

Machine Learning-based Beamforming Design for Millimeter Wave IRS Communications with Discrete Phase Shifters

Wencai Yan, Gangcan Sun, Wanming Hao, Zhengyu Zhu, Zheng Chu, and Pei Xiao

Abstract—In this paper, we investigate an intelligent reflecting surface (IRS)-assisted millimeter-wave multiple-input single-output downlink wireless communication system. By jointly calculating the active beamforming at the base station and the passive beamforming at the IRS, we aim to minimize the transmit power under the constraint of each user's signal-to-interference-plus-noise ratio. To solve this problem, we propose a low-complexity machine learning-based cross-entropy (CE) algorithm to alternately optimize the active beamforming and the passive beamforming. Specifically, in the alternative iteration process, the zero-forcing (ZF) method and CE algorithm are applied to acquire the active beamforming and the passive beamforming, respectively. The CE algorithm starts with random sampling, by the idea of distribution focusing, namely shifting the distribution towards a desired one by minimizing CE, and a near optimal reflection coefficients with adequately high probability can be obtained. In addition, we extend the original one-bit phase shift at the IRS to the common case with high-resolution phase shift to enhance the effectiveness of the algorithms. Simulation results verify that the proposed algorithm can obtain a near optimal solution with lower computational complexity.

Index Terms—Intelligent reflecting surface, mmWave, discrete phase shifts, machine learning.

I. INTRODUCTION

Millimeter-wave (mmWave) and intelligent reflecting surface (IRS) have recently been considered as two promising technologies for supporting the high-speed data transmission in future wireless communications [1]-[3]. Because of the small wavelength of the mmWave, a large antenna array can be easily deployed in a compact form to achieve higher array and multiplexing gains. Moreover, the high directive mmWave beam reduces the multi-user interference. However, the severe path loss of the mmWave signals results in limited coverage [3]. To tackle this problem, the IRS, a passive planar surface composing of massive intelligent reflector units, is developed, and the signal transmission direction can be changed by designing the amplitude/phase of the reflector elements so as to improve the coverage. In addition, the IRS is also energy efficient without requiring any extra radio frequency (RF) chain [1]. As such the mmWave IRS communications will be a promising enabling technology for future wireless networks [6]-[10].

One of the main challenges in the IRS communications is beamforming optimization. Primitively, the perfect reflector surface is considered, namely the continuous phase shifts at the IRS and the active beamforming at the base station (BS)

are jointly optimized to minimize the transmit power [7], and a semidefinite relaxation (SDR)-based alternating optimization method is proposed. Instead of single IRS, the authors in [8] consider a cooperative IRS communication scenario. Based on this, an SDR-based bisection method of cooperative passive beamforming at the IRSs and active beamforming at the BS is developed to maximize the minimum signal-to-interference-plus-noise ratio (SINR). Later, considering the practical hardware limit, only the discrete phase shifts at the IRS are applied. The authors of [4] assume that the phase resolution is infinite to make the targeted problem more tractable. The authors of [5] first solve problem with all discrete optimization variables relaxed to their continuous counterparts and then directly quantize each of the obtained continuous phase shifts to its nearest discrete value. The authors in [9] propose a successive refinement method to formulate the effective passive beamforming by adjusting the discrete phase shifts. Similarly, the authors of [11] also consider the finite-level phase shifts and propose an exhaustive search algorithm. However, for the above schemes, when the number of the reflective elements becomes larger, the computational complexity is extremely high.

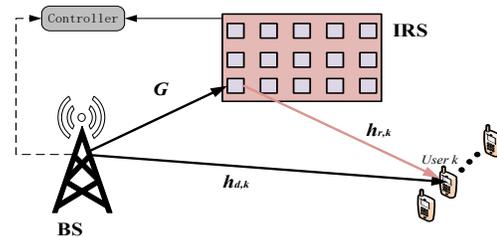


Fig. 1: The downlink mmWave IRS communication systems

To reduce the computational complexity of optimizing the discrete phase shifts, in this paper, a low-complexity machine learning-based optimization scheme in the mmWave IRS communications is proposed. By jointly optimizing the active beamforming at the BS and passive beamforming at the IRS, we first consider a one-bit phase shift at the IRS, and formulate a problem of minimizing the transmit power under the constraint of user's SINR. Specifically, on the basis of the probability distributions of the reflection coefficients at the IRS, we first randomly generate several groups of IRS reflection coefficients. Then, based on the generated reflection coefficients, the zero-forcing (ZF) method is applied to obtain the active beamforming at the BS. Next, we calculate the transmit power and select the reflection coefficients corresponding to minimum transmit power to update the probability distributions of the reflection coefficients at the IRS by minimizing CE. The above procedure starts with random initialization, but by iteratively minimizing

