

Offset Learning based Channel Estimation for Intelligent Reflecting Surface-Assisted Indoor Communication

Zhen Chen, *Member, IEEE*, Jie Tang, *Senior Member, IEEE*, Xiu Yin Zhang, *Senior Member, IEEE*, Qingqing Wu, *Member, IEEE*, Yuxin Wang, Daniel So, *Senior Member, IEEE*, Shi Jin, *Senior Member, IEEE*, and Kai-Kit Wong, *Fellow, IEEE*

Abstract—The emerging intelligent reflecting surface (IRS) can significantly improve the system capacity, and it has been regarded as a promising technology for the beyond fifth-generation (B5G) communications. For IRS-assisted multiple input multiple output (MIMO) systems, accurate channel estimation is a critical challenge. This severely restricts practical applications, particularly for resource-limited indoor scenario as it contains numerous scatterers and parameters to be estimated, while the number of pilots is limited. Prior art tackles these issues and associated optimization using mathematical-based statistical approaches, but are difficult to solve as the number of scatterers increase. To estimate the indoor channels with an affordable piloting overhead, we propose an offset learning (OL)-based neural network for channel estimation. The proposed OL-based estimator can dynamically trace the channel state information (CSI) without any prior knowledge of the IRS-assisted channel structure as well as indoor statistics. In addition, inspired by the powerful learning capability of convolutional neural network (CNN), CNN-based inversion blocks are developed in the offset estimation module to build the offset estimation operator. Numerical results show that the proposed OL-based estimator can achieve more accurate indoor

This work has been supported in part by Key Research and Development Project of Guangdong Province under Grant 2019B010156003, in part by National Key Research and Development Project under Grant 2019YFB1804100, in part by the National Natural Science Foundation of China under Grant 61971194, 62001171, in part by the Macau Science and Technology Development Fund, Macau SAR, under Grant 0119/2020/A3SKL-IOTSC-2021-2023 and 0108/2020/A, in part by the Natural Science Foundation of Guangdong Province under Grant 2019A1515011607, 2021A1515011966, in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University (No. 2019D06, 2021D15), in part by the Fundamental Research Funds for the Central Universities under Grant 2019JQ08, and in part by the Research Fund Program of Guangdong Key Laboratory of Aerospace Communication and Networking Technology under Grant 2018B030322004. (*Corresponding author: Jie Tang.*)

Z. Chen, X. Zhang and Y. Wang are with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China. (e-mail: chenz@scut.edu.cn; zhangxiuyin@scut.edu.cn and wang.yuxin@zte.com.cn).

J. Tang is with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China, and also with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: eejtang@scut.edu.cn).

Q. Wu is with the State Key Laboratory of Internet of Things for Smart City, University of Macau, Macau, 999078, China (email: qingqingwu@um.edu.mo).

D. K. C. So is with the School of Electrical and Electronic Engineering, University of Manchester, Manchester, United Kingdom. (e-mail: d.so@manchester.ac.uk).

S. Jin is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China, (e-mail:jinshi@seu.edu.cn).

K.-K Wong is with the Department of Electronic and Electrical Engineering, University College London, London, United Kingdom. (e-mail: kai-kit.wong@ucl.ac.uk).

CSI with a lower complexity as compared to the benchmark schemes.

Index Terms—Indoor 5G, indoor channel estimation, massive MIMO, deep learning, intelligent reflecting surface (IRS)

I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) as a key technology of the fifth-generation (5G) wireless communications system has received considerable research interests in both academia and industry [1], [2]. Although massive MIMO with beamforming is able to improve the spectral efficiency and boost the system throughput, the channel channels between the transceivers lack line-of-sight (LOS) links and suffer from deep fading and shadowing due to unfavorable propagation conditions e.g. offices, factories and indoor environments.

Recently, intelligent reflecting surface (IRS) has been developed to enhance coverage, user throughput, and system capacity with low hardware cost and energy consumption [3]. Due to the low-cost and passive advantage of reflecting materials, IRS operates without transmit power consumption and can be deployed properly to establish extra links between the access point (AP) and user equipments (UEs). By reconfiguring IRS's reflecting elements according to the surrounding radio environment, IRSs are able to control passive beamforming through adjusting the phase shifts of IRS to increase the received signal power, suppress interference and achieve high beamforming gain [4]. It is worth noting that high beamforming gains of IRS require accurate channel state information (CSI) of both the AP-IRS link and IRS-UE link. Although channel estimation has been widely studied in the conventional wireless communication system, channel estimation of an IRS-assisted wireless communication is still challenging due to the following reasons. First, as a passive device, IRS does not possess any radio frequency (RF) chains, which renders it incapable of performing pilot signaling functionality [5]. Thus the conventional pilot-aided channel estimation may not be applied to IRS-assisted channel estimation directly. Second, a large number of reflecting elements should be optimized in a limited channel training time, which may fail to support the mobility of practical indoor scenarios, such as supermarket, commercial towers and shopping mall [6]. This motivates us to devote our endeavour to developing efficient indoor channel estimation scheme for IRS-assisted beyond 5G (B5G) system.

A. Recent Works

Various IRS-assisted strategies have been studied to efficiently estimate the mmWave channels [7]–[13]. In [7], the authors proposed the channel estimation for IRS-assisted energy transfer, in which the physical channel was converted into a sequence of cascade channel estimation by adjusting the reflective elements. It is worth noting that the training overhead linearly increases with the number of IRS elements and may be prohibitively large in practice. To reduce the training overhead, compressed sensing (CS)-based channel estimation was developed in IRS-assisted mmWave systems by exploiting the sparsity structure of mmWave channels [8]. In [9], a sparse matrix factorization-based channel estimation framework was proposed for IRS-assisted MIMO by exploiting sparse matrix factorization and matrix completion. However, such a scheme requires to switch ON and OFF some reflecting elements, which is costly to implementation as it needs to control the amplitude of each reflecting element separately [10]. In [11], the parallel factor decomposition was investigated and the corresponding alternating least squares (ALS) algorithm was adapted to estimate the cascaded channels. In [14], a direction-of-arrival (DOA)-based channel estimation approach was considered by using the limited RF chains. In [15], two channel models were investigated for robust beamforming design in IRS-assisted MISO systems, which consider the bounded CSI error and the statistical CSI error, respectively.

Meanwhile, in 5G/B5G application scenarios such as dense factories, industrial and ultra-stringent office building circumstances, the radio scattering conditions change across time rapidly. The existing works with the adopted mathematical model are closely dependent on hypothetical statistical models, which may result in serious channel mismatch and offset error. To address this issue, deep learning (DL)-based solution has been investigated to implement the channel estimation tasks and achieve superior performance [16]–[20]. In [21], the authors proposed a DL-based channel estimation framework for OFDM system by exploiting the time-frequency grid of the channel response model. The proposed scheme has a competitive performance and is more robust than conventional channel estimator. In [22], the DL-based channel estimator was proved more robust than conventional channel estimation algorithms when small amounts of training pilots are used. In addition, authors also proposed a DL-based channel estimation for beamspace MIMO system [23]. A similar network model was utilized in [24], where the authors presented deep channel estimator with untrained deep neural network to reduce pilot contamination. To tackle the time-varying Rayleigh fading channel, a sliding bidirectional gated recurrent unit channel estimator was designed to improve the channel estimation performance [25]. In [26], a residual learning based deep neural network (DNN) was designed for channel estimation. Although DNNs were used to implement a joint scheme of pilot training and channel estimation, the number of linear layers in DNN was sensitive to neurons size and the length of input data. Despite data can be segmented into patches to address this issue, segmented patches will cause the discontinuity of estimation results. Recently, some unfolded architecture of

the iterative procedure is developed to form the framework of the deep learning network [27]–[30]. In [27], the architecture of unfolding FISTA is developed to form a deep network, where the parameters are learned through end-to-end training. Subsequently, a new DL-based network architecture named alternating direction method of multipliers (ADMM)-Net was proposed, which was derived from the iterative procedures in the ADMM algorithm [28]–[30]. Due to its flexibility in parameter learning, the network framework can also be potentially applied to channel estimation through jointly estimate the LOS and non-LOS (NLOS) channels (also call reflected paths). However, these kind of joint LOS and NLOS channel estimation network may not be suitable for indoor environments (including a classroom), since the characteristics of channel depends on the LOS and reflected channels that are differences greatly. To illustrate this observation, a typical indoor measurements were conducted in [31], which provide large-scale path loss and temporal statistics. Results indicates that in LOS environments, the close-in free space reference distance path loss model with respect to 1 m free space path loss distance resulted in the path loss exponent of 1.7, while in NLOS cases, diffraction was a significant propagation mechanism that the path loss through walls was very high, and hence the measured path loss exponents is 3.1. The result of measurements thoroughly verified that the characteristics of LOS and NLOS channel are obviously different, which can not be estimated with conventional ADMM-CSNet scheme directly to achieve satisfactory performance.

B. Contributions

The combination of IRS and DL is envisioned as an emerging and enabling technique to enhance the coverage of the upcoming B5G wireless communications, and therefore has drawn significant attention recently. The previous literature on IRS-assisted channel estimation mainly focus on the design of outdoor scenario. In fact, scatterer-rich indoor scenarios are a common scenario in micro- or pico-cell deployments, such as those corresponding to UEs in industrial areas being served by small APs, which is as important as outdoor. However, a realistic scenario, indoor channels typically have a large of short-distance scatterer-rich path, which yields increased channel dimension, diversity and computation complexity in channel estimation. As the dimension of space or frequency increases, existing mathematical-based deterministic approaches were limited on the joint optimization of indoor channel estimation and reflection elements of IRS. Furthermore, under IRS-assisted indoor channels scenario, the propagation characteristics can dramatically change even for short displacements of the UEs; thus making it difficult for indoor channel estimation. To satisfy the requirements of indoor B5G communication, we explore novel channel estimation techniques with deep learning architecture to reliably estimate the IRS-assisted channels with a reduced number of pilot overhead in indoor communication scenario. To address the computationally expensive optimization problem with mathematical-based deterministic algorithm, we unroll the iterative procedures of the ADMM-based deterministic algorithm to a supervised

model-driven network. All the parameters (e.g., sparse transforms, regularization operator, penalty parameters, etc.) in the deep architecture can be discriminatively learned from training pairs of pilot signal and estimated signal over the data flow graph. The main contributions of our work are summarized as follows.

