Beam Alignment for Massive MIMO Transmission Based on Multi-dimensional and Multi-particle Quantum Walks

Maki Arai
College of Engineering
Shibaura Institute of Technology
Tokyo, Japan
m-arai@shibaura-it.ac.jp

Tomoki Yamagami
Graduate School of Science and Engineering
Saitama University
Saitama, Japan
tyamagami@mail.saitama-u.ac.jp

Takatomo Mihana
Graduate School of Information Science and Technology
The University of Tokyo
Tokyo, Japan
takatomo_mihana@ipc.i.u-tokyo.ac.jp

Ryoichi Horisaki
Graduate School of Information Science and Technology
The University of Tokyo
Tokyo, Japan
horisaki@g.ecc.u-tokyo.ac.jp

Mikio Hasegawa
Dept. of Electrical Engineering, Faculty of Engineering
Tokyo University of Science
Tokyo, Japan
hasegawa@ee.kagu.tus.ac.jp

***Abstract*—This paper proposes a multi-armed bandit (MAB) algorithm based on a quantum walk (QW) with multiple dimensions and multiple particles. By extending our fast beam selection method using one-dimensional discrete-time QW to multi-dimensional and multi-particle systems, we aim to improve the search efficiency for optimal beam combinations for multiple users and enhance the total channel capacity. Numerical analysis demonstrates that the proposed method improves the average channel capacity even when the number of users is three or more. The proposed method provides an effective framework for achieving low latency and large capacity in multi-user beam allocation problems.**

***Index Terms*—multi-armed bandit, beam selection, beamforming, array antenna**

I. INTRODUCTION

In the contemporary age, quantum computing holds significant promise for addressing sophisticated and large-scale computational problems. One approach to utilizing quantum phenomena as search algorithms involves the use of quantum walks (QWs) [1]. The QW is a quantum version of random walks, where the walker moves according to the principle of quantum superposition [2], [3]. The application of the two properties of QW, namely localization and linear spread, to search problems has been investigated to accelerate classical search quadratically [4].

On the other hand, in massive MIMO (Multiple-input multiple-output) targeting 6G, the combination space for optimal beam search and allocation expands exponentially due to the large number of beam candidates and simultaneous multi-user connections [5]. The massive MIMO technology enhances the efficiency of spatial resources for accommodating a large number of users by high-gain beamforming with array antennas, while also addressing challenges posed by propagation losses in high-frequency bands [6]. However, as the number of beams and the number of simultaneously connected users increase, the combination space for searching optimal beam allocation is expected to grow exponentially.

In the conventional procedure of sequentially switching beams while transmitting and receiving pilot signals to perform beam allocation, a greater number of signal transmissions and receptions are required, thereby increasing communication overhead. Furthermore, delayed responses to environmental changes result in beam misalignment. These factors result in a reduction of the channel capacity of wireless communications, thereby requiring the identification of the optimal beams for multiple users more rapidly.

To address the aforementioned issue, a methodology for beam selection using a multi-arm bandit (MAB) algorithm has been proposed. The MAB algorithm determines the optimal choice based on the empirical communication status without detailed state information [7], [8]. Furthermore, in the MAB algorithm based on QW [9], which applies linear spreading and localization to exploration and exploitation in classical MAB, the optimal solution can be found faster.

This paper proposes a multi-armed bandit algorithm based on QW with multiple dimensions or multiple particles. By extending our fast beam selection method using QW [10] to multi-dimensional and multi-particle systems, we aim to improve the search efficiency for optimal beam combinations for multiple users and enhance the channel capacity. Numerical analysis demonstrates that the proposed method improves the average channel capacity even when the number of users is three or more.

This research provides a novel approach for achieving low-latency beam alignment in multi-user environments for the 6G era. The contributions of this paper are as follows:

- First, by mapping the beam selection problem for each user to the dimension or the particle of QW, the proposed method enables further reduction in search time compared to conventional MAB algorithms.
- Second, the proposed method enables the appropriate selection of a combination of beams and multiple users with a reduced number of searches. This approach ensures not only the increase of the desired signal power but also the reduction of interference between users, enhancing the total channel capacity.

The structure of this paper is as follows. Section II provides an overview of related research on beam selection using MAB and discusses the contributions of this study. Section III introduces the principles of QW, and Section IV describes a massive MIMO system model. Section V proposes a beam selection method that utilizes multi-dimensional QW and multi-particle QW. Section VI provides the numerical analysis to confirm the validity of the proposed method. The conclusion of this paper is presented in Section VII.