W. Yan, G. Sun, W. Hao and Z. Zhu are with the School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China. (E-mail: yanwencai001@163.com, {jegcsun, iewmhao, iezyzhu}@zzu.edu.cn)

Z. Chu and P. Xiao is with the 5G Innovation Centre, Institute of Communication Systems, University of Surrey, Guildford GU2 7XH, U.K. (E-mail: andrew.chuzheng7@gmail.com, p.xiao@surrey.ac.uk).

the CE distance, and it becomes a highly effective learning algorithm. The above steps are repeated until convergence, such that near optimal reflection coefficients and active beamforming are obtained. Finally, we extend the CE algorithm to the common case with high-resolution phase shifts. Experimental results verify that the proposed algorithm strikes a good balance between performance and computational complexity.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

As shown in Fig. 1, we consider a downlink mmWave IRS communication system, which comprises an M -antenna BS, an N -element IRS and K single-antenna users. Meanwhile, the BS and the IRS are deployed as uniform planer array (UPA). We assume that the CSIs of all links can be obtained using existing channel estimation schemes [7], [15]. The mmWave channels of the BS-IRS link, the IRS-user link and the BS-user link are, respectively denoted by $\mathbf{G} \in \mathbb{C}^{N \times M}$, $\mathbf{h}_{r,k} \in \mathbb{C}^{N \times 1}$ and $\mathbf{h}_{d,k} \in \mathbb{C}^{M \times 1}$. \mathbf{G} can be expressed as [15]

$$\mathbf{G} = \sqrt{\frac{MN}{L_G}} \sum_{l_1}^{L_G} \alpha_{l_1}^{G_i} \mathbf{b}(\vartheta_{l_1}^{G_i}, \psi_{l_1}^{G_i}) \mathbf{a}(\vartheta_{l_1}^{G_r}, \psi_{l_1}^{G_r})^T, \quad (1)$$

where L_G denotes the paths number, $\alpha_{l_1}^{G_i}$, $\vartheta_{l_1}^{G_i}$ ($\psi_{l_1}^{G_i}$) and $\vartheta_{l_1}^{G_r}$ ($\psi_{l_1}^{G_r}$) denote the complex gain, the azimuth (elevation) angle of departure, and the azimuth (elevation) angle of arrival for the l_1 th path, respectively. Similarly, $\mathbf{h}_{r,k}$ and $\mathbf{h}_{d,k}$ are given by

$$\mathbf{h}_{r,k} = \sqrt{\frac{N}{L_{r,k}}} \sum_{l_2=1}^{L_{r,k}} \alpha_{l_2}^{r,k} \mathbf{a}(\vartheta_{l_2}^{r,k}, \psi_{l_2}^{r,k}), \quad (2)$$

$$\mathbf{h}_{d,k} = \sqrt{\frac{M}{L_{d,k}}} \sum_{l_3=1}^{L_{d,k}} \alpha_{l_3}^{d,k} \mathbf{b}(\vartheta_{l_3}^{d,k}, \psi_{l_3}^{d,k}), \quad (3)$$

where $L_{r,k}$ denotes the number of paths between the the IRS and the k th user, $L_{d,k}$ represents the number of paths between the BS and the k th user, $\mathbf{a}(\vartheta, \psi) \in \mathbb{C}^{N \times 1}$ and $\mathbf{b}(\vartheta, \psi) \in \mathbb{C}^{M \times 1}$ represent array steering vectors at the IRS and the BS, respectively. For the IRS, we use a UPA with N_1 elements in horizon and N_2 elements in vertical ($N = N_1 N_2$), and thus $\mathbf{a}(\vartheta, \psi)$ is expressed as

$$\mathbf{a}(\vartheta, \psi) = \frac{1}{\sqrt{N}} \left[e^{-j2\pi d \sin(\vartheta) \cos(\psi) \mathbf{n}_1 / \lambda} \right] \otimes \left[e^{-j2\pi d \sin(\psi) \mathbf{n}_2 / \lambda} \right], \quad (4)$$

where $\mathbf{n}_1 = [0, 1, \dots, N_1 - 1]$ and $\mathbf{n}_2 = [0, 1, \dots, N_2 - 1]$, λ represents the signal wavelength, d denotes the element spacing and usually set as $d = \lambda/2$. $\mathbf{b}(\vartheta, \psi)$ has the similar expression, and we omit it due to the limited space.