- We characterize the 5G indoor channel model with random number of clusters/scatterers in indoor scenarios. The model includes many physical characteristics such as shadowing effects, LOS probability, and shared scatterers, which may result in high computational complexity of the channel estimation solutions. To address this issue, we investigate IRS-assisted channel estimation for the indoor communication system by exploiting deep neural networks. By utilizing the properties of CS theory, the complicated indoor channel model can be reformulated as a cascade channel, which is a joint task of the IRS-to-BS channel matrix and the IRS-to-UE channel matrix estimation.
- Since the reflected channels are different from the LOS channel, there are many random clusters/scatterers, which includes many physical characteristics such as shadowing effects, shared scatterers. And thus, the existing ADMM-CSNet is difficult to estimate the optimize LOS path and NLOS pathes (also call reflected links) simultaneously. To tackle this challenge, a DL-based framework named offset learning neural network (OLNN-Net) channel estimator is proposed to learn the indoor fading channel. Moreover, offset learning module is introduced to improve the precision of channel estimation. We also optimize the number of hidden nodes to reduce the number of neurons from the order of millions to a few thousand, which demonstrates the potential benefit to network training.
- In the proposed OLNN-Net, the conjugate gradient (CG) optimization scheme is developed to solve the $(\rho\mathbf{I} + \Phi^H\Phi)^{-1}$ in (22) that can avoid calculating matrix inversion of the $(\rho\mathbf{I} + \Phi^H\Phi)$, which is desirable for large-scale or ill-conditioned inverse problems. Moreover, there are no trainable parameters within intermediate results, and hence a large number of CG scheme are performed almost without memory overhead during training.
- The performance analysis of the proposed OLNN-Net is provided in different cases. The simulation results show the proposed OLNN-Net estimator achieves higher spectral efficiency and lower normalized mean square error (NMSE) performance compared with conventional approaches and other DL-based methods. Besides, it demonstrated that the proposed channel estimator can be potentially applied in indoor 5G communications.

C. Organization and Notation

The remaining of this paper is organized as follows. In Section II, we describe the system model of an IRS-assisted system along with formulation of the channel estimation optimization problem. In Section III, we present the basic idea and the detailed optimization procedure. In Section IV, simulations are provided to validate the theoretical findings, followed by the conclusions in Section V.

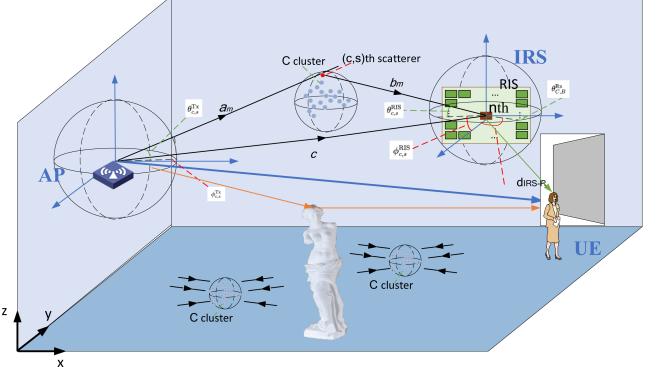


Fig. 1: Generic indoor communication with clusters between AP-IRS and an IRS mounted in the wall.

Notations: Scalar is denoted by lower-case letters; lower- and upper-case boldface letters are used for vectors and matrices, respectively; “*” denotes the convolution operator; the superscripts $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and conjugate transpose, respectively; \otimes stands for the Kronecker product; \mathbf{x}^H are the conjugate of a complex signal \mathbf{x} and H would be complex conjugate transpose. $(\cdot)_j$ is denoted by the j -th column of a matrix; $(\cdot)_{i,j}$ denotes i, j -th element of a matrix; $\|\cdot\|_2$, $\|\cdot\|_F$ and $\mathbb{E}[\cdot]$ denote the Euclidean norm, Frobenius norm and expectation operators, respectively.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. IRS-Assisted mmWave MIMO system

We consider an IRS-assisted mmWave indoor communication, in which an AP with N transmit antennas is serving a single-antenna UE. The IRS is deployed in the form of an $L \times L$ uniform rectangular array (URA) to assist the communication between the UE and the AP. For the requirement in indoor scenario, the clustered statistical MIMO channel model is studied, which is generally used in 3GPP standardization [32]. As shown in Fig. 1, there are M scatterers between the AP and the IRS, while assuming a LOS channel link from the IRS to the UE. We assume the UE transmit power as τ , i.e., $\mathbb{E}[|s_b|^2] = \tau, \forall t$. Denote a_m , b_m and c as the distances from AP to the m -th scatterer, from the m -th scatterer to the IRS, and from the IRS element to UE, respectively. Over the time duration T , the transmitted signal vector associated with the b -th ($1 \leq b \leq B$) subcarrier can be given by

$$\mathbf{x}_b = \mathbf{P}_b \mathbf{s}_b, \quad (1)$$

where $\mathbf{P} \in \mathbb{C}^{N \times T}$ denotes the digital precoding matrix at the AP, $\mathbf{s}_b = [s_{b,1}, \dots, s_{b,T}]^T$ denotes the training sequence of UE, $s_{b,t}$ is the training symbol of UE associated with the t -th time slot. Thus, the received signal $\hat{\mathbf{y}} \in \mathbb{C}^{K \times 1}$ at the UE can be written as

$$\hat{\mathbf{y}}_b = \mathbf{W}_b (\mathbf{H}_{UA,b}^H + \mathbf{G}^H \boldsymbol{\Theta} \mathbf{H}_{IA}) \mathbf{x}_b + \mathbf{n}_b, \quad (2)$$

where $\mathbf{H}_{UA,b} \in \mathbb{C}^{N \times K}$ denotes the channels between the AP and the UE, $\mathbf{H}_{IA} \in \mathbb{C}^{L \times N}$ and $\mathbf{G} \in \mathbb{C}^{L \times K}$ denote the channels from the AP to the IRS and from the IRS to the

UE, respectively. $\mathbf{W}_b \in \mathbb{C}^{M \times K}$ denotes antenna selectors. $\Theta \in \mathbb{C}^{L \times L}$ is diagonal phase-shifting matrix of the IRS, which can be expressed as

$$\Theta = \text{diag}(\beta_1 e^{j\theta_1}, \dots, \beta_L e^{j\theta_L}), \quad (3)$$

and $\theta_l \in [0, 2\pi]$, $\beta_l \in [0, 1]$, $l = 1, \dots, L$ being the reflection amplitude and phase shift associated with the l -th reflecting element of the IRS, respectively. $\mathbf{n}_b \in \mathbb{C}^{M \times 1}$ is the additive white Gaussian noise (AWGN) vector whose entries are with zero mean and variance σ^2 . From (2), three transmission channels are considered in IRS-assisted system including the LOS subchannel \mathbf{H}_{UA} from AP to the UE, the reflected subchannel \mathbf{G} from IRS to the UE and the reflected subchannel \mathbf{H}_{IA} from AP to IRS.

B. IRS-Assisted channel model

1) *AP-IRS Channel*: It is assumed that the total number of M scatterers are described with C clusters. Each of cluster is consists of S_c sub-rays for $c = 1, \dots, C$, that is $M = \sum_{c=1}^C S_c$ and there are sharing the same spatial and temporal distribution characteristics [33]. Then, the AP-IRS channel link $\mathbf{H}_{IA} \in \mathbb{C}^{L \times K}$ can be expressed as a clustered model as follows

$$\mathbf{H}_{IA} = \mathbf{H}_{IA,NLOS} + \mathbf{H}_{IA,LOS}, \quad (4)$$

where $\mathbf{H}_{IA,NLOS}$ can be expressed as

$$\mathbf{H}_{IA,NLOS} = \gamma \sum_{c=1}^C \sum_{s=1}^{S_c} \beta_{c,s} \sqrt{G_e(\theta_{c,s}^{IRS}) L_{c,s}^{IRS}} \mathbf{a}(\varphi_{c,s}^{AP}) \mathbf{a}(\phi_{c,s}^{IRS}, \theta_{c,s}^{IRS})^H \quad (5)$$

and $\gamma = \sqrt{\frac{1}{\sum_{c=1}^C S_c}}$ is a normalization factor used in clustered channel models [32]. $\varphi_{c,s}^{AP}$ is the corresponding azimuth angle-of-arrival (AoA) to the AP; $L_{c,s}^{IRS}$ denotes the path gains of the AP-IRS channel associated with the (c, s) -th propagation path. $\phi_{c,s}^{IRS}$ (or $\theta_{c,s}^{IRS}$) denotes the corresponding angle-of-departure (AoD) of the IRS. The steering vector $\mathbf{a}(\varphi_{c,s}^{AP})$ (or $\mathbf{a}(\phi_{c,s}^{IRS}, \theta_{c,s}^{IRS})$) of the AP (or IRS) can be written as

$$\begin{aligned} \mathbf{a}(\varphi_{c,s}^{AP}) &= f_M (\sin(\varphi_{c,s}^{AP})), \\ \mathbf{a}(\phi_{c,s}^{IRS}, \theta_{c,s}^{IRS}) &= f_{L_2} (-\cos(\phi_{c,s}^{IRS}) \cos(\varphi_{c,s}^{IRS})) \\ &\quad \otimes f_{L_1} (\sin(\phi_{c,s}^{IRS}) \cos(\varphi_{c,s}^{IRS})), \end{aligned} \quad (6)$$

where \otimes denotes the Kronecker product and

$$f_N(x) \triangleq \frac{1}{\sqrt{N}} \left[1, e^{-j\frac{2\pi}{\varrho} dx}, \dots, e^{-j\frac{2\pi}{\varrho} d(N-1)x} \right], \quad (7)$$

and ϱ denotes the mmWave wavelength, d is the Euclidean distance between any two adjacent antennas, and $G_e(\theta_{c,s}^{IRS})$ denotes the rotationally symmetric IRS element pattern with the (c, s) -th directed scatterer [34]. Inspired by [34], the \cos^q pattern is used to model the reflect arrays of IRS

$$G_e(\theta_{c,s}^{IRS}) = 2(2q+1)\cos^{2q}(\theta_{c,s}^{IRS}), \quad -\frac{\pi}{2} < \theta_{c,s}^{IRS} < \frac{\pi}{2}, \quad (8)$$

where $2(2q+1)$ is defined as energy conservation and $q = 0.25G_e(0) - 0.5$ is consistent with in [35]. For the attenuation of the (c, s) -th path, the close-in free space reference distance in the 5G standard protocol is considered to model the

frequency-dependent of path loss, which is widely used for the indoor hotspot and urban microcellular scenarios [36]:

$$\begin{aligned} L_{c,s}^{IRS} &= -20 \log_{10} \left(\frac{4\pi}{\lambda} \right) - X_\sigma \\ &\quad - 10n \left(1 + \varpi \left(\frac{f - f_0}{f_0} \right) \right) \log_{10}(d_{c,s}), \end{aligned} \quad (9)$$

where $d_{c,s} = a_{c,s} + b_{c,s}$ denotes the length of the (c, s) -th path, and f_0 denotes a fixed reference frequency of the path loss model, ϖ denotes a system parameter and $X_\sigma \sim \mathcal{N}(0, \sigma^2)$ is the shadow fading factor in dB, which is the same in [37].