II. Related works

The MAB algorithm contributes to identifying the optimal option that maximizes cumulative reward using efficient exploration and exploitation, even in circumstances where prior learning or information gathering is not feasible [7], [8]. Several beam alignment methods based on the MAB algorithms have been investigated, which aim to improve the communication quality without estimating detailed channel state information (CSI) [11]–[14]. This makes them effective for reducing communication overhead and improving adaptability to variable environments.

The authors have also proposed and investigated MAB-based beam selection, which further accelerates exploration and exploitation by incorporating optical and quantum operating principles into decision-making. In references [10] and [15], we applied the QW-based MAB algorithm [9] to vertical and horizontal beam selection, demonstrating that it achieves higher capacity than conventional MABs and the beam sweeping with overhead. Furthermore, the reference [16] proposed a beam tracking method using the Laser Chaos Decision Maker (LCDM) based on semiconductor laser chaos dynamics. However, it should be noted that our prior analyses were constrained to a maximum of two users. Therefore, this paper investigates the extension to simultaneous beam alignment for three or more users using two cases: the multi-dimensional QW and the multi-particle QW.

III. Principle of quantum walk

This section describes a formulation that extends the circular discrete-time quantum walk to multiple dimensions and multiple particles. The fundamental formulation concerning one-dimensional and three-state circular discrete quantum walks was described in [10].

A. Multi-dimensional quantum walk

First, we explain the d-dimensional QW, as illustrated in Fig. 1a. Several examples of multi-dimensional QWs have been introduced in references [1], [17]. The space of each dimension for a walker's moving is defined as a cycle consisting of the set of vertices $V_M := \{1, 2, \cdots, M\}$ and edges $E_M := \{\{x_i, x_i + 1\} \mid x_i \in V_M\}$ for $i \in \{1, 2, \cdots, d\}$. The position Hilbert space is expressed as the tensor product of spaces for each dimension, as follows:

$$\mathcal{H}_{\mathrm{P}} = \mathrm{span}\left\{|x_1, \cdots, x_d\rangle \mid \{x_1, \cdots, x_d\} \in \otimes^d V_M\right\} \quad (1)$$

$$|x_1, \cdots, x_d\rangle = |x_1\rangle \otimes |x_2\rangle \otimes \cdots \otimes |x_d\rangle \quad (2)$$

$$\otimes^d V_M \triangleq V_M \otimes V_M \otimes \cdots \otimes V_M, \quad (3)$$

where $|\cdot\rangle$ represents the bra-ket notation, which is a column vector that denotes quantum states in the Hilbert space [18]. The coin Hilbert space, which indicates the internal state of the walker, is defined as follows:

$$\mathcal{H}_{\mathrm{C}} = \mathrm{span}\left\{|\epsilon_1, \cdots, \epsilon_d, O\rangle\right\}, \quad (4)$$

Here, the positive or negative direction of the walker's movement is expressed as $\epsilon_i = \pm$ for the i-th dimension, and O represents the state of staying. Consequently, the total number of states is $2d{+}1$. According to the aforementioned definitions, the total Hilbert space of QW is derived as follows:

$$\mathcal{H} = \mathcal{H}_{\mathrm{P}} \otimes \mathcal{H}_{\mathrm{C}}. \quad (5)$$

The time evolution of QW at the timestep t is expressed by two operations, such as a shift S and a coin $C^{(t)}$, given by

$$\left|\Psi^{(t+1)}\right\rangle = U^{(t)} \left|\Psi^{(t)}\right\rangle \quad (6)$$

$$U^{(t)} = S(C^{(t)} \otimes I) \quad (7)$$

$$\left|\Psi^{(t)}\right\rangle = \sum_{\{x_1,\cdots,x_d\} \in \otimes^d V_M} |x_1, \cdots, x_d\rangle \otimes \left|\psi^{(t)}(x_1, \cdots, x_d)\right\rangle \quad (8)$$

where $\left|\psi^{(t)}(x_1, \cdots, x_d)\right\rangle$ denotes the amplitude vector of the walker's existence probability at the timestep t. I is the identity operator of the Hilbert space.

