374, Jul. 2021.
- [14] "D-wave systems—The practical quantum computing company," 2021. Accessed Dec. 23, 2021. [Online]. Available: <https://www.dwavesys.com>
- [15] T. Inagaki et al., "A coherent Ising machine for 2000-node optimization problems," *Science*, vol. 354, no. 6312, pp. 603–606, Nov. 2016.
- [16] P. L. McMahon et al., "A fully programmable 100-spin coherent Ising machine with all-to-all connections," *Science*, vol. 354, no. 6312, pp. 614–617, Nov. 2016.
- [17] R. Hamerly et al., "Experimental investigation of performance differences between coherent Ising machines and a quantum annealer," *Sci. Adv.*, vol. 5, no. 5, pp. 1–10, May 2019.
- [18] T. Honjo et al., "100,000-spin coherent Ising machine," *Sci. Adv.*, vol. 7, no. 40, pp. 1–10, Sep. 2021.
- [19] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: A comprehensive survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 334–366, 2021.
- [20] A. Lucas, "Ising formulations of many NP problems," *Front. Phys.*, vol. 2, Feb. 2014, Art. no. 5.
- [21] M. Hasegawa, H. Ito, H. Takesue, and K. Aihara, "Optimization by neural networks in the coherent Ising machine and its application to wireless communication systems," *IEICE Trans. Commun.*, vol. 104, no. 3, pp. 210–216, 2021.
- [22] H. Ito, Y. Jiang, H. Yasuda, H. Takesue, K. Aihara, and M. Hasegawa, "High-speed optimization method for resource allocation in wireless communication systems by coherent Ising machine," in *Proc. IEEE Int. Conf. Artif. Intell. Inf. Commun.*, 2020, pp. 093–097.
- [23] S. Kobayashi, Y. Murata, K. Aihara, and M. Hasegawa, "Fast optimization of scheduling based on quantum neural network in distributed antenna system," in *Proc. IEICE Soc. Conf., N-2-1*, 2019.
- [24] K. Kurasawa et al., "A high-speed channel assignment algorithm for dense IEEE 802.11 systems via coherent Ising machine," *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1682–1686, Aug. 2021.
- [25] Y. Haribara et al., "A coherent Ising machine for maximum cut problems: Performance evaluation against semidefinite programming relaxation and simulated annealing," *Princ. Meth. Quantum Inf. Technol.*, vol. 911, pp. 251–262, Feb. 2016.
- [26] I. F. Akyildiz, A. Kak, S. Nie, and L. Zhao, "6G and beyond: The future of wireless communications systems," *IEEE Access*, vol. 8, pp. 133995–134030, 2020.
- [27] H. Takesue, T. Inagaki, K. Inaba, and T. Ikuta, and T. Honjo, "Large-scale coherent Ising machine," *J. Phys. Soc. Jpn.*, vol. 88, Mar. 2019, Art. no. 061014.

