Channel Estimation Based Deep Learning Using IRS-Assisted MISO Systems with Correlated Channel

Zainab Ali Alsalman * and Hayder Almosa
Department of Electronic and Communications Engineering, Faculty of Engineering, University of Kufa, Najaf, Iraq
Email: zainaba.azawi@student.uokufa.edu.iq (Z.A.A.); zaynabshaney@gmail.com (Z.A.A.);hayder.almusa@uokufa.edu.iq (H.A.)
*Corresponding author

***Abstract*—Intelligent reflecting surface (IRS) is a developing technology that can significantly enhance the efficiency of wireless communications. It achieves this by smartly adjusting the signal reflections at several passive reflecting elements. The channel estimation is a crucial problem in implementing a viable IRS-assisted communication system. Deep learning (DL) has attracted considerable attention for tackling physical layer design issues because of its capacity to acquire intricate patterns from data with less computing complexity and enhanced resilience. This paper proposes a channel estimation model with a correlated channel based on DL that employs the Minimum Mean-Squared Error (MMSE) criterion. Specifically, we design and train the Convolutional Neural Network (CNN) architecture using received signals to simultaneously estimate both the direct channel between the transmitter and receiver, as well as the cascaded channel that incorporates the IRS's reflection. Furthermore, the numerical results demonstrate that incorporating IRS and the proposed DL-based channel estimation technique leads to substantial performance gains over conventional channel estimation methods. Specifically, at low SNRs (-5 dB), the DL-based approach exhibits NMSE values of approximately 0.7659, whereas at higher SNRs (25 dB), NMSE values decrease to around 0.0022. These findings underscore the efficacy of the proposed solution in mitigating the adverse effects of channel impairments.**

***Keywords*—deep learning, intelligent reflecting surface (IRS), channel estimation, CNN, MISO system**

I. Introduction

In recent years, there has been a significant increase in the popularity of utilizing reconfigurable metasurfaces for wireless communication systems. Intelligent Reflecting Surfaces (IRS) offer adjustable degrees of freedom to modify the propagation characteristics of problematic channels. This makes them a valuable asset for preserving and improving network users' Quality of Service (QoS) [1–3]. An IRS typically comprises many passive reflecting elements, each capable of reconfiguration and autonomous control to modify its phase shift based on the present environment. This enables the alteration of the reflection of incoming signals. It is possible to get a desired reflection pattern by simultaneously altering the phase shifts of all the passive components. This pattern establishes an advantageous wireless channel that improves the quality of transmission while decreasing the power consumption of the system [4,5].

However, most techniques that take advantage of this ability require Channel State Information (CSI) to and from the IRS elements. This is difficult because the number of IRS elements can be significant. They are usually constructed as passive devices without active transceivers or computational resources. Since the IRS is not active, the estimation of the CSI needs to be done by devices, often a Base Station (BS) or access point, that is not located in the same place as the IRS. For instance, the BS receives training signals from the user equipment after they are reflected by the IRS, possibly through a direct path to the BS. These signals are utilized to estimate CSI. The performance of the system architecture in both conventional and IRS-assisted large MIMO scenarios is highly dependent on the precision of the instantaneous CSI. Therefore, the accuracy of channel estimate is crucial in developing analog and digital beamformers in traditional massive MIMO [6,7], as well as in the design of phase shifts for reflecting beamformers in RIS-assisted scenarios.

II. Literature Review

A significant amount of published research on CSI estimates for IRS-based systems has emerged recently. The pilot overhead can become high due to the number of cascaded channel coefficients among the BS, the IRS, and the user, which is the product of the number of BS antennas, IRS elements, and users. Hu, Chen, *et al.* [8] suggested a two-timescale channel estimation. More significantly, a dual-link pilot transmission method is used in which the BS transmits downlink pilots and receives uplink pilots reflected by the IRS. Therefore, a method based on coordinate descent is suggested for BS-IRS channel recovery. Jensen, Lindstrøm, *et al.* [9], presented a Minimum Variance Unbiased estimator (MMU) estimator to lower the variance. Despite the high pilot overhead, this technique achieved good estimation accuracy. A method for channel estimation was proposed

Manuscript received August 11, 2024; revised September 30, 2024; accepted October 15, 2024; published December 23, 2024.

by de Araújo, Gilderlan T, *et al.* [10] that uses the Parallel Factor (PARAFAC) of received signal tensor modeling. This approach handled the less-than-ideal scenario, in which the receiver is unaware of the IRS phase shift, and phase disturbance is present.