The received signal at user k is formulated as

$$y_k = (\mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G} + \mathbf{h}_{d,k}^H) \sum_{j=1}^K \mathbf{w}_j s_j + n_k, \quad k \in [0, 1, \dots, K], \quad (5)$$

where $\mathbf{\Phi} = \text{diag}(\varphi_1, \dots, \varphi_N)$ is the $N \times N$ diagonal reflection matrix of the IRS, reflection coefficient φ_n is the n th element of the IRS and defined as $\varphi_n = \beta_n e^{j\theta_n}$, in which $\beta_n \in [0, 1]$ and $\theta_n \in [0, 2\pi)$ respectively denote the amplitude reflection coefficient and the reflection phase shift, $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ is the active

beamforming at the BS for user k , s_k represents the message intended to be received by user k satisfying $E[|s_k|^2] = 1$, $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ is the additive zero average white Gaussian noise (AWGN) with variance of σ_k^2 at user k . For the IRS, we only optimize the phase shifts and set $\beta_n = 1, n \in [0, 1, \dots, N]$ for maximizing the reflection efficiency [9]. Considering the practical hardware limit, we assume that Q -bit finite-level phase shift is deployed at each element and employ the uniform quantization method to produce 2^Q discrete phase-shift values. Therefore, the set of discrete phase-shift values θ is given by

$$\mathcal{S}_\theta = \{0, \Delta\theta, \dots, (2^Q - 1)\Delta\theta\}, \quad (6)$$

where $\Delta\theta = 2\pi/2^Q$. The set of discrete reflection coefficient φ values is denoted by

$$\mathcal{S}_\varphi = \{e^{j0}, e^{j\Delta\theta}, \dots, e^{j(2^Q-1)\Delta\theta}\}. \quad (7)$$

The SINR of user k is formulated as

$$\text{SINR}_k = \frac{\left| (\mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G} + \mathbf{h}_{d,k}^H) \mathbf{w}_k \right|^2}{\sum_{j \neq k}^K \left| (\mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G} + \mathbf{h}_{d,k}^H) \mathbf{w}_j \right|^2 + \sigma_k^2}. \quad (8)$$

B. Problem Formulation

We aim to calculate the active beamforming \mathbf{W} at the BS and the reflection coefficients φ at the IRS to minimize the transmit power of the BS, where $\varphi = [\varphi_1, \dots, \varphi_N]$ and $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_K] \in \mathbb{C}^{M \times K}$. The optimization problem is given by

$$\text{P1} : \min_{\mathbf{W}, \varphi} \sum_{k=1}^K \|\mathbf{w}_k\|^2 \quad (9a)$$

$$\text{s.t.} \quad \frac{\left| (\mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G} + \mathbf{h}_{d,k}^H) \mathbf{w}_k \right|^2}{\sum_{j \neq k}^K \left| (\mathbf{h}_{r,k}^H \mathbf{\Phi} \mathbf{G} + \mathbf{h}_{d,k}^H) \mathbf{w}_j \right|^2 + \sigma_k^2} \geq \gamma_k, \forall k, \quad (9b)$$

$$\varphi_n \in \mathcal{S}_\varphi, \quad n \in [0, 1, \dots, N], \quad (9c)$$

where γ_k is the SINR requirement of user k . Note that the optimal solution to P1 is intractable due to the non-convex constraints. The non-convexity is caused by the coupled variables (i.e., \mathbf{W} and φ) and the discrete values of the reflection coefficient φ . To solve it, one generally adopted method is iterative optimization algorithm. Specifically, for any given reflection coefficient φ , the active beamforming \mathbf{W} is obtained by employing the minimum mean squared error (MMSE) or the suboptimal ZF-based method with low computational complexity. Then, \mathbf{W} and φ are updated alternatively until convergence. However, existing algorithms for optimizing reflection coefficient φ are usually based on the exhaustive search or successive refinement [9], which results in higher computational complexity. To work around this issue, we will develop a machine learning-based CE algorithm with low complexity to obtain the reflection coefficient φ .

III. PROPOSED SOLUTION

In this section, we propose a machine learning-based CE algorithm and utilize an alternatively iterative method to solve problem P1. We first randomly generate several reflection coefficients $\boldsymbol{\varphi}$ based on the probability distributions. Based on the generated reflection coefficients $\boldsymbol{\varphi}$, we solve the active beamforming \mathbf{W} by applying the ZF scheme for obtaining a tradeoff between performance and complexity. Next, we update the probability distributions of the reflection coefficient at the IRS to regenerate the reflection coefficients $\boldsymbol{\varphi}$. Repeating above steps, \mathbf{W} and $\boldsymbol{\varphi}$ are optimized alternately until convergence.