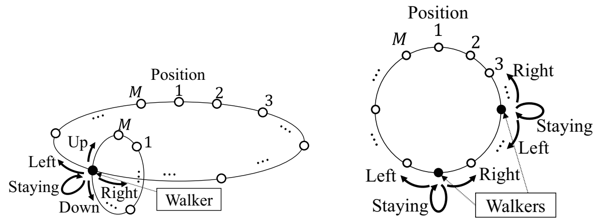
In this paper, the following unitary matrix is employed for the coin operation.

$$C(x_1, \cdots, x_d) = 2\,|D(x_1, \cdots, x_d)\rangle\,\langle D(x_1, \cdots, x_d)| - I \quad (9)$$

$$|D(x_1, \cdots, x_d)\rangle = [s, \cdots, s, c, s, \cdots, s]^{\mathrm{T}} \quad (10)$$

$$c \triangleq \cos\left(\frac{\theta(x_1, \cdots, x_d)}{2}\right) \quad (11)$$

$$s \triangleq \frac{\sin\left(\frac{\theta(x_1,\cdots,x_d)}{2}\right)}{\sqrt{2d}}, \quad (12)$$



(a) Two-dimensional QW. (b) Two-particle QW.

Fig. 1: Examples of multi-dimensional and multi-particle QWs.

where $\theta(x_1, \cdots, x_d) \in [0, 2\pi)$ is determined at each position. This parameter affects the walker's time evolution and its existence probability.

B. Multi-particle quantum walk

Next, we describe a multi-particle QW, as shown in Fig. 1b, where multiple walkers exist simultaneously within the position space. Several studies have presented the formulation of the multi-particle QW [19], [20]. This paper assumes that multiple walkers, that is, multiple particles, are non-interacting and distinguishable from each other.

When walkers are non-interacting, the total Hilbert space containing N multiple walkers is expressed by extending the single-walker model, as follows:

$$\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_N \tag{13}$$

$$\mathcal{H}_n \triangleq \mathcal{H}_{\mathrm{P}n} \otimes \mathcal{H}_{\mathrm{C}n}. \tag{14}$$

Note that the walkers are distinguishable and $n \in \{1, \cdots, N\}$ represents the index of the walker.

The operator with the shift and the coin in each timestep is also given by

$$U^{(t)} = U_1^{(t)} \otimes \cdots \otimes U_N^{(t)}. \tag{15}$$

Therefore, in the case of non-interacting and distinguishable particles, it can be considered a combination of multiple independent single-particle QWs.

IV. System model of massive MIMO

A massive MIMO system model with one base station (BS) and K user equipment units (UEs) is considered. The received signal of the k-th UE is expressed as follows:

$$y_k = \mathbf{h}_k \mathbf{W}_{\mathrm{a},k} \mathbf{w}_{\mathrm{d},k} s_k + \sum_{l \neq k} (\mathbf{h}_k \mathbf{W}_{\mathrm{a},l} \mathbf{w}_{\mathrm{d},l} s_l) + n_k. \tag{16}$$

Here, $\mathbf{h}_k$, s_k, and n_k denote a channel matrix between BS and the k-th UE, a transmit signal, and a noise, respectively. $\mathbf{W}_{\mathrm{a},k}$ and $\mathbf{w}_{\mathrm{d},k}$ are analog beamforming weights and digital pre-coding weights for the k-th UE.

When the UEs receive signals from the BS simultaneously using the same frequency band, interference occurs. The signal-to-interference-plus-noise ratio (SINR) is derived as follows:

$$SINR_k = \frac{\|\mathbf{h}_k \mathbf{W}_{\mathrm{a},k} \mathbf{w}_{\mathrm{d},k}\|^2 P_{\mathrm{s},k}}{\sum_{l \neq k} \|\mathbf{h}_k \mathbf{W}_{\mathrm{a},l} \mathbf{w}_{\mathrm{d},l}\|^2 P_{\mathrm{s},l} + P_{\mathrm{n},k}}, \tag{17}$$

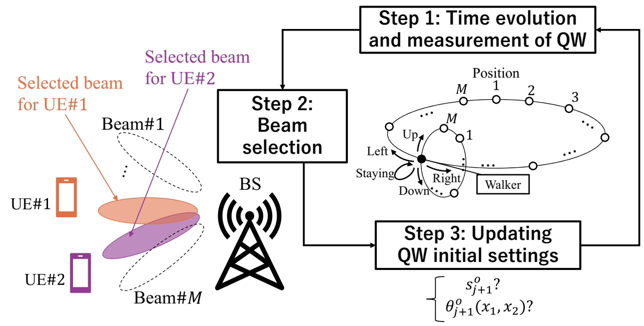


Fig. 2: Beam selection procedure using multi-dimensional QW for two UEs.

where $P_{\mathrm{s},k}$ and $P_{\mathrm{n},k}$ represent a transmit power and a noise power for the k-th UE, respectively.