To tackle the problems with channel estimation or beamforming, data-driven DL methods have recently been proposed [11]. To improve the efficiency of model-based compressive channel estimation for mm Wave IRS systems in Orthogonal Frequency Division Multiplexing (OFDM) systems. Due to the high-dimensional cascaded MIMO channels and passive reflecting components. Liu, Slicing, *et al.* [12] present a deep denoising neural network to assist in compressive channel estimation for mm Wave IRS systems, hence minimizing training overhead. The result demonstrates enhanced representational capabilities beyond real-valued ones, leading to improved performance. In addition, Elbir, Ahmet M. *et al.* [13] proposed a twin Convolutional Neural Network (CNN) architecture that is developed and trained using the received pilot signals to estimate the direct and cascaded channels accurately. Within a multiuser setting, every individual user can utilize CNN to calculate and determine their channel. The performance of the proposed DL strategy is assessed and compared to state-of-the-art DL-based techniques, showcasing its exceptional performance. Moreover, Elbir, Ahmet M., *et al.* [14] use a Federated Learning (FL) architecture for channel estimation to minimize the communication overhead associated with data collection. A CNN was developed and trained using users' local datasets without sending the datasets to a central server BS.

Furthermore, the proposed architecture exhibits lower estimation errors than state-of-the-art Machine Learning (ML) based schemes. In other research, Shtaiwi, Eyad, *et al.* [15] employed a CNN-based methodology for mm-wave channel estimation to decrease the number of active users during the training phase. Besides channel estimation, CNN can be used in performance analysis, including Bit Error Rate (BER) or Symbol Error Rate (SER) for IRS-based communication. The difference between these studies and our study is in the system model. In our paper, we derive the channel model of our system and propose an efficient channel estimation scheme to estimate both the BS-user direct channel and BS-IRS-user cascade channel simultaneously. Still, all of these studies estimate cascade and direct channels of IRS separately or assume the direct channel is a blockage and do not consider it. Therefore, ignoring the direct link in IRS systems is typically ineffective. It can result in the underestimation of the power received and the disregard for the advantages of variety. We focus on combining two main IRS channels in a single vector and then analyze channel estimators.

This paper proposes a DL approach for channel estimation in IRS-assisted-MISO systems. In the proposed DL framework, the CNN is specifically developed to estimate the direct and cascade channels directly in the presence of a correlated Rayleigh fading channel, considering the received observation signals as input. The performance and efficiency of DL-based channel estimation are assessed and contrasted with classic channel estimation techniques, namely the Least Squares (LS) and Minimum Mean Square Error (MMSE) estimators. The DL estimator, designed by utilizing Deep Neural Networks (DNNs) with Rectified Linear Unit (ReLU) activation, efficiently reached a performance similar to that of the MMSE estimator when a significant amount of training data is accessible. The DL model is trained on a large number of channel realizations to attain reliable performance of the estimator. The trained model is assessed using test data generated separately during the prediction stage. The subsequent sections of this paper are organized in the following ways. Section III illustrates the system model of the proposed IRS system, and the channel estimation scheme, which uses two different methods, is introduced. Section IV presents the proposed channel estimation scheme via deep learning and the CNN architecture model for the system. The simulation findings are presented in Section V, while the conclusions are supplied in Section IV.

Notation: Lowercase and uppercase "**a**" and "**A**" represent a vector and a matrix by boldface letters, respectively. The symbols A^H and A^T represent the conjugate and transpose of matrix A. The expression diag(x) represents a diagonal matrix in which the vector x is positioned along its diagonal. The Kronecker product of a and b denotes a $\otimes$ b. Vec (A) is vectorizing the matrix A. Finally, CN $\sim(\mu,\sigma^2)$ represents the complex Gaussian distribution, where μ is the mean and σ^2 is the variance.

III. System Model

Consider an IRS-assisted MISO communication system, as depicted in Fig.1. The BS is presumed to have M antennas to serve K individual single antenna users. The beam steering is facilitated by an IRS consisting of N passive reflecting components. Within the IRS-assisted communication scheme, every element of the IRS causes the arriving signal from the BS to a phase shift. Changing the PIN diodes allows one to change the phase of every individual IRS element. The IRS controller controls the diodes, and the backhaul link is connected to the BS [16, 17]. One may characterize the received signal $\boldsymbol{y} \in \mathbb{C}^{M\times 1}$ at time t, t= 1… T_P as:

$$\boldsymbol{y}_{\mathrm{t}} = \sqrt{p}(\boldsymbol{h_d} + \boldsymbol{G}\emptyset_t \boldsymbol{h_r})\mathrm{x}_{\mathrm{t}} + \boldsymbol{n}_t. \quad (1)$$

where p denotes the transmit power, $\boldsymbol{G} \in \mathbb{C}^{M\times N}$ represents the channel matrix between the BS and IRS, $\boldsymbol{h_r} \in \mathbb{C}^{N\times 1}$ expresses the channel between the IRS and the user, and the channel connection between the BS and the user is denoted as $\boldsymbol{h}_d \in \mathbb{C}^{M\times 1}$. The signal-to-noise ratio (SNR) is represented as p/σ^2. x_t and $\boldsymbol{n}_t$ denote the transmitted pilot signal of the user and Additive White Gaussian Noise (AWGN) at the receiver with CN~(μ,σ^2) respectively. The correlation model for spatial correlation is used at the BS and IRS elements expressed as:

$$\boldsymbol{h_r} \in \mathbb{C}N(0, \boldsymbol{R}_N) \quad (2)$$

$$\boldsymbol{G} \in \mathbb{C}N(0, \boldsymbol{R}_{MN}) \quad (3)$$

$$\boldsymbol{h_d} \in \mathcal{C}N(0, \boldsymbol{R_M}) \quad (4)$$

where $\boldsymbol{R}_N, \boldsymbol{R}_{MN}$, and $\boldsymbol{R}_M$ in Eqs.(2–4) represent the channel correlation matrix between IRS-user, IRS-BS, and BS-user, respectively. The calculation of $\boldsymbol{R}_N, \boldsymbol{R}_{MN}$, and $\boldsymbol{R}_M$ are obtained in Section V.

