

Enhanced Deep Residual Shrinkage Network Based Channel Estimation in RIS Communication System

Yuanyuan Ma^{*}, Zhiquan Bai^{*,†}, Bangwei He^{*}, Jinqui Zhao^{*}, Abeer Mohamed^{*}, Ke Pang^{*}, KyungSup Kwak^{**}

^{*}Shandong Provincial Key Laboratory of Wireless Communication Technologies,

School of Information Science and Engineering, Shandong University, Qingdao, Shandong, China

^{**}Graduate School of Information Technology and Telecommunications, INHA University, Incheon, Korea

Email: myy98@mail.sdu.edu.cn^{*}, zqbai@sdu.edu.cn^{*,†}, hbw017@mail.sdu.edu.cn^{*}, 202020373@mail.sdu.edu.cn^{*},
abeermohamed@mail.sdu.edu.cn^{*}, pangke9612@163.com^{*}, kskwak@inha.ac.kr^{**}

Abstract—Reconfigurable intelligent surface (RIS) composed of a large number of passive components can significantly save hardware cost and energy. However, its channel estimation is challenging since RIS commonly operates as a passive reflecting device and cannot acquire the channel state information (CSI) independently. In this paper, a data-driven approach for achieving highly accurate channel estimation in RIS assisted multi-user system is proposed. First, the channel estimation problem is modeled as a residual noise cancellation problem, and then the channel matrix with noise is obtained by a scaled least square (SLS) channel estimation. Finally, an enhanced deep residual shrinkage network (EDRSN) is designed to reduce the noise and further improve the accuracy of channel estimation.

Index Terms—channel estimation, reconfigurable intelligent surface, EDRSN, denoising block

I. INTRODUCTION

The sixth generation wireless communication system has a great demand for the improvement of data rate and the expansion of coverage. Reconfigurable intelligent surface (RIS) with the advantages of low energy consumption has been considered as an enabling technology to increase the system capacity and coverage. Specifically, RIS is a plane composed of a large number of low-cost passive reflecting elements, in which each passive element can reflect the received signal independently. The optimal design of RIS system, such as reflecting phase shift design and resource allocation, relies heavily on accurate channel state information (CSI). However, RIS is unable to independently estimate the CSI since it is a kind of passive device and cannot process the information actively. Additionally, as the number of reflecting elements in RIS system increases, the complexity of the channel estimation will increase dramatically, which poses a great challenge to acquire the CSI.

The channel estimation for RIS system can be divided into two scenarios. One is the integration of RIS and active devices. In this case, RIS is generally equipped with radio frequency (RF) chains, which can be used to estimate the channel matrix from user to RIS (UE-RIS) and from RIS to base station (RIS-BS) [1]. Another scenario is for the passive RIS, where the cascaded channel matrix of UE-RIS-BS can be estimated [2]. [3] employed compressed sensing (CS) to estimate the cascaded channel but with an obvious system overhead. In addition, the

offline Deep Neural Networks (DNN) can be used in RIS system for channel estimation to maximize the system data rate [4]. In [5], a channel estimation method by combining deep learning (DL) and CS was proposed for better estimation performance. However, despite the plentiful results about channel estimation in RIS system, low complexity and accurate channel estimation is still necessary and challenging.

The pilot based channel estimation can be modeled as a residual noise cancellation problem [6]. Meanwhile, deep residual shrinkage network (DRSN) has been proposed as a potential method for denoising [7]. Unlike the previous DL-based methods that employ DNN to recover the channel matrices, we generate a noisy channel matrix using the scaled least square (SLS) estimation [8] and propose an enhanced residual shrinkage network (EDRSN) to further denoise the channel matrix and achieve accurate channel estimation. In EDRSN, a convolutional block attention module (CBAM) as in [9] is specially designed to explore the channel features from both channel and spatial dimensions. The CBAM module enables the neural network to focus more on the decisive feature regions and ignore the irrelevant ones, and the attention allocation mechanism in both dimensions enhances the model performance. Moreover, soft thresholding is introduced on the basis of residual network (ResNet) to improve the system performance in feature extraction and denoising, which leads to better estimation accuracy. Simulation results demonstrate the superiority of the proposed EDRSN scheme in channel estimation.