Consequently, the channel capacity of each UE is expressed, accounting for interference, as follows:

$$C_k = B \log_2 \det(1 + SINR_k), \tag{18}$$

where B denotes a bandwidth.

V. Proposed beam selection based on multi-dimensional quantum walk

This section describes our proposed beam selection method based on the principles of multi-dimensional and multi-particle QWs. For example, Fig. 2 illustrates the beam selection procedure using multi-dimensional QW for two UEs.

A. Case of multi-dimensional QW

1) Time evolution of QW and selecting beams: In this paper, the dimensionality of the multi-dimensional QW is defined as the number of UE, as $d = K$, and the position is expressed by $x_1, \cdots, x_K$.

The initial settings for the time evolution of QW are defined as follows:

$$\left|\Psi_1^{(0)}\right\rangle = |s_1^o\rangle \otimes |O\rangle \tag{19}$$

$$\theta_1(x_1, \cdots, x_K) = \theta^o \in [0, 2\pi), \quad \forall \{x_1, \cdots, x_K\} \in \otimes^K V_M. \tag{20}$$

The time evolution of multi-dimensional QW is performed following the principle described in the previous section. After T timesteps have elapsed, the walker's position is measured and determined as $\{\hat{x}_{j,1}, \cdots, \hat{x}_{j,K}\}$. The beam index $\hat{m}_{j,k}$ for the k-th UE at the j-th decision is determined corresponding to the measured position $\hat{x}_{j,k}$ in the k-th dimension at the j-th decision, as follows:

$$\hat{m}_{j,k} = \hat{x}_{j,k}. \tag{21}$$

Subsequently, the transmission and reception are conducted between BS and UEs using the beams corresponding to the measured position.

2) Reward determination: After the transmission and reception, a reward in the j-th decision is defined as the mean of rewards for all UEs, as follows:

$$r_j = \frac{\sum_{k=1}^{K} r_{j,k}}{K} \tag{22}$$

$$r_{j,k} = \begin{cases} 1 & (\text{for } C_{j,k} > C_{\text{req}}) \\ 0 & (\text{otherwise}). \end{cases} \tag{23}$$

Here, $r_{j,k}$ is obtained by whether the channel capacity of Eq. (18) exceeds the required value C_{req} for the j-th decision, as $j \in \{1, \cdots, J\}$. J is the total number of decisions per UE. $C_{j,k}$ denotes the channel capacity of the k-th UE in the j-th decision.

The empirical reward probability for the k-th UE is calculated and updated repeatedly with each decision as follows:

$$\hat{p}_j(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K}) = \frac{\hat{R}_j(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K})}{\hat{N}_j(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K})} \tag{24}$$

$$\hat{R}_j(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K}) = \hat{R}_{j-1}(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K}) + r_j \tag{25}$$

$$\hat{N}_j(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K}) = \hat{N}_{j-1}(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K}) + 1, \tag{26}$$

where $\hat{R}_j(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K})$ represents the number of decisions that result in a reward for the position $\{\hat{x}_{j,1}, \cdots, \hat{x}_{j,K}\}$ until the j-th decision. $\hat{N}_j(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K})$ denotes the number of decisions that select the beam corresponding to the position $\{\hat{x}_{j,1}, \cdots, \hat{x}_{j,K}\}$ until the j-th decision.

3) Updating initial states: The initial parameters for the next decision are updated in accordance with the empirical reward probability, given by

$$\left|\Psi_{j+1}^{(0)}\right\rangle = \left|s_{j+1}^{o}\right\rangle \otimes |O\rangle \tag{27}$$

$$s_{j+1}^{o} = \underset{x_0, \cdots, x_d \in \otimes^d V_M}{\arg\max} \hat{p}_j(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K}) \tag{28}$$

$$\theta_{j+1}(x_1, \cdots, x_K) = \theta^o \exp\Big(-a \cdot (\hat{p}_j(\hat{x}_{j,1}, \cdots, \hat{x}_{j,K}))^b\Big). \tag{29}$$

After setting the initial states in accordance with the equations above, the time evolution of QW, beam determination, the transmission and reception between BS and UEs, and reward determination are repeated to converge toward the optimal choice of the beams.