We first introduce the machine learning-based CE algorithm, and the aim is to obtain the (approximate) optimal performance for a learning task [12]-[14]. It comprises two phases: 1) Generating several random data on the basis of a specified probability distribution; 2) Based on certain criteria (e.g., minimum transmit power in this work), selecting several best data as “elite” to update the probability distribution parameters by minimizing the CE. Repeating such procedure, the probability distribution will be refined to generate a solution close to the optimal one with a sufficiently high probability. Next, the specific steps of the proposed algorithm as follows:

When $Q = 1$, the phase shift $\theta_n \in \mathcal{S}_\theta = \{0, \pi\}$, the reflection coefficient $\varphi_n \in \mathcal{S}_\varphi = \{-1, +1\}$. First, we define the probability parameter $\mathbf{p} = [p_1, \dots, p_N]$ as a $1 \times N$ vector, where $0 \leq p_n \leq 1$ denotes the probability of reflection coefficient $\varphi_n = 1$. Because there is no priori information, we initialize the parameter $\mathbf{p}^{(0)} = \frac{1}{2} \times \mathbf{1}_{1 \times N}$, where $\mathbf{1}$ is a vector with all-one elements. In other words, in the initialization phase, we assume that all the reflection coefficients belong to $\{-1, +1\}$ with equal probability. At the i th iteration, we generate S candidate reflection coefficients $\{\boldsymbol{\varphi}^{[s]}\}_{s=1}^S$ according to the probability distribution $\mathcal{T}(\mathcal{S}_\varphi; \mathbf{p}^{(i)})$. Under any given reflection coefficients $\boldsymbol{\varphi}^{[s]}$, the transmit power minimization problem P1 can be formulated as

$$\text{P2} : \min_{\mathbf{W}} \sum_{k=1}^K \|\mathbf{w}_k\|^2 \quad (10a)$$

$$\text{s.t.} \quad \frac{\left| (\mathbf{h}_{r,k}^H \boldsymbol{\Phi} \mathbf{G} + \mathbf{h}_{d,k}^H) \mathbf{w}_k \right|^2}{\sum_{j \neq k} \left| (\mathbf{h}_{r,k}^H \boldsymbol{\Phi} \mathbf{G} + \mathbf{h}_{d,k}^H) \mathbf{w}_j \right|^2 + \sigma_k^2} \geq \gamma_k, \forall k. \quad (10b)$$

We apply the low-complexity ZF technique to acquire the active beamforming \mathbf{W} , which can be computed as

$$\mathbf{W} = \mathbf{H}^H (\mathbf{H} \mathbf{H}^H)^{-1} \mathbf{U}^{\frac{1}{2}}, \quad (11)$$

where $\mathbf{H} = \mathbf{H}_r^H \boldsymbol{\Phi} \mathbf{G} + \mathbf{H}_d^H$, $\mathbf{H}_r^H = [\mathbf{h}_{r,1}, \dots, \mathbf{h}_{r,K}]^H$ and $\mathbf{H}_d^H = [\mathbf{h}_{d,1}, \dots, \mathbf{h}_{d,K}]^H$, $\mathbf{U} = \text{diag}(u_1, \dots, u_K)$ is the power allocation matrix. According to [9], the constraint (10b) can be reformed as $u_k / \sigma_k^2 \geq \gamma_k, \forall k$. To obtain the optimal solution, we have $u_k = \sigma_k^2 \gamma_k, \forall k$. The transmit power of BS can be computed as

$$\sum_{k=1}^K \|\mathbf{w}_k\|^2 = \text{tr} \left(\mathbf{U}^{\frac{1}{2}} (\mathbf{H} \mathbf{H}^H)^{-1} \mathbf{U}^{\frac{1}{2}} \right) = \text{tr} \left(\mathbf{U} (\mathbf{H} \mathbf{H}^H)^{-1} \right) = \mathcal{P}(\boldsymbol{\varphi}). \quad (12)$$