The diagonal matrix $\boldsymbol{\emptyset_t} \in \mathbb{C}^{N\times N}$ represents the properties of the IRS. The amplitude reflection coefficient α belongs to the interval [0,1] and represents the on/off state of the system. The phase shift variables [$\emptyset_1, \ldots, \emptyset_N$] where $\boldsymbol{\emptyset_t} \in [0,2\pi)$, are optimized to enhance the performance of the IRS. The cascade channel **V** can be expressed mathematically as:

$$\boldsymbol{\emptyset_t} = \alpha \text{ diag } (e^{j\emptyset 1}, e^{j\emptyset 2}, \ldots, e^{j\emptyset \mathrm{N}}). \quad (5)$$

It is expedient to express Eq. (1) equivalently as:

$$\boldsymbol{y}_{\mathrm{t}} = \sqrt{\mathrm{p}}(\boldsymbol{h_d} + \boldsymbol{V}\boldsymbol{\emptyset_t})\mathrm{x_t} + \boldsymbol{n}_t. \quad (6)$$

This scenario utilizes the channels in a correlated Rayleigh fading, in which the channel coefficients remain constant inside a specific time block known as a coherence time $\tau > \mathrm{T}$. Where T is the time of the training period but undergoes independent fluctuations between blocks. We utilize Time-Division Duplexing (TDD) for both the transmission of data from the user to the network (uplink) and the transfer of data from the network to the user (downlink) depending on channel reciprocity to acquire the channel state information at the IRS in both directions of the transmission.

In communication systems assisted by IRS, channel estimation aims to determine the cascaded channel **V**, denoted as M (N + 1), which can be represented by the Khatri-Rao product.

$$\text{vec}(\boldsymbol{h}_d + \boldsymbol{V}\boldsymbol{\emptyset_t}) = \boldsymbol{h_d} + \text{vec } (\boldsymbol{V}\ \boldsymbol{\emptyset_t}) \quad (7)$$

With $\boldsymbol{h}_d$ = vec ($\mathbf{h}_d$). By using the properties of the vectorization product in Eq (7) and applying vec (ABC) = ($\mathrm{C^T} \otimes$ A) vec(B) , we obtain:

$$vec(\boldsymbol{y}_{\,t}) = (x^T \otimes \boldsymbol{I}_M)(\boldsymbol{h}_d + (\boldsymbol{\emptyset_t} \otimes \boldsymbol{I}_M\ vec\ (\boldsymbol{V})) + \boldsymbol{n}_t \quad (8)$$

$$= (x^T \otimes \boldsymbol{I}_M)\ [\boldsymbol{I}_M\ \boldsymbol{\emptyset_t} \otimes \boldsymbol{I}_M\ ] \begin{bmatrix} \boldsymbol{h}_d \\ vec(\boldsymbol{V}) \end{bmatrix} + \boldsymbol{n}_t \quad (9)$$

$$= (x^T \otimes \boldsymbol{I}_M)\ [\ \boldsymbol{I}_M\ \boldsymbol{\emptyset_t} \otimes \boldsymbol{I}_M]\ \boldsymbol{h} + \boldsymbol{n}_t \quad (10)$$

$$\boldsymbol{y} = \boldsymbol{Z}\ \boldsymbol{\Psi}\ \boldsymbol{h} + \boldsymbol{n} \quad (11)$$

Assume **X**=**Z*Ψ***, Where **Z** =diag ([$\mathbf{Z}_1, \ldots, \boldsymbol{Z}_M$]) using a known pilot matrix **X**. The received signal is generally written as shown in Eq. (11).

$$\mathbf{y} = \sqrt{\mathrm{p}}\,\boldsymbol{X}\,\mathbf{h} + \mathbf{n} \quad (12)$$

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{y}_T \end{bmatrix} \boldsymbol{n} = \begin{bmatrix} \boldsymbol{n}_1 \\ \boldsymbol{n}_2 \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{n}_T \end{bmatrix}, \boldsymbol{h} = \begin{bmatrix} \boldsymbol{h}_d \\ \boldsymbol{v}_1 \\ \cdot \\ \cdot \\ \cdot \\ \boldsymbol{v}_{MN} \end{bmatrix} \quad (13)$$

The IRS channel estimation problem involves the estimation of the vector **h**, which comprises the direct and cascading channels from Eq. (13), which include the observation vector, pilot matrix, channel vector, and noise vector in that order. Two methods are commonly used to estimate **h**, both based on Eq. (12).

