Channel Estimation via Direct Calculation and Deep Learning for RIS-Aided mmWave Systems

Fangqing Jiang, Liang Yang, Daniel Benevides da Costa, and Qingqing Wu

Abstract—This paper proposes a novel reconfigurable intelligent surface (RIS) architecture which enables channel estimation of RIS-assisted millimeter wave (mmWave) systems. More specifically, two channel estimation methods, namely, direct calculation (DC) and deep learning (DL) methods, are proposed to skillfully convert the overall channel estimation into two tasks: the channel estimation and the angle parameter estimation of a small number of active elements. In particular, the direct calculation method calculates the angle parameters directly through the channel estimates of adjacent active elements and, based on it, the DL method reduces the angle offset rate and further improves the accuracy of angle parameter estimation. Compared with the traditional methods, the proposed schemes reduce the complexity of the RIS channel estimation while outperforming the beam training method in terms of minimum square error, achievable rate, and outage probability.

Index Terms—RIS, mmWave, channel estimation, direct calculation, deep learning.

I. INTRODUCTION

As one of the key technologies envisaged to future wireless communication systems, reconfigurable intelligent surface (RIS), also known as intelligent reflecting surface (IRS)) has recently received a huge attention from both the academia and industry [1-3]. The main feature behind RIS is to adjust the electromagnetic response of its impinging signals in order to adapt to applications of interest. In particular, the adjustment of phase shifts is usually related to the channel state information (CSI) knowledge and how to acquire full CSI is still a critical challenge in an RIS-based communication system as it involves the estimation of a cascaded channel [4-5].

To address the channel estimation in RIS-aided systems, the current literature is mainly divided into four kinds of methods. More specifically, the first method uses beam training to avoid channel estimation. In the training phase, the RIS traverses the codewords in a given codebook, and the user feeds back the best codewords to the base station [6-7]. One drawback of this method is that it consumes considerably resources in the training phase, which renders it unfeasible to be implemented for large codebooks. The second method is to model the two parts of the cascaded channel as one channel [8-10] and then to employ traditional channel estimation methods, such as least square (LS) and minimum mean square error (MMSE). The third method is to assume that part of the channel information is known and computable [12], and then to combine codebook and compressive sensing technologies to estimate the channel based on sparse characteristics [11-12]. However, these two methods cannot fundamentally solve the problem of estimating all channels. The last approach is to use the alternative minimization method to solve the cost function. For instance, assuming a cost function with two variables, the first one is solved alternately and the other one is fixed [13-15]. One disadvantage of this method is that it does not necessarily attain the optimal solution, but requires demanding experimental consumption and calculation costs.

Motivated by the deficiencies inherent to the existing channel estimation approaches, herein it is proposed a novel RIS framework which enables channel estimation of RIS-assisted millimeter wave (mmWave) systems. Specifically, two channel estimation methods, namely, direct calculation (DC) and deep learning (DL) methods, are proposed to achieve the separate estimation of the two parts of the cascaded channel. In the proposed strategies, the channel estimation is cleverly converted into the channel estimation and angle estimation of a small number of active elements without the need of codebook design and additional pilot overhead. Thus, compared with the traditional methods, the proposed schemes reduce the complexity of the RIS channel estimation while achieving superior performance, for instance, outperforms the beam training method in terms of minimum square error, achievable rate, and outage probability.

II. SYSTEM/CHANNEL MODELS AND PROBLEM DESCRIPTION

A. System and Channel Models

Consider an RIS-assisted mmWave communication system, which consists of a transmitter (T), a receiver (R), and an RIS, as shown in Fig. 1. We assume that the RIS is composed of $M$ reflecting elements, while both T and R are single-antenna devices. Let $h_T$ and $h_R$ represent the cascade channel between T and RIS, and between RIS and R, respectively. There is no direct link between T and R due to obstacles. Moreover, we assume that the channels satisfy the uplink and downlink reciprocity. In an RIS-based mmWave system with orthogonal frequency-division multiplexing (OFDM), the transmitter T sends a pilot signal and the received signal at the receiver R can be expressed as

$$y_k = h_{R,k}^T \Psi h_{T,k} s_k + n_k,$$  \hspace{1cm} (1)