II. SYSTEM MODEL

In this section, we first introduce the model of RIS assisted multi-user system and the uplink estimation problem is formulated as a residual noise cancellation problem. Then, the pilot transmission protocol and the classical channel estimation are given.

A. Pilot Transmission Protocol

In this paper, time division duplex (TDD) mode is considered. Thus, uplink and downlink channels become reciprocal. RIS assisted multi-user system as shown in Fig. 1, where RIS is

This work was supported in part by the Innovation and Development Joint Foundation of Shandong Provincial Natural Science Foundation under Grant ZR2021LZH003, in part by the National Natural Science Foundation of China under Grant 61771291, and in part by the National Research Foundation of Korea-Grant Funded by the Korean Government (MSIT-NRF-2020R1A2B5B02002478). [†]Corresponding author.

equipped with N reflecting elements, BS has N_r receiving antennas, and there exist K single antenna users.

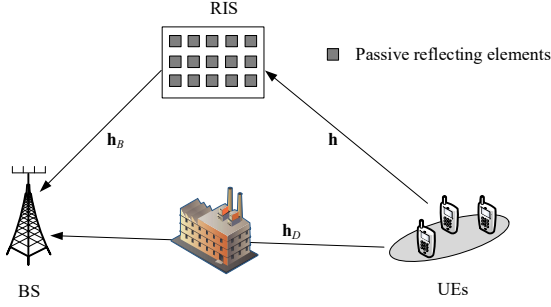


Fig. 1. RIS assisted multi-user system model.

Let $\mathbf{h}_D \in \mathbb{C}^{N_r \times 1}$ represents the channel matrix of UE-BS and $\mathbb{C}$ denotes Complex vector set, $\mathbf{h} \in \mathbb{C}^{N \times 1}$ denotes the channel matrix of UE-RIS, $\mathbf{h}_B \in \mathbb{C}^{N_r \times N}$ is the channel matrix of RIS-BS. We consider the passive RIS in this paper. Therefore, the RIS cannot separately estimate the channel matrices $\mathbf{h}$ and $\mathbf{h}_B$. The objective of our work is to estimate the cascaded channel matrix $\mathbf{h}_G = \mathbf{h}_B \text{diag}(\mathbf{h})$ and $\mathbf{h}_D$. The reflecting phase matrix is given by $\Phi = \text{diag}(\mathbf{r}) = \text{diag}([e^{j\varphi_1}, \dots, e^{j\varphi_N}]^T) \in \mathbb{C}^{N \times N}$, where $0 \leq \varphi_n \leq 2\pi$ is the reflecting phase at the n th reflecting element.

The pilot signals that contain the channel information of all the users will be sent to BS to estimate the CSI. We assume that the RIS may generate $D=N+1$ different phase shift modes when performing the channel estimation. In order to distinguish the different users, the pilot sequences of all the K users are mutually orthogonal and have the same power. For the k th user, its pilot sequence is denoted as $\mathbf{x}_k$ with length L and the transmitting power of each pilot symbol is P . The received pilot signal at BS in the d th phase shift mode can be expressed as

$$\mathbf{y}_d = \sum_{k=1}^K \mathbf{H}_k \mathbf{p}_d \mathbf{x}_k^H + \mathbf{V}_d \in \mathbb{C}^{N_r \times L} \quad (1)$$

where $\mathbf{H}_k = [\mathbf{h}_D, \mathbf{h}_G]$ is the channel matrix of k th user to be estimated, $\mathbf{p}_d = [1, \mathbf{r}_d]^T \in \mathbb{C}^{(N+1) \times 1}$ and $\mathbf{r}_d$ represents reflecting phase matrix of the d th phase shift mode, $\mathbf{V}_d$ is additive white Gaussian noise. Due to the orthogonality between pilot sequences, we can rewrite the received signal in the d th phase shift mode as

$$\mathbf{y}_{d,k} = \frac{1}{PL} \mathbf{y}_d \mathbf{x}_k = \mathbf{H}_k \mathbf{p}_d + \mathbf{V}_{d,k}. \quad (2)$$