B. Case of multi-particle QW

1) Time evolution of QW and selecting beams: This section explains the beam selection procedure employing K particles, and a one-dimensional QW, where each particle corresponds to the selection of each UE. When all particles are non-interacting and distinguishable, based on the principle described in Section III, each particle can be considered to move simultaneously and independently in one-dimensional space.

The initial settings of QW corresponding to the k-th particle are defined as follows:

$$\left|\Psi_{1,k}^{(0)}\right\rangle = \left|s_{1,k}^{o}\right\rangle \otimes |O\rangle \tag{30}$$

$$\theta_{1,k}(x) = \theta^o \in [0, 2\pi), \forall x \in V_M. \tag{31}$$

Using the above settings, the time evolution of multi-particle QW is performed. After T timesteps have elapsed, the k-th walker's position is measured and determined as $\hat{x}_{j,k}$ and the beam index $\hat{m}_{j,k}$ for the k-th UE at the j-th decision is determined, as follows:

$$\hat{m}_{j,k} = \hat{x}_{j,k}. \tag{32}$$

2) Reward determination: A reward of the k-th UE in the j-th decision is defined as $r_{j,k}$, and calculated by the channel capacity of each UE and the required value, as follows:

$$r_{j,k} = \begin{cases} 1 & (\text{for } C_{j,k} > C_{\text{req}}) \\ 0 & (\text{otherwise}). \end{cases} \tag{33}$$

The empirical reward probability for the k-th UE is also calculated and updated repeatedly as follows:

$$\hat{p}_{j,k}(\hat{x}_j) = \frac{\hat{R}_{j,k}(\hat{x}_j)}{\hat{N}_{j,k}(\hat{x}_j)} \tag{34}$$

$$\hat{R}_{j,k}(\hat{x}_j) = \hat{R}_{j-1,k}(\hat{x}_j) + r_{j,k} \tag{35}$$

$$\hat{N}_{j,k}(\hat{x}_j) = \hat{N}_{j-1,k}(\hat{x}_j) + 1, \tag{36}$$

where $\hat{R}_{j,k}(\hat{x}_j)$ represents the number of decisions that result in a reward for the position $\hat{x}_j$ until the j-th decision. $\hat{N}_{j,k}(\hat{x}_j)$ denotes the number of decisions that select the beam corresponding to the position $\hat{x}_j$ until the j-th decision.

3) Updating initial states: The initial parameters for the next decision are updated according to the empirical reward probability, as follows:

$$\left|\Psi_{j+1,k}^{(0)}\right\rangle = \left|s_{j+1,k}^{o}\right\rangle \otimes |O\rangle \tag{37}$$

$$s_{j+1,k}^{o} = \underset{x \in V_M}{\arg\max}\, \hat{p}_{j,k}(\hat{x}_{j,k}) \tag{38}$$

$$\theta_{j+1,k}(x) = \theta^o \exp\Big(-a \cdot (\hat{p}_{j,k}(\hat{x}_{j,k}))^b\Big). \tag{39}$$

Similar to the process of the multi-dimensional QW, the decision-making is repeated while the initial parameters are updated, and the choice of the beams converges toward the optimal one.

VI. Numerical analysis

By numerical analysis, the validity of the proposed beam selection method is confirmed.

A. Analysis condition

The analysis condition is described in Table I. All UEs are positioned on a horizontal plane, with a uniform height of 1.5 meters. The distance from BS to each UE is set to a radius $r = 40$ meters centered on BS, and the horizontal angle between adjacent UEs is fixed at 40 degrees, as illustrated in Fig. 3.

This paper compares the proposed MAB algorithms based on multi-dimensional QW and multi-particle QW, with conventional MAB algorithms and methods using sequential beam sweeping. The comparison methods are as follows: ϵ-greedy [8], UCB1-tuned [7], all combination search for $(N_{\text{BS,H}} \cdot N_{\text{BS,V}})^K$ combinations, and all beam search per user for $(N_{\text{BS,H}} \cdot N_{\text{BS,V}})$ beams. Table II shows the hyperparameters

TABLE I: Analysis condition.

