Near-Field Channel Estimation for Extremely Large-Scale Array Communications: A model-based deep learning approach

Xiangyu Zhang, *Student Member, IEEE,* Zening Wang, Haiyang Zhang, *Member, IEEE,* and Luxi Yang, *Senior Member, IEEE*

***Abstract*—Extremely large-scale massive MIMO (XL-MIMO) has been reviewed as a promising technology for future wireless communications. The deployment of XL-MIMO, especially at high-frequency bands, leads to users being located in the near-field region instead of the conventional far-field. This letter proposes efficient model-based deep learning algorithms for estimating the near-field wireless channel of XL-MIMO communications. In particular, we first formulate the XL-MIMO near-field channel estimation task as a compressed sensing problem using the spatial gridding-based sparsifying dictionary and then solve the resulting problem by applying the Learning Iterative Shrinkage and Thresholding Algorithm (LISTA). Due to the near-field characteristic, the spatial gridding-based sparsifying dictionary may result in low channel estimation accuracy and a heavy computational burden. To address this issue, we propose a novel sparsifying dictionary learning-LISTA (SDL-LISTA) algorithm that formulates the sparsifying dictionary as a neural network layer and embeds it into LISTA neural network. The numerical results show that our proposed algorithms outperform non-learning benchmark schemes, and SDL-LISTA achieves better performance than LISTA with ten times atoms reduction.**

***Index Terms*—Near-field channel estimation, Extremely large MIMO, Model-based deep learning.**

I. Introduction

Extremely large-scale MIMO (XL-MIMO) has been viewed as a promising physical technology for future sixth-generation (6G) communications [1]. The main difference between XL-MIMO systems and the traditional multiple-antenna communication system is that the user and scatterers likely locate in the near field region of the XL-MIMO transmitter, where the electronic waves follow the spherical wave model instead of the far-field planner model [2]. This shift brings new opportunities to benefit wireless communications, such as the beam-focusing effect, which enables improving the transmission rate of near-field multiuser communications [3], as well as enhancing the energy transfer efficiency of near-field wireless power transfer systems [4].

The benefit of near-field beam focusing relies on accurate channel estimation, which is an important and novel task. Due to the new channel model, the near-field channel estimation needs to utilize the sparse in both angle and distance dimensions, which is totally different from the traditional far-field channel estimation that utilizes the sparse in the angle dimension. So far, very limited literature has investigated the near-field channel estimation problem [5], [6]. In [5], a hybrid near and far field channel is formulated and represented in the polar domain. As the channel can be treated as a spatial sparse signal, the authors in [5] apply the orthogonal matching pursuit (OMP) to estimate the hybrid-field channel parameters. In [6], the authors consider a non-stationary near-field channel model and estimate the near-field channel with the same algorithm as [5].

The existing XL-MIMO channel estimation methods are the same as the algorithms for the conventional far-field hybrid antenna array, i.e., introducing the channel sparse representation method and applying the compressed sensing (CS) algorithms such as OMP to estimate channel parameters [7]. The common sparse representation method is spatial gridding, which partitions the spacing into the grid and assumes the signal incidents from these grids. In general, due to the limited granularity of grids, this assumption limits the upper boundary of channel estimation accuracy [8]. Besides, this sparse representation method suffers from a huge dictionary due to the near-field space needing to be partitioned in both the angle axis and the distance axis, resulting in high computational complexity. Moreover, the sensing matrix utilizing the spatial gridding dictionary has high coherence because the array response vector in the distance axis shows a high correlation. In this case, the recovery accuracy of CS algorithms will be unavoidably degraded [9].

In this letter, we investigate the near-field channel estimation problem for multi-user XL-MIMO communication systems. The XL-MIMO base station (BS) estimates the wireless channels between itself and multiple signal-antenna users using the collected orthogonal pilot signals sent by users. We consider communication systems operating in high-frequency bands. Thus, the wireless channel can be assumed to be sparse. For such a scenario, we propose model-based deep learning approaches to solve the channel estimation problem with high accuracy and efficiency. In particular, the main contributions

Manuscript received November 30, 2022; revised January 19, 2023; accepted February 13, 2023. This work was supported by the National Natural Science Foundation of China under Grants U1936201 and 61971128, and the National Key Research and Development Program of China under Grant 2020YFB1804901. *(Corresponding authors: L. Yang and H. Zhang.)*

X. Zhang, and L. Yang are with the School of Information Science and Engineering, and the National Mobile Communications Research Laboratory, Southeast University, and also with the Pervasive Communications Center, Purple Mountain Laboratories, Nanjing, China (e-mail: xyzhang@seu.edu.cn; lxyang@seu.edu.cn); Z. Wang is with the AsiaInfo Technologies Limited, Beijing, China (e-mail:zeningw2022@126.com); H. Zhang is with the School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China. (e-mail: haiyang.zhang@njupt.edu.cn).

