Codebook Based Antenna Configuration: A New Network Planning Paradigm for mmWave Mobile Communication Systems

Linzhi Shen, Yifang Zhang, and Shaowei Wang

***Abstract*—To configure the azimuths and downtilts of massive antennas optimally is of great importance for the mmWave mobile communication systems to support diverse services, however, it is always computationally forbidden due to the huge number of possible combinations in real-world mobile networks. In this paper, we develop an effective and efficient method to deal with this intractable optimization task. First, we explore the most promising azimuth and downtilt configurations of each antenna with a branch-and-bound procedure followed by a semi-definite relaxation algorithm, which finally generates a codebook with finite elements to reduce the unnecessary searches for similar or unpromising configurations during global search; second, a Monte Carlo tree search method is introduced to iteratively optimize the azimuths and downtilts of all antennas, aiming at alleviating the mutual interference of adjacent antennas to maximize the signal-to-interference ratio of the target service area, during which the generated antenna configuration codebook guides the iterative search process efficiently. Experiment results in real urban scenarios show that our proposed method can produce the best signal-to-interference ratio coverage as compared with state-of-the-art ones. Moreover, the proposed algorithm can scale up straightforwardly, making it a competitive choice for large-scale network optimization.**

***Index Terms*—Codebook, mmWave mobile communication systems, Monte Carlo tree search, semi-definite relaxation.**

I. Introduction

With the explosive growth in mobile data traffic and increasing number of user equipments in the fifth generation (5G) and beyond mobile communications, great demands have been placed for operators to provide users with guaranteed quality of service (QoS) in the service region [2, 3]. Radio network optimization (RNO) is a crucial step to improve network performance, during which the tunable parameters associated with base stations are configured in an optimal manner to provide both high power coverage and system throughput for the mobile networks [4–6].

Manuscript received October 17, 2022; revised February 9, 2023; accepted March 8, 2023. This work was partially supported by the National Natural Science Foundation of China under Grants 61931023 and U1936202. Part of this work has been presented at the IEEE Global Communications Conference 2022 [1], Rio de Janeiro, Brazil, December 4-8, 2022. The associate editor coordinating the review of this article and approving it for publication was W. Wang. *(Linzhi Shen and Yifang Zhang are co-first authors.) (Corresponding author: Shaowei Wang.)*

The authors are with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China (e-mail: DG20230072@smail.nju.edu.cn, mg21230076@smail.nju.edu.cn, wangsw@nju.edu.cn).

The RNO problem has been investigated extensively in the literature, which generally involves site selection, frequency planning and radio parameter configuration. Site selection is usually implemented in the early design stage for the mobile network, during which the locations of base stations (BSs) are selected to fullfill coverage and capacity requirements [7–9]. However, it is increasingly harder to obtain capacity or QoS gains in this way as the deployment of BSs becomes denser [10]. In the process of frequency planning, adjacent cells are allocated different frequencies to avoid inter-cell interference [11]. Nevertheless, cells can share the same frequency band with the help of advanced signal processing techniques such as coordinated multi-point communication and inter-cell interference coordination from the fourth generation (4G) system [12]. Thus, frequency planning is no longer necessary for most of the mobile communication systems. Radio parameter configuration involves adjustments to transmit power and antenna parameters, among which the angle setting of antenna plays an essential role and continues to catch attention from industrial projects [13]. An advantageous angle setting not only expands the power coverage but also mitigates the interference to neighboring cells so as to improve the system throughput [14].

In the 4G systems, angle tuning mainly focuses on downtilt adjustments, while azimuth adjustments are neglected due to the wide antenna beam [15]. With the evolution of cellular network, the 5G and beyond systems introduce mmWave technology to get more bandwidth resources [16], which results in further challenges for the antenna tuning. Specifically, the large-scale antenna arrays in mmWave systems synthesize highly directional beams to cope with the severe fading caused by the shorter wavelength, which allows the antenna to be equipped with more flexible pattern that varies with its azimuth and downtilt [17, 18]. Therefore, in contrast to the 4G, adjusting the downtilt and the azimuth jointly is more crucial for the mmWave systems. Indeed, recent studies have shown that tuning the azimuths and downtilts of directional antennas properly can improve the spectrum utilization efficiency in the spatial domain [19, 20].

Despite its great significance, searching for the optimal configuration of azimuths and downtilts for massive antennas is always an intractable combinatorial optimization problem [21]. In the existing literature, quite a few optimization methods have been proposed to obtain promising antenna configurations. In [22], the solution space is explored from a set of feasible configurations, and a memory of previous exploration is maintained to avoid cycling around local optimal

solutions. In [23], a series of feasible antenna configurations are generated through mutation and crossover, while poorly performing configurations are eliminated according to the fitness function. In [24], the antenna tilt angles are optimized by a primal-dual method. Similarly, a stochastic gradient descent method is proposed in [25] to accelerate the gradient calculation and produce near optimal solutions. In [26], the antenna downtilts and the transmission power are tuned by Bayesian method, where the power coverage and capacity objective are treated as black-box functions. The mentioned methods are usually computationally efficient, however, they are either restricted to handcrafted rules or easily trapped in local optima.

Recently, reinforcement learning (RL) has attracted widespread attention due to its high flexibility to different network scenarios, and arises as a promising solution for antenna configuration. In [27], antenna downtilts are adjusted successively based on fuzzy Q-learning. The main limitation of this method is that the remaining antennas are fixed when configuring the given one, which leads to a short-sighted and suboptimal solution. In [28], a mean field RL approach is utilized to learn the cumulative interference from adjacent cells, where all the neighboring cells are treated as a virtual one. However, such an approximation is oversimplified and results in imprecise SINR estimates. In [29], a deep Q-learning method is proposed to jointly optimize the configuration of all antennas. This method can be hardly applied in real-world scenarios due to its huge cost of training a reliable deep neural network. In [30], a Monte Carlo tree search (MCTS) method is proposed to overcome the above mentioned limitations, where the effect of neighboring cells on antenna configurations is averaged via multiple Monte Carlo simulations.

Unfortunately, the MCTS method also suffers from the curse of dimensionality brought by the vast combinations of antenna configurations. Consider that a minor angle alteration of one antenna affects little on its radiation pattern [31], one potential way is to compress the number of candidate angle settings of each antenna. In other words, only the most promising angle settings are retained to constitute a codebook for the antenna, which serves as a set of candidate settings for the subsequent antenna configuration process. This approach significantly reduces the number of configurations explored during the optimization phase, thus enhancing search efficiency. It should be noted that the selection of elements in the codebook is crucial and aims to concurrently achieve two objectives: Firstly, the power coverage range of the selected antenna setting should be sufficiently expansive; secondly, a significant disparity in power coverage between the selected antenna settings is necessary to avoid redundant searches for analogous configurations.

In this paper, we propose a codebook based Monte Carlo tree search (CB-MCTS) scheme to configure the azimuths and downtilts of massive antennas efficiently. The proposed scheme contains two parts: codebook generation and antenna configuration. During codebook generation, a finite-size codebook is constituted for each antenna by exploring the promising angle settings, and the formulated optimization problem is solved by a branch-and-bound procedure followed by a semi-definite relaxation algorithm. For antenna configuration, a Monte Carlo tree search method is introduced to iteratively optimize the angle settings of antennas under the guidance of the generated codebook, aiming at maximizing both system throughput and power coverage. The main contributions of this paper are summarized as follows:

- We introduce an innovative method for selecting candidate angle settings, designed to identify promising configurations for each antenna. This approach forms a codebook that, to the best of our knowledge, represents the first study in the RNO focused on expediting antenna configuration optimization by pre-filtering favorable angle settings. Our proposed codebook strategy minimizes unnecessary exploration within the solution space, significantly decreasing search overhead.
- We cast the codebook creation process as a binary quadratic programming problem, which we efficiently tackle using a branch and bound method. Furthermore, we address the challenge of determining the upper bound of the branch, typically difficult to obtain due to its non-convex objective function, by transforming the original problem into a semi-definite programming problem.
- We present a Monte Carlo tree search technique, guided by the generated codebooks, to configure the azimuths and downtilts of all antennas within the target area. This approach produces promising antenna configurations while conserving computational resources.
- We assess the proposed CB-MCTS method through comprehensive experiments. Numerical results indicate that the CB-MCTS approach surpasses existing state-of-the-art solutions, striking a balance between system performance and search load. In summary, our proposed strategy offers valuable insights into optimally and efficiently configuring antenna angle settings for practical mmWave communication systems.

The remainder of this paper is organized as follows. In Section II, we give system model and optimization task. In Section III, the CB-MCTS scheme is illustrated in detail. Section IV provides numerical results and discussions. Section V concludes this paper.