where $k$ denotes the $k$-th subcarrier, $s_k$ stands for the signal sent by the transmitter on the $k$-th subcarrier, $n_k \sim N_c(0, \delta_n^2)$.
represents the additive white Gaussian noise (AWGN), \( \Psi \) is the reflecting coefficient matrix of the RIS, which is an \( M \times M \) diagonal matrix, namely, \( \Psi = \text{diag}(\psi) \). Each reflecting element \( m \) \( (m = 1, \ldots, M) \) multiplies the incident signal by the factor \( \psi_m = \exp(j\psi_m) \) with the aim to adjust the phase shift of the signal to ultimately optimize certain performance metrics, such as the achievable rate and outage probability.

In this work, we adopt a wideband geometric channel model to characterize the respective channels [15]. In particular, the antenna array of the RIS is assumed to be a uniform planar array (UPA) and the number of reflecting elements is set as \( M = X \times Y \), where \( X \) and \( Y \) represent the number of rows and columns, respectively. Then, the channels between \( T \) (or \( R \)) and the RIS can be defined as

\[
\begin{align*}
    h &= \sqrt{M} \frac{L}{\rho_T} \sum_{i=1}^{L} \alpha_i p(\tau - \tau_i) a(\theta_i, \phi_i), \\
    a(\theta_i, \phi_i) &= \frac{1}{\sqrt{M}} [1, \ldots, e^{-j\frac{2\pi}{N}d(x \sin(\phi) \sin(\theta) + y \cos(\theta))}],
\end{align*}
\]

where \( h \) means the time domain channel response, \( \alpha_i \) denotes the path gain, \( a(\theta_i, \phi_i) \) symbolizes the array response of the RIS. \( L \) represents the number of resolvable paths, which is usually around 3-5 in mmWave systems [16], \( \theta_i \) and \( \phi_i \) are the azimuth and elevation angles of arrival, respectively, \( d \) is the inter-element spacing, \( \lambda \) is the wavelength of the carrier, and \( (x, y) \) stands for the coordinate of the reflecting elements. Let \( \rho_T \) denote the path loss between \( T \) and the RIS, and \( \rho \) represents the pulse shaping filter function with the sampling space \( T_s \). A block-fading channel model is considered, which means that the channel parameters are assumed to remain unchanged over the channel coherence time.

\section*{B. Problem Description}

The appropriate reflection coefficient configuration, power allocation, and signal detection in RIS-based systems rely on the CSI of the respective channels. However, the traditional methods assume that all reflecting elements of the RIS are passive, which makes it very challenging to estimate \( h_T \) and \( h_R \) separately. On the other hand, if all reflecting elements of the RIS are active, the deployment cost and complexity will be greatly increased which may deviate from its original goal to facilitate practical implementation.

Motivated by these constraints, a novel RIS-based system architecture will be investigated in this paper, as shown in Fig. 1. In such a configuration, a small number of adjacent reflecting elements are set to be activated, and each one can receive and process the incident signal. Since the number of active elements is very small, it does not significantly increase the deployment cost and complexity. On the contrary, it increases the degree of freedom of the system and makes easier the estimation of the two parts of the cascade channel.

In Fig. 1, blue areas represent the active elements that are close to each other. Theoretically, the overall channel estimation can be obtained based on the channel estimation of the active elements and detailed analysis will be provided next.

\section*{III. DIRECT CALCULATION METHOD}

As discussed previously, the channel can be modeled as the product of the path gain and the array response. Let \( h_{x,y} \) denote the \( l \)-th channel response of the reflecting element at coordinate \((x, y)\), and \( d = 0.5\lambda, u = \cos(\theta), v = \sin(\phi) \sin(\theta) \). Then, the relationship between \( h_{x1,y2} \) and \( h_{x2,y2} \) can be defined as

\[
    h_{x2,y2} = h_{x1,y1} e^{j\pi(x_2 - x_1) \sin(\phi) \sin(\theta) + j\pi(y_2 - y_1) \cos(\theta)}
    = h_{x1,y1} e^{j\pi(x_2 - x_1)u + j\pi(y_2 - y_1)v}.
\]