Algorithm 1: The proposed machine learning-based CE algorithm

- 1 **Input:** Channel matrix \mathbf{G} , $\mathbf{h}_{r,k}$, $\mathbf{h}_{d,k}$; Number of candidates and elites S , S_{elite} ; Number of iterations I .
 - 2 **Initialization:** $i = 0$, $\mathbf{p}^{(0)} = \frac{1}{2} \times \mathbf{1}_{N \times 1}$.
 - 3 **for** $i = 0$ **to** I **do**
 - 4 Generate S random candidate reflection coefficients $\{\boldsymbol{\varphi}^{[s]}\}_{s=1}^S$ according to the $\mathcal{T}(\mathcal{S}_\varphi; \mathbf{p}^{(i)})$;
 - 5 Obtain the active beamforming \mathbf{W} based on (11);
 - 6 Calculate the S corresponding transmit power $\{\mathcal{P}(\boldsymbol{\varphi}^{[s]})\}_{s=1}^S$ based on (12);
 - 7 Sort $\{\mathcal{P}(\boldsymbol{\varphi}^{[s]})\}_{s=1}^S$ in a descend order and select S_{elite} reflection coefficients with the lowest transmit power as elite samples;
 - 8 Update $\mathbf{p}^{(i+1)}$ based on (17);
 - 9 **end**
 - 10 **Output:** \mathbf{W} , $\boldsymbol{\varphi}$.
-

Therefore, the problem P2 can be transformed to

$$\text{P3} : \min_{\boldsymbol{\varphi}} \mathcal{P}(\boldsymbol{\varphi}) \quad (13a)$$

$$\text{s.t.} \quad \boldsymbol{\varphi} \in \{\boldsymbol{\varphi}^{[1]}, \dots, \boldsymbol{\varphi}^{[S]}\}. \quad (13b)$$

Then, we compute S transmit power $\{\mathcal{P}(\boldsymbol{\varphi}^{[s]})\}_{s=1}^S$ corresponding to the candidate reflection coefficients $\{\boldsymbol{\varphi}^{[s]}\}_{s=1}^S$ based on (12) and sort $\{\mathcal{P}(\boldsymbol{\varphi}^{[s]})\}_{s=1}^S$ in a descend order. After that, we choose S_{elite} reflection coefficients with the lowest transmit power as elite samples to update $\mathbf{p}^{(i+1)}$ by minimizing CE

$$\mathbf{p}^{(i+1)} = \arg \max_{\mathbf{p}^{(i)}} \frac{1}{S} \sum_{s=1}^{S_{\text{elite}}} \ln \mathcal{T}(\mathcal{S}_\varphi^{[s]}; \mathbf{p}^{(i)}). \quad (14)$$

Note that the n th element $\varphi_n^{[s]}$ of $\boldsymbol{\varphi}^{[s]}$ complies with Bernoulli distribution. We set the probability of $\varphi_n^{[s]} = +1$ as $p_n^{(i)}$ and the probability of $\varphi_n^{[s]} = -1$ as $1 - p_n^{(i)}$. Thus the probability distribution $\mathcal{T}(\mathcal{S}_\varphi^{[s]}; \mathbf{p}^{(i)})$ is given by

$$\mathcal{T}(\mathcal{S}_\varphi^{[s]}; \mathbf{p}^{(i)}) = \prod_{n=1}^N (p_n^{(i)})^{\frac{1}{2}(1+\varphi_n^{[s]})} (1 - p_n^{(i)})^{\frac{1}{2}(1-\varphi_n^{[s]})}. \quad (15)$$

By substituting (15) into (14), the first derivative of (14) with respect to $p_n^{(i)}$ can be calculated as

$$\frac{1}{S} \sum_{s=1}^{S_{\text{elite}}} \left(\frac{1 + \varphi_n^{[s]}}{2p_n^{(i)}} - \frac{1 - \varphi_n^{[s]}}{2(1 - p_n^{(i)})} \right). \quad (16)$$

Setting (16) as zero, $p_n^{(i+1)}$ is updated by

$$p_n^{(i+1)} = \frac{\sum_{s=1}^{S_{\text{elite}}} (\varphi_n^{[s]} + 1)}{2S_{\text{elite}}}. \quad (17)$$

The new probability distributions is employed to regenerate S candidate reflection coefficients. The above steps are repeated until the number of iterations $i = I$, where I is the number of iterations required for convergence. Finally, near optimal active beamforming \mathbf{W} and reflection coefficients $\boldsymbol{\varphi}$ with minimum

transmit power are obtained. We summarize the above steps in Algorithm 1. Several best data can be selected as “elite” to update the probability distribution parameters at each iteration, and it guarantees that the solution tends to be optimal. Moreover, it has been shown in [13] that the CE method terminates with probability 1 in a finite number of iterations.