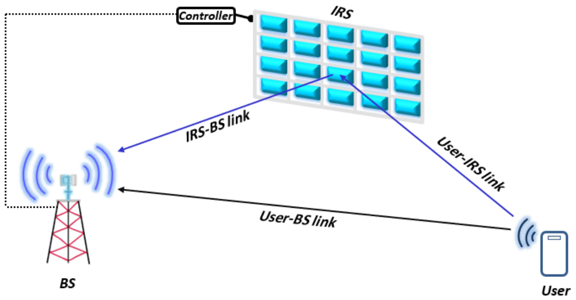

Fig. 1. System model of IRS-assisted wireless communication.

A. LS Estimation Approach

The LS estimator is the simplest way to estimate **h**. Also, this approach can connect the gap between the present coding theory and mathematics statistics. [18] may be expressed as:

$$\hat{\mathbf{h}} = \|\ \mathbf{y} - \hat{\mathbf{y}}\ \|^2 \quad (14)$$

$$\hat{\mathbf{h}} = \|\ \mathbf{y} - \sqrt{\mathbf{p}}\ \mathbf{X}\ \mathbf{h}\ \|^2. \quad (15)$$

The solution is an optimization problem to optimize the CSI. LS estimator of single k user and M transmit antenna with MISO system is given by:

$$\hat{\mathbf{h}}_{\mathrm{LS}} = \frac{1}{\sqrt{\mathrm{p}}}\,(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\ \mathbf{y}. \quad (16)$$

Proof. See Appendix A.

Eq. (16) represents the LS estimators applied to our system. The random variable **n** is distributed according to a multivariate normal distribution with a mean of zero and a covariance matrix, denoted as CN ~ (0,$\boldsymbol{\sigma}^2\boldsymbol{I}_M$) assuming that the noise is uncorrelated.

B. MMSE Estimation Approach

The MMSE estimator is predicated on a stochastic model. Conversely, the LS technique presupposes a channel that functions consistently and does not depend on prior information. Estimating the values $\mathbf{h}_r$, **G** and $\mathbf{h}_d$ is often done by utilizing coupled Rayleigh fading and prior knowledge of second-order statistics. The optimal value of **h** that minimizes the Mean Squared Error (MSE) can be obtained using the following method.

$$\hat{\mathbf{h}} = \min \|\ \mathbf{y} - \mathbf{X}\mathbf{h}\|^2. \quad (17)$$

By minimizing the received signal in Eq. (17), the MMSE estimator of single k user and M transmit antenna with MISO system is given by:

$$\hat{\mathbf{h}}_{\mathbf{MMSE}} = \frac{1}{\sqrt{\mathrm{p}}}\,\mathbf{R}_{\mathrm{h}}(\mathbf{R}_{\mathrm{h}} + \frac{\sigma^2}{\mathrm{P}}\,\mathbf{I}_{\mathrm{M}})^{-1}\,\mathbf{X}^{\mathbf{T}}\,\mathbf{Y}. \quad (18)$$

Proof. See Appendix B.

The matrices in Eq. (18) are associated with $\mathbf{R}_h$ are independent of the data and can be computed and saved offline as $\mathbf{R}_h$ varies at a relatively slow rate.

IV. Channel Estimation via Deep Learning

Deep learning involves using deep neural networks (DNNs) and has demonstrated significant efficacy in diverse domains. This includes computer vision and natural language processing. It has been used in numerous communication systems, covering both the physical and network levels. These disciplines involve the distribution of electricity [19] and the assessment of channels [20]. The communication system operates as a black box, employing an end-to-end deep learning structure to transmit and receive information. The effectiveness of the beamformer design depends significantly on the comprehension of channel information [21]. The IRS-assisted systems consist of multiple communication lines, including a direct channel that connects BS directly to users and a cascaded channel that connects BS to users through the IRS.

A. Deep Learning Approach

This section presents an efficient approach for implementing channel estimation using CNN, a deep learning system designed to analyze visual data in various computer vision tasks. It is intended to emulate the function of the human visual brain. CNN is made up of layers that process input data. CNN comprises layers that analyze the incoming data. The convolutional layers utilize filters to extract information from the input data. CNN utilizes parameter sharing and provides spatial invariance, allowing for object recognition irrespective of position or orientation. They acquire hierarchical representations of features, ranging from low-level to high-level. CNN is trained using labeled datasets, modifying weights to enhance performance. They have attained remarkable outcomes in activities such as picture classification, object identification, and image segmentation. CNN are specialized algorithms that facilitate automatic feature extraction and precise identification of visual patterns. When training the algorithms, while highly effective for image processing, is primarily designed to operate on real-valued data. However, communication systems often employ complex-valued signals, introducing challenges when directly applying CNN to such data. To address this, a common approach involves separating the real and imaginary components of the complex-valued signal. Subsequently, these sequences can be reshaped into two-dimensional images, which can then be fed into the CNN as separate inputs. By processing these images independently and combining their outputs in a subsequent layer, the CNN can effectively capture the complex-valued nature of the original signal. In the following sections, we will present the architectural design and channel estimation process using CNN.