Note that the overall channel response of the RIS can be calculated by the channel response \( h_{x,y} \) of any reflecting element and angle parameter \( \{u, v\} \). The channel response of the active element can be estimated by using traditional methods, such as LS, and \( \{u, v\} \) can be calculated by the difference of the channel responses of the adjacent active elements. Let \( h_{x,y}, h_{x,y + 1}, \) and \( h_{x + 1,y} \) denote the channel responses of three adjacent active elements, respectively. Then, their relationships can be expressed as

\[
    \begin{align*}
        h_{x,y + 1} &= h_{x,y} e^{j\pi u}, \\
        h_{x + 1,y} &= h_{x,y} e^{j\pi v}. \\
    \end{align*}
\]

From (4), we have

\[
    u = \arctan\left(\frac{\text{imag}(\Delta h)}{\text{real}(\Delta h)}\right) / \pi, \quad \Delta h = \frac{h_{x,y+1}}{h_{x,y}},
\]

where \( \text{imag}(\cdot) \) and \( \text{real}(\cdot) \) represent the imaginary and real parts of the complex number, respectively. Similarly, \( v \) can be obtained.

Due to noise and interference, it is difficult to obtain the accurate channel estimation of the active elements. Let \( \hat{h}_{x,y}, \hat{u}, \hat{v}, \) and \( \hat{\Delta h} \) denote the estimation of \( h_{x,y}, u, v, \) and \( \Delta h \), respectively. Then, \( \hat{\Delta h} \) can be expressed as

\[
    \hat{\Delta h} = \frac{\hat{h}_{x,y+1}}{\hat{h}_{x,y}} = \frac{h_{x,y+1}}{h_{x,y}} + \varepsilon = \Delta h + \varepsilon,
\]

where \( \varepsilon \) represents the error caused by the estimated channel. Finally, \( \hat{u} \) can be estimated by substituting \( \hat{\Delta h} \) into (5).

The channel estimation process can be summarized as follows. Firstly, channel responses of the active elements are
estimated by LS. Then, the angle parameters are calculated by the above method. Finally, the overall channels can be estimated by the angle parameters and the channel response of one reference active element. This process does not require codebook design and additional pilot overhead.

IV. DEEP LEARNING METHOD

From (5) and (6), it is easily observed that a small error may lead to a significant change of its estimation value when the real or imaginary part of $\Delta h$ has a small value. This results in a high estimation error of $\hat{u}$, which is referred to as angle estimation offset.

The DC method is simple to be implemented, but the angle estimation offset problem severely affects the estimation values of $\{u, v\}$. If the angle estimation offset can be avoided by some methods, the estimation performance can be further improved. However, according to the previous analysis, the fundamental cause of the angle offset is the channel estimation error of the active element, which is inevitable. Therefore, it is difficult to avoid the angle estimation offset problem by using traditional channel estimation methods. Motivated by this, we introduce the DL method to correct the angle estimation offset problem. The DL model is used to explore the potential relationship between the angle estimation offset and the received signals through its powerful fitting ability, so as to correct the angle estimation offset.

A. Model Structure

The model structure is shown in Fig. 2, which consists of four parts: input layer, encoding layer, presentation layer, and output layer. Let $[\hat{u}, \hat{v}]$ denotes the output of the model corresponding to the correction to $\{\hat{u}, \hat{v}\}$. The input consists of two parts: the calculated $\{\hat{u}, \hat{v}\}$ and the received pilots of the active elements, where $\{\hat{u}, \hat{v}\}$ is used as the expert knowledge to accelerate the training convergence and the received pilot is used to correct the angle estimation offset. The encoding layer consists of two parts with different numbers of fully connected (FC) layers, which are used to learn the encoding representations of $\{\hat{u}, \hat{v}\}$ and the received pilot, respectively. The presentation layer consists of an interaction layer and one FC layer, where the interaction layer concatenates the two parts of the output of the encoding layer to establish the relationship between $\{\hat{u}, \hat{v}\}$ and the received pilot. The output layer consists of a FC layer, and finally outputs the prediction results $[\bar{u}, \bar{v}]$.