II. System Model and Problem Formulation

Consider an urban area served by a set $\mathbb{B} = \{1, \cdots, B\}$ of base stations. Each base station of the network contains three directional antennas. The set of antennas is represented by $\mathbb{A}$, where $|\mathbb{A}| = 3B$. For each antenna $a \in \mathbb{A}$, its azimuth θ_a can be adjusted within $[0°, 360°)$ and its downtilt ϕ_a can be adjusted within $[-90°, 90°]$. Let $\boldsymbol{\theta} = \{\theta_a\}_{a\in\mathbb{A}}$ and $\boldsymbol{\phi} = \{\phi_a\}_{a\in\mathbb{A}}$ denote the azimuth set and downtilt set of all antennas, respectively. The target area is divided into 5m × 5m square grids [32], as shown in Fig. 1. The center of each grid is termed as a traffic demand point (TDP) and the set of TDPs is represented by $\mathbb{U}$.

Let $P_{a,u}$ denote the power of reference signal received by TDP $u \in \mathbb{U}$ from antenna $a \in \mathbb{A}$, which is given as follows:

$$P_{a,u} = P^T + G_{a,u} + G_u - L_{a,u}, \tag{1}$$

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2023.3259435

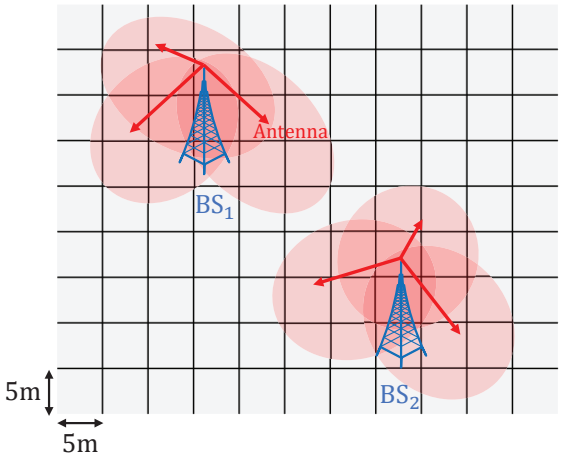


Fig. 1. Illustration of the network scenario.

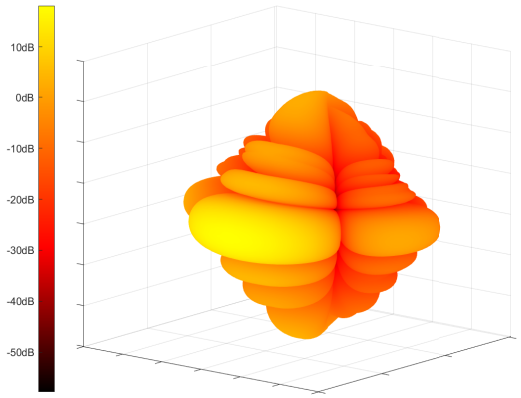


Fig. 2. Radiation pattern of the antenna.

where P^T is the transmit power, $G_{a,u}$ is the antenna transmit gain, G_u is the mobile terminal gain, and $L_{a,u}$ is the path loss between a and u. In addition, the radiation pattern of the antenna is shown in Fig. 2. The transmit gain $G_{a,u}$ varies with the angle between the relative direction of u and the main lobe direction of a. For given a and u, the relative direction of u is fixed, while the main lobe direction of a can be determined by θ_a and ϕ_a. Therefore, $G_{a,u}$ and $P_{a,u}$ are functions of θ_a and ϕ_a, respectively.

Note that each TDP is always associated with the antenna that provides the strongest average reference signal. The valid signal power received by u is defined as

$$P_u(\boldsymbol{\theta}, \boldsymbol{\phi}) = \max_{a \in \mathbb{A}} P_{a,u}(\theta_a, \phi_a), \tag{2}$$

and the signal-to-interference ratio (SIR) at u can be expressed as

$$\rho_u(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{P_u(\boldsymbol{\theta}, \boldsymbol{\phi})}{\sum_{a \in \mathbb{A}} P_{a,u}(\theta_a, \phi_a) - P_u(\boldsymbol{\theta}, \boldsymbol{\phi})}. \tag{3}$$

The power coverage, which refers to the ratio of the TDPs with maximum signal strength exceeding a certain threshold T^p to the total number of TDPs, is calculated by

$$\eta^p(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{\sum_{u \in \mathbb{U}} \mathbb{1}(P_u(\boldsymbol{\theta}, \boldsymbol{\phi}) > T^p)}{|\mathbb{U}|}, \tag{4}$$

where $|\mathbb{U}|$ is the cardinality of the set $\mathbb{U}$, $\mathbb{1}(x)$ is an indicator denoted by

$$\mathbb{1}(x) = \begin{cases} 1 & , \text{ if } x \textit{ is } \textbf{true}, \\ 0 & , \text{ otherwise.} \end{cases} \tag{5}$$

The capacity coverage η^c, which refers to the ratio of the TDPs with SIR exceeding a certain threshold T^c to the total number of TDPs, is expressed as

$$\eta^c(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{\sum_{u \in \mathbb{U}} \mathbb{1}(\rho_u(\boldsymbol{\theta}, \boldsymbol{\phi}) > T^c)}{|\mathbb{U}|}. \tag{6}$$

The aim is to provide both high η^p and high η^c in the target area by adjusting $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$. Therefore, a new indicator called valid coverage ratio is introduced as the metric of network performance, which is defined as

$$\eta^e(\boldsymbol{\theta}, \boldsymbol{\phi}) = \frac{\sum_{u \in \mathbb{U}} \min(\mathbb{1}(P_u(\boldsymbol{\theta}, \boldsymbol{\phi}) > T^p), \mathbb{1}(\rho_u(\boldsymbol{\theta}, \boldsymbol{\phi}) > T^c))}{|\mathbb{U}|}. \tag{7}$$

The proposed indicator is designed to monitor the tradeoff between signal strength and SIR, which represents the ratio of TDPs whose signal quality meet the demand. More specifically, the signal received by TDP u is considered to meet the requirement only if both the signal strength $P_u(\boldsymbol{\theta}, \boldsymbol{\phi})$ and SIR $\rho_u(\boldsymbol{\theta}, \boldsymbol{\phi})$ are above the thresholds, T^p and T^c, respectively.

The optimization problem can be formulated as follows:

$$\begin{aligned} \max_{\boldsymbol{\theta}, \boldsymbol{\phi}} \quad & \eta^e(\boldsymbol{\theta}, \boldsymbol{\phi}) \\ \text{s.t.} \quad & \theta_a \in \{0°, 5°, 10°, ..., 355°\} \quad \forall a \in \mathbb{A} \\ & \phi_a \in \{-90°, -89°, -88°, ..., 90°\} \quad \forall a \in \mathbb{A}, \end{aligned} \tag{8}$$

where the angle range of azimuths is discretized into $5°$ intervals and the angle range of downtilts is discretized into $1°$ intervals. The azimuth and downtilt configuration combinations for antenna a is $\Gamma_a = \{(\theta_a^i, \phi_a^i)\}_{i=1}^D$, where D is the number of combinations, i.e., 72×181 in this case.

III. Codebook Based RNO Scheme

The optimization task (8) defines a combinatorial optimization problem which is intractable even for medium-sized number of antennas. Specifically, for a single antenna, 72×181 angle settings need to be considered to find out the optimal one. The number of configuration combinations for 30 antennas is up to $(72 \times 181)^{30}$ and cannot be enumerated even with huge computing resources.

We propose a codebook based antenna configuration scheme to handle this problem efficiently. First, we present a codebook generation algorithm to select a set of candidate angle settings for each antenna and constitute the codebook. Next, we propose a Monte Carlo tree search based antenna configuration method, where the azimuths and downtilts of antennas are optimized based on the predefined codebooks.

A. Codebook Generation

As mentioned above, the number of feasible azimuth and downtilt configuration combinations for each antenna is $D = 72 \times 181$ in our case, which is redundant in the process of massive antenna configuration and requires considerably high search overhead. Given that minor changes in the azimuth or downtilt have negligible effects on the radiation pattern, a feasible solution to reduce the searching complexity is to select some promising angle settings for each individual antenna. These angle settings constitute a codebook which is then leveraged to guide the configuration of massive antennas.

Due to the effect of the surrounding environment including terrain and buildings, identical angle settings may yield divergent power coverage outcomes for distinct antennas. Therefore, a site-specific codebook design scheme should be conducted to generate a unique codebook for each antenna. In other words, instead of constructing a general codebook for all antennas, each antenna is given its own specifically designed codebook, where the elements are determined based on the surrounding propagation environment of the antenna. Note that the antennas in a BS share a common codebook since their surroundings are the same.