The received signal in all phase shift modes can be expressed as

$$\mathbf{Y}_k = \mathbf{H}_k \mathbf{P} + \mathbf{V} \quad (3)$$

Where $\mathbf{P}$ represents the reflecting phase matrix with D phase shift modes, as Equation (4), which can be given by a discrete Fourier transform (DFT) [10], and $\mathbf{V} \in \mathbb{C}^{N_r \times D}$ is the noise matrix

whose element $v_{n_r,d}$ follows the complex Gaussian distribution denoted as $CN(0, \sigma_v^2)$ with zero mean and variance σ_v^2 .

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & w & \dots & w^{D-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & w^N & \dots & w^{N(D-1)} \end{bmatrix} \in \mathbb{C}^{(N+1) \times D} \quad (4)$$

B. Classical Channel Estimation

1) Least square channel estimation

Least square (LS) is designed according to the least square criterion. With known reflecting phase matrix $\mathbf{P}$ and received signal $\mathbf{Y}_k$, the channel matrix estimated by LS algorithm can be denoted by

$$\hat{\mathbf{H}}_{k, \text{LS}} = \mathbf{Y}_k \mathbf{P}^\dagger \quad (5)$$

where $\mathbf{P}^\dagger = \mathbf{P}^H (\mathbf{P} \mathbf{P}^H)^{-1}$ represents the pseudoinverse of $\mathbf{P}$ and $(\cdot)^H$ means Hermitian transpose. The channel estimation error J_{LS} is denoted as

$$J_{\text{LS}} = \mathbb{E} \left\{ \left\| \mathbf{H}_k - \hat{\mathbf{H}}_{k, \text{LS}} \right\|_F^2 \right\} = \sigma^2 N_r \text{tr} \left\{ \left(\mathbf{P} \mathbf{P}^H \right)^{-1} \right\}. \quad (6)$$

2) SLS channel estimation

LS estimation is easy to be implemented. However, it is very sensitive to noise and the estimation performance may decrease dramatically with the decrease of the signal-to-noise ratio (SNR). Therefore, it is important to optimize and update LS estimation for lower estimation error. SLS estimation can further reduce the mean square error (MSE) by allowing for bias [8]. Based on the analysis of LS estimation, the channel estimation error of SLS estimation is given by

$$\begin{aligned} & \mathbb{E} \left\{ \left\| \mathbf{H}_k - \gamma \hat{\mathbf{H}}_{k, \text{LS}} \right\|_F^2 \right\} \\ &= \left(J_{\text{LS}} + \text{tr} \{ \mathbf{R}_H \} \right) \cdot \left(\gamma - \frac{\text{tr} \{ \mathbf{R}_H \}}{J_{\text{LS}} + \text{tr} \{ \mathbf{R}_H \}} \right)^2 + \frac{J_{\text{LS}} \text{tr} \{ \mathbf{R}_H \}}{J_{\text{LS}} + \text{tr} \{ \mathbf{R}_H \}} \end{aligned} \quad (7)$$

where $\mathbf{R}_H$ represents the autocorrelation matrix of $\mathbf{H}_k$ and γ is the scaling factor. For $\gamma = \gamma_0 = \frac{\text{tr} \{ \mathbf{R}_H \}}{J_{\text{LS}} + \text{tr} \{ \mathbf{R}_H \}}$, the mean square error (MSE) of the channel estimation is minimum. The channel matrix estimated by SLS estimation can be rewritten as

$$\hat{\mathbf{H}}_{k, \text{SLS}} = \gamma_0 \hat{\mathbf{H}}_{k, \text{LS}} = \frac{\text{tr} \{ \mathbf{R}_H \}}{\sigma^2 r \text{tr} \{ (\mathbf{P} \mathbf{P}^H)^{-1} \} + \text{tr} \{ \mathbf{R}_H \}} \mathbf{Y} \mathbf{P}^\dagger \quad (8)$$

where $\text{tr} \{ \mathbf{R}_H \} = \text{tr} \{ \hat{\mathbf{H}}_{k, \text{LS}}^H \hat{\mathbf{H}}_{k, \text{LS}} \}$ is the trace of $\mathbf{R}_H$. Compared with LS, the SLS estimation further considers the effect of noise, and thus achieves a high accuracy even at low SNR regions.