B. Network Architecture

The architectural design of the CNN network is shown in Fig. 2. This paper employed a CNN model architecture to assess the NMSE of single-user MISO signals throughout repeated training with the generated dataset. The complex signal was initially divided into its constituent real and imaginary components. Next, the CNN model is provided with both the numerical values and the label. In the proposed model, the input layer is linked to the observation vector, and the input size is determined by the number of features in the input data. The input size is defined by the number of users and the total number of antennas utilized in the system. The connectivity between the layers is established by a series of five convolutional layers, an activation function called ReLU, and regression layers. Table I presents the number of filters and the size of the convolutional layers. The convolutional layer consists of a grid of neurons arranged in a rectangular pattern, where each neuron receives inputs from a similar rectangular area in the previous layer.

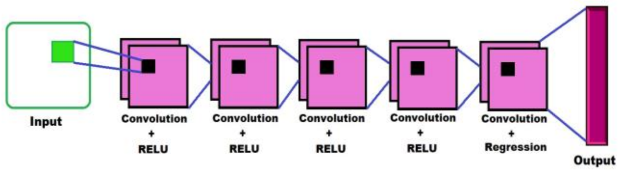

Fig. 2. Architecture of CNN model.

TABLE I. Hyper Parameter of the Proposed CNN

Layer	Size	Operation
1	512(9×9×4)	convolutional+RelU
2	512(9×9×512)	convolutional+RelU
3	128(9×9×512)	convolutional+RelU
4	128(9×9×128)	convolutional+RelU
5	1(5×5×128)	convolutional+Regreesion

In our MISO system-based IRS, creating synthetic data for different SNR levels is the first step in the training process for channel estimation. The direct channel from the access point to the user and the channel matrices from the IRS to the user and from the access point to the IRS are generated randomly for every SNR level. Noise is added to mimic the transmission environment. A CNN uses the actual channel as the label and the MMSE channel estimate as the input. The real and imaginary components of the channel estimates are divided into two image sets, which are then concatenated and fed into the CNN. The network consists of several convolutional layers with ReLU activation, designed to learn the mapping between the estimated and true channel responses, with the final layer outputting the channel estimate.

TABLE II. Simulation Parameters

Parameter	Value
IRS Reflecting Element	32
Training Data Size	10,000
Transmit Antenna	7
Number Of Epochs	5
Learning Rate	0.01
Optimizer	ADAM
Minibatch Size	32

The model based on CNN is trained using both the label and the data. The produced channel data is inputted into the neural network, which utilizes an ADAM optimizer for optimization. 10,000 data sets were used to construct the model, with 0.25% allocated for validation purposes and

0.75% designated for training. Table II indicates the simulation parameters.

V. Simulation Result

This section discusses the simulation results of the proposed method compared to the conventional estimation techniques in terms of NMSE. The NMSE metric can be expressed as $E\{\frac{\|\mathbf{h}-\hat{\mathbf{h}}\|^2}{\|\mathbf{h}\|^2}\}$. Where E denotes the expectation operation and $\hat{\boldsymbol{h}}, \boldsymbol{h}$ are the estimated and true channels, respectively. The simulation assumes the number of transmit antennas at the BS, M = 10, and the number of passive elements in IRS, N = 32, with a single user. IRS phase values are DFT matrix. The transmit signal x_t generated using a mathematical formula based on Zadoff-Chu (ZC) Sequences. The exponential correlation model for spatial correlation is used at the BS and IRS elements and expressed as follows [22].

$$[\mathbf{R}_{\mathrm{IRS-BS}}] = (\rho_1 e^{j\emptyset_k})^{|i-j|} \tag{19}$$

$$[\mathbf{R}_{\mathrm{BS-user}}] = (\rho_2 e^{j\emptyset_k})^{|i-j|} \tag{20}$$

$$[\mathbf{R}_{\mathrm{IRS-user}}] = (\rho_3 e^{j\emptyset_k})^{|i-j|}. \tag{21}$$

where $0< \rho_1, \rho_{2,}\rho_1 <1$ denotes the level of spatial correlation between the BS antenna and IRS elements. We set $\rho_1 = \rho_1 = \rho_1 = 0.5$. The SNR is defined as $\frac{P}{\sigma^2}$. The symbols received at the receiver are decomposed into their real and imaginary components and the associated label using 10,000 data sets for the model. The network is trained with a widely recognized Adam optimizer with a learning rate 0.001 and minibatch = 32. In addition, we set up L=100 Monte-Carlo experiments are conducted to assess the NMSE.

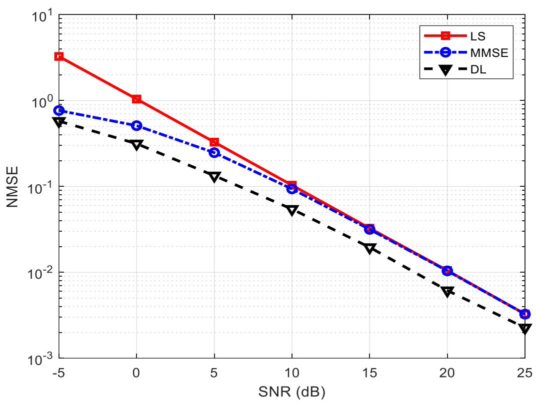

Fig. 3. NMSE against SNR for *M*= 10, *N* = 32, *T*=*N*+1.