The procedure of constructing a codebook for an antenna is as follows. Consider antenna a with a set of all feasible angle settings $\Gamma_a = \{(\theta_a^i, \phi_a^i)\}_{i=1}^D$. Without loss of generality, let $\boldsymbol{C}$ denote the matrix that indicates the discrepancies between the elements of Γ_a. The element $c_{i,j}$ in $\boldsymbol{C}$, which represents the discrepancy of power coverage between the two angle settings (θ_a^i, ϕ_a^i) and (θ_a^j, ϕ_a^j), is defined as

$$c_{i,j} \triangleq \frac{\sum_{u\in\mathbb{U}}\left(\mathbb{1}(P_{a,u}(\theta_a^i,\phi_a^i) > T^p) \oplus \mathbb{1}(P_{a,u}(\theta_a^j,\phi_a^j) > T^p)\right)}{|\mathbb{U}|}, \tag{9}$$

where $\oplus$ is the exclusive-or operator, i.e.,

$$x \oplus y = \begin{cases} 0 & , \quad \text{if } x = y, \\ 1 & , \quad \text{otherwise.} \end{cases} \tag{10}$$

To select typical angle settings for the codebook, it is essential to maximize the disparity in power coverage among these settings. Hence, the optimization objective in codebook generation is given by

$$\begin{aligned} \max_{\tilde{\boldsymbol{x}}} \quad & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} = \sum_{i=1}^{D}\sum_{j=1}^{D} c_{i,j}x_i x_j \\ \text{s.t.} \quad & \boldsymbol{e}^T\boldsymbol{x} = K \\ & \boldsymbol{x} \in \{0,1\}^D, \end{aligned} \tag{11}$$

where $x_i \in \{0,1\}$ is a decision variable for each angle setting ($x_i = 1$ if the i-th angle setting is included in the codebook, and $x_i = 0$ if is not), $\boldsymbol{C}$ is a $D \times D$ symmetric matrix, $\boldsymbol{e}$ is a D-dimensional all one vector, K is the predefined size of codebook. It is worth noting that the coverage of each setting is relatively extensive although the objective in (11) is to maximize the coverage discrepancy, since the difference between two settings with small coverage range is usually smaller than the difference between settings with large coverage.

Observe that (11) is a linearly constraint binary quadratic programming (BQP) problem, which is known to be a non-deterministic polynomial-time hard problem [33]. Branch-and-bound (B&B) is an extension of brute-force and is typically used to find out the optimal solution for the BQP problem. A B&B procedure usually performs two main tasks namely branching and bounding [34]. The branching step divides the solution space into two or more subspaces to be investigated in the subsequent iteration and dynamically generates a search tree. The bounding step finds the upper bound for the subspace and compares it to the current best solution. If the upper bound of a subspace is lower than the best solution obtained so far, the whole subspace is discarded. Otherwise it is checked whether a leaf node is reached, and if so, the best solution is updated if the value of this solution is greater than the current best one. This procedure terminates until all the subspaces are visited or discarded.

One key of the B&B procedure is to obtain a tight upper bound for each subspace so that it can be decided whether to explore this subspace or not. However, the matrix $\boldsymbol{C}$ in (11) is not semi-definite since its diagonal elements equal to 0, which makes the objective function non-convex. As a result, the upper bound is hard to obtain by linear programming relaxation. Semi-definite relaxation (SDR) is an efficient approximation technique for host of non-convex optimization problems, which relaxes the original problem to a semi-definite programming problem so that polynomial methods can be utilized to solve it [35]. In this work, the SDR technology is embedded in the B&B procedure to solve the codebook generation problem.

Several transforms for (11) are formulated before SDR to make the upper bound tight. Define $\boldsymbol{z} = 2\boldsymbol{x} - \boldsymbol{e}$, the problem (11) can be reformulated by

$$\begin{aligned} \max_{\boldsymbol{z}} \quad & \frac{1}{4}(\boldsymbol{z}^T\boldsymbol{C}\boldsymbol{z} + 2\boldsymbol{z}^T\boldsymbol{C}\boldsymbol{e} + \boldsymbol{e}^T\boldsymbol{C}\boldsymbol{e}) \\ \text{s.t.} \quad & \boldsymbol{e}^T\boldsymbol{z} = 2K - D \\ & \boldsymbol{z} \in \{-1,1\}^D, \end{aligned} \tag{12}$$

which is equal to the following optimization problem

$$\begin{aligned} \max_{\tilde{\boldsymbol{z}}} \quad & \tilde{\boldsymbol{z}}^T\tilde{\boldsymbol{C}}\tilde{\boldsymbol{z}} \\ \text{s.t.} \quad & \tilde{\boldsymbol{z}}^T\tilde{\boldsymbol{C}}_0\tilde{\boldsymbol{z}} = 4K - 2D \\ & \tilde{\boldsymbol{z}} \in \{-1,1\}^{D+1}, \end{aligned} \tag{13}$$

where $\tilde{\boldsymbol{z}} = \begin{pmatrix} z_0 \\ \boldsymbol{z} \end{pmatrix}$, and the symmetric matrices $\tilde{\boldsymbol{C}}$ and $\tilde{\boldsymbol{C}}_0$ are represented by

$$\tilde{\boldsymbol{C}} = \frac{1}{4}\begin{bmatrix} \boldsymbol{e}^T\boldsymbol{C}\boldsymbol{e} & \boldsymbol{e}^T\boldsymbol{C} \\ \boldsymbol{C}\boldsymbol{e} & \boldsymbol{C} \end{bmatrix}, \quad \tilde{\boldsymbol{C}}_0 = \begin{bmatrix} 0 & \boldsymbol{e}^T \\ \boldsymbol{e} & \boldsymbol{0} \end{bmatrix}.$$

The equivalence between (12) and (13) can be easily proved as follows: On one hand, if $\boldsymbol{z}^*$ is optimal in (12), then $\begin{pmatrix} 1 \\ \boldsymbol{z}^* \end{pmatrix}$ and $\begin{pmatrix} -1 \\ -\boldsymbol{z}^* \end{pmatrix}$ are both feasible in (13) with the same objective value. On the other hand, if $\begin{pmatrix} z_0 \\ \boldsymbol{z}^* \end{pmatrix}$ is optimal in (13), then $z_0\boldsymbol{z}^*$ is feasible in (12) with the same objective value. Thus, (12) and (13) share the same optimal objective value. □

A crucial step to derive an SDR is to observe that

$$\tilde{\boldsymbol{z}}^T\tilde{\boldsymbol{C}}\tilde{\boldsymbol{z}} = \mathrm{Tr}(\tilde{\boldsymbol{z}}^{\mathrm{T}}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{z}}) = \mathrm{Tr}(\tilde{\boldsymbol{C}}\tilde{\boldsymbol{z}}\tilde{\boldsymbol{z}}^{\mathrm{T}}),$$
$$\tilde{\boldsymbol{z}}^T\tilde{\boldsymbol{C}}_0\tilde{\boldsymbol{z}} = \mathrm{Tr}(\tilde{\boldsymbol{z}}^{\mathrm{T}}\tilde{\boldsymbol{C}}_0\tilde{\boldsymbol{z}}) = \mathrm{Tr}(\tilde{\boldsymbol{C}}_0\tilde{\boldsymbol{z}}\tilde{\boldsymbol{z}}^{\mathrm{T}}).$$

By introducing a new variable $\tilde{\boldsymbol{Z}} = \tilde{\boldsymbol{z}}\tilde{\boldsymbol{z}}^T$ and noticing that $\tilde{\boldsymbol{Z}}$ is a rank one symmetric positive semi-definite matrix whose diagonal elements are all 1, we obtain the following equivalent formulation of (13):

$$\begin{aligned}
\max_{\tilde{Z}} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}\tilde{\boldsymbol{Z}}) \\
\text{s.t.} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}_0\tilde{\boldsymbol{Z}}) = 4K - 2D \\
& \mathrm{Tr}(\boldsymbol{E}_i\tilde{\boldsymbol{Z}}) = 1 \quad \forall i \in \{1, 2, ..., D\} \\
& \tilde{\boldsymbol{Z}} = \tilde{\boldsymbol{Z}}^T \\
& \tilde{\boldsymbol{Z}} \succeq \boldsymbol{0} \\
& \mathrm{Rank}(\tilde{\boldsymbol{Z}}) = 1, \qquad (14)
\end{aligned}$$

where $\boldsymbol{E}_i$ is a $(D+1) \times (D+1)$ matrix comprised of all zero entries, except for a distinct value of 1 located at (i, i). The only difficult constraint in (14) is the rank one constraint, which is non-convex. Thus, we may drop it to obtain a relaxed version:

$$\begin{aligned}
\max_{\tilde{Z}} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}\tilde{\boldsymbol{Z}}) \\
\text{s.t.} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}_0\tilde{\boldsymbol{Z}}) = 4K - 2D \\
& \mathrm{Tr}(\boldsymbol{E}_i\tilde{\boldsymbol{Z}}) = 1 \quad \forall i \in \{1, 2, ..., D\} \\
& \tilde{\boldsymbol{Z}} = \tilde{\boldsymbol{Z}}^T \\
& \tilde{\boldsymbol{Z}} \succeq \boldsymbol{0}. \qquad (15)
\end{aligned}$$

The problem (15) is known as an SDR of (14).