Besides, for SLS estimation, there is no need to know the prior channel statistical properties. The error of SLS estimation can be computed as

$$J_{\text{SLS}} = \frac{J_{\text{LS}} \text{tr}\{\mathbf{R}_{\text{H}}\}}{J_{\text{LS}} + \text{tr}\{\mathbf{R}_{\text{H}}\}} < J_{\text{LS}}. \quad (9)$$

From Equation (9), it is seen that the error of SLS estimation is always lower than that of LS estimation. However, the estimation accuracy of SLS scheme is not sufficiently high and can still be improved. In the following, we design an EDRSN to learn the channel characteristics. In this way, the channel matrix estimated by SLS algorithm under the influence of noise is made closer to the real channel matrix, thus achieving the result of noise cancellation.

III. EDRSN BASED CHANNEL ESTIMATION

In this section, the channel estimation in RIS assisted multi-user system can be modeled as a residual noise cancellation problem and we propose an EDRSN based channel estimation scheme to improve the noise feature extraction and cancellation by introducing soft thresholding and convolutional block attention module (CBAM). Therefore, the EDRSN scheme can further increase the channel estimation accuracy. In the following, the EDRSN scheme is described in terms of residual network, CBAM, and soft thresholding.

A. Residual Network

The structure of the residual network unit is shown in Fig. 2, where $F(x)$ is the pre-summing network mapping, and $H(x)$ is the input of the post-summing network mapping. As illustrated in the ResNet, the identity mapping as well as the residual mapping can effectively alleviate the problem of gradient disappearance and performance degradation during backpropagation by transcendental input mapping and summation.

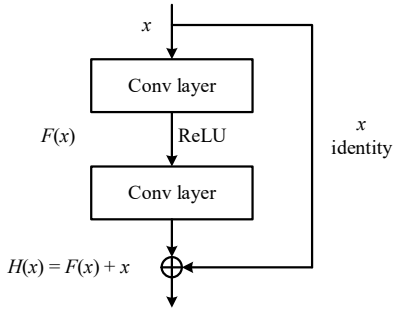


Fig. 2. Diagram of the residual network unit.

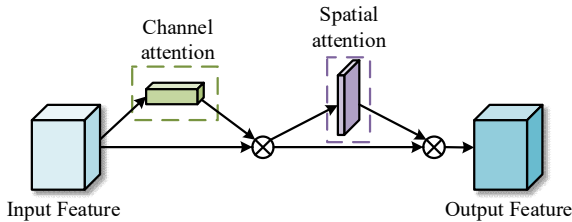


Fig. 3. Structure of CBAM.

B. Convolutional Block Attention Module

In this subsection, we further improve the residual deep shrinkage network by introducing the CBAM shown in Fig. 3. As a common attention module embedded in neural networks, the CBAM sequentially passes the input feature map through two independent dimensions, the channel dimension and the space dimension, to obtain the attention map. Then the original input feature map is multiplied with the attention map for adaptive feature optimization. The CBAM is a lightweight attention module and can improve the model performance with low complexity.

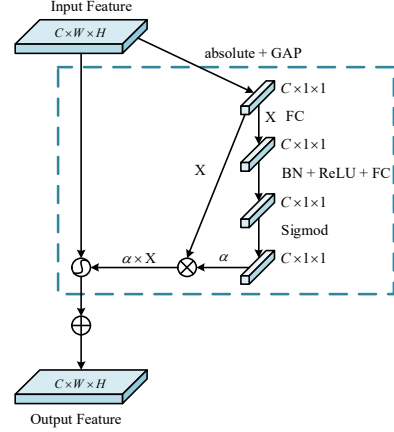


Fig. 4. Model of the soft thresholding.