On the other hand, there is an existing LOS link from the AP to IRS in (4). To this end, the LOS channel link attenuation for each sub-ray is introduced from the AP to IRS, which is calculated by

$$\mathbf{H}_{IA,LOS} = \Gamma(d_{IRS}) \sqrt{G_e(\theta_{LOS}^{IRS}) L_{LOS}^{IRS}} \eta \mathbf{a}(\varphi_{IRS}^{AP}) \mathbf{a}(\phi_{LOS}^{IRS}, \theta_{LOS}^{IRS})^H, \quad (10)$$

where L_{LOS}^{IRS} denotes the attenuation of the LOS link to be calculated with (9), $G_e(\theta_{LOS}^{IRS})$ denotes the corresponding channel gain of IRS, $\eta \sim \mathcal{U}[0, 2\pi]$ is the random phase and $\mathbf{a}(\varphi_{IRS}^{AP})$ (or $\mathbf{a}(\phi_{LOS}^{IRS}, \theta_{LOS}^{IRS})$) denotes the array response of the AP (or IRS). $\Gamma(d_{c,s})$ is a Bernoulli random variable taking value between 0 and 1. Let $p(\Gamma(d_{IRS}))$ be the frequency independent LOS probability. Then, we can resort to the 5G channel model [36] to achieve

$$p(\Gamma(d_{IRS})) = \begin{cases} 1 & d_{c,s} \leq 1.2 \\ e^{-\frac{d_{c,s}-1.2}{4.7}} & 1.2 < d_{c,s} \leq 6.5 \\ 0.32e^{-\frac{d_{c,s}-6.5}{32.6}} & d_{c,s} > 6.5. \end{cases} \quad (11)$$

2) *IRS-UE Channel*: In indoor communication scenarios, assuming that the IRS is close to the UE and there is a clear LOS channel link between the IRS and the UE without noticeable NLOS channel link. According to (11), it is not difficult to notice that if the distance between the IRS and UE is less 4.5m, an LOS link probability of greater than 50% is achieved. Therefore, this distance generally takes place in indoor communication. To be specific, the IRS-UE channel \mathbf{G} can be formulated as

$$\mathbf{G} = \sqrt{G_e(\theta_{UE}^{IRS}) L_{LOS}^{IRS-R}} e^{j\eta} \mathbf{a}(\phi_{UE}^{IRS}, \theta_{UE}^{IRS}) \mathbf{a}(\varphi_{LOS}^{UE})^H, \quad (12)$$

where $\mathbf{a}(\phi_{UE}^{IRS}, \theta_{UE}^{IRS})$ denotes the array response of IRS to be calculated by the azimuth and elevation departure angles ϕ_{UE}^{IRS} and θ_{UE}^{IRS} , $\mathbf{a}(\varphi_{LOS}^{UE})$ denotes the array response of the UE, and L_{LOS}^{IRS-R} denotes the attenuation of the path loss along with channel related parameters between the IRS and UE that is obtained with (9). The gain of IRS $G_e(\theta_{UE}^{IRS})$ with respect to the departure elevation angle θ_{UE}^{IRS} is obtained by (8).

3) *AP-UE Channel*: As shown in [36], if the distance from the IRS to the UE is less than the correlation distance, there are common clusters in indoor environments. This motives us to model the channel between AP and the UE by exploiting

$$\begin{aligned}
\text{vec}(\mathbf{H}_b) &= \text{vec}((\mathbf{F}_{LOS,1}\boldsymbol{\Gamma}_{LOS}\mathbf{F}_{LOS,2}^H + \mathbf{F}_{NLOS,1}\boldsymbol{\Gamma}_{NLOS}\mathbf{F}_{NLOS,2}^H)\boldsymbol{\Theta}\mathbf{R}_1\boldsymbol{\Lambda}_G\mathbf{R}_2^H) \\
&= \text{vec}(\mathbf{F}_{LOS,1}\boldsymbol{\Gamma}_{LOS}\mathbf{F}_{LOS,2}^H\boldsymbol{\Theta}\mathbf{R}_1\boldsymbol{\Lambda}_G\mathbf{R}_2^H) + \text{vec}(\mathbf{F}_{NLOS,1}\boldsymbol{\Gamma}_{NLOS}\mathbf{F}_{NLOS,2}^H\boldsymbol{\Theta}\mathbf{R}_1\boldsymbol{\Lambda}_G\mathbf{R}_2^H) \\
&\stackrel{(a)}{=} \mathbf{Q}_{LOS}\text{vec}(\boldsymbol{\Gamma}_{LOS}\mathbf{F}_{LOS,2}^H\boldsymbol{\Theta}\mathbf{R}_1\boldsymbol{\Lambda}_G) + \mathbf{Q}_{NLOS}\text{vec}(\boldsymbol{\Gamma}_{NLOS}\mathbf{F}_{NLOS,2}^H\boldsymbol{\Theta}\mathbf{R}_1\boldsymbol{\Lambda}_G) \\
&\stackrel{(b)}{=} \mathbf{Q}_{LOS}(\boldsymbol{\Lambda}_G^T \otimes \boldsymbol{\Gamma}_{LOS})(\mathbf{R}_1^H \otimes \mathbf{F}_{LOS,2}^H)\mathbf{v} + \mathbf{Q}_{NLOS}(\boldsymbol{\Lambda}_G^T \otimes \boldsymbol{\Gamma}_{NLOS})(\mathbf{R}_1^H \otimes \mathbf{F}_{NLOS,2}^H)\mathbf{v} \\
&\stackrel{(c)}{=} ((\mathbf{R}_G^H \otimes \mathbf{F}_{LOS,2}^H)\mathbf{v})^T \otimes \mathbf{Q}_{LOS}\mathbf{h}_{LOS,b} + ((\mathbf{R}_G^H \otimes \mathbf{F}_{NLOS,2}^H)\mathbf{v})^T \otimes \mathbf{Q}_{NLOS}\mathbf{h}_{NLOS,b},
\end{aligned} \tag{13}$$

shared clusters [38]. Therefore, the AP-UE channel can be expressed as

$$\mathbf{H}_{UA,b} = \gamma \sum_{c=1}^C \sum_{s=1}^{S_c} \beta_{c,s} \sqrt{G_e(\theta_{c,s}^{UA}) L_{c,s}^{IRS}} \mathbf{a}(\varphi_{c,s}^{UE}) \mathbf{a}(\phi_{c,s}^{IRS}, \theta_{c,s}^{IRS}), \tag{14}$$

where the parameters γ , C , $L_{c,s}^{IRS}$ and S_c are defined as in (5) and the common clusters are considered for the AP-UE channel and the AP-IRS channel. By turning off the IRS reflecting elements, i.e., $\beta_{c,s} = 0$, the channel $\mathbf{H}_{UA,b}$ can be estimated by considered the conventional estimation schemes.

Combining (4), (12) with (14) and assuming the pilot signal $\mathbb{E}[\mathbf{s}_b\mathbf{s}_b^H] = \mathbf{I}$, we can further obtain (2) by using the vectorization operation

$$\mathbf{y}_b = (\mathbf{P}_b\mathbf{s}_b \otimes \mathbf{W}_b)\text{vec}(\mathbf{H}_b) + \mathbf{n}_b = \Phi\mathbf{h}_b + \mathbf{n}_b, \tag{15}$$

where $\mathbf{y}_b = \hat{\mathbf{y}}_b - \Phi\mathbf{H}_{UA,b}$, $\Phi = (\mathbf{P}_b\mathbf{s}_b) \otimes \mathbf{W}_b$ is the measurement matrix, $\mathbf{h}_b = \text{vec}(\mathbf{H}_b) = \text{vec}(\mathbf{H}_{IA}\boldsymbol{\Theta}\mathbf{G})$ is the vectorized channel. According to the definition of \mathbf{H}_b , the cascade channel \mathbf{H}_b can be further represented as

$$\begin{aligned}
\mathbf{H}_b &= \mathbf{H}_{IA}^H \boldsymbol{\Theta} \mathbf{G} \\
&\stackrel{(a)}{=} (\mathbf{H}_{IA,NLOS} + \mathbf{H}_{IA,LOS})\boldsymbol{\Theta} \mathbf{G} \\
&= (\mathbf{F}_{LOS,1}\boldsymbol{\Gamma}_{LOS}\mathbf{F}_{LOS,2}^H + \mathbf{F}_{NLOS,1}\boldsymbol{\Gamma}_{NLOS}\mathbf{F}_{NLOS,2}^H) \boldsymbol{\Theta}\mathbf{R}_1\boldsymbol{\Lambda}_G\mathbf{R}_2^H
\end{aligned} \tag{16}$$

where (a) is from (4), $\mathbf{F}_{NLOS,1}$ and $\mathbf{F}_{NLOS,2}$ denote the overcomplete matrices and each of its columns has a form of $\mathbf{a}(\varphi_{c,s}^{UE})$ and $\mathbf{a}(\phi_{c,s}^{IRS}, \theta_{c,s}^{IRS})$, respectively; $\mathbf{F}_{LOS,1}$, $\mathbf{F}_{LOS,2}$ are defined with each of its columns having a form of $\mathbf{a}(\varphi_{c,s}^{AP})$, and $\mathbf{a}(\phi_{c,s}^{IRS}, \theta_{c,s}^{LOS})$, respectively. Similarly, \mathbf{R}_1 and \mathbf{R}_2 are formed from columns $\mathbf{a}_{IRS,2}(\varphi_{c,s})$ and $\mathbf{a}_{AP}(\phi_{c,s})$, respectively.

Furthermore, the vectorization of \mathbf{H}_b can be expressed in (13), where (a) is from $\mathbf{Q}_{LOS} = (\mathbf{R}_2^* \otimes \mathbf{F}_{LOS,1})$, $\mathbf{Q}_{NLOS} = (\mathbf{R}_2^* \otimes \mathbf{F}_{NLOS,1})$, (b) comes from $\mathbf{v} = \text{vec}(\boldsymbol{\Theta})$ and Kronecker product of matrices, and (c) is from $\mathbf{h}_{LOS,b} = \text{vec}(\boldsymbol{\Lambda}_G^T \otimes \boldsymbol{\Gamma}_{LOS})$ and $\mathbf{h}_{NLOS,b} = \text{vec}(\boldsymbol{\Lambda}_G^T \otimes \boldsymbol{\Gamma}_{NLOS})$.

In the channel estimation phase, to estimate the b -th subcarrier of the channel, the received signal can be expressed as

$$\begin{aligned}
\mathbf{y}_b &= (\mathbf{P}_b\mathbf{s}_b \otimes \mathbf{W}_b)\text{vec}(\mathbf{H}_b) + \mathbf{n}_b \\
&= \Phi_{LOS}\mathbf{h}_{LOS,b} + \Phi_{NLOS}\mathbf{h}_{NLOS,b} + \mathbf{n}_b,
\end{aligned} \tag{17}$$

where $\Phi_{LOS} = (\mathbf{P}_b\mathbf{s}_b \otimes \mathbf{W}_b)((\mathbf{R}_G^H \otimes \mathbf{F}_{LOS,2}^H)\mathbf{v})^T \otimes \mathbf{Q}_{LOS}$ and $\Phi_{NLOS} = (\mathbf{P}_b\mathbf{s}_b \otimes \mathbf{W}_b)((\mathbf{R}_G^H \otimes \mathbf{F}_{NLOS,2}^H)\mathbf{v})^T \otimes \mathbf{Q}_{NLOS}$.