Location (BS)	$(x, y, z) = (0, 0, 25)$
The number of antenna elements (BS)	$N_{\text{BS,H}} = 4$ $N_{\text{BS,V}} = 4$
Antenna element spacing (BS)	$l_{\text{spacing}} = 0.5\lambda_0$
The number of antenna elements (UE)	1
The number of UEs	$K \in \{1, 2, 3, 4\}$
Carrier frequency	$f_0 = 28$ [GHz]
Channel model	Ricean, $K_{\text{rice}} = 13$ [dB]
Transmit power per UE	$P_s = 27$ [dBm]
The number of decisions	$J = 1000$
The number of averaging	$N_{\text{average}} = 10$

TABLE II: Hyperparameters.

(a) Multi-dimensional QW.

	a	b	θ^o	T
$K = 1$	5	2	$3\pi/16$	4
$K = 2$	11	8	$11\pi/16$	4
$K = 3$	9	8	$7\pi/16$	2
$K = 4$	3	2	$5\pi/16$	4

(b) Multi-particle QW.

	a	b	θ^o	T
$K = 1$	5	2	$3\pi/16$	4
$K = 2$	9	4	$7\pi/16$	12
$K = 3$	5	8	$11\pi/16$	16
$K = 4$	11	2	$5\pi/16$	12

(c) ϵ-greedy.

	ϵ
$K = 1$	0.1
$K = 2$	0.1
$K = 3$	0.1
$K = 4$	0.2

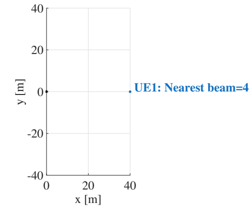
used in each MAB algorithm. These values are optimized to maximize the total average channel capacity for each condition of the placement of UEs.

The antenna equipped in BS is a 4×4 uniform planar array antenna, and analog beamforming weights based on the DFT matrix described in [21], [22] are used. The digital pre-coding weights are defined by using Zero-forcing [23].

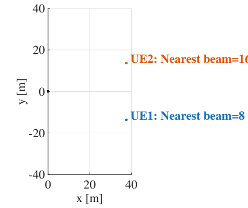
B. Total average channel capacity

Figure 4 shows the numerical analysis results for the total average channel capacity as a function of the number of UEs. The QW-based MAB algorithms that we proposed exhibited greater capacity, even as the number of UEs increased.

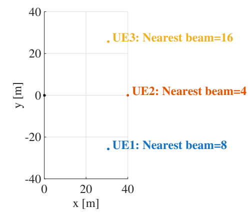
In addition, it was confirmed that the multi-particle QW method achieves larger capacity than the multi-dimensional QW method. In the case of $K = 4$, the multi-particle QW method achieved 86 % of the theoretical maximum channel capacity that is obtained through the consistent selection of the optimal combination in all decisions. This was due to beam selection for each UE independently and concurrently, thereby facilitating the exploration of beam combinations more efficiently. The desired signal power received at each UE correlates with the adjacent relationship of the selected beams. On the other hand, the interference power from other UEs is not affected only by the adjacent beam relationship, but rather by the combination of all selected beams. Consequently, in the beam selection for multiple users, it is necessary to obtain the optimal solution for the combination of beams for all users. The results of the analysis indicate that the multi-particle QW



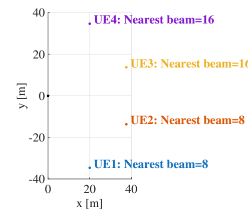
(a) $K = 1$.



(b) $K = 2$.



(c) $K = 3$.



(d) $K = 4$.

Fig. 3: Locations of UEs in xy-plane.

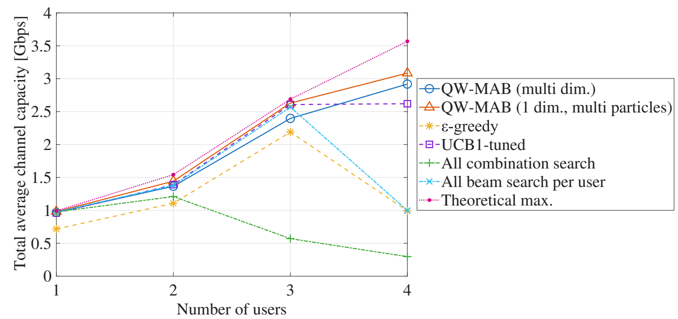


Fig. 4: Total average channel capacity for each number of users.