C. Soft Thresholding

As shown in Fig. 4, the core of soft thresholding is to obtain a threshold value to eliminate non-significant features. First, we take absolute values of the input feature and get feature X through global pooling (GAP) layer. Subsequently, feature X is fed into a small fully connected (FC) network consisting of batch normalization (BN), ReLU and Sigmoid activation function. The output is then normalized to obtain a normalization factor given by α . The final threshold value can be formulated as $\alpha \times X$. Thus, threshold value obtained by this module is the average of the normalized input features X . Therefore, the threshold value is not only guaranteed to be positive, but it can be generated differently for different samples. Thus, soft thresholding can be viewed as a special kind of attention mechanism that distinguishes between the main and the irrelevant features and achieves noise elimination for the data by setting the irrelevant features to be 0.

As a typical denoising method [9], soft thresholding function is expressed as

$$y = \begin{cases} x - \tau & x > \tau \\ 0 & -\tau \leq x \leq \tau \\ x + \tau & x < -\tau \end{cases} \quad (9)$$

where x indicates the input feature, y denotes the output feature, and τ is the positive threshold.

We design the EDRSN as a modification of the residual network (ResNet) and it consists of one input layer and one

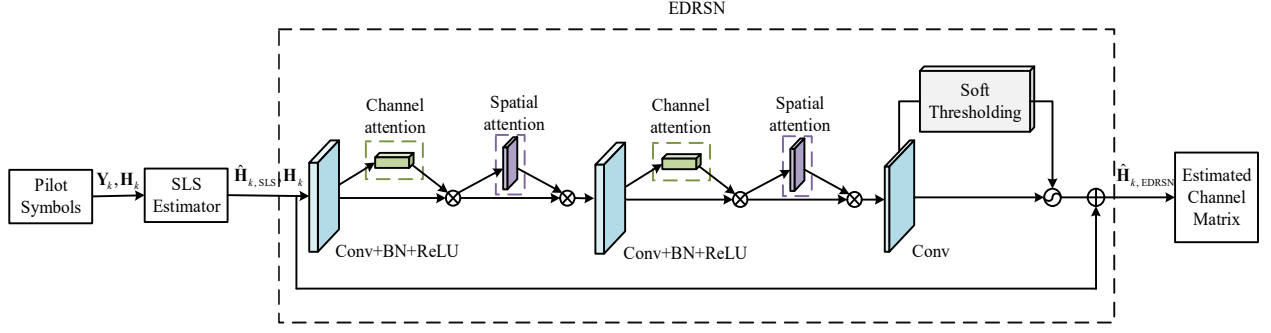


Fig. 5. The structure of the EDRSN.

residual shrinkage module with CBAM as shown in Fig 5. The residual shrinkage module includes three convolutional layers of size 3×3 , two BN layers, two ReLU activation functions and a soft thresholding module, where the BN layers are designed to speed up the convergence of EDRSN, the RELU activation functions are used to obtain the nonlinear output and the soft thresholding module is used to normalize the noise to zero. In addition, the CBAM module is embedded between the convolutional layers of the residual shrinkage module to enhance the learning capability of EDRSN for channel features. We first input $\mathbf{Y}_k$ and $\mathbf{P}$ into SLS estimation to obtain the rough channel estimation result $\hat{\mathbf{H}}_{k,\text{SLS}}$. Then, $\hat{\mathbf{H}}_{k,\text{SLS}}$ and the real channel matrix $\mathbf{H}_k$ are fed into the EDRSN as input data and labels. At last, the EDRSN optimizes the network parameters by means of a moment estimation gradient descent algorithm to make the predicted channel matrix $\hat{\mathbf{H}}_{k,\text{EDRSN}}$ more closely resemble the real channel matrix $\mathbf{H}_k$.