For the optimization problem of a subspace with some elements of $\boldsymbol{x}$ fixed, the formulation can be written by

$$\begin{aligned}
\max_{\boldsymbol{x}} \quad & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{e}^T\boldsymbol{x} = K \\
& x_i = q_i \quad \forall i \in \mathbb{I} \\
& \boldsymbol{x} \in \{0, 1\}^D, \qquad (16)
\end{aligned}$$

where $\mathbb{I}$ is the index set of fixed elements in $\boldsymbol{x}$, and q_i is a constant with a value of 0 or 1. The corresponding SDR of (16) is expressed as

$$\begin{aligned}
\max_{\tilde{Z}} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}\tilde{\boldsymbol{Z}}) \\
\text{s.t.} \quad & \mathrm{Tr}(\tilde{\boldsymbol{C}}_0\tilde{\boldsymbol{Z}}) = 4K - 2D \\
& \mathrm{Tr}(\boldsymbol{E}_i\tilde{\boldsymbol{Z}}) = 1 \quad \forall i \in \{1, 2, ..., D\} \\
& \tilde{z}_{i+1,j+1} = (2q_i - 1)(2q_j - 1) \quad \forall i, j \in \mathbb{I} \\
& \tilde{\boldsymbol{Z}} = \tilde{\boldsymbol{Z}}^T \\
& \tilde{\boldsymbol{Z}} \succeq \boldsymbol{0}. \qquad (17)
\end{aligned}$$

This problem can be solved by using existing software such as MATLAB with the convex optimization toolbox CVX [36]. Therefore, the upper bound of a subspace can be calculated according to (17).

Algorithm 1 Bound initialization

Initialization: Codebook size K, $\boldsymbol{x} = \boldsymbol{0}$.
for $k = 1 : K$ **do**
  if $k == 1$ **then**
    $i^* \leftarrow \arg\max_i \sum_{u \in \mathbb{U}} \mathbb{1}(P_{a,u}(\phi_{a,i}) > T^p)$
  else
    $i^* \leftarrow \arg\max_i (\boldsymbol{x} + \boldsymbol{e_i})^T\boldsymbol{C}(\boldsymbol{x} + \boldsymbol{e_i})$
  end if
  $\boldsymbol{x} \leftarrow \boldsymbol{x} + \boldsymbol{e_{i^*}}$
end for
return $\boldsymbol{x}$

Algorithm 2 DFS($\boldsymbol{x}$,i,$\boldsymbol{x}^*$,V^*)

if $\|\boldsymbol{x}\|_1 > K$ **then**
  return $\boldsymbol{x}^*$,V^*
end if
for $x_i = 0 : 1$ **do**
  if $i == D$ **then**
    if $\boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} > V^*$ **then**
      $\boldsymbol{x}^* \leftarrow \boldsymbol{x}$
      $V^* \leftarrow \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x}$
    end if
  else
    $V \leftarrow$ value of (17) using CVX
    if $V \leq V^*$ **then**
      return $\boldsymbol{x}^*$,V^*
    else
      $\boldsymbol{x}^*$,$V^* \leftarrow$ DFS($\boldsymbol{x}$,i+1,$\boldsymbol{x}^*$,V^*)
    end if
  end if
end for
return $\boldsymbol{x}^*$,V^*

With the above bounding strategy, the B&B procedure to constitute the codebook for antenna $a \in \mathbb{A}$ is then designed as follows:

Bound initialization: Before performing the branching strategy, an initial bound should be determined to guide the discard. Therefore, we propose a heuristic algorithm to obtain an initial solution for the B&B algorithm, which gives a tight upper bound. The heuristic algorithm is based on greedy strategy. At the beginning, the codebook is set to empty and the corresponding $\boldsymbol{x}$ is a D-dimensional zero vector. Then, the antenna setting with the largest coverage ratio is put into the codebook, while the corresponding element of $\boldsymbol{x}$ is changed to 1. Subsequently, the setting that differs the most from those in the codebook is iteratively put into the codebook until the predefined size is reached. The vector $\boldsymbol{x}$ is iteratively updated at the same time. When the initialization process has been completed, the latest version of $\boldsymbol{x}$ is the initialized codebook we need. The detailed procedure of initialization is illustrated in Algorithm 1. Note that $\boldsymbol{e_i}$ in the algorithm is a D-dimensional vector where the i-th element is 1 and the other elements are 0.

Codebook generation: The codebook generation procedure

Algorithm 3 Codebook generation

Initialization: Root node n_0,
The codebook $\Psi_a = \varnothing$.
$\boldsymbol{x}^* \leftarrow$ Bound initialization
$V^* \leftarrow \boldsymbol{x}^{*T}\boldsymbol{C}\boldsymbol{x}^*$
$\boldsymbol{x}^*, V^* \leftarrow$ DFS($\boldsymbol{x}$,1,$\boldsymbol{x}^*$,V^*)
for $i = 1 : D$ **do**
  if $x_i^* == 1$ **then**
    $\Psi_a \leftarrow \Psi_a \cup \{(\theta_a^j, \phi_a^j)\}$
  end if
end for
return Ψ_a

is performed based on a depth-first search tree. Specifically, the initial solution $\boldsymbol{x}^*$ is obtained through Algorithm 1 and the initialized bound V^* is calculated by $\boldsymbol{x}^{*T}\boldsymbol{C}\boldsymbol{x}^*$. It starts from the root node where no element in $\boldsymbol{x}$ is determined. Two branches originate from the root node: In the first branch, $x_1 = 1$, signifying that the first antenna setting in set Γ_a is put into the codebook; whereas, in the second branch, $x_1 = 0$, indicating that the first setting is not selected. The upper bound of each branch is calculated by solving the SDR of the corresponding optimization problem. If the obtained upper bound is lower than the initial bound V^*, the corresponding branch is discarded and all the children branches of it are also removed. If the obtained upper bound is higher than V^*, the corresponding branch is chosen to proceed in the next layer. This procedure is recursively executed by the depth first search (DFS) function illustrated in Algorithm 2. The value of x^* and V^* are updated when reaching a leaf node where all elements of $\boldsymbol{x}$ are determined and $\boldsymbol{x}^T C \boldsymbol{x} > V^*$. It terminates until all the nodes are visited or discarded. Ultimately, the optimal solution $\boldsymbol{x}^*$ is obtained, and the codebook Ψ_a is constituted. The whole codebook generation procedure is summarized in Algorithm 3.

B. Antenna Configuration

The number of candidate settings for each antenna is reduced from D to K with the help of the obtained codebooks. Consequently, the number of setting combinations for all antennas becomes $K^{|\mathbb{A}|}$, which is considerably less than $D^{|\mathbb{A}|}$. However, finding the optimal configuration for massive antennas from these candidate settings still requires a high searching overhead if exhaustive search is adopted.

Reinforcement learning methods have demonstrated exceptional ability in handling Markov decision process (MDP) [37]. Fortunately, the configuration of antennas can also be formulated as an MDP, which provides a significant opportunity to solve it using RL methods. To convert the antenna configuration problem to an MDP, we define the agent, state, action and reward accordingly:

Agent: Each antenna in set $\mathbb{A}$ is treated as an agent.

Action: Each action corresponds to one possible configuration of an antenna based on its codebook.

State: The state is defined as the signal power distribution and the SIR distribution in the target area, which is measured by all TDPs.

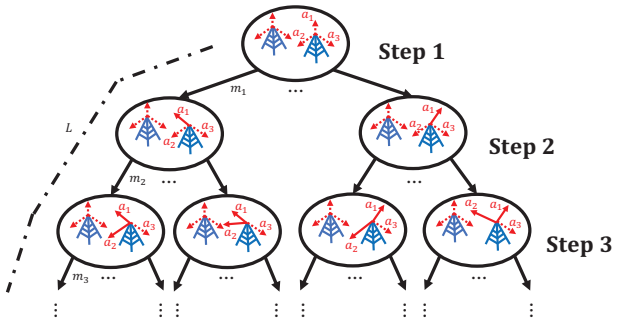


Fig. 3. Search tree for RNO.

Reward: The reward is the valid coverage ratio η^e of the network calculated by (7).

The MDP derived from the antenna configuration problem is formed as a tree search process where the antennas are configured successively, as shown in Fig. 3. The antenna to be configured at step j is represented by a_j, and the state with antennas $\{a_1, a_2, ..., a_{j-1}\}$ already configured is represented by s_j. Note that s_1 refers to the initial state where no antennas have been configured. For each step j, antenna a_j can take K available actions based on the corresponding codebook Ψ_{a_j}. Given the current state s_j, action ψ_j is selected by a_j from Ψ_{a_j}, which leads to a new state s_{j+1}. Then it is the turn of the next antenna a_{j+1} to choose its action. When all the antennas have been configured step by step, the reward η^e can be worked out with the action trajectory $L = [\psi_1, \psi_2, ..., \psi_{|\mathbb{A}|}]$. Clearly, the trajectory with the largest reward is the optimal antenna configuration we want to find.