By stacking all the subcarriers, the received signal can be given by

$$\mathbf{Y} = \Phi_{LOS}\mathbf{H}_{LOS} + \Phi_{NLOS}\mathbf{H}_{NLOS} + \mathbf{N}, \tag{18}$$

where $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_B]^T$, $\mathbf{H}_{LOS} = [\mathbf{h}_{LOS,1}, \mathbf{h}_{LOS,2}, \dots, \mathbf{h}_{LOS,B}]^T$ is the LOS channel link, $\mathbf{H}_{NLOS} = [\mathbf{h}_{NLOS,1}, \mathbf{h}_{NLOS,2}, \dots, \mathbf{h}_{NLOS,B}]^T$ is the NLOS channel link and $\mathbf{N} = [\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_B]^T$.

It is observed from (18) that the cascaded indoor channel between the AP and UE can be of extremely high dimension such that many parameters will need to be estimated. Although most conventional strategies can employ regularized-based iterative algorithms to estimate the channel estimation problem, such as ADMM. These traditional recovery methods have high computational cost and many penalty parameters, the Lagrange multipliers and fixed sparse transformation need to be emphasized in iterative processes. To overcome above deficiency, the DL-based framework with offset learning will be proposed to solve (18) in the next section.

III. DEEP LEARNING BASED CASCADED CHANNEL ESTIMATION

In this section, we present the channel estimation problem formulation together with its detailed algorithmic solution. Due to the angular-domain sparsity nature of mmWave MIMO channels, the high-dimensional channel estimation problem can be solved by applying CS theory to reduce pilot overhead [39]. Many CS-based algorithms are developed to tackle this problem. Inspired by previous work [40], the augmented Lagrangian function of a given CS-based channel model is proposed, which splits the variables into several blocks and can be optimized alternatively. By introducing an auxiliary variable \mathbf{U} , the problem (18) can be acquired by solving the following optimization problem

$$\begin{aligned}
&\min_{\mathbf{H}, \mathbf{U}} \frac{1}{2} \|\mathbf{Y} - \Phi\mathbf{H}\|_F^2 + \lambda_a g_a(\mathbf{H}_{LOS}) \\
&\quad + \lambda_b g_b(\mathbf{H}_{NLOS}) + \frac{\rho}{2} \|\mathbf{H} - \mathbf{U}\|_F^2 \\
&\text{s.t. } \mathbf{U} = \mathbf{U}_{LOS} + \mathbf{U}_{NLOS},
\end{aligned} \tag{20}$$

where λ_a and λ_b denote regularization parameters, ρ is the penalty parameter and $g_a(\cdot)$, $g_b(\cdot)$ are the regularization function derived from the prior knowledge, such as l_p -norm ($0 \leq p \leq 2$).

To optimize the ill-posed inversion problem (20), ADMM has proven to be an excellent variable splitting algorithm with convergence guarantee. However, the approach has a

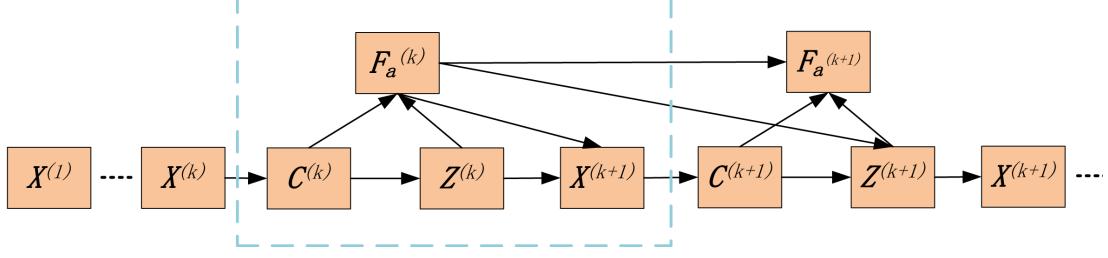


Fig. 2: The data flow graph for the ADMM algorithm

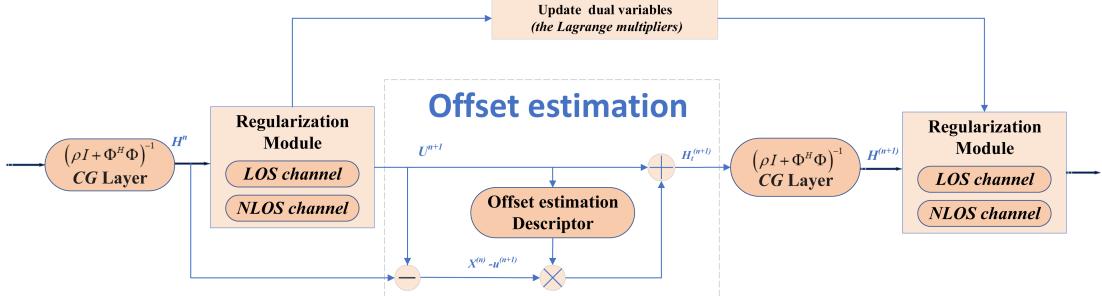


Fig. 3: Overview of the offset learning based indoor channel estimation approach.

$$\mathcal{S}_{PF} \left(c_1^{(n,k)}; \{ p_i, q_i^{(n,k)} \}_{i=1}^{N_c} \right) = \begin{cases} c_1^{(n,k)} + q_i^{(n,k)} - p_1, & c_1^{(n,k)} < p_1 \\ c_1^{(n,k)} + q_{N_c}^{(n,k)} - p_{N_c}, & c_1^{(n,k)} > p_{N_c} \\ q_i^{(n,k)} + \frac{(c_1^{(n,k)} - p_i)(q_i^{(n,k+1)} - q_i^{(n,k)})}{p_{i+1} - p_i}, & p_1 \leq c_1^{(n,k)} \leq p_{N_c}. \end{cases} \quad (19)$$

high computational complexity and difficult to determine λ_a and λ_b in iterative processes. To alleviate this deficiency, the unfolded architecture of the ADMM-based iterative procedure is developed to form the framework of the deep learning network. As shown in Fig. 2, it is observed that each nodes of the graph corresponding to different operations of the ADMM algorithm.

According to the basic idea of ADMM framework, we design a linear offset estimation module to ADMM for channel offset estimation processing. Thus, the solution to the problem in (20) can be written as follows:

$$\left\{ \begin{array}{l} \mathbf{H}^{(k+1)} = \arg \min_{\mathbf{H}} \frac{1}{2} \|\mathbf{Y} - \Phi \mathbf{H}\|_F^2 + \frac{\rho}{2} \|\mathbf{H} - \mathbf{H}_t^{(k)}\|_F^2, \\ \mathbf{H}_{LOS}^{(k+1)} = \arg \min_{\mathbf{H}_{LOS}} \frac{1}{2} \|\mathbf{H}^{(k+1)} - \mathbf{H}_{LOS} - \mathbf{F}_a^{(k)}\|_F^2 + \lambda_a g_a(\mathbf{H}_{LOS}), \\ \mathbf{H}_{NLOS}^{(k+1)} = \arg \min_{\mathbf{H}_{NLOS}} \frac{1}{2} \|\mathbf{H}^{(k+1)} - \mathbf{H}_{NLOS} - \mathbf{F}_b^{(k)}\|_F^2 + \lambda_b g_b(\mathbf{H}_{NLOS}), \\ \mathbf{H}_t^{(k+1)} = \mathbf{U}^{(k+1)} + \mathcal{T}^{(k+1)} \otimes (\mathbf{H}^{(k+1)} - (\mathbf{H}_{LOS}^{(k+1)} + \mathbf{H}_{NLOS}^{(k+1)})), \\ \mathbf{F}_a^{(k+1)} = \mathbf{F}_a^{(k)} + \varrho_a (\mathbf{H}^{(k+1)} - \mathbf{H}_{LOS}^{(k+1)}), \\ \mathbf{F}_b^{(k+1)} = \mathbf{F}_b^{(k)} + \varrho_b (\mathbf{H}^{(k+1)} - \mathbf{H}_{NLOS}^{(k+1)}), \end{array} \right. \quad (21)$$

where \mathbf{F}_a and \mathbf{F}_b are dual variables (the Lagrange multipliers) adding the constraints of (20) to the cost function, \mathcal{T} is

the offset descriptor determined by a factor V (Details are described in Appendix A), $\mathbf{H}_t^{(k)}$ is the k -th iteration of the offset estimation result, \otimes is a Kronecker product operation. It is worth mentioning that the optimization problem (21) commonly needs to run dozens of iterations to obtain a satisfactory estimation result. Furthermore, it is challenging to determine the shrinkage operator for the general regularization functions $g_a(\cdot)$ and $g_b(\cdot)$. To alleviate these difficulties, we first map the iterative procedures in Eq. (21) to design an ADMM-based deep learning framework, which can discriminatively learn all the above transforms, offset descriptor, CG optimization operator. The optimization of the parameters in proposed OLNN-net is expected to improve the precision of channel estimation.

As shown in Fig. 3, the proposed deep learning framework consists of four main modules: CG optimization module, an offset estimation module, regularization module and dual variables module. The descriptor \mathcal{T} in offset estimation module is similar to a special filter, which filters out undesired feature of the residual channel and keeps the desired channel. Finally, the desired feature of the channel can be added back to the estimated channel.

A. Solution via OLNN-based deep learning network

In this section, an OLNN-based model is proposed, which contains the network framework designing and network pa-

parameter learning. For the network framework designing stage, the architecture of OLNN-Net is formed by four successive modules to tackle the problem (21).

To facilitate the description of the proposed deep learning framework, the graph of network architecture is unrolled as the iterative scheme at the k -th stage, is shown in Fig. 4(a). There are four modules mapped from five operations, i.e., CG optimization operation $(\mathbf{H}^{(k)})$, a regularization operation $(\mathbf{Z}_{LOS}^{(k)})$ (the details are shown in Fig. 4(b)), a regularization operation $(\mathbf{Z}_{NLOS}^{(k)})$ (the details are shown in Fig. 4(c)), an offset estimation operation $(\mathbf{R}^{(k)})$ and dual variables operation $(\mathbf{F}^{(k)}(\mathbf{F}_a, \mathbf{F}_b))$. A whole data flow graph is formed as a multiple repetition of the iterations corresponding to above stages, as depicted in Fig. 4. It is observed that the proposed OLNN-Net framework can be regarded as the concatenation of many stages, where each stage includes four network modules. To well-described and analyze this framework, the k -th stage of the proposed OLNN-Net is provided as an example to depict four types of modules as follows.