IV. SIMULATION RESULTS

In this section, we provide the simulation results of the proposed EDRSN channel estimation in RIS assisted multi-user system and assume that BS has $N_r = 8$ receiving antennas and there are $K = 8$ users each with a single transmitting antenna in the uplink scenario. In addition, RIS is equipped with $N = 32$ reflecting elements. The total number of epochs is set to be 200. The training will be stopped when the accuracy of the validation sets does not change for 5 epochs. We compare four channel estimation schemes, the proposed EDRSN scheme, the conventional LS scheme, the SLS estimation, and the linear minimum mean square error (LMMSE) algorithm, by evaluating the normalized mean squared error (NMSE) performance.

TABLE I. SIMULATION PARAMETERS.

Parameter	Value
RIS elements	32, 64
No. of users	8
Learning rate	0.0001
Activation functions	ReLU, Sigmoid
Loss function	NMSE
Optimizer	Adam

The RIS-BS channel follows the Rician distribution, while the channels of UE-RIS and UE-BS follow Rayleigh distribution. The distances from UE to BS, from UE to RIS, and from RIS to BS are 100, 90, and 16 meters, respectively. We assume that there are 80,000 samples, where 75% and 25 % of them are used to train and validate the EDRSN scheme, respectively. After the training stage of the EDRSN scheme, 600,000 Monte Carlo simulations are performed at each SNR to test the NMSE performance of the EDRSN scheme. The simulation parameters of the EDRSN scheme are shown in Table I.

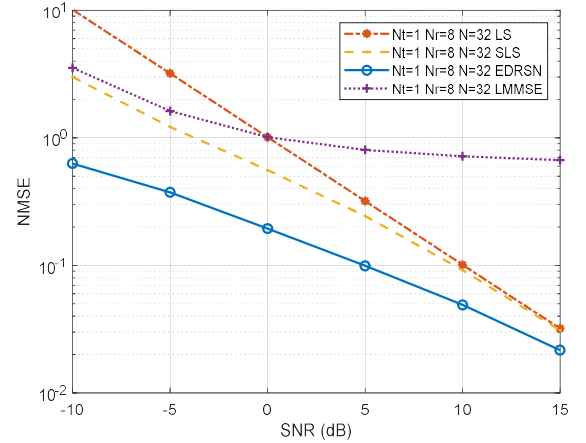


Fig. 6. NMSE performance versus SNR (dB).

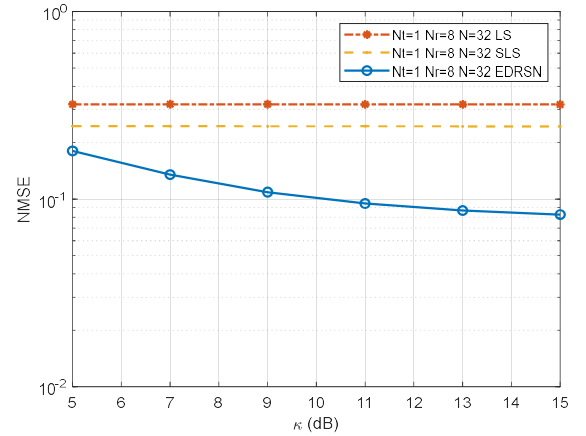


Fig. 7. NMSE performance versus Rician factor κ (dB) under SNR = 5 dB.

The NMSE performance of the proposed EDRSN scheme, LS estimation, SLS algorithm, and LMMSE scheme is shown in Fig. 6. We find that the EDRSN scheme outperforms the traditional LS, SLS and LMMSE estimation because the EDRSN scheme is a nonlinear estimation that can learn the channel features and improve the NMSE performance. In addition, the EDRSN scheme outperforms the conventional schemes especially in the case of low SNR because it can get more accurate channel matrix to eliminate the signal noise at the receiver.

In Fig. 7, we investigate the NMSE performance of the proposed EDRSN scheme under different Rician factors at $SNR = 5$ dB. The values of the Rician factor κ affect the degree of frequency selective fading, which in turn influences the channel state. We can observe that LS and SLS estimation schemes are not affected by the variation of Rician factor κ . This is because LS and SLS scheme are appropriate for arbitrary frequency-selective fading channels. In addition, since the EDRSN model is trained at $\kappa = 10$ dB, it is seen that the NMSE performance of the EDRSN scheme degrades with $\kappa < 10$ dB, due to the fact that the channel state of RIS-BS link becomes worse with small κ . However, the proposed scheme still outperforms the LS and SLS schemes and its estimation accuracy becomes better as κ increases. Therefore, when the channel state changes, the proposed EDRSN scheme can still estimate the channel matrix with high accuracy.