of this letter are summarized as follows:

- We first formulate the near-field channel estimation as a CS problem using the spatial gridding-based sparsifying dictionary, which allows for estimating the sparse channel from a small number of observed pilot signals. Then, we apply the model-based deep learning approach, Learning Iterative Shrinkage and Threshold Algorithm (LISTA), to solve this sparse parameter estimation problem efficiently.
- Next, we prove that the LISTA is limited to achieving high channel estimation accuracy and suffers high computational complexity due to the spatial gridding dictionary. In response, we propose a new and dedicated deep unrolling algorithm Sparsifying Dictionary Learning-LISTA (SDL-LISTA), to optimize the sparsifying dictionary. SDL-LISTA unfolds the sparse representation process as a neural network layer and embeds it inside each layer of LISTA. Thus, a sparsifying dictionary can be optimized at the training stage of LISTA, which exploits the physical property of near-field channels.
- We provide numerical examples to validate the effectiveness of our proposed algorithms. In particular, we show that our proposed algorithms outperform non-learning benchmark schemes, including OMP and Fast ISTA (FISTA), and SDL-LISTA achieves better performance than LISTA with ten times atoms reduction.

The rest of the letter is organized as follows. Section II introduces the considered XL-MIMO systems and the formulated near-field channel estimation problem. In Section III, we present our algorithms, i.e., LISTA and SDL-LISTA. Simulation results and conclusions are provided in Section IV and Section V, respectively.

II. System Model

As shown in Fig. 1, we consider a multi-user XL-MIMO communication system consisting of an XL-array BS and M single-antenna users, operating at high-frequency bands. The users are located in the near-field region of XL-array BS. We consider the XL-MIMO BS adopts a hybrid antenna architecture, where N antenna elements are connected to N_{RF} RF chains through a phase shifter-based analog combiner network [8]. The number of RF chains is less than the number of antenna numbers, i.e., $N_{\text{RF}} < N$. For such a scenario, we study the uplink near-field channel estimation problem. The users send mutual orthogonal pilots to the BS, and the BS estimates the near-field channels according to the processed pilot signals of the phase shifter-based analog combiner. Since users send mutual orthogonal pilots, the channel between BS and each user can be estimated individually. In addition, we denote the $\lambda, c, \Delta d = \frac{\lambda}{2}$ and $D = (N-1)\Delta d$ as the wavelength, the speed of light, antenna spacing, and antenna aperture, respectively.

A. Near-Field Region and Channel Model

A typical signal propagation area includes two regions: near-field (Fresnel area) and far-field (Fraunhofer area), which are separated by the classical Rayleigh distance $r_{\text{Ray}} = \frac{2D^2}{\lambda}$. Due to the utilization of the larger antenna aperture and the higher

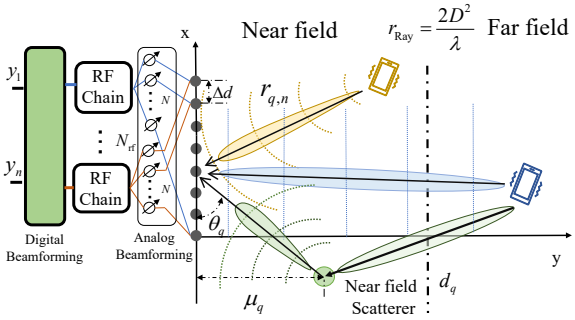

Fig. 1: The system layout of extremely large arrays and the signal propagation paths

communication frequency, the near-field area will significantly expand. Thus users/scatterers will likely be located within the near-field region. As in [2], [5], we consider a scattering environment in which the Q scatterers are located either in the near field or far field. The received signal is assumed as the signal incident from several direct and scattering paths. Based on the bistatic radar equation, the channel between the specific user and the n-th antenna element of BS can be formulated as

$$h_n = \sum_{q=1}^{Q} \frac{\lambda\sqrt{\rho}}{(4\pi)^{3/2} r_{q,n}} e^{-jkr_{q,n}}, \tag{1}$$

where $\beta_{q,n} = \frac{\lambda}{(4\pi)^{3/2} r_{q,n}}$ represents the free space path loss; ρ represents the scattering cross sections; For the direct path, $\rho = 1$. $k = 2\pi/\lambda$. $r_{q,n} = d_q + \sqrt{\mu_q^2 + (n\Delta d)^2}$ is the length of the propagation path, where μ_q is the y-axis coordinate of the last-hop scatter of the path q; d_q is the distance from the user to the last-hop scatter. In the far-field, plane wavefront always assumes $\mu_q \gg n\Delta d$ and simplifies $r_{q,n} \simeq d_q + \mu_q - n\Delta d \sin\theta_q$, where θ_q is the incident angle. However, it will not be supported in the near field. In the near-field, we simplify $r_{q,n} \simeq d_q + \mu_q + \frac{(n\Delta d)^2}{2\mu_q} - n\Delta d \sin\theta_q$ according to [10].