Enumerating all trajectories in the search tree is infeasible in practical scenarios, however, the exploration can be more efficient if the branches with lower rewards are visited as few as possible. The Monte Carlo tree search (MCTS) provides a promising procedure to evaluate the rewards of branches in an efficient way. It usually consists of the following four steps:

Selection: The selection starts at the root node, and the child nodes are selected successively according to a predefined selection policy until a node with unexplored branches is reached.

Expansion: One random child node is expanded from the node chosen in the selection step. A statistic set containing the estimated reward and the number of visits of the child node is initialized according to the feasible actions.

Simulation: Simulation is executed to evaluate the reward of the node initialized in the expansion step. Monte Carlo method can implement such a simulation, which determines the reward repeatedly through random samplings.

Backup: The reward obtained in the simulation step is backed up through the tree to update the statistic set of the branch with regard to its corresponding ancestor nodes.

The MCTS is utilized to iteratively optimize the azimuths and downtilts of all antennas with the guidance of the generated antenna setting codebooks, which is called codebook based MCTS (CB-MCTS). Denote $\{Q(s,\psi), V(s,\psi)\}$ as the statistic set of the branch (s,ψ), where s is the current state, ψ is the action, $Q(s,\psi)$ is the evaluation reward and $V(s,\psi)$

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2023.3259435

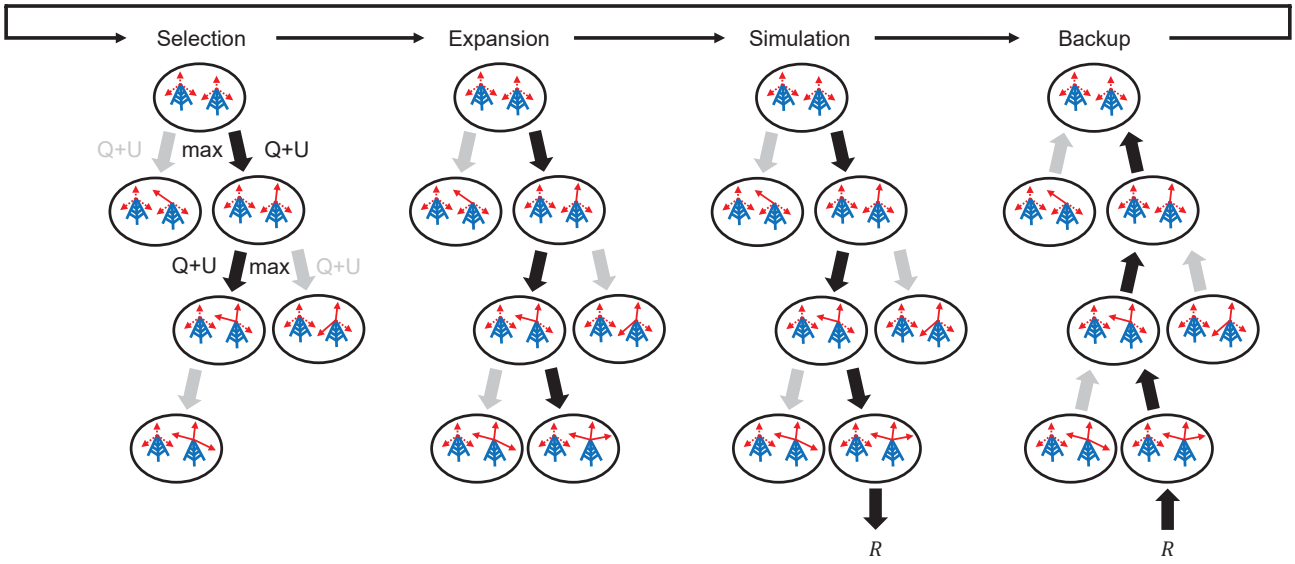


Fig. 4. Monte Carlo tree search Procedure.

is the visit count. The algorithm proceeds by iterating over the above mentioned four phases (in Fig. 4).

The selection step begins at the root node s_1, and finishes when reaching a node s_{l_e} with unexplored branches in step l_e. At each step $l < l_e$, an action ψ_l is chosen from the codebook Ψ_{a_l}, expressed as

$$\psi_l = \arg\max_{\psi \in \Psi_{a_l}} (Q(s_l, \psi) + U(s_l, \psi)), \tag{18}$$

where $Q(s, \psi)$ is the evaluation reward indicating the instance reward for current state and action, while $U(s_l, \psi)$ denotes the upper confidence bound [38], which is given by

$$U(s_l, \psi) = c \frac{\sqrt{\sum_{\psi' \in \Psi_{a_l}} V(s_l, \psi')}}{1 + V(s_l, \psi)}, \tag{19}$$

where c is a parameter determining the preference on exploration and $V(s, \psi)$ is the number of times the action ψ has been visited so far. The term $U(s_l, \psi)$ here balances between exploiting the antenna configuration that looks currently the best and the exploration of currently suboptimal configurations, so as to avoid missing good alternatives due to early evaluation errors. It can be seen that the selection policy proposed in (18) prefers the action with low visit count $V(s_l, \psi)$ in the early stage of exploration, but asymptotically tends to select the action with high estimated value $Q(s_l, \psi)$ when more evaluations are executed.

If the node is not fully expanded, the expansion step is performed to obtain a successor state s_{l_e}, which is determined by the previous state and the randomly chosen action ψ_{l_e}. The statistic set of branch (s_{l_e}, ψ_{l_e}) is initialized to

$$\{Q(s_{l_e}, \psi_{l_e}) = 0, V(s_{l_e}, \psi_{l_e}) = 0\}.$$

Obviously, $Q(s_{l_e}, \psi_{l_e})$ requires to be estimated as accurate as possible so that a good decision can be made in the selection step.

In the simulation step, the reward R of (s_{l_e}, ψ_{l_e}), which refers to the valid coverage ratio η^e worked out through (s_{l_e}, ψ_{l_e}), is obtained by Monte Carlo method. Specifically, for those antennas that need to be configured after step l_e, actions are selected at random until reaching a leaf node. Multiple simulations are executed in parallel, meanwhile, the valid coverage ratio η^e of each simulation is calculated and R values for the maximal η^e of these simulated configurations.

After finishing the simulation step, the reward R will be backed up to the previous branches. In each step $l \le l_e$, R is fed back to (s_l, ψ_l) and the statistic set of the corresponding branch is updated as follows:

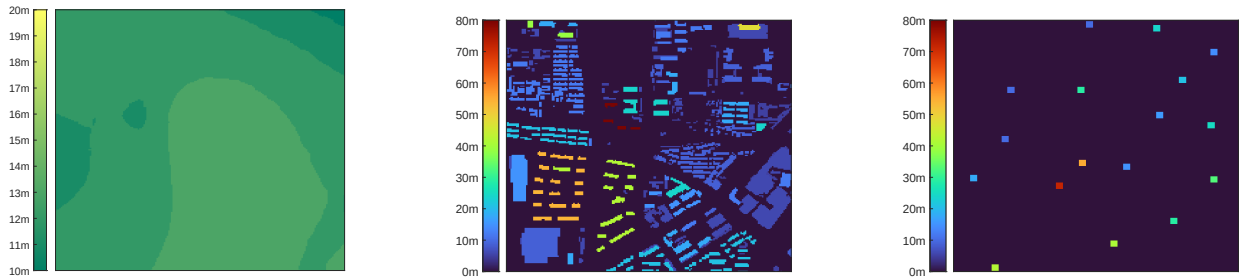
$$V(s_l, \psi_l) = V(s_l, \psi_l) + 1, \tag{20}$$

$$Q(s_l, \psi_l) = \max(Q(s_l, \psi_l), R). \tag{21}$$

The iteration is running over selection, expansion, simulation and backup steps until reaching a predefined computational budge, such as the maximum number of iterations Y in our case. When the computational budge is reached, the search will be halted and the antenna configuration ψ_h^* for the root state s_h is determined based on the policy $\pi(\psi_h|s_h)$, which is expressed as:

$$\pi(\psi_h|s_h) = \frac{V(s_h, \psi_h)^\tau}{\sum_{\psi_h' \in M(s_h)} V(s_h, \psi_h')^\tau}, \tag{22}$$

where τ denotes a temperature parameter that controls the level of exploration, ψ_h is the available antenna configuration in codebook $M(s_h)$. Then the child node corresponding to (s_h, ψ_h^*) becomes the new root node, and the subtree below this node is retained along with its statistic set. This process terminates until one of the leaf node becomes the root node [39]. By this time, all the antennas have been configured and $[\psi_1^*, \psi_2^*, ..., \psi_{|\mathbb{A}|}^*]$ denotes the final antenna configuration. The CB-MCTS procedure is summarized in Algorithm 4.



(a) Terrain information. (b) Building layout. (c) BS locations and heights.

Fig. 5. Map information.