The NMSE performance of the proposed EDRSN scheme and the conventional LS estimation is tested in Fig. 8 with different numbers of reflecting elements, $N = 32$ and 64 . It can be seen that the channel matrix becomes more sparse as N increases, so the NMSE performance of the EDRSN scheme at $N = 64$ is better than that at $N = 32$. Besides, the EDRSN scheme is superior to the conventional LS estimation at both $N = 32$ and $N = 64$, which well demonstrates the superiority of the EDRSN scheme with different N .

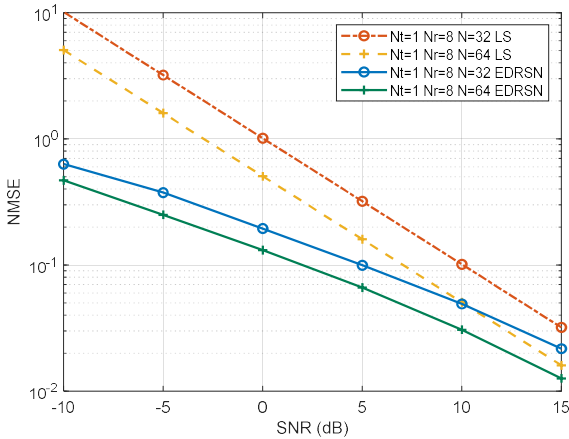


Fig. 8. NMSE performance with $N = 32$ and 64 .

V. CONCLUSION

In this paper, we propose an EDRSN channel estimation scheme based on SLS algorithm, where the SLS algorithm considers the impact of the noise on channel estimation to

improve the channel pre-estimation accuracy. In addition, CBAM and soft thresholding are introduced into the residual network to further extract the channel features and eliminate the noise. Simulation results demonstrate that the EDRSN scheme is more accurate compared with the other conventional channel estimation schemes. Our future research may focus on the design of lightweight networks to reduce the complexity while maintaining the accuracy of channel estimation.

REFERENCES

- [1] G. C. Alexandropoulos and E. Vlachos, "A hardware architecture for reconfigurable intelligent surfaces with minimal active elements for explicit channel estimation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Apr. 2020, pp. 9175-9179.
- [2] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Apr. 2019, pp. 4659-4663.
- [3] Q. U. A. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent reflecting surface-assisted multi-user MISO communication: channel estimation and beamforming design," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 661-680, May 2020.
- [4] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, "Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2018, pp. 270-278.
- [5] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," *IEEE Access*, vol. 9, pp. 44304-44321, Mar. 2021.
- [6] S. Hu, Z. Wei, Y. Cai, D. W. K. Ng, and J. Yuan, "Sum-rate maximization for multiuser MISO downlink systems with self-sustainable IRS," in *Proc. IEEE Global Commun. Conf.*, Dec. 2020, pp. 1-7.
- [7] M. Zhao, S. Zhong, X. Fu, B. Tang, and M. Pecht, "Deep residual shrinkage networks for fault diagnosis," *IEEE Trans. Ind. Inform.*, vol. 16, no. 7, pp. 4681-4690, Jul. 2020.
- [8] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884-893, Mar. 2006.
- [9] S. Woo, J. Park, J.-Y. Lee, I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vision*, pp. 3-19, Aug. 2018.
- [10] T. L. Jensen and E. D. Carvalho, "An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Apr. 2020, pp. 5000-5004.
- [11] C. Liu, X. Liu, D. W. K. Ng, and J. Yuan, "Deep residual network empowered channel estimation for IRS-assisted multi-user communication systems," in *Proc. IEEE Int. Conf. Commun.*, Aug. 2021, pp. 1-7.