We ignore the difference of path loss on the different antennas, i.e., setting $\beta_q = \beta_{q,0} = \cdots = \beta_{q,N}$. Then, the near-field channel vector can be represented by $\mathbf{h} = [h_0, \cdots, h_{N-1}]^{\text{T}} = \mathbf{A}_Q \boldsymbol{\alpha}_Q$, where $\mathbf{A}_Q \in \mathbb{C}^{N\times Q}$, and the q-th column vector $\mathbf{a}_q$ is the array response vector of q-th path, given by $\mathbf{a}_q(\mu_q, \theta_q) = [\cdots, e^{-jk\left(\frac{(n\Delta d)^2}{2\mu_q} - n\Delta d \sin\theta_q\right)}, \cdots]^{\text{T}}$, $\boldsymbol{\alpha}_Q$ represents the path loss, whose elements is $\alpha_q = \beta_q e^{-jk(\mu_q + d_q)}$.

To represent the channel in the sparse domain, we adopt the method mentioned in [5], known as spatial partition. The spatial partition discretizes the interesting areas along the distance dimension μ and the angle dimension ψ and generates G grids. By assuming the channel is incident from these grids, the channel can be represented as

$$\mathbf{h} \simeq \mathbf{A}\boldsymbol{\alpha}, \tag{2}$$

where $\mathbf{A} \in \mathbb{C}^{N\times G}$ is the sparsifying dictionary, given by $\mathbf{A} = [\mathbf{a}(\mu_1, \theta_1), \ldots, \mathbf{a}(\mu_G, \theta_G)]$.

$\boldsymbol{\alpha} \in \mathbb{C}^{G\times 1}$ is the sparse channel gain corresponding to $\mathbf{A}$. When there is no channel path in g-th grid, the g-th element of $\boldsymbol{\alpha}$, denoted by α_g is equal to zero, i.e., $\alpha_g = 0$. As $Q \ll N \ll G$, $\boldsymbol{\alpha}$ is highly sparse. [11]

B. Problem Formulation

The observed pilot signal of the BS over T pilot slots, denoted by $\mathbf{Y} \in \mathbb{C}^{N_{\mathrm{RF}} \times T}$, is given by $\mathbf{Y} = \mathbf{Whs} + \mathbf{Wn}$, where $\mathbf{W} \in \mathbb{C}^{N_{\mathrm{RF}} \times N}$ is the configurable analog combiner matrix of the hybrid antenna; $\mathbf{s} \in \mathbb{C}^{1 \times T}$ denotes the orthogonal pilot sequence of the specific user, satisfying $\mathbf{ss}^{\mathrm{H}} = 1$; $\mathbf{n} \in \mathbb{C}^{N \times T}$ represents the noise matrix with each element following the distribution $\mathcal{N}(0, \sigma^2)$. As the orthogonal pilot is known at the BS side, we have

$$\mathbf{y} = \mathbf{Ys}^{\mathrm{H}} = \mathbf{Wh} + \mathbf{Wns}^{\mathrm{H}} \simeq \mathbf{WA}\boldsymbol{\alpha} + \hat{\mathbf{n}}, \tag{3}$$

where $\hat{\mathbf{n}} = \mathbf{Wns}^{\mathrm{H}}$ is the equivalent noise.

Our objective is to estimate the near-field channel $\mathbf{h}$ from $\mathbf{y}$, which is an underdetermined problem because of $N_{\mathrm{RF}} \ll N$. To tackle this challenge, we utilize the sparse representation in (2) that turns the estimation of channel $\mathbf{h}$ to estimate the sparse signal $\boldsymbol{\alpha}$ equivalently. The estimation of $\boldsymbol{\alpha}$ is a typical CS problem [9], which is formulated as,

$$\min_{\boldsymbol{\alpha}} \ \|\mathbf{y} - \mathbf{WA}\boldsymbol{\alpha}\|_2, \quad \text{s.t} \ \ \|\boldsymbol{\alpha}\|_0 < \epsilon, \tag{4}$$

where ϵ is a given threshold parameter related to σ^2.

Problem (4) is non-convex due to the non-convex ℓ_0 norm constraint. In the next section, we propose two model-based deep learning algorithms to solve it efficiently.

III. Proposed Algorithm

A. LISTA-based Near Field Channel Estimation

By relaxing the ℓ_0 norm constraint to ℓ_1 norm constraint, we transfer (4) into a tractable form, given by

$$\min_{\boldsymbol{\alpha}} \ \|\mathbf{y} - \boldsymbol{\Psi}\boldsymbol{\alpha}\|_2 + \xi\|\boldsymbol{\alpha}\|_1, \tag{5}$$

where $\boldsymbol{\Psi} = \mathbf{WA}$