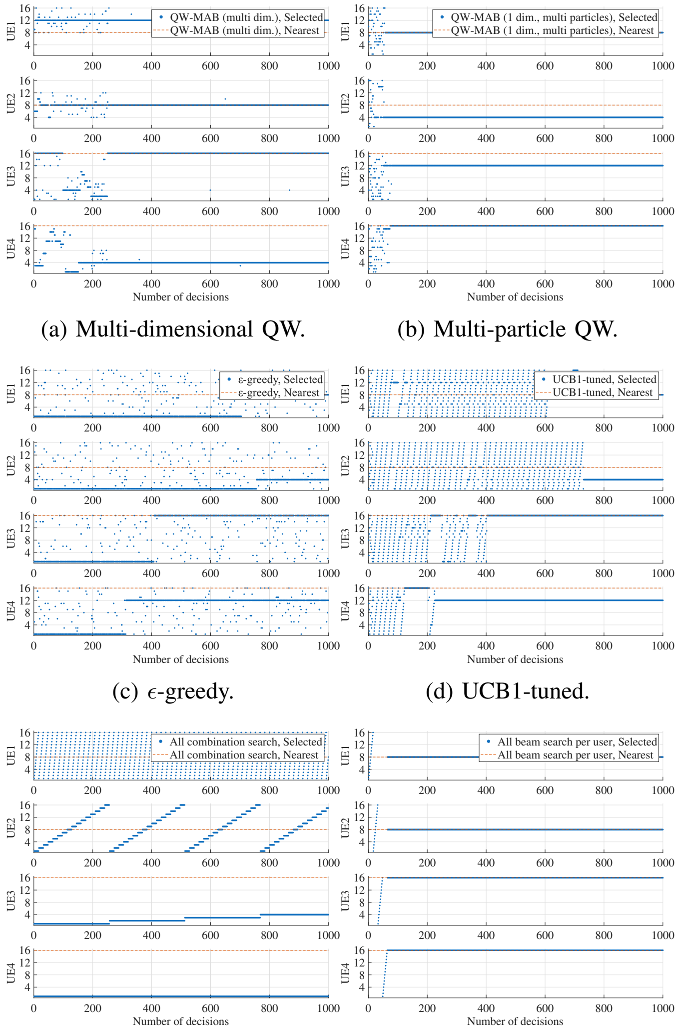
method is an effective approach for simultaneous exploration of combinations of multiple options.

C. Selected beam index

The beam selection results at each decision in the case of $K = 4$ are depicted in Fig. 5. Since the distance between users is comparatively minimal in relation to the beam width, the nearest beam indices are identical for two users. Conversely, the beam selection results indicate that distinct beams are utilized by all users. In the MIMO system, the interference and the channel correlation between UEs increase when multiple UEs utilize a common beam, resulting in the degradation of the capacity. The MAB algorithms, especially those based on QW, employ a strategy of selecting different beams for all users, resulting in the enhancement of channel capacity. Utilizing the proposed method leads to an appropriate combination of beams that achieves both objectives: increasing the desired signal power and reducing interference between users.

VII. Conclusion

This paper proposed beam selection methods based on multi-dimensional and multi-particle QWs for the multi-user massive MIMO system. The beam selection problem for



(a) Multi-dimensional QW. (b) Multi-particle QW.

(c) ϵ-greedy. (d) UCB1-tuned.

(e) All combination search. (f) All beam search per user.

Fig. 5: Selected beam index in the case of $K = 4$.

each user can be associated with dimensions or particles of QW, enabling efficient beam exploration and exploitation for multiple users. Numerical analysis confirmed that the proposed method can improve the average channel capacity of multiple users and thus demonstrated its effectiveness. The QW-based decision-making method proposed in this paper is expected to contribute to improving the communication performance of future wireless communication systems.

Future issues include evaluating on larger-scale system models, comparing the proposed method to other beam selection algorithms, examining implementation methods, and considering specific communication protocols.

Acknowledgment

This work was supported partly by JST [Moonshot R&D][Grant Number JPMJMS2216] and JSPS KAKENHI Grant Numbers P22H05197 and JP25K17613.

References

[1] R. Portugal, *Quantum walks and search algorithms*. Cham: Springer International Publishing, 2018.