B. Model Training and Prediction

The proposed deep learning method is operated in two phases, namely, training phase and prediction phase, which are summarized as follows:

Phase 1 (Training Phase): Based on the communication model described in Section II, the simulation data are generated with different noise and channel parameters for training the deep learning model. The training phase consists of three steps: training data generation, data preprocessing, and model training, which are elaborated next.

1) Step 1: It refers to the training data generation, in which the transmitter/receiver sends uplink pilots to the RIS. Then, the active elements of the RIS perform LS-based channel estimation on the received pilots. In Fig. 2, let $\text{rx}\_\text{pilots}$ denote the received pilots, $\{u, v\}$ denote the real value of the angle parameters, and $\{\hat{u}, \hat{v}\}$ denote the angle parameters estimated through the DC method. Finally, the training data is constructed as: input = $[[\hat{u}, \hat{v}], \text{rx}\_\text{pilots}]$, target = $[u, v]$, input is the training sample, and target is the corresponding label. We repeat this process until the required amount of data is generated.

2) Step 2: It refers to the data preprocessing. Thus, in order to facilitate calculation, $\{\hat{u}, \hat{v}\}$ and $\{u, v\}$ are normalized.

3) Step 3: It is the model training. The generated training data is divided into a training set and a validation set. The training and validation sets are used to train and verify the model, respectively. The training objective is to minimize the mean square error (MSE) between the model output and the target. Therefore, the loss function can be defined as

$$loss = \frac{1}{2L} \sum_{l=0}^{L} ((u_l - \hat{u}_l)^2 + (v_l - \hat{v}_l)^2). \quad (7)$$

Phase 2 (Prediction Phase): The trained model is used for prediction. Unlike the training phase, there are no labels and parameters that need to be updated.

V. SIMULATION RESULTS

In this section, we evaluate the performance of both DC and DL methods in terms of the MSE performance of the angle parameter estimation, angle offset rate, achievable rate, and outage probability.

For comparison purposes, we also consider the case that all reflecting elements of the RIS are passive. For this scenario, the beam training (BT) method is used to estimate the angle parameters. Instead of explicit estimation, the angle parameters can be obtained through an over the air beam training process. More specifically, this process searches the optimal codeword by testing the RIS codebook that maximizes the received signal.
power. The optimal codeword corresponds to the angle parameter estimation. This exhaustive training process, however, incurs very large training overhead for the RIS systems [17].

The system and model parameters used in the simulation results are shown in Table 1 and Table 2, respectively.

### A. MSE Performance

In Fig. 3, we evaluate the MSE performances of the two proposed methods to estimate the angle parameter \(\{u, v\}\). Our results show that both proposed methods are effective for estimation. When the signal-to-noise ratio (SNR)>15 dB, the DC method performs better than the BT method with far fewer training overhead. However, due to the limitation of codebook, the BT method cannot achieve satisfactory MSE performance even at high SNR. The DL based method outperforms the other methods significantly when SNR>10 dB, which is attributed to the correction of the angle offset. Therefore, compared with the traditional BT method, the proposed schemes achieve significant performance improvement on the MSE with far fewer training overhead.

### B. Angle Offset Rate

In order to analyze the angle offset problem, we use the angle offset rate to measure it, which can be defined as

\[
r = \frac{\sum_{n=1}^{N} \left| u - \tilde{u} \right|}{N},
\]

where \(r\) represents the angle offset rate, \(n\) stands for the number of tests, and \(\epsilon\) means the threshold value. Assuming \(\epsilon = 0.5\) and \(n = 1024\), Fig. 4 plots the angle offset rate performance for the two methods. Our results show that the DC method has a higher angle offset rate than the DL one. In particular, at SNR=20 dB, the angle offset rate of the DL method is reduced to the \(10^{-3}\) level, which is a large improvement compared to the DC method. This proves that the angle estimation error can be reduced by the DL method, so as to improve the performance of channel estimation.