1) *CG optimization Module ($\mathbf{H}^{(k)}$)*: The input of this module are \mathbf{Y} and $\mathbf{H}_t^{(k-1)}$. The output of this module is defined as

$$\mathbf{H}^{(k)} = \left(\rho^{(k)} \mathbf{I} + \Phi^H \Phi \right)^{-1} \left(\rho^{(k)} \mathbf{H}_t^{(k-1)} + \Phi^H \mathbf{Y} \right). \quad (22)$$

The operator $(\rho^{(k)} \mathbf{I} + \Phi^H \Phi)^{-1}$ is not analytically invertible for complex operators such as IRS-assisted MIMO system. To solve this issue, we exploit the CG optimization scheme to solve (22) that can avoid calculating matrix inversion of the $(\rho^{(k)} \mathbf{I} + \Phi^H \Phi)$, which is desirable for large-scale or ill-conditioned inverse problems. The CG sub-blocks involve several CG steps to enforce the data-consistency constraint, which offers a faster reduction of cost per iteration.

2) *Regularization Module ($\mathbf{Z}_{LOS}^{(k)}$)*: To train a more practicable *LOS* channel operator and improve the network capacity, the CNN-based convolution layers have introduced into the regularization module ($\mathbf{Z}_{LOS}^{(k)}$). As shown in Fig. 4(b), the corresponding module is divided into four layers: a united layer $\mathbf{U}^{(n,k)}$, the convolution layers $\mathbf{C}_1^{(n,k)}$ and $\mathbf{C}_2^{(n,k)}$, and a nonlinear transform layer $\mathbf{S}^{(n,k)}$. The detailed realization procedure of each step is detailed as follows.

a) *United Layer ($\mathbf{U}^{(n,k)}$)*: The inputs of this layer are $\mathbf{U}^{(n,k-1)}$, $\mathbf{H}^{(n)}$ and $\mathbf{C}_2^{(n,k)}$. The corresponding output result is expressed as

$$\mathbf{U}^{(n,k)} = \mu_1^{(n,k)} \mathbf{U}^{(n,k-1)} + \mu_2^{(n,k)} \mathbf{H}^{(n)} - \mathbf{C}_2^{(n,k)}, \quad (23)$$

where the initial value of $\mathbf{U}^{(n,0)}$ is set as $\mathbf{H}^{(n)}$.

b) *Convolution Layer ($\mathbf{C}_1^{(n,k)}$)*: The input of this layer is $\mathbf{U}^{(n,k)}$ and corresponding the output result can be given by

$$\mathbf{C}_1^{(n,k)} = \mathbf{w}_1^{(n,k)} * \mathbf{U}^{(n,k)} + \mathbf{b}_1^{(n,k)}. \quad (24)$$

c) *Nonlinear Transform Layer ($\mathbf{S}^{(n,k)}$)*: $\mathbf{C}_1^{(n,k)}$ is used as the input of this layer, and the corresponding output result can be given by

$$\mathbf{S}^{(n,k)} = \mathcal{S}_{PF} \left(\mathbf{C}_1^{(n,k)}; \left\{ p_i, q_i^{(n,k)} \right\}_{i=1}^{N_c} \right), \quad (25)$$

where a piecewise linear function $\mathcal{S}_{PF}(\cdot)$ is determined by a fixed point set $\left\{ p_i, q_i^{(n,k)} \right\}_{i=1}^{N_c}$. The corresponding result can be expressed in (19).

d) *Convolution Layer ($\mathbf{C}_2^{(n,k)}$)*: By calculating (25), the corresponding result $\mathbf{S}^{(n,k)}$ can be used as the input of this layer, and the output of this layer is expressed as

$$\mathbf{C}_2^{(n,k)} = \mathbf{w}_2^{(n,k)} * \mathbf{S}^{(n,k)} + \mathbf{b}_2^{(n,k)}. \quad (26)$$

3) *Regularization Module ($\mathbf{Z}_{NLOS}^{(k)}$)*: Since the signal power of the *NLOS* channels are generally weaker than *LOS* channel, the deconvolution module is used to learn the parameters by exploiting the learnable Wiener Filtering layer [41]. In case of multiple regularization kernels, we formulate the *NLOS* channels objective function as:

$$\mathbf{H}_{NLOS}^{(k+1)} = \arg \min_{\mathbf{H}_{NLOS}} \frac{1}{2} \left\| \mathbf{H}^{(k+1)} - \mathbf{H}_{NLOS} - \mathbf{F}_b^{(k)} \right\|_F^2 + \lambda_b \|\mathcal{G} \mathbf{H}_{NLOS}\|_2^2, \quad (27)$$

where \mathcal{G} denotes the multiple regularization kernel that plays the role of multiple regularizer filters.

Inspired by [41], the optimization problem (27) can be solved by Wiener deconvolution technique. To learn a more suitable regularization kernel and parameters in (27), the Wiener filtering layer is developed as follows

$$\mathbf{H}_{NLOS}^{(k+1)} = \mathcal{F}(\mathbf{H}, \mathbf{F}, \mathcal{G}, \sigma, \vartheta), \quad (28)$$

where ϑ denotes the weights of trainable regularization kernels by gradient descent scheme in the Wiener deconvolution network. The training processing for the proposed deconvolution module is shown in Fig. 5.

4) *Offset Estimation Module ($\mathbf{R}^{(n)}$)*: The inputs of this module have two sets $\mathbf{U}^{(n)}$, $\mathbf{H}^{(n)}$, and the corresponding output result is given by

$$\mathbf{H}_t^{(n)} = \mathbf{U}^{(n)} + \mathcal{T}^{(n)} \otimes \left(\mathbf{H}^{(n)} - (\mathbf{H}_{LOS}^{(n)} + \mathbf{H}_{NLOS}^{(n)}) \right), \quad (29)$$

where $\mathcal{T}^{(n)}$ denotes offset operator at the n -th of stage.

5) *Dual Variables Operation Module ($\mathbf{F}^{(n)}$)*: This module aims to update the variables \mathbf{F}_a and \mathbf{F}_b . The outputs of this layer are

$$\mathbf{F}_a^{(k+1)} = \mathbf{F}_a^{(k)} + \varrho_a (\mathbf{H}^{(k)} - \mathbf{H}_{LOS}^{(k)}) \quad (30)$$

and

$$\mathbf{F}_b^{(k+1)} = \mathbf{F}_b^{(k)} + \varrho_b (\mathbf{H}^{(k)} - \mathbf{U}_{NLOS}^{(k)}). \quad (31)$$

B. OL-based Neural Network Training

In the training procedure, our task aims to learn the parameters of each module Ξ as follows: $\rho^{(n)}$ in the CG module; offset estimation module $\mathbf{w}_1^{(n,k)}$, $\mathbf{w}_2^{(n,k)}$ and $V^{(n)}$ in offset operator $T^{(n)}$; biases $b_1^{(n,k)}$, $b_2^{(n,k)}$ in the regularization module with respect to $\mu_1^{(n,k)}$, $\mu_2^{(n,k)}$, ϱ_a , ϱ_b in variables operation module and $\left\{ q_i^{(n,k)} \right\}_{i=1}^{N_c}$. Based on the parameters Ξ , the received signal \mathbf{Y} is generated as the input through a specially designed the parameters Ξ . By updating the parameters Ξ , the training procedure is devoted to learning the final output result $\mathbf{H}(\mathbf{Y}, \Xi)$

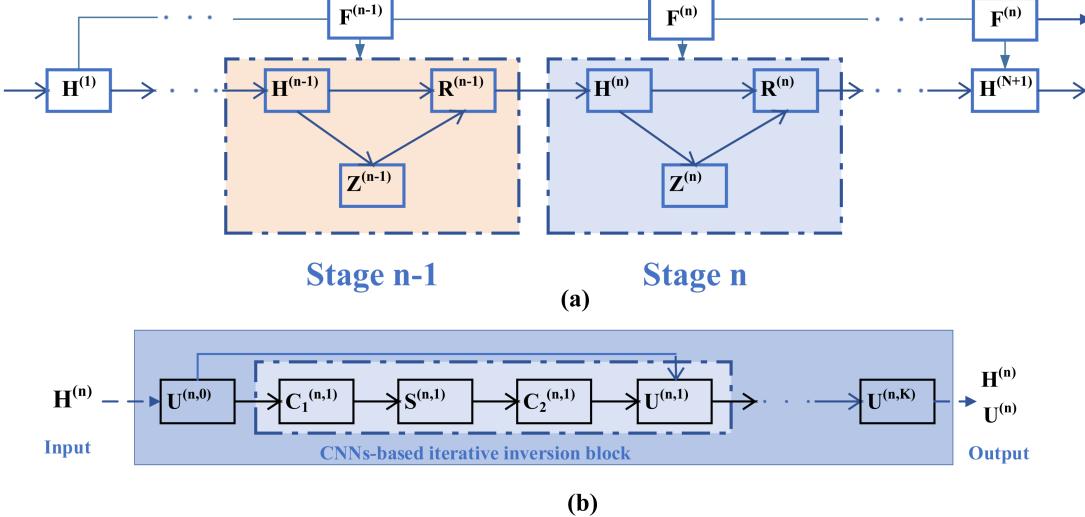


Fig. 4: Framework of the OLNN-Net channel estimation approach.

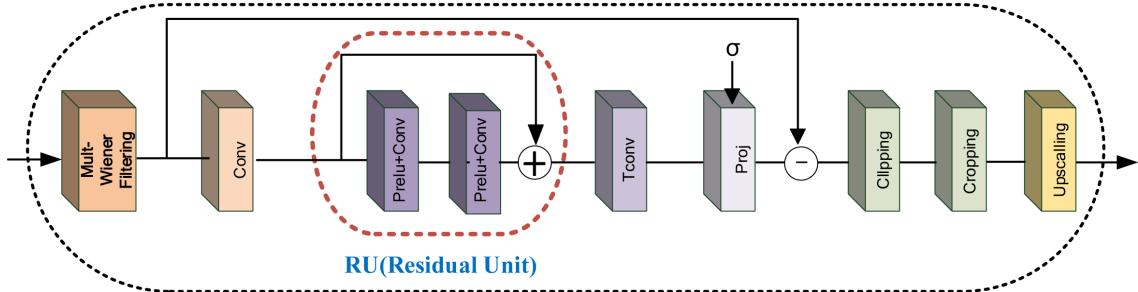


Fig. 5: Framework of the deconvolution module ($Z_{\text{NLOS}}^{(k)}$).

as close as possible to the true channel \mathbf{H}_0 . For the given pairs of training data $\psi = \{\mathbf{Y}, \mathbf{H}_0\}$ and Ξ , the NMSE is used as the loss function, which can be rewritten as

$$E(\Xi) = \frac{1}{|\psi|} \sum_{(\mathbf{Y}, \mathbf{H}_0) \in \psi} \frac{\sqrt{\|\hat{\mathbf{H}}(\mathbf{Y}, \Xi) - \mathbf{H}_0\|_2^2}}{\sqrt{\|\mathbf{H}_0\|_2^2}}, \quad (32)$$

where $|\psi|$ denotes the number of channel pairs in ψ . To achieve the optimized parameters, the gradients of the loss function $E(\Xi)$ with regard to parameters Ξ can be obtained by using the back-propagation (BP) algorithm.