- [28] Z. Ding, Z. Yang, P. Fan, and H. V. 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.
- [29] S. Utsunomiya, K. Takata, and Y. Yamamoto, "Mapping of Ising models onto injection-locked laser systems," *Opt. Exp.*, vol. 19, no. 19, pp. 18091–18108, 2011.
- [30] J. J. Hopfield and D. W. Tank, "Neural computation of decisions in optimization problems," *Biol. Cybern.*, vol. 52, pp. 141–152, Jul. 1985.
- [31] Z. Wang, A. Marandi, K. Wen, R. L. Byer, and Y. Yamamoto, "Coherent Ising machine based on degenerate optical parametric oscillators," *Phys. Rev. A*, vol. 88, Dec. 2013, Art. no. 063853.
- [32] H. Nakahara, "Convolutional neural network implementation on an FPGA," *J. Inst. Electron., Inf. Commun. Eng.*, vol. 103, no. 5, pp. 501–506, May 2020.
- [33] E. H. L. Aarts and J. H. M. Korst, "Boltzmann machines for travelling salesman problems," *Eur. J. Oper. Res.*, vol. 39, pp. 79–95, Mar. 1989.
- [34] B. Hajek, "Cooling schedules for optimal annealing," *Math. Operations Res.*, vol. 13, pp. 311–329, May 1988.
- [35] D. Ito, T. Ueta, and K. Aihara, "Bifurcation analysis of eight coupled degenerate optical parametric oscillators," *Physica D: Nonlinear Phenomena*, vol. 372, pp. 22–30, Jun. 2018.
- [36] H. Ye, G. Y. Li, and B. -H. F. Juang, "Deep reinforcement learning based resource allocation for V2V communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3163–3173, Apr. 2019.
- [37] 3rd Generation Partnership Project, "Technical specification group radio access network: Study LTE-based V2X services: (Release 14)," 3GPP, Sophia Antipolis, France, Tech. Rep. TR 36.885, Jun. 2016.
- [38] J. Wang, H. Xu, L. Fan, B. Zhu, and A. Zhou, "Energy-efficient joint power and bandwidth allocation for NOMA systems," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 780–783, Apr. 2018.
- [39] J. Kim, W. Lee, and H. Song, "Performance enhancement using receive diversity with power adaptation in the NOMA system," *IEEE Access*, vol. 7, pp. 102867–102875, 2019.
- [40] M. Gong and Z. Yang, "The application of antenna diversity to NOMA with statistical channel state information," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3755–3765, Apr. 2019.



Teppei Otsuka received the B.Eng. and M.Eng. from the Department of Electrical Engineering, Faculty of Engineering, Tokyo University of Science, Tokyo, Japan, in 2021 and 2023, respectively. His research interests include NOMA, quantum annealing, and coherent Ising machine. He is currently with NTT Docomo, Inc. Japan.



Aohan Li (Member, IEEE) received the Ph.D. degree from Keio University, Tokyo, Japan, in 2020. From 2020 to 2022, she was an Assistant Professor with the Tokyo University of Science, Tokyo, Japan, where she is currently a Visiting Researcher. She is also an Assistant Professor with the University of Electro-Communications, Tokyo, Japan. Her research interests include resource management, quantum annealing, machine learning, and the Internet of Things.



Hiroki Takesue received the B.Eng., M.Eng., and Ph.D. degrees in engineering science from Osaka University, Japan, in 1994, 1996, and 2002, respectively. In 1996, he joined NTT Laboratories. Since then, he has been engaged in research on quantum communication and photonic computation. He is currently a Senior Distinguished Researcher of NTT, and the group Leader of Quantum Optical State Control Research Group, NTT Basic Research Laboratories.



Kensuke Inaba received the B.E., M.E., and Dr.Eng. from the Department of Applied Physics, Osaka University, Japan, in 2003, 2005, and 2007, respectively. Since 2008, he has been a Senior Research Scientist with Quantum Optical State Control Research Group, NTT Basic Research Laboratories, where he has been theoretically studying many body systems and also photonic computation.



Kazuyuki Aihara received the B.E. degree of electrical engineering and the Ph.D. degree of electronic engineering from the University of Tokyo, Japan, in 1977 and 1982, respectively. He is currently a University Professor and Professor Emeritus of the University of Tokyo. His research interests include mathematical modeling of complex systems, parallel distributed processing with spatio-temporal neurodynamics, and time series analysis of complex data.



Mikio Hasegawa (Member, IEEE) received the B.Eng., M.Eng., and Dr.Eng. degrees from the Tokyo University of Science, Tokyo, Japan, in 1995, 1997, and 2000, respectively. From 1997 to 2000, he was a Research Fellow with the Japan Society for the Promotion of Science. From 2000 to 2007, he was with the Communications Research Laboratory (CRL), Ministry of Posts and Telecommunications, which was reorganized as the National Institute of Information and Communications Technology (NICT) in 2004. He is currently a Professor with the Department of Electrical Engineering, Faculty of Engineering, Tokyo University of Science. His research interests include mobile networks, cognitive radio, neural networks, machine learning, and optimization techniques. He is a Senior Member of IEICE.