The computational complexity of Algorithm 1 is mainly due to: 1) In step 5, we obtain the active beamforming \mathbf{W} based on (11) with the complexity $\mathcal{O}(NSK^2)$. 2) In step 6, we calculate S corresponding transmit power $\{\mathcal{P}(\varphi^{[s]})\}_{s=1}^S$ according to (12), which involves the complexity $\mathcal{O}(S)$. 3) The final part is step 8, we update $\mathbf{p}^{(i+1)}$ on the basis of (17), thus the complexity of this procedure is $\mathcal{O}(NS_{elite})$. After I iterations, the total computational complexity of Algorithm 1 is $\mathcal{O}(INSK^2)$. It can be observed from simulations that a good performance can be obtained even for a small I and S .

Next, we investigate a more common scenario where $Q > 1$. In this case, the phase shift $\theta_n \in \mathcal{S}_\theta = \{0, \Delta\theta, \dots, (2^Q - 1)\Delta\theta\}$, the reflection coefficient $\varphi_n \in \mathcal{S}_\varphi = \{e^{j0}, e^{j\Delta\theta}, \dots, e^{j(2^Q-1)\Delta\theta}\}$, we can generate a random sample $\{\varphi_n\}_{n=1}^N$, φ_n is independently drawn from the set \mathcal{S}_φ and follows discrete distribution $\{p_{q,n}^{(i)}\}_{q=1}^{2^Q}$, where $\{p_{q,n}^{(i)}\}$ is the probability of the q th element in set \mathcal{S}_φ being selected as φ_n . Here, $\{p_{q,n}^{(i)}\}$ satisfies $\sum_{q=1}^{2^Q} p_{q,n}^{(i)} = 1$. We still assume that all the reflection elements belong to the set \mathcal{S}_φ with equal probability in the initialization phase.

The next step involves employing the selected elite samples to update $p_{q,n}^{(i+1)}$ by minimizing CE

$$p_{q,n}^{(i+1)} = \arg \max_{p_{q,n}^{(i)}} \frac{1}{S} \sum_{s=1}^{S_{elite}} \ln \mathcal{T}(\mathcal{S}_\varphi^{[s]}; p_{q,n}^{(i)}), \quad (18)$$

where $\mathcal{T}(\mathcal{S}_\varphi^{[s]}; p_{q,n}^{(i)})$ is calculated as

$$\mathcal{T}(\mathcal{S}_\varphi^{[s]}; p_{q,n}^{(i)}) = \prod_{n=1}^N \sum_{q=1}^{2^Q} p_{q,n}^{(i)} \mathbb{1}_{\{\varphi_n^{[s]} = \mathcal{S}_{\varphi,q}\}}, \quad (19)$$

where $\mathcal{S}_{\varphi,q}$ denotes the q th element in set \mathcal{S}_φ , and the function $\mathbb{1}_{\{\bullet\}} = 1$ when the constraint $\{\bullet\}$ is satisfied, otherwise $\mathbb{1}_{\{\bullet\}} = 0$. Furthermore, to satisfy the constraint of $\sum_{q=1}^{2^Q} p_{q,n}^{(i)} = 1$, we introduce the Lagrange multiplier $\{\mathcal{L}_n\}_{n=1}^N$ into (18) as

$$p_{q,n}^{(i+1)} = \arg \max_{p_{q,n}^{(i)}} \frac{1}{S} \sum_{s=1}^{S_{elite}} \ln \mathcal{T}(\mathcal{S}_\varphi^{[s]}; p_{q,n}^{(i)}) + \sum_{n=1}^N \mathcal{L}_n \left(\sum_{q=1}^{2^Q} (p_{q,n}^{(i)} - 1) \right). \quad (20)$$

Taking the first derivative of (20) with respect to $p_{q,n}^{(i)}$ and setting the result as zero, yield

$$\frac{1}{S} \sum_{s=1}^{S_{elite}} \mathbb{1}_{\{\varphi_n^{[s]} = \mathcal{S}_{\varphi,q}\}} + \mathcal{L}_n p_{q,n}^{(i)} = 0. \quad (21)$$

By adding all of (21) for $q = 1, 2, \dots, 2^Q$, we have $\mathcal{L}_n = -\frac{1}{S} S_{elite}$. Then, by substituting (21) into (20), we obtain

$$p_{q,n}^{(i+1)} = \frac{\sum_{s=1}^{S_{elite}} \mathbb{1}_{\{\varphi_n^{[s]} = \mathcal{S}_{\varphi,q}\}}}{S_{elite}}. \quad (22)$$