Note that the BP scheme requires the evaluation of the terms $\nabla E_{\mathbf{H}^{(k)}}(\Xi)$, $k = 1, \dots, |\psi| - 1$. Thus, we focus on how to evaluate these terms by exploiting the CG blocks. We have

$$\nabla E_{\mathbf{H}^{(k)}}(\Xi) = \mathbf{J}_{\mathbf{H}^{(k)}} \left(\mathbf{H}_t^{(k+1)} \right)^T \nabla E_{\mathbf{H}_t^{(k+1)}}(\Xi), \quad (33)$$

where $\mathbf{H}^{(k)}$ is the output of the network at the k -th iteration. Let $\mathcal{C} = (\rho \mathbf{I} + \Phi^H \Phi)$, the (22) can be rewritten as $\mathbf{H}_t^{(k+1)} = \mathbf{C} \mathbf{H}^{(k)} - \Phi^H \mathbf{Y}$. And the Jacobian matrix $\mathbf{J}_{\mathbf{H}^{(k)}}(\mathbf{H}_t^{(k+1)})$ is expressed as

$$\mathbf{J}_{\mathbf{H}^{(k)}} \left(\mathbf{H}_t^{(k+1)} \right) = (\rho \mathbf{I} + \Phi^H \Phi)^{-1}. \quad (34)$$

It follows that the above expression is evaluated using the CG scheme until convergence. Since \mathcal{C} is symmetric, the gradients can be back-propagated as

$$\nabla E_{\mathbf{H}_t^{(k)}}(\Xi) = \mathbf{J}_{\mathbf{H}_t^{(k)}}(\mathbf{H}^{(k+1)})^T \nabla E_{\mathbf{H}^{(k+1)}}(\Xi). \quad (35)$$

Therefore, the network parameters and regularization parameter ρ are updated with the BP scheme. For regularization module and offset estimation module, the parameters of deep OLNN-Net are optimized using gradient-based algorithm.

C. Network Structure Analysis

The proposed network is non-conventional deep architecture with layers inspired by the ADMM algorithm for channel estimation. Compared with the conventional ADMM-CSNet and its variants, our network architecture have the following similar and distinct characteristics.

First, for the IRS-assisted indoor channel model, it often includes LOS path and NLOS pathes, which are difficult to optimize simultaneously using the existing ADMM-CSNet. In particular, the characteristics of the NLOS channel are different from the LOS channel, which leads to an inescapable performance loss by using the existing ADMM-CSNet. Sparked by above issue, we build a novel deep learning optimization framework, which named OLNN-Net. In the designed optimization, the deconvolution module is developed

to estimate the NLOS channels by exploiting the learnable Wiener Filtering layer, while the LOS channel is trained by using the CNN-based network. Different from the ADMM-CSNet in [29], both NLOS and LOS channels in this network are separately trained. Specifically, we first construct an unified residual CNN-based network for pilot design, and different noise sigma are utilized to improve a practical CNN-based channel estimation for real applications. In addition, compared to ADMM-CSNet, we also develop offset estimation module to enhance the network capability. With the fine structure preserving ability of the offset estimation module, the proposed OLNN-Net can also be used for small dataset training, while adding trainable CNN blocks. The architecture of the proposed OLNN-Net has the ability to learn the characteristics of IRS-assisted channels that may suffer from nonlinear interference and distortion.

Second, the conjugate gradient (CG) optimization scheme is developed to solve the $(\rho^{(k)}\mathbf{I} + \Phi^H\Phi)^{-1}$ in (22) that can avoid calculating matrix inversion of the $(\rho^{(k)}\mathbf{I} + \Phi^H\Phi)$, which is desirable for large-scale or ill-conditioned inverse problems. Different from ADMM-CSNet, proximal gradient algorithm is often adopted to solve the (22), which alternates between steepest descent scheme and CNN-based blocks. It is observed that such algorithm may result in a large number of iterations/unfolding steps, since the trained network with the proximal gradient scheme is limited, especially for the GPUs with limited on-board memory. In addition, the intermediate results and layer need to be stored to perform backpropagation, which limits the performance of the ADMM-based network during training. By contrast, the CG sub-blocks involve several CG optimization steps to enforce the data-consistency constraint, which offers a faster reduction of cost per iteration. It is worth noting that there are no extra parameters generating within intermediate results and need not be stored to perform backpropagation. This implies that a large number of CG optimization steps can be performed almost without memory overhead during training processing.

TABLE I: The proposed OLNN-Net framework parameters for simulation

Parameters	Estimator
Network architecture	OLNN-Net
Window length	40 symbols
Activation function	Rule for hidden layers
Loss function	mean square error (MSE)
Learning rate	0.001
Batch size	128
Train SNR	0, 5, 10, 15, 20, 25 dB
Train SNR	0, 5, 10, 15, 20, 25, 30 dB
Train number	100000
Validation number	10000
Test number	10000

IV. SIMULATION RESULTS

In this section, Monte Carlo simulations are provided to demonstrate the performance of the IRS-assisted deep channel estimator under the indoor communication scenario. The default parameters of the proposed OLNN-Net is firstly described. Then, the IRS-assisted deep channel estimator is investigated and verified.

A. Simulation Setting

We consider a typical IRS-assisted indoor communication, in which the AP is equipped with $N_t = 36$ antennas and the IRS is equipped with 36 reflecting elements. Two operating frequencies, 28 and 73 GHz, are adapted to verify the feasibility and validity of the proposed OLNN-Net. Meanwhile, all reported simulation results are obtained from averaging 20000 independent Monte Carlo realizations. To adequately collect the training samples, different direction of unit signals are generated from the transmitter to the corresponding received signals. Based on this, the training dataset comprises 150000 examples, while 20000 samples is considered as a validation dataset for each signal to noise ratio (SNR) level. Following the rules in [35], the IRS element gain $G_e(0)$ is calculated from $G_e(0) = 4\pi A_e(0)/\lambda^2$, where the physical area of an IRS element is expressed as $A_e(0) = (\lambda/2)^2$.

In the simulations, the NMSE and achievable spectral efficiency (ASE) performance are adopted to provide quantitative evaluation of the considered CSI estimation techniques. The NMSE is defined as follows:

$$\text{NMSE} \triangleq \mathbb{E} \left\{ 10 \log_{10} \frac{\|\hat{\mathbf{H}} - \mathbf{H}\|_F^2}{\|\mathbf{H}\|_F^2} \right\}, \quad (36)$$

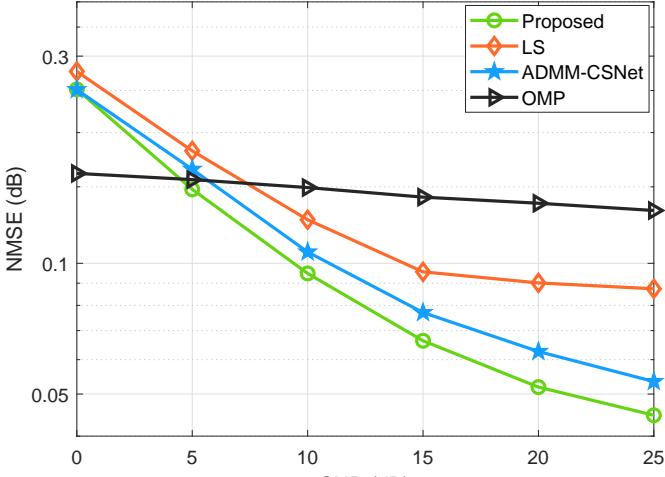
where $\hat{\mathbf{H}}$ denotes the estimation channel and \mathbf{H} is the true channel. In addition, the ASE is adapted as follows

$$\text{ASE} \triangleq \mathbb{E} \left\{ \log_2 \det \left(\mathbf{I} + (N_t N_r (\sigma^2 + \text{NMSE}))^{-1} \mathbf{H} \mathbf{H}^H \right) \right\}, \quad (37)$$

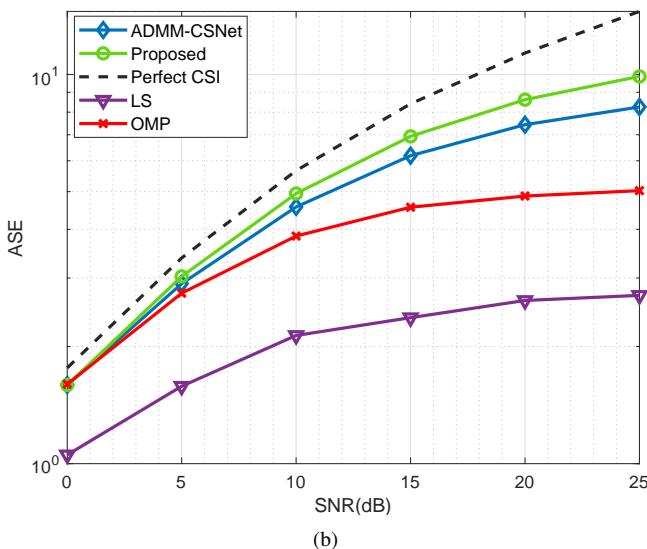
where N_t and N_r denote the number of transmit and receive antennas, respectively.

According to the [36], the LOS probabilities of Eq. (11) is provided based on various indoor office environments. When the heights of UEs are between 1 and 1.5m, the channel gains decays very fast with transmission distance increasing. For the channel link between the AP and the IRS, the IRS is located below the level of AP ($z_{IRS} < z_{AP}$) regardless of the distance $d_{c,s}$. Therefore, when the height of the IRS is no smaller than that of the AP, a clear LOS path is considered in this simulation, which is a reasonable assumption for arbitrary distance $d_{c,s}$.

For the first simulation, the performance of proposed OLNN-Net estimator is evaluated for two channel models, namely the LS estimator in [7] and OMP [42]. Fig. 6 depicts the impact of the ASE/NMSE of all considered estimation techniques in the IRS-assisted indoor system. We investigate the NMSE of the proposed OLNN-Net scheme against the SNR levels, where the training sequences $L = 120$. From Fig. 6, it is also evident that the proposed OLNN-Net generally achieves better performance than the OMP and LS channel



(a)



(b)

Fig. 6: (a) NMSE performance of LS, OMP and OLNN-Net estimators; (b) ASE performance of LS, OMP and OLNN-Net estimators at the 73Hz.

estimators. In particular, the proposed OLNN-Net scheme substantially improves the estimation performance in terms of the NMSE value compared with the LS, which benefits from the offset learning-based module that can learn the channel error in indoor scenarios. We conclude that the proposed OLNN-Net scheme is an efficient channel estimation approach among all the compared methods. Additionally, it is further observed that the proposed OLNN-Net outperforms that of ADMM-CSNet in terms of NMSE, which benefits from the proposed OLNN-Net that can realize LOS and NLOS channel estimation separately. This favorable result indicating that the characteristics of channel are different, which is necessary to estimate separately.