[2] N. Konno, "Quantum random walks in one dimension," *Quantum Inf. Process.*, vol. 1, no. 5, pp. 345–354, Oct. 2002.

[3] S. E. Venegas-Andraca, "Quantum walks: a comprehensive review," *Quantum Inf. Process.*, vol. 11, no. 5, pp. 1015–1106, Oct. 2012.

[4] A. Ambainis *et al.*, "One-dimensional quantum walks," in *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, ser. STOC '01. New York, NY, USA: Association for Computing Machinery, 2001, p. 37–49.

[5] Q. Xue *et al.*, "A survey of beam management for mmwave and thz communications towards 6g," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 3, pp. 1520–1559, 2024.

[6] H. Tataria *et al.*, "6G Wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proceedings of the IEEE*, vol. 109, no. 7, pp. 1166–1199, 2021.

[7] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine Learning*, vol. 47, pp. 235–256, 2002.

[8] R. Sutton and A. Barto, *Reinforcement learning: An introduction*, 2nd ed., ser. Adaptive Computation and Machine Learning series. Cambridge: MIT Press, 1998.

[9] T. Yamagami, E. Segawa, T. Mihana, A. Röhm, R. Horisaki, and M. Naruse, "Bandit algorithm driven by a classical random walk and a quantum walk," *Entropy*, vol. 25, no. 6, p. 843, 2023.

[10] M. Arai, T. Yamagami, T. Mihana, R. Horisaki, and M. Hasegawa, "Two-dimensional beam selection by multiarmed bandit algorithm based on a quantum walk," *IEEE Transactions on Quantum Engineering*, vol. 6, pp. 1–16, 2025.

[11] N. Gulati and K. R. Dandekar, "Learning state selection for reconfigurable antennas: A multi-armed bandit approach," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 3, pp. 1027–1038, 2014.

[12] Y. Wei, Z. Zhong, V. Y. F. Tan, and C. Wang, "Fast beam alignment via pure exploration in multi-armed bandits," in *2022 IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 1886–1891.

[13] E. M. Mohamed, S. Hashima, K. Hatano, and S. A. Aldossari, "Two-stage multiarmed bandit for reconfigurable intelligent surface aided millimeter wave communications," *Sensors*, vol. 22, no. 6, 2022.

[14] Y. Song, C. Liu, W. Zhang, Y. Liu, H. Zhou, and X. Shen, "Two stage beamforming in massive MIMO: A combinatorial multi-armed bandit based approach," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 5, pp. 6794–6799, 2023.

[15] M. Arai, T. Yamagami, T. Mihana, R. Horisaki, and M. Hasegawa, "A study on optimal beam selection using a multi-armed bandit algorithm based on a five-state quantum walk," in *The International Symposium on Physics and Applications of Laser Dynamics 2024*, 2024.

[16] H. Asano, M. Arai, and M. Hasegawa, "Fast beam tracking in massive MIMO by laser chaos decision maker," *NOLTA*, vol. 16, no. 3, pp. 495–505, 2025.

[17] T. D. Mackay, S. D. Bartlett, L. T. Stephenson, and B. C. Sanders, "Quantum walks in higher dimensions," *Journal of Physics A: Mathematical and General*, vol. 35, no. 12, p. 2745, mar 2002.

[18] P. A. M. Dirac, "A new notation for quantum mechanics," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 35, no. 3, p. 416–418, 1939.

[19] P. P. Rohde, A. Schreiber, M. Štefaňák, I. Jex, and C. Silberhorn, "Multi-walker discrete time quantum walks on arbitrary graphs, their properties and their photonic implementation," *New Journal of Physics*, vol. 13, no. 1, p. 013001, jan 2011.

[20] L. Rigovacca and C. Di Franco, "Two-walker discrete-time quantum walks on the line with percolation," *Sci. Rep.*, vol. 6, no. 1, p. 22052, Feb. 2016.

[21] D. Yang, L.-L. Yang, and L. Hanzo, "Dft-based beamforming weight-vector codebook design for spatially correlated channels in the unitary precoding aided multiuser downlink," in *2010 IEEE International Conference on Communications*, 2010, pp. 1–5.

[22] M. Arai, K. Sakaguchi, and K. Araki, "A study on optimal beam patterns for single user massive MIMO transmissions," *IEICE Transactions on Communications*, 2019.

[23] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, 2004.