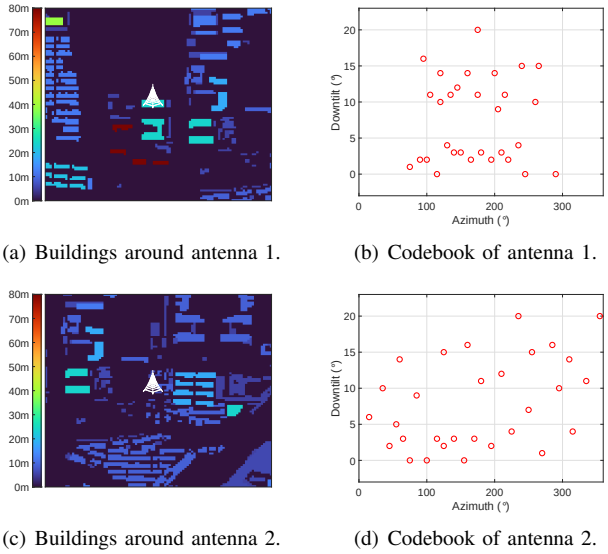
Algorithm 4 Codebook based Monte Carlo tree search

Initialization: root node s_1.
Counter $h \leftarrow 1$
for $h = 1 : |\mathbb{A}|$ **do**
  for $y = 1 : Y$ **do**
    $s_l \leftarrow s_h$
    while all children of s_l are visited **do**
      Select ψ_l using (18)
      Update $s_l \leftarrow$ child of s_l with ψ_l
    end while
    Expand s_l by a random ψ_l
    $Q(s_l, \psi_l) = 0, V(s_l, \psi_l) = 0$
    $R \leftarrow$ simulation outcome of (s_l, ψ_l)
    do
      $V(s_l, \psi_l) \leftarrow V(s_l, \psi_l) + 1$
      $Q(s_l, \psi_l) \leftarrow \max(Q(s_l, \psi_l), R)$
      $\psi_l \leftarrow$ action connecting s_l and its parent
      $s_l \leftarrow$ parent of s_l
    while $s_l \neq s_h$
  end for
  Generate policy $\pi(\psi|s_h)$ using (22)
  $\psi_h^* \leftarrow \pi(\psi|s_h)$
  $s_{h+1} \leftarrow (s_h, \psi_h^*)$
  $h \leftarrow h + 1$
end for
return $[\psi_1^*, \psi_2^*, ..., \psi_{|\mathbb{A}|}^*]$

IV. Numerical Results

Consider a 1km×1km urban area whose terrain information and building layout are shown in Fig. 5(a) and Fig. 5(b), respectively. The network in this area is composed of 17 BSs, whose locations and heights are depicted in Fig. 5(c). The transmit power P^T is set to 15.4dBm, and the mobile terminal gain G_u is set to 0dBm.

We adopt the COST 231 model [40] as the path loss model to validate our proposed method. Line-of-sight (LOS) and none-line-of-sight (NLOS) cases are separately considered.



(a) Buildings around antenna 1. (b) Codebook of antenna 1.

(c) Buildings around antenna 2. (d) Codebook of antenna 2.

Fig. 6. Building layout and codebook of different antennas.

The path loss can be calculated by

$$L_{a,u} = \begin{cases} 11.85 + 47.69 \lg(d_{a,u}) + 5.83 \lg(h_a) \\ \quad + B_u - 6.55 \lg(d_{a,u}) \lg(h_a) \quad , \text{ if LOS}, \\ 16.24 + 47.69 \lg(d_{a,u}) + 5.83 \lg(h_a) \\ \quad + B_u - 6.55 \lg(d_{a,u}) \lg(h_a) \quad , \text{ if NLOS}, \end{cases} \tag{23}$$

where $d_{a,u}$ denotes the distance between antenna a and TDP u, and h_a denotes the height of a. The building penetration loss B_u is simplified as a discrete variable, given by

$$B_u = \begin{cases} 14\text{dB}, & \text{if u is indoor}, \\ 0\text{dB}, & \text{if u is outdoor}. \end{cases} \tag{24}$$

The building layouts and the corresponding codebooks for antenna 1 and 2 are depicted in Fig. 6. It is noteworthy that the codebook of antenna 1 significantly differs from that of antenna 2, primarily attributed to the distinct spatial distributions of buildings surrounding the two antennas. Consequently, the specifically designed codebook captures the characteristics of the propagation environment in the vicinity of the respective antenna. Also, only typical angle settings are retained while analogous ones are eliminated.

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2023.3259435

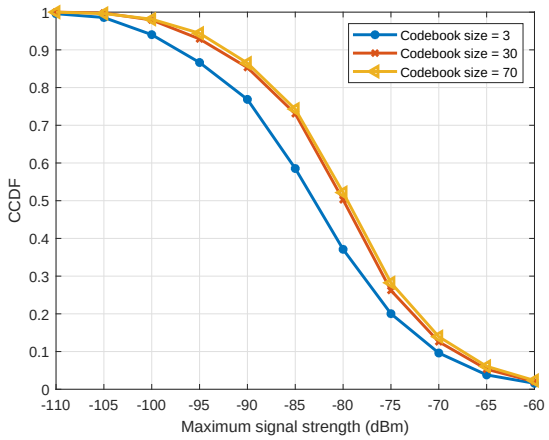


Fig. 7. Complementary cumulative distribution of maximum signal strength.

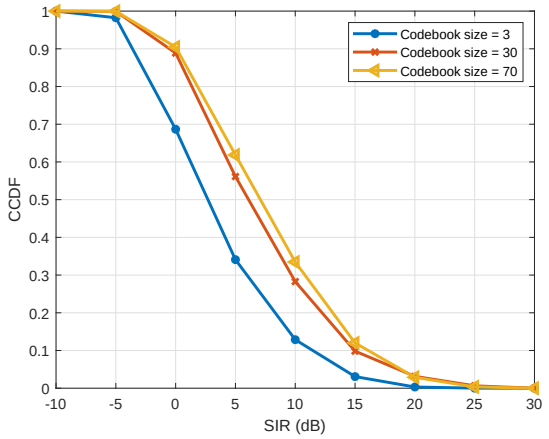


Fig. 8. Complementary cumulative distribution of SIR.

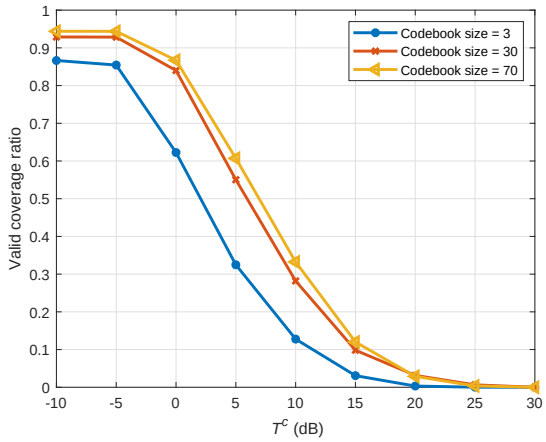


Fig. 9. Valid coverage ratio as a function of SIR threshold.

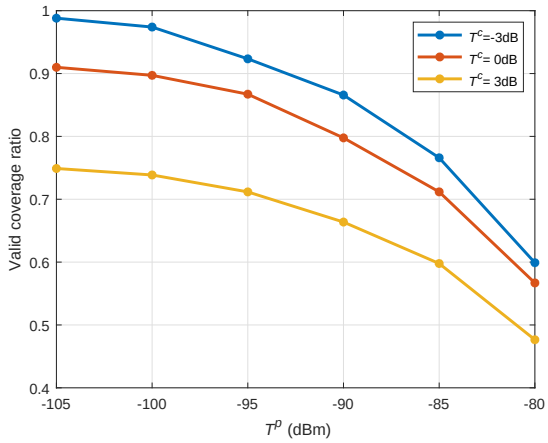


Fig. 10. Valid coverage ratio under different power thresholds and SIR thresholds.