Fig. 3 demonstrates the performance comparison of three distinct methods, LS, MMSE, and DL, for channel estimation across different SNRs. The metric employed for comparison is the NMSE. The LS technique exhibits the greatest NMSE among all SNR values. These findings suggest that LS exhibits the lowest level of performance compared to the other two approaches. As SNR grows, the NMSE drops. However, the NMSE is still considerably greater compared to the different approaches, particularly at lower SNRs.

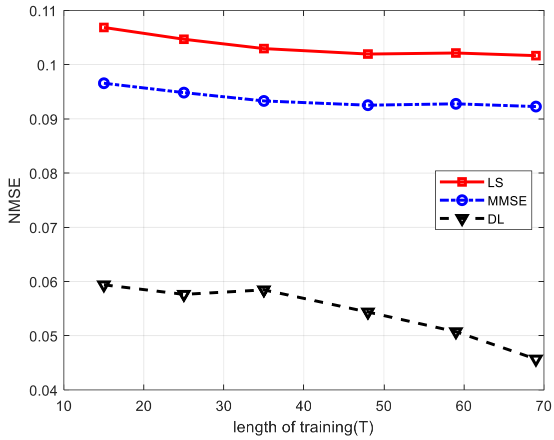

Fig. 4. NMSE against length of training T for M= 10, N=32.

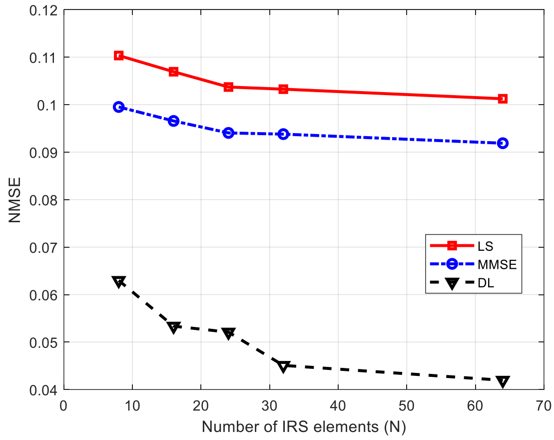

Fig. 5. NMSE against the number of passive elements for *M*= 10, *T*=*N*+1.

When SNR is higher, DL approaches demonstrate superior efficacy in managing noise and delivering precise channel estimates. Additionally, the MMSE method outperforms the LS method. It demonstrates a reduced NMSE for all SNR values. The NMSE for MMSE drops faster than LS when the SNR improves, suggesting a higher level of resilience to noise. The DL approach surpasses LS and MMSE in performance for all SNR values. The NMSE for DL is the lowest, especially when the SNR is higher. DL approaches demonstrate superior efficacy in managing noise and delivering accurate channel estimates.

Fig. 4 shows the NMSE with the number of IRS elements N examined for three different estimation techniques. The total number of IRS elements ranges from 8 to 65. The superiority of the DL approach over both the LS and MMSE methods is apparent for all values of N. The NMSE for the DL approach, represented by a dashed black line, begins at roughly 0.06 for a value of N equal to 10 and gradually decreases to approximately 0.04 as the number of elements IRS reaches 60. The significant reduction in the NMSE highlights the greater ability of the DL technique to utilize the additional information from extra IRS parts.

Fig. 5 depicts the relationship between the NMSE and the length of training (T) for three distinct estimation methodologies. The performance of each technique is monitored throughout different lengths of training data, ranging from 10 to 70.

It is clear from the data that the DL approach consistently performs better than both the LS and MMSE methods, regardless of the training period. The superiority of the DL approach is evident in its lower NMSE values. The NMSE starts at around 0.06 for a training length of 10 and decreases to approximately 0.05 as the training length grows. The decrease in NMSE suggests that the DL approach greatly benefits from additional training data, improving prediction accuracy.

VI. Conclusions

This paper proposed a DL approach to address the channel estimation for IRS single user with correlated Rayleigh fading channels in MISO communication systems. Specifically, a CNN architecture is designed to accurately predict both the direct and cascading channels of the IRS based on the received signal. Our simulation result is the CNN outperforms standard MMSE and LS estimators by achieving almost optimal MMSE performance at different SNR levels. Moreover, the study examines how the duration of pilot training and the size of the IRS affect estimation accuracy, offering valuable insights for designing practical systems.