In the next simulation, the NMSE performance against training epochs is investigated in Fig. 7. It is observed that the NMSE performance is enhanced as the training epochs increases, illustrating that the proposed OLNN-Net can obtain a near-optimal performance. In addition, we can further ob-

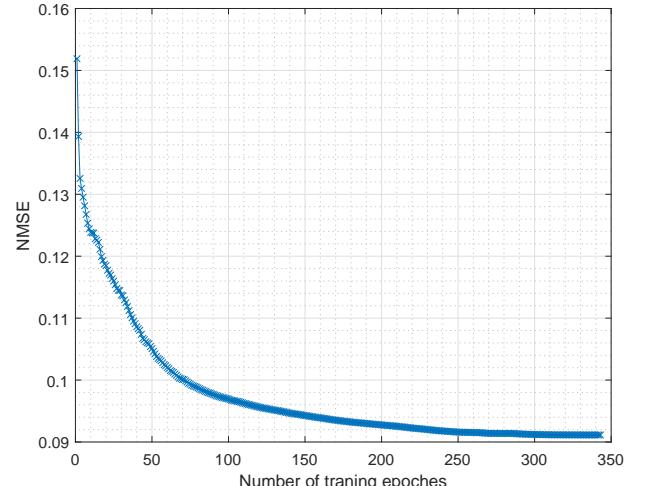


Fig. 7: MSE performance against training epochs.

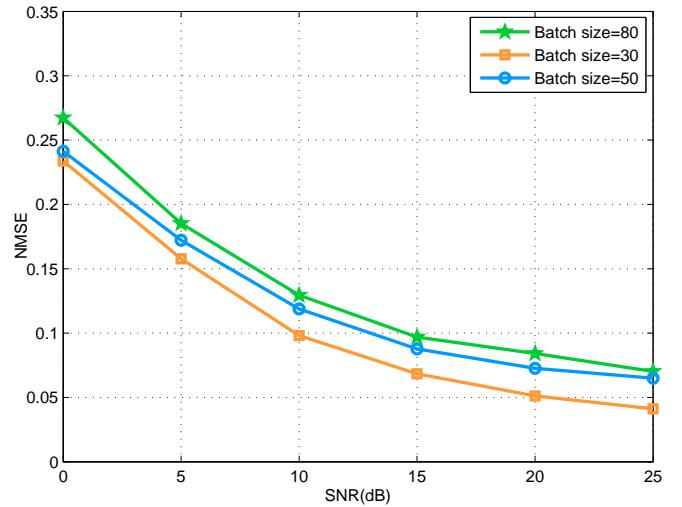
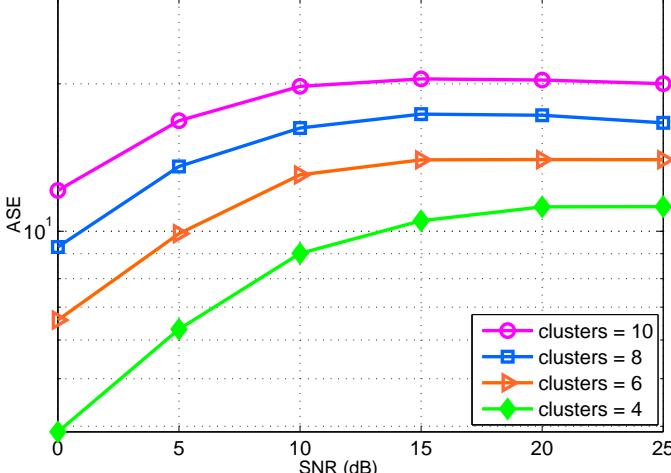


Fig. 8: NMSE performance against the patch size.

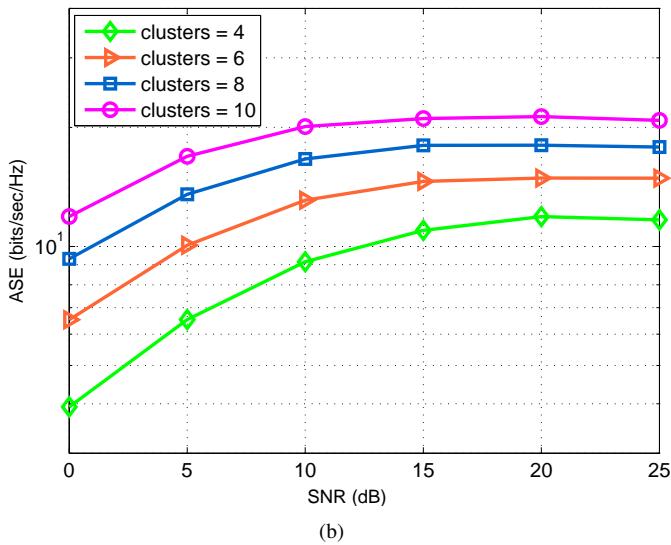
serve from Fig. 7 that the NMSE performance becomes stable when the training epochs over 300, which demonstrates the convergence behavior of the proposed OLNN-Net for the IRS-assisted channel estimation.

Fig. 8 provides a NMSE performance comparison of the proposed OLNN-Net with different batch size, where the batch size are set as 30, 50 and 80. It can be observed from Fig. 8 that the value of NMSE decreases as SNR levels increasing. This result is attributed to the fact that a smaller batch size leads to a slower convergence speed, while a large batch size for the training procedure will cause very large epochs. Thus, there exists a non-trivial tradeoff between the the accuracy of CSI and the stability of the OLNN-Net when optimizing the batch size to improve the network performance.

To show the effect of number of clusters on the IRS-assist channel estimation, we also conduct the achievable rates of different schemes versus the number of channel clusters in Fig. 9. It is observed that there exists a tradeoff between the the precision of IRS-assist channel estimation and training overhead. This is because a small number of clusters M is



(a)

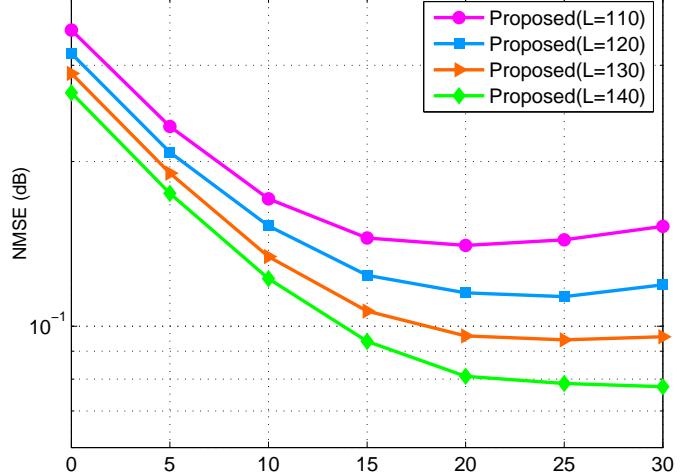


(b)

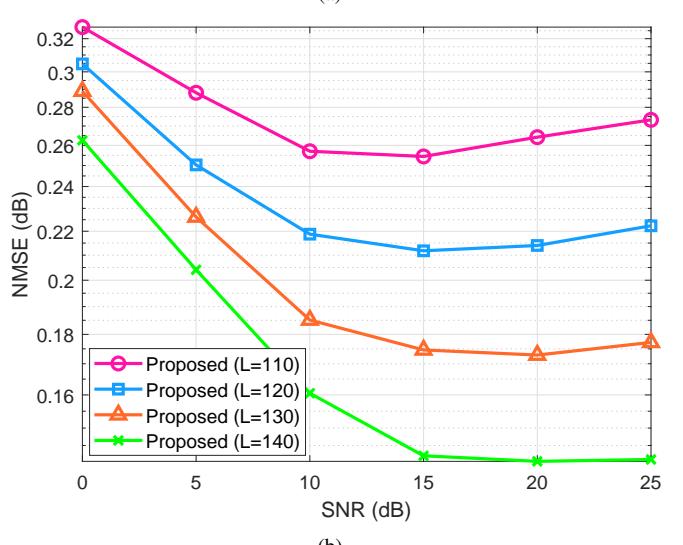
Fig. 9: (a) ASE performance against the number of clusters at 73 GHz; (b) ASE performance against the number of clusters M at 28 GHz.

not sufficiently accurate for achieving high gain of IRS-based channel, while too many clusters results in computationally expensive in channel estimation process. Moreover, it is observed that the proposed channel estimator with the operating frequencies at 73Hz are slightly better than the 28Hz, at the same number of clusters. It is worth noting that the proposed scheme with OLNN-Net framework can effectively escape computing both system memory and time cost.

In the next simulation, the NMSE performance of the proposed OLNN-Net estimator for different amount of pilots L is depicted in Fig. 10. In the same way, two operating frequencies, 28 and 73 GHz, are adapted to observe the impact of an increased number of pilots by varying $L = 110, 120, 130$ and 140 under the different SNR values. Specifically, from Fig. 10, it is clearly seen that the NMSEs of all schemes are enhanced when employing longer pilot sequences. The main reason for this is because the training time is proportional to the length of the pilot sequences. However, when pilot



(a)



(b)

Fig. 10: (a) NMSE of the proposed estimator for different amount of pilots L at 73 GHz; (b) NMSE of the proposed estimator for different amount of pilots L at 28 GHz

sequences L is sufficiently large, the time available will decrease, which leads to a smaller capacity. As shown in Fig. 10, when the pilots $L > 120$, the NMSE value of the proposed OLNN-Net is less than 0.2, which means that the OLNN-Net can be trained without taking longer the pilot sequence.

To further quantify the pilot contamination performance of our estimator, we conduct the NMSE performance of the OLNN-Net in different number of the RF chains N . According to Fig. 11, the NMSE values of the channel estimation are reducing with the number of pilots increasing, and then it becomes stable gradually until the number of pilots are sufficiently large. This indicates that the precision of the channel estimation can be improved when providing longer pilots sequences. In addition, one can observe significant improvement of the proposed OLNN-Net by increasing the number of RF chains N . Moreover, from Fig. 11(a) and Fig. 11(b), it is observed that the NMSE performance of

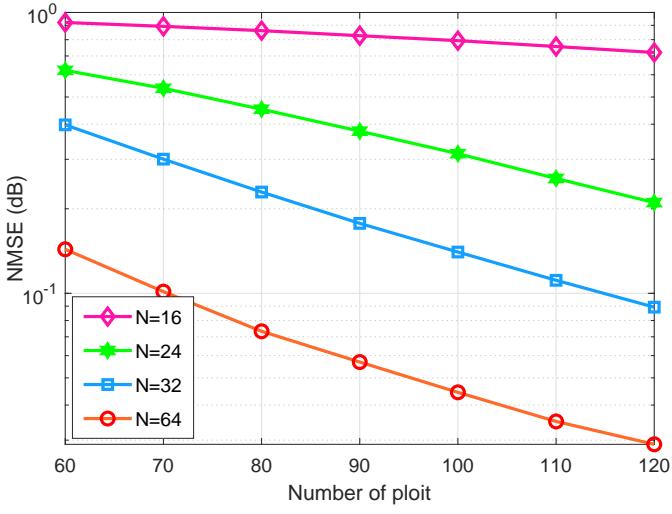


Fig. 11: (a) NMSE performance against the number of pilot at 73GHz; (b) NMSE performance against the number of pilot at 28GHz.

channel estimation can be reduced with the number of the number of RF chains N in. It is interesting to note that the different operating frequencies also influence the performance of channel estimation.