IV. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed machine learning-based CE algorithm. In the simulation, we set $L_G = 4$, $L_{r,k} = 5$ and $L_{d,k} = 3$ for $\forall k$. The distance-dependent channel path loss of BS-IRS link, IRS-user link and BS-user link are $10^{-3}d_{BR}^{-2.2}$, $10^{-3}d_{RU}^{-2.8}$, $10^{-3}d_{BU}^{-3.5}$, respectively. d_{BR} , d_{RU} , d_{BU} respectively denote the distance of BS-IRS link, IRS-user link and BS-user link. We deploy $d_{BR} = 50m$, $d_{RU} = 2m$, $d_{BU} = 60m$ for $\forall k$. We set that the noise power to $\sigma_k^2 = -90\text{dBm}$ for $\forall k$ and all users have the same SINR requirement. Besides, we deploy UPA at both the BS and IRS. The antenna spacing is set as $d = \lambda/2$. The carrier frequency is 30 GHz.

First, we evaluate the convergence of the proposed algorithm. Fig. 2(a) shows the BS transmit power versus the iteration number under different number of candidates with $Q = 1$, when $S_{elite}/S = 0.2$, $M = 64$, $N = 625$, $K = 4$, and $\gamma = 20\text{dB}$. One can observe that the transmit power first decreases and then tends to be stabilized upon a number of iterations, which shows the effectiveness of the proposed algorithm. Meanwhile, it can be seen that a larger S can obtain a lower transmit power. However, it is sufficient to obtain a near optimal performance when $S = 200$.

Fig. 2(b) compares the computational complexity between our proposed algorithm and the successive refinement algorithm [9], where we adopt $S = 200$, $S_{elite} = 40$, $M = 64$, $K = 4$. According to [9], the computational complexity of the successive refinement algorithm is $\mathcal{O}(I_{iter}2^Q(K^3 + K^2M + KMN))$, where I_{iter} is the number of iterations and $I_{iter} = 10 \times N$ according to our simulations, which means that the number of iterations to reach convergence has close relation to the number of the IRS elements. From Fig. 2(b), we can observe that when the number of the IRS elements is large, the computational complexity of the proposed algorithm is much lower than that of the successive refinement algorithm. The main reason is that the successive refinement algorithm needs to optimize all the elements of IRS one by one. In addition, the computational complexity of our algorithm is affected by Q , but the impact is slight compared with the successive refinement algorithm [9].

Fig. 2(c) shows the transmit power versus SINR for different algorithms. The proposed algorithm is compared with the following schemes: 1) the successive refinement algorithm [9]; 2) the optimal exhaustive search algorithm. In addition, we compared the performance between the second-order cone program (SOCP) and the ZF method for solving the active beamforming. Due to the high complexity of the optimal exhaustive search algorithm, we set $M = 4$, $K = 2$, $N = 8$, $S = 10$, $S_{elite} = 2$. One can observe that the proposed algorithm can obtain close to optimal solution with very low computational complexity.

Fig. 2(d) plots the transmit power versus SINR with $Q = 1$, $Q = 2$ for our proposed algorithm and the successive refinement algorithm [9]. We set $S = 200$, $S_{elite} = 40$, $M = 64$, $N = 625$, $K = 4$, and $d_{RU} = 10m$. As shown in Fig. 1(d), the transmit power is higher under our proposed algorithm than that under the successive refinement algorithm. However, from Fig. 2(b), when $N = 625$, we can see that the computational complexity of the successive refinement algorithm is much higher than