Appendix A

proof of theorem 1

LS derivative

$$\text{MINIMIZE } \mathrm{L}(\mathrm{Y},\mathrm{H})\ \Delta \text{ MIN } |\mathbf{y} - \mathbf{xh}|^{2}$$

An estimate $\mathbf{h}$ can be formed by minimizing.

$$||\ \mathbf{y} - \mathbf{Xh}\ ||^{2}$$

$$||\ \mathbf{y} - \mathbf{Xh}\ ||^{2} \leftarrow \text{square of norm error } (\mathbf{y} - \mathbf{xh})$$

find h, which has the least square error

$$\mathrm{L}(\mathbf{y},\mathbf{h}) = ||\ \mathbf{y} - \mathbf{Xh}\ ||^{2}$$

$$||\ \mathbf{y} - \mathbf{xh}\ ||^{2} = (\mathbf{y} - \mathbf{Xh})^{\mathbf{T}}\ (\mathbf{y} - \mathbf{Xh})$$

$$(\mathbf{y}^{\mathbf{T}}_{-}\mathbf{X}^{\mathbf{T}}\mathbf{h}^{\mathbf{T}})\ (\mathbf{y} - \mathbf{Xh})$$

$$\mathbf{d}(\mathbf{y},\mathbf{h}) = (\mathbf{yy}^{\mathbf{T}} - \mathbf{y}^{\mathbf{T}}\mathbf{Xh} - \mathbf{X}^{\mathbf{T}}\mathbf{h}^{\mathbf{T}}\mathbf{y} + \mathbf{X}^{\mathbf{T}}\mathbf{h}^{\mathbf{T}}\mathbf{Xh})$$

$$\text{we know } \mathrm{v}^{\mathrm{T}}\mathrm{u} = \mathrm{u}^{\mathrm{T}}\mathrm{v}\ \text{ scalar}$$

$$(\mathrm{v}^{\mathrm{T}}\mathrm{u})^{\mathrm{T}} = \mathrm{u}^{\mathrm{T}}\mathrm{v}\quad \text{dot product}$$

$$||\ \mathbf{y} - \mathbf{Xh}\ ||^{2} = \mathbf{yy}^{\mathbf{T}} - 2\mathbf{X}^{\mathbf{T}}\mathbf{h}^{\mathbf{T}}\mathbf{y} + \mathbf{X}^{\mathbf{T}}\mathbf{h}^{\mathbf{T}}\mathbf{Xh} = \mathrm{f}(\mathbf{h})$$

$$= -2\mathbf{X}^{\mathbf{T}}\mathbf{y} + 2\mathbf{X}^{\mathbf{T}}\mathbf{h}^{\mathbf{T}}\mathbf{Xh} = \mathbf{0}$$

$$\rightarrow \mathbf{X}^{\mathbf{T}}\mathbf{Xh} = \mathbf{X}^{\mathbf{T}}\mathbf{y}$$

$$\hat{\mathbf{h}}_{\mathrm{LS}} = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\ \mathbf{X}^{\mathbf{T}}\mathbf{y}$$

Appendix B

Proof of Theorem 2

MMSE Derivation:

Let's consider the received signal

$$\mathbf{y} = \mathbf{Xh} + \mathbf{n}$$

processed to find MMSE estimate h:

$$\mathbf{E}[\mathbf{n}^{\mathbf{T}}] = \mathbf{0}$$

$$\mathbf{E}[\mathbf{y}] = \mathbf{E}[\mathbf{Xh} + \mathbf{n}]$$

$$\mathbf{E}[\mathbf{h}] + \mathbf{E}[\mathbf{n}]$$

$$\therefore\quad \mathbf{h}^{\wedge} = \mathbf{R}_{\mathbf{hy}}\,\mathbf{R}_{\mathbf{yy}}^{-1} \qquad (\mathbf{1}) \quad :\ \text{MMSE estimate}$$

$$\mathbf{R}_{\mathbf{yy}} = \mathbf{E}[\mathbf{yy}^{\mathbf{T}}] = \mathbf{E}\left[(\mathbf{Xh} + \mathbf{n})(\mathbf{xh} + \mathbf{n})^{\mathbf{T}}\right]$$

$$= \mathbf{E}\left[(\mathbf{Xh} + \mathbf{n})(\mathbf{h}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}} + \mathbf{n}^{\mathbf{T}})^{\mathbf{T}}\right]$$

$$= \mathbf{X}\ \mathrm{E}(\mathbf{hh}^{\mathbf{T}})\mathbf{x}^{\mathbf{T}} + \mathbf{X}\ \mathrm{E}[\mathbf{hn}^{\mathbf{T}}] + \mathrm{E}[\mathbf{hn}^{\mathbf{T}}]\mathbf{X}^{\mathbf{T}} + \mathrm{E}[\mathbf{nn}^{\mathbf{T}}]$$

$$\therefore \mathbf{R}_{\mathbf{yy}} = \boldsymbol{\sigma}_{\mathbf{n}}^{2}(\mathbf{X}\,\mathbf{X}^{\mathbf{T}} + \boldsymbol{\sigma}^{2}\mathbf{I}_{\mathbf{n}}) \quad \mathbf{(2)}$$

By using the Gaussian noise vector.

$$\mathbf{E}[\mathbf{nn}^{\mathbf{T}}] = \boldsymbol{\sigma}^{2}\mathbf{I}$$

$$\mathbf{R}_{\mathbf{hy}} = \mathbf{E}[\mathbf{h}\ \mathbf{y}^{\mathbf{T}}]$$

$$= \mathrm{E}[\mathbf{h}(\mathbf{hX} + \mathbf{n})^{\mathbf{T}}]$$

$$\mathbf{R}_{\mathbf{hy}} = \mathrm{E}\ [\mathbf{h}\ (\mathbf{h}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}} + \mathbf{n}^{\mathbf{T}})] \qquad (3)$$

Now, substituting Eqs. (2) & (3) in Eq. (1)

$$\mathbf{h}_{\mathbf{mmse}}^{\wedge} = \frac{\mathbf{1}}{\sqrt{\mathbf{P}}}\ (\mathbf{X}\,\mathbf{X}^{\mathbf{T}} + \boldsymbol{\sigma}^{2}\mathbf{I}_{\mathbf{n}})^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{y}$$

Conflicts of Interest

There do not appear to be any competing interests pertaining to this study that we are aware of.