Finally, we investigate the validation error against the number of iteration in Fig. 12. As the depth of OLNN-Net increasing, trained accuracy gets saturated (which might be unsurprising) and then could degrades. Thus, the validation error curve for the model selection should be considered. Fig. 12 shows that the 20-layer OLNN-Net performs better than the 38-layer one (by 1.48%). More importantly, the 20-layer OLNN-Net exhibits considerably lower training error. This indicates that the degradation problem is well addressed in this setting. Therefore, the 20-layer OLNN-Net is adopted as the constructed solution of the proposed model.

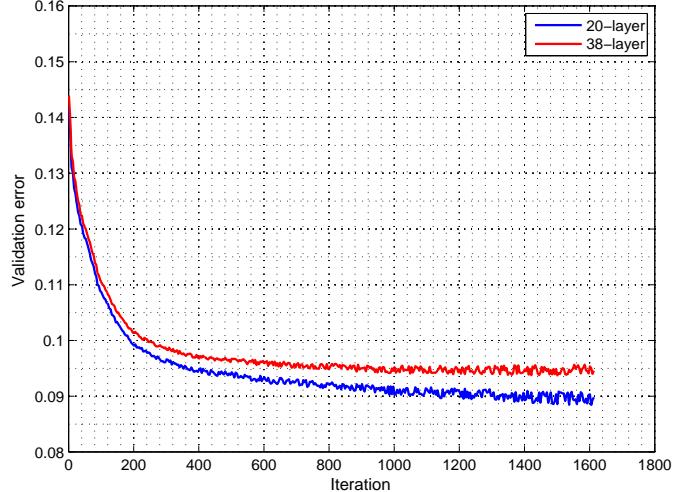


Fig. 12: Validation error against the number of iteration.

V. CONCLUSION

In this paper, an OL-based neural network is proposed to estimate CSI of an IRS-assisted mmWave MIMO system, which is a challenging problem for B5G indoor deployment. To improve the performance in an indoor environment, the offset learning-based module is integrated into the proposed OLNN-Net framework. The proposed network architecture is derived from ADMM, which unrolls the iterative procedures to a supervised model-driven network. Theoretical and simulation analysis demonstrate that in the constructed indoor 5G system, the proposed channel estimator exhibits superior performance compared to LS and OMP estimators without any prior knowledge of the IRS-assisted channel and pilot contamination. With the same pilot overhead, the proposed OLNN-Net provides improved results compared with existing algorithms, which is beneficial for improving the spectrum utilization of 5G system.

In the future work, the proposed OLNN-Net is planned to be implemented in an universal software radio peripheral (USRP)-based hardware platform. Furthermore, as a general framework, the proposed OLNN-Net structure will be enhanced by further employing the channel correlation in time, frequency, and spatial domain.

APPENDIX A

Inspired by [19], a Gaussian filter is used to obtain the feature descriptor $T(u)$ by employing the Gaussian filter on u . Assume that p and q are two local patches extracted from u and u_b respectively, the feature descriptor \mathcal{T} can be represented as

$$\mathcal{T}(u) = 1 - |c(p, q)s(p, q)| = 1 - \left| \frac{2\sigma_{pq} + V}{\sigma_p^2 + \sigma_q^2 + V} \right| \quad (38)$$

where $c(p, q)$ is the degraded operation that aims to obtain the reduction of contrast variation, and $s(p, q)$ denotes the structural correction of local patches between the p and q . The other variables are given by $\sigma_p = (\sum_{i=1}^{N_u} (p_i - \mu_p)^2)^{1/2}$, $\sigma_q = (\sum_{i=1}^{N_u} (q_i - \mu_q)^2)^{1/2}$ and $\sigma_{pq} = \sum_{i=1}^{N_u} (p_i - \mu_p)(q_i - \mu_q)^2$, where $\mu_p = \sum_{i=1}^{N_u} p_i$, $\mu_q = \sum_{i=1}^{N_u} q_i$ and N_u is the pixel number of u . Since the constant V has a great influence on

the estimation result, the descriptor \mathcal{T} is tuned by using the proposed OLNN-Net to train the constant V .

REFERENCES

- [1] B. Li, D. Zhu, and P. Liang, "Small cell in-band wireless backhaul in massive MIMO systems: A cooperation of next-generation techniques," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 7057–7069, 2015.
- [2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.
- [3] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, 2020.
- [4] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.
- [5] Z. Q. He, H. Liu, X. Yuan, Y. J. A. Zhang, and Y. C. Liang, "Semi-blind cascaded channel estimation for reconfigurable intelligent surface aided massive MIMO." [Online]. Available: <https://arxiv.org/abs/2101.07315>. dec., 2021.
- [6] Z. Chen, J. Tang, X. Y. Zhang, D. K. C. So, S. Jin, and K.-K. Wong, "Hybrid evolutionary-based sparse channel estimation for IRS-assisted mmwave MIMO systems," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2021.
- [7] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4659–4663, 2019.
- [8] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Processing Letters*, vol. 27, pp. 905–909, 2020.
- [9] Z. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive MIMO," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 210–214, 2020.
- [10] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Communications Letters*, vol. 9, no. 4, pp. 518–522, 2020.
- [11] L. Wei, C. Huang, G. C. Alexandropoulos, and C. Yuen, "Parallel factor decomposition channel estimation in RIS-assisted multi-user MISO communication," in *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 1–5, 2020.
- [12] X. Guan, Q. Wu, and R. Zhang, "Anchor-assisted intelligent reflecting surface channel estimation for multiuser communications," in *GLOBECOM 2020 IEEE Global Communications Conference*, pp. 1–6, 2020.
- [13] N. J. Myers and R. W. Heath, "Message passing-based joint CFO and channel estimation in mmwave systems with one-bit ADCs," *IEEE Transactions on Wireless Communications*, vol. 18, no. 6, pp. 3064–3077, 2019.
- [14] X. Chen, J. Shi, Z. Yang, and L. Wu, "Low-complexity channel estimation for intelligent reflecting surface-enhanced massive MIMO," *IEEE Wireless Communications Letters*, (Early Access) 2021.
- [15] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5092–5106, 2020.
- [16] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8549–8560, 2018.
- [17] Y. Yang, F. Gao, X. Ma, and S. Zhang, "Deep learning-based channel estimation for doubly selective fading channels," *IEEE Access*, vol. 7, pp. 36579–36589, 2019.
- [18] H. Hirose, T. Ohtsuki, and G. Gui, "Deep learning-based channel estimation for massive MIMO systems with pilot contamination," *IEEE Open Journal of Vehicular Technology*, vol. 2, pp. 67–77, 2021.
- [19] Y. Liu, Q. Liu, M. Zhang, Q. Yang, S. Wang, and D. Liang, "IFR-Net: Iterative feature refinement network for compressed sensing MRI," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 434–446, 2020.
- [20] X. Ma and Z. Gao, "Data-driven deep learning to design pilot and channel estimator for massive MIMO," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5677–5682, 2020.
- [21] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, "Deep learning-based channel estimation," *IEEE Communications Letters*, vol. 23, no. 4, pp. 652–655, 2019.
- [22] H. Ye, G. Y. Li, and B. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114–117, 2018.
- [23] H. He, C. Wen, S. Jin, and G. Y. Li, "Deep learning-based channel estimation for beamspace mmwave massive MIMO systems," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 852–855, 2018.
- [24] E. Balevi, A. Doshi, and J. G. Andrews, "Massive MIMO channel estimation with an untrained deep neural network," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2079–2090, 2020.
- [25] Q. Bai, J. Wang, Y. Zhang, and J. Song, "Deep learning-based channel estimation algorithm over time selective fading channels," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 125–134, 2020.
- [26] L. Li, H. Chen, H. Chang, and L. Liu, "Deep residual learning meets OFDM channel estimation," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 615–618, 2020.
- [27] J. Xiang, Y. Dong, and Y. Yang, "FISTA-Net: Learning a fast iterative shrinkage thresholding network for inverse problems in imaging," *IEEE Transactions on Medical Imaging*, vol. 40, no. 5, pp. 1329–1339, 2021.
- [28] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-net for compressive sensing MRI," in *Advances in Neural Information Processing Systems* (D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, eds.), vol. 29, pp. 1–9, Curran Associates, Inc., 2016.
- [29] Y. Yang, J. Sun, H. Li, and Z. Xu, "ADMM-CSNet: A deep learning approach for image compressive sensing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 3, pp. 521–538, 2020.
- [30] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 18–44, 2021.
- [31] S. Ghassemzadeh, R. Jana, C. Rice, W. Turin, and V. Tarokh, "Measurement and modeling of an ultra-wide bandwidth indoor channel," *IEEE Transactions on Communications*, vol. 52, no. 10, pp. 1786–1796, 2004.
- [32] "3GPP TR 38.901 V16.1.0 - study on channel model for frequencies from 0.5 to 100 ghz," 2019.
- [33] I. A. Hemadet, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 2, pp. 870–913, 2018.
- [34] P. Nayeri, F. Yang, and A. Z. Elsherbeni, *Reflectarray Antennas: Theory, Designs, and Applications*. 2018.
- [35] S. W. Ellingson, "Path Loss in Reconfigurable Intelligent Surface-Enabled Channels." [Online]. Available: <https://arxiv.org/pdf/1912.06759.pdf>. dec., 2019.
- [36] "5G channel model for bands up to 100 GHz." [Online]. Available: http://www.5gworkshops.com/5GCMSIG_White\%20Paper_r2dot3.pdf. Dec., 2016.
- [37] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3029–3056, 2015.
- [38] M. Alonso, S. Buzzi, A. Zappone, and C. DElia, "Energy-efficient power control in cell-free and user-centric massive MIMO at millimeter wave," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 3, pp. 651–663, 2019.
- [39] Z. Gao, L. Dai, S. Han, C. I, Z. Wang, and L. Hanzo, "Compressive sensing techniques for next-generation wireless communications," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 144–153, 2018.
- [40] Z. Chen, C. Huang, and S. Lin, "A new sparse representation framework for compressed sensing MRI," *Knowledge-Based Systems*, vol. 188, p. 104969, 2020.
- [41] R. M. Umer, G. L. Foresti, and C. Micheloni, "Deep super-resolution network for single image super-resolution with realistic degradations," *the 13th International Conference*, pp. 1–10, 2019.
- [42] K. Venugopal, A. Alkhateeb, N. Gonzlez Prelic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1996–2009, 2017.