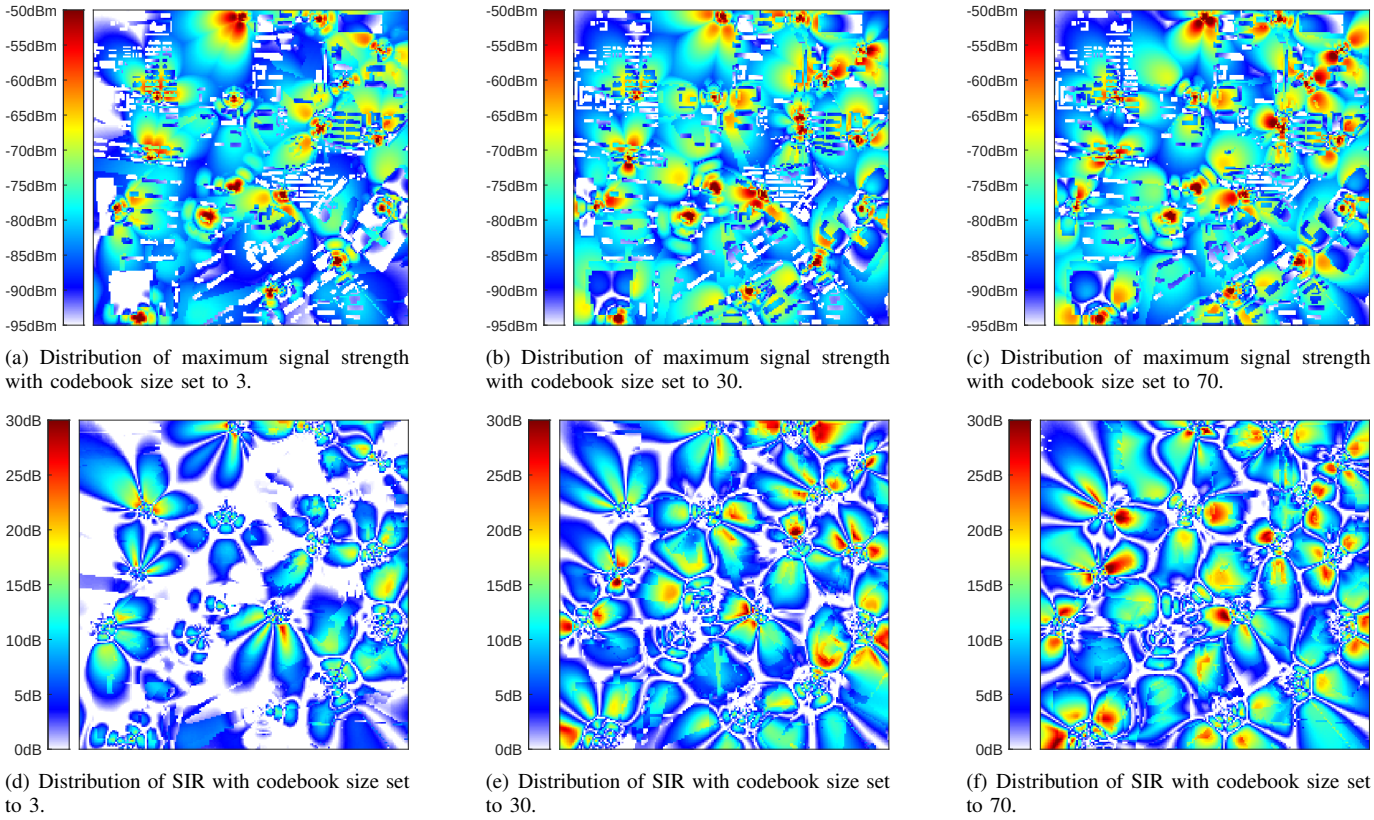
Fig. 7 shows the complementary cumulative distribution function (CCDF) of the maximum signal strength received at TDPs under different codebook sizes. The power threshold T^p is set to -95dBm and the SIR threshold T^c is set to 0dB. Observe that when the codebook size is set to 70, the proportion of TDPs with signal strength exceeding the threshold is about 96%. Even when the codebook size is 3, more than 85% of the TDPs can meet the power threshold. It suggests that most of the TDPs can receive strong enough signal in a densely deployed network scenario even if the antenna parameters of BSs are not optimized.

Similarly, the CCDF of SIR under different codebook sizes is shown in Fig. 8. The power threshold T^p is -95dBm and the SIR threshold T^c is 0dB. As depicted in Fig. 8, 90.47% of the TDPs can meet T^c with codebook size 70, on the contrast, the ratio is only 68.66% for codebook size 3. As the codebook size expands, it encompasses an increased number of candidate antenna configurations, thereby enhancing the likelihood of containing an advantageous antenna configuration. It is observed that the two curves corresponding to codebook size of 30 and 70 are close to each other. This arises since the elements in the codebook are adequate enough for the antenna configuration once the size of the codebook surpasses a threshold. In this case, the discrepancy between different antenna configurations in the codebook is too small to cause a significant difference in capacity coverage. Thus, it is unnecessary to set the size of the codebook too large considering the search cost.

Fig. 9 shows the valid coverage ratio as a function of SIR threshold T^c under different codebook sizes, where the power threshold T^p is fixed at -95dBm. When the SIR threshold is 0dB and the codebook size is 70, the CB-MCTS scheme achieves a valid coverage ratio of 86.71%. More specifically, the valid coverage ratio is positively correlated with the SIR distribution, which results in the consist trend of the corresponding curves in Fig. 8 and Fig. 9.

Fig. 10 shows the valid coverage ratio under different power thresholds and SIR thresholds, where the size of codebook is 70. We can see that when the SIR threshold T^c is fixed, the valid coverage ratio decreases as the power threshold T^p increases. Similarly, when the power threshold T^p is fixed, the valid coverage ratio decreases as the SIR threshold T^c increases. Besides, observe that the curve of higher SIR threshold declines slower than that of the lower one. This can

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2023.3259435



(a) Distribution of maximum signal strength with codebook size set to 3.

(b) Distribution of maximum signal strength with codebook size set to 30.

(c) Distribution of maximum signal strength with codebook size set to 70.

(d) Distribution of SIR with codebook size set to 3.

(e) Distribution of SIR with codebook size set to 30.

(f) Distribution of SIR with codebook size set to 70.

Fig. 11. The distribution of signal strength and SIR of the CB-MCTS scheme under different codebook sizes.

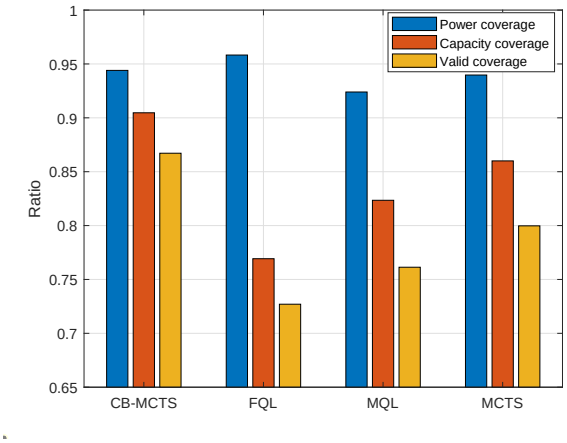


Fig. 12. Network performance of different algorithms.

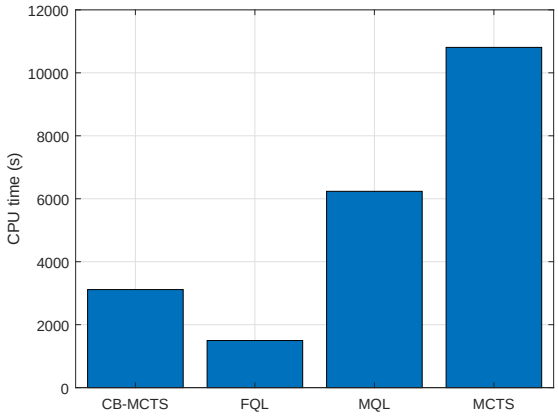


Fig. 13. CPU time of different algorithms.

be explained as follows: Under the same power threshold T^p, the TDPs which meet higher SIR threshold T^c are more likely to have higher received signal strength. Therefore, when the power threshold T^p increases, for those TDPs meeting the high T^c, the number of TDPs becoming invalid is smaller than that of the TDPs meeting the low T^c. Consequently, the higher the SIR threshold, the slower the curve declines.

Further, the detailed performance of the CB-MCTS in terms of the distribution of maximum signal strength and SIR under different codebook sizes are shown in Fig. 11(a)-(c) show that the size of codebook has little influence on the power coverage, since the demand on signal strength is easy to meet as long as the BSs are dense enough. Observe that most TDPs with poor signal strength are located inside buildings, which enlightens us to deploy additional network support devices indoors to meet the coverage requires. Contrast to the maximum signal strength distribution, the SIR distribution is more likely to be affected by the codebook size especially when the size is small, which can be seen from Fig. 11(d)-(f). It is because the small number of candidate configurations restricts the flexibility of antenna patterns, which leads to a high probability of interference between antennas.

The coverage ratio of our proposed scheme CB-MCTS is compared with the following ones: the fuzzy Q-learning (FQL) [27], the mean field based Q-learning (MQL) [28] and the MCTS [30]. The performance of different methods are illustrated in Fig. 12. The codebook size of the CB-MCTS is 70. The blue bar refers to the power coverage ratio, the red bar refers to the capacity coverage ratio, and the orange bar refers to the valid coverage ratio. The CB-MCTS scheme performs better than FQL and MQL since the latter two algorithms address only one antenna without considering the interference between adjacent antennas. On the other hand, ineffective configurations are not filtered before exploration for the MCTS method, resulting in a performance loss as compared to the CB-MCTS. Last, we compare the CPU time of the algorithms in Fig. 13. The computational overhead of the CB-MCTS is about 30% of that of the MCTS due to the introduction of codebooks. From Fig. 12 and Fig. 13, we can conclude that the CB-MCTS achieves a good tradeoff between the valid coverage and the computational complexity.