Author Contributions

Each author's contribution to this work is acknowledged by us. While the second author supervised the research, the first author handled simulation and paper writing.

References

[1] S. Gong, L. Xiao, H. D. Thai, N. Dusit *et al*., "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Communications. Surveys*, vol. 22, no. 4, pp. 2283-2314, 2020.

[2] Z. Jun, "A survey of intelligent reflecting surfaces (IRSs): Towards 6G wireless communication," arXiv preprint arXiv:1907-04789, 2020.

[3] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 162-169, 2018.

[4] R. M. Di, M. Debbah, D.-Thuy Phan-Huy, Al. Zappone, M. Alouini, Ch., V.Sciancalepore "Smart radio environments empowered by reconfigurable AI meta-surfaces," *EURASIP Journal on Wireless Communications and Networking*, no. 1, pp 1-20, 2018

[5] C. Huang, S. Hu, G. C. Al. A. Zappone, C. Yuen, R. Zhang, M. D. Renzo, and M. Debbah "Holographic MIMO surfaces for 6G

wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Communications*, vol. 27, no. 5, pp. 118-125, 2020.

[6] Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun*., vol. 64, no. 5, pp. 1801–1818, 2016.

[7] F. Sohrabi and W. Yu, "Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1432–1443, 2017.

[8] H. Chen *et al.*, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7736-7747, 2021.

[9] J. T. Lindstrøm and E. De Carvalho, "An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator," in *Proc. ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2020.

[10] G. T. de Ara'ujo and A. L. F. de Almeida, "PARAFAC-based channel estimation for intelligent reflective surface assisted MIMO system," in *Proc. IEEE Sens. Array Multichannel Signal Process. Workshop (SAM)*, pp. 1–5, Hangzhou, China, Jun. 2020

[11] M. Elbir and K. V. Mishra, "A survey of deep learning architectures for intelligent reflecting surfaces," arxiv.org, 2009.02540, 2020

[12] S. Liu, Z. Gao, J. Zhang, M. Di Renzo, and M.-S. Alouini, "Deep denoising neural network assisted compressive channel estimation for mmwave intelligent reflecting surfaces," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 9223–9228, Aug. 2020.

[13] M. Elbir, A. Papazafeiropoulos, P. Kourtessis, and S. Chatzinotas, "Deep channel learning for large intelligent surfaces aided mm-wave massive MIMO systems," *IEEE Wireless Communications Letters*,vol. 9, no. 9, pp. 21447-1451, 2020.

[14] M. Elbir and S. Coleri, "Federated Learning for channel estimation in conventional and RIS-Assisted massive MIMO," *IEEE Transactions on Wireless Communication*, vol. 21, no. 6, pp. 4255-4268,2021

[15] S. Eyad *et al*., "RIS-assisted mmWave channel estimation using convolutional neural networks," in *Proc. 2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, 2021

[16] M. A. S. Sejan, M. H. Rahman, and H. K. Song, "Demod-CNN: A robust deep learning approach for intelligent reflecting surface-assisted multiuser MIMO Communication" *MDPI*, vol. 22, no. 16, 5971.2022.

[17] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications," *IEEE Wireless Communications and Networking Conference WCNC*, vol. 19, no. 10, pp.6607-6620, 2020.

[18] S. Kumar, T. Daniya, M. Geetha, B. Santhosh Kumar, and R. Cristin, "Least square estimation of parameters for linear regression," *International Journal of Control and Automation,* vol. 13, no. 2, pp. 447-452, 2020

[19] S. Khan, K. S. Khan, and S. Y. Shin, "Symbol denoising in high order m-qam using residual learning of deep CNN," in *Proc. 2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC). Las Vegas, NV, USA IEEE*, 2019, pp. 1–6.

[20] S. Khan and S. Y. Shin, "Deep learning aided transmit power estimation in mobile communication system," *IEEE Communications Letters*, vol. 23, no. 8, pp. 1405–1408, 2019.

[21] C. Huang, G. C. Alexandropoulos, C. Yuen, and M. Debbah, "Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces," in *Proc. 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC),* July 2019.

[22] H. Almosa, S. Mosleh, E. Perrins, and L. Liu. "Downlink channel estimation with limited feedback for FDD multiuser massive MIMO with spatial channel correlation," in *Proc. International Conference on Communications (ICC)*, IEEE, 2018, pp. 1-6.