### C. Achievable Rate and Outage Probability

Now, our objective is to design the RIS reflection factor to maximize the achievable rate and minimize the outage probability. For simplicity, we assume that the channel is dominated by the line-of-sight (LoS) path, that is, \(|\alpha_0| \gg |\alpha_l|, l = 1, \ldots, L\). Therefore, only the LoS path is considered when calculating the optimal reflection factor in the simulation results. In particular, the achievable rate can be defined as

\[
R = E\left( \log\left(1 + \frac{|h_T^\Psi h_R|^2}{\sigma_n^2}\right) \right)
\]

where \(E[\cdot]\) denotes expectation. Accordingly, the outage probability can be mathematically defined as 

\[
P_{out} = \Pr(R < R_{min}).
\]

Figs. 5 and 6 plot the achievable rate and outage probability, respectively, for different estimation conditions. Specifically, “perfect CSI” means that the achievable rate obtained when the reflection factor is designed based on the real CSI, while “without CSI” means that the reflection factor is set to a

---

**TABLE I**

| Parameter       | Value                                                                 |
|-----------------|----------------------------------------------------------------------|
| RIS             | \(M=32\times32, 2\times2\) active elements                          |
| OFDM            | 64 FFT points, \(1/4\) Cyclic prefix, QAM                           |
| Channel         | Wideband geometric channel, \(L = 5, d = 0.5\lambda, \alpha_l \sim N(0, 1), \theta_l, \emptyset_l \in [-\pi/2, \pi/2]\) |

**TABLE II**

| Parameter       | Value                                                                 |
|-----------------|----------------------------------------------------------------------|
| Encoding layer  | For \([u, v]\): 1 FC layer with 128 units                           |
|                 | For rx pilots: 3 FC layers with 128 units                           |
| Presentation layer | 2 FC layers with 128 units                                   |
| Output layer    | 1 FC layer with 10 units                                           |
| Training set size | 100K                                                   |
| Activation functions | Output layer: linear, others: tanh                                   |
constant. Fig. 5 shows that the two proposed methods can achieve better achievable rate performance than the BT method when SNR > 10 dB. In this case, as expected, the DL method achieves the best achievable rate performance. Specially, when SNR > 20 dB, its performance is close to the system with perfect CSI. Fig. 6 shows that the two proposed methods can achieve lower outage probability than the BT method. Thus, it can be concluded that the proposed methods can achieve better communication rate and reliability than the traditional BT method.

VI. CONCLUSIONS

In this paper, we studied the channel estimation problem of the RIS-aided mmWave systems. We proposed a new RIS architecture along with two kinds of channel estimation methods. Based on the innovative design of the RIS architecture, the proposed schemes were shown to reduce the complexity of the RIS channel estimation by realizing the separation into two parts of the cascaded channel estimation. While the DC method has moderate performance but is easy and simple, the DL method can achieve superior performance with relatively high complexity. The simulation results confirmed the advantages of the proposed schemes compared to traditional schemes on MSE, communication rate and reliability. Furthermore, the proposed schemes did not require any codebook design and beam training process. All these demonstrated the effectiveness and superiority of the proposed schemes. For future works, it is interesting to study more efficient deep learning models for other various system setups.

REFERENCES

[1] L. Yang, Y. Yang, M. O. Hasna and M. Alouini, “Coverage, probability of SNR gain, and DOR analysis of RIS-aided communication systems,” IEEE Wireless Commun. Lett., to be published, doi: 10.1109/LWC.2020.2987798.
[2] L. Yang, Y. Jinxia, W. Xie, M. Hasna, T. Tsiftsis and M. Di Renzo, “Secrecy performance analysis of RIS-aided wireless communication systems,” IEEE Trans. Veh. Technol., to be published, doi: 10.1109/TVT.2020.3007521.
[3] Q. Wu and R. Zhang, “Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming,” IEEE Trans. Wireless Commun., 2019.
[4] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” IEEE Trans. Wirel. Commun., vol. 18, no. 8, pp. 4157-4170, Aug. 2019.
[5] Q. Wu and R. Zhang, “Towards smart and reconfigurable environment: Intelligent reflecting surface assisted wireless network,” IEEE Commun. Mag., vol. 58, no. 1, pp. 106-112, Jan. 2020.

Fig. 5. Achievable rate performance comparison.

Fig. 6. Outage probability performance comparison.