V. Conclusion

In this paper, we proposed a codebook-based antenna angle optimization scheme for mmWave mobile communication systems, aiming to enhance system throughput and power coverage. We designed a set of candidate azimuth and downtilt configurations, referred to as the codebook, for each antenna based on its surrounding propagation environment, thereby avoiding unnecessary exploration of poorly performing and similar antenna configurations. Subsequently, we introduced a Monte Carlo tree search method to iteratively optimize the azimuths and downtilts of all antennas, guided by the generated codebooks. Numerical results demonstrate that our proposed scheme can produce promising antenna configurations with limited computing resources, making it suitable for large-scale networks.

Acknowledgment

The authors would like to thank the editors and the anonymous reviewers, whose invaluable comments helped improve the presentation of this paper substantially.

References

[1] L. Shen and S. Wang, "An efficient codebook based radio parameter optimization method for mobile networks," in *Proc. IEEE GLOBECOM'22*, Rio de Janeiro, Brazil, Dec. 2022.

[2] M. Shafi *et al.*, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, pp. 1201–1221, Jun. 2017.

[3] Y. Heng *et al.*, "Six key challenges for beam management in 5.5G and 6G systems," *IEEE Commun. Mag.*, vol. 59, pp. 74–79, Jul. 2021.

[4] S. Wang and C. Ran, "Rethinking cellular network planning and optimization," *IEEE Wireless Commun.*, vol. 23, pp. 118–125, Apr. 2016.

[5] W. Zhang *et al.*, "Toward intelligent network optimization in wireless networking: An auto-learning framework," *IEEE Wireless Commun.*, vol. 26, pp. 76–82, Jun. 2019.

[6] L. Liu *et al.*, "Joint optimization of scheduling and power control in wireless networks: Multi-dimensional modeling and decomposition," *IEEE Trans. Mobile Comput.*, vol. 18, pp. 1585–1600, Jul. 2019.

[7] S. Wang, W. Zhao, and C. Wang, "Budgeted cell planning for cellular networks with small cells," *IEEE Trans. Veh. Technol.*, vol. 64, pp. 4797–4806, Oct. 2015.

[8] W. Zhao *et al.*, "Approximation algorithms for cell planning in heterogeneous networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1561–1572, Feb. 2017.

[9] W. Zhao and S. Wang, "Traffic density-based rrh selection for power saving in c-ran," *IEEE J. Sel. Areas in Commun.*, vol. 34, no. 12, pp. 3157–3167, Dec. 2016.

[10] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, pp. 2522–2545, May 2016.

[11] S. Hurley, "Planning effective cellular mobile radio networks," *IEEE Trans. Veh. Technol.*, vol. 51, pp. 243–253, Mar. 2002.

[12] N. Himayat *et al.*, "Interference management for 4G cellular standards [WIMAX/LTE UPDATE]," *IEEE Commun. Mag.*, vol. 48, pp. 86–92, Aug. 2010.

[13] A. Engels *et al.*, "Autonomous self-optimization of coverage and capacity in LTE cellular networks," *IEEE Tran. Veh. Technol.*, vol. 62, pp. 1989–2004, Jun. 2013.

[14] S. Berger *et al.*, "Joint downlink and uplink tilt-based self-organization of coverage and capacity under sparse system knowledge," *IEEE Trans. Veh. Technol.*, vol. 65, pp. 2259–2273, April 2016.

[15] A. Imran, "Self organization of tilts in relay enhanced networks: A distributed solution," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 764–779, Feb. 2014.

[16] Y. Niu *et al.*, "A survey of millimeter wave communications (mmwave) for 5G: Opportunities and challenges," *Wireless Netw.*, vol. 21, pp. 2657–2676, Apr. 2015.

[17] S. Razavizadeh, M. Ahn, and I. Lee, "Three-dimensional beamforming: A new enabling technology for 5G wireless networks," *IEEE Signal Process. Mag.*, vol. 31, pp. 94–101, Oct. 2014.

[18] Q. Li *et al.*, "Physical layer enhancement for next-generation railway communication systems," *IEEE Access*, vol. 10, pp. 83152–83175, Aug. 2022.

[19] M. Rebato *et al.*, "Stochastic geometric coverage analysis in mmwave cellular networks with realistic channel and antenna radiation models," *IEEE Trans. Commun.*, vol. 67, pp. 3736–3752, May 2019.

[20] W. He *et al.*, "Intelligent optimization of base station array orientations via scenario-specific modeling," *IEEE Trans. Commun.*, vol. 70, pp. 2117–2130, Mar. 2022.

[21] K. Hassan *et al.*, "Channel estimation techniques for millimeter-wave communication systems: Achievements and challenges," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1336–1363, Aug. 2020.

[22] S. Hurley *et al.*, "Modelling and planning fixed wireless networks," *Wireless Netw.*, vol. 16, pp. 577–592, Apr. 2010.

[23] Y. Yoon and Y. Kim, "An efficient genetic algorithm for maximum coverage deployment in wireless sensor networks," *IEEE Trans. Cybern.*, vol. 43, pp. 1473–1483, Oct. 2013.

[24] B. Partov, D. Leith., and R. Razavi, "Utility fair optimization of antenna tilt angles in LTE networks," *IEEE/ACM Trans. Netw.*, vol. 23, pp. 175–185, Feb. 2015.

[25] Y. Liu *et al.*, "An efficient stochastic gradient descent algorithm to maximize the coverage of cellular networks," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 3424–3436, Jul. 2019.

[26] R. Dreifuerst *et al.*, "Optimizing coverage and capacity in cellular networks using machine learning," in *Proc. IEEE ICASSP'21*, Toronto, ON, Canada, Jun. 2021.

[27] V. Buenestado *et al.*, "Self-tuning of remote electrical tilts based on call traces for coverage and capacity optimization in LTE," *IEEE Trans. Veh. Technol.*, vol. 66, pp. 4315–4326, May 2017.

[28] E. Balevi and J. Andrews, "Online antenna tuning in heterogeneous cellular networks with deep reinforcement learning," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, pp. 1113–1124, Dec. 2019.

[29] S. Marco *et al.*, "Cellular network capacity and coverage enhancement with MDT data and deep reinforcement learning," *Comput. Commun.*, vol. 195, pp. 403–415, Sept. 2022.

[30] L. Shen and S. Wang, "Monte Carlo Tree Search for network planning for next generation mobile communication networks," in *Proc. IEEE GLOBECOM'21*, Madrid, Spain, Dec. 2021.

[31] S. Mabrouki *et al.*, "Codebook designs for millimeter-wave communication systems in both low- and high-mobility: Achievements and challenges," *IEEE Access*, vol. 10, pp. 25786–25810, 2022.

[32] Y. Zhang and S. Wang, "K-nearest neighbors gaussian process regression for urban radio map reconstruction," *IEEE Commun. Lett.*, vol. 26, pp. 3049–3053, Sept. 2022.

[33] M. Garey and D. Johnson, *Computers and Intractabilityy: A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman, 1979.

[34] G. Kochenberger *et al.*, "The unconstrained binary quadratic programming problem: a survey," *J. Combinatorial Optim.*, vol. 28, pp. 58–81, Apr. 2014.
[35] Z. Luo *et al.*, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, 2010.
[36] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," 2008, available: https://web.stanford.edu/~boyd/software.html.
[37] J. Chen *et al.*, "Intelligent massive MIMO antenna selection using Monte Carlo tree search," *IEEE Trans. Signal Process.*, vol. 67, pp. 5380–5390, Oct. 2019.
[38] C. Rosin, "Multi-armed bandits with episode context," *Ann. Math. Artif. Intell.*, vol. 61, pp. 203–230, Aug. 2011.
[39] D. Silver *et al.*, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, pp. 484–489, Jan. 2016.
[40] E. Damosso and L. Correia, *COST Action 231: Digital Mobile Radio Towards Future Generation Systems: Final Report*. Brussels, Belgium: European Commission, 1999.

Linzhi Shen received the B.S. degree from Nanjing University, Nanjing, China, in 2018, where he is currently pursuing the Ph.D. degree with the School of Electronic Science and Engineering. His current research interests include wireless network optimization, operations research and machine learning.

Yifang Zhang received the B.S. degree from Huazhong University of Science and Technology, Wuhan, China, in 2021. She is currently pursuing the M.S. degree with the School of Electronic Science and Engineering, Nanjing University, Nanjing, China. Her research interests include wireless network optimization and machine learning.

Shaowei Wang (S'06-M'07-SM'13) received the Ph.D. degree from Wuhan University, Wuhan, China, in 2006. He joined the School of Electronic Science and Engineering, Nanjing University, Nanjing, China, as a Faculty Member in 2006, where he is currently a Full Professor. From 2012 to 2013, he was a Visiting Scholar/a Professor with Stanford University, Stanford, CA, USA, and The University of British Columbia, Vancouver, BC, Canada. His research interests include communications and networking, operations research and machine learning.