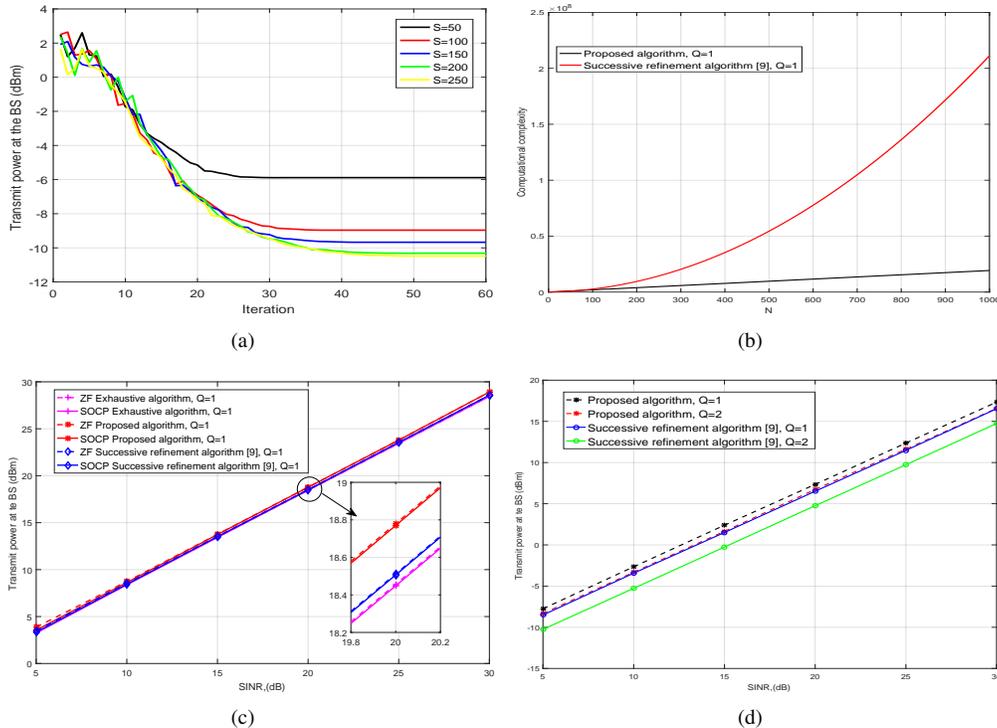


Fig. 2: (a) Transmit power versus iterations I . (b) Computational complexity comparison versus the number of IRS elements. (c) Transmit power versus SINR, $M = 4$, $K = 2$, $N = 8$. (d) Transmit power versus SINR, $M = 64$, $K = 4$, $N = 625$.

the proposed algorithm. This demonstrates that our proposed algorithm strikes a good balance between performance and computational complexity. Moreover, the proposed algorithm can achieve a near optimal performance with lower complexity.

V. CONCLUSION

In this paper, under a downlink mmWave IRS communication system, we proposed a low-complexity machine learning-based CE algorithm and utilized an iterative method to alternately optimize the active beamforming at the BS and the passive beamforming at the IRS, and the target is to minimize the transmit power of the BS. We first derived the probability update expression in the CE algorithm when deploying one-bit phase shift at the IRS and then extended it to high-resolution phase shift setup. The analysis of complexity and simulation results demonstrated that the proposed algorithm can obtain a near optimal performance with a low computational complexity.

REFERENCES

- [1] Q. Wu, S. Zhang, et al., "Intelligent reflecting surface aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313-3351, May. 2021.
- [2] Q. Wu, R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Magazine*, vol. 58, no. 1, pp. 106-112, Jan. 2020.
- [3] M. Xiao, S. Mumtaz, et al., "Millimeter wave communications for future mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909-1935, Sep. 2017.
- [4] J. Ye, S. Guo, et al., "Joint reflecting and precoding designs for ser minimization in reconfigurable intelligent surfaces assisted mimo systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5561-5574, Aug. 2020.
- [5] M. M. Zhao, Q. Wu, et al., "Intelligent reflecting surface enhanced wireless networks: Two-timescale beamforming optimization," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 1, pp. 2-17, Jan. 2021.
- [6] Y. Cao, T. Lv, "Intelligent reflecting surface aided multi-user millimeter-wave communications for coverage enhancement," *arXiv preprint arXiv:1910.02398*, 2019.
- [7] Q. Wu, R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394-5409, Nov. 2019.
- [8] B. Zheng, C. You, et al., "Double-IRS assisted multi-user MIMO: cooperative passive beamforming design," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4513-4526, July. 2021.
- [9] Q. Wu, R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838-1851, Mar. 2020.
- [10] J. Qiao, M. S. Alouini, "Secure transmission for intelligent reflecting surface-assisted mmWave and Terahertz systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1743-1747, Oct. 2020.
- [11] X. Tan, Z. Sun, et al., "Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays," in *Proc. IEEE INFOCOM 2018-IEEE Conf. Computer Commun.*, pp. 270-278, Apr. 2018.
- [12] G. Chaslot, M. Winands, et al., "Cross-entropy for monte-carlo tree search," *ICGA Journal*, vol. 31, no. 3, pp. 145-156, Sep. 2008.
- [13] P. Boer, D. P. Kroese, et al., "A tutorial on the cross-entropy method," *Annals of Operations Research*, vol. 134, pp. 19-67, Feb. 2005.
- [14] X. Gao, L. Dai, et al., "Machine learning inspired energy-efficient hybrid precoding for mmWave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 1-6, May. 2017.
- [15] X. Wei, D. Shen, et al., "Channel estimation for RIS assisted wireless communications: part II—an improved solution based on double-structured sparsity," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1403-1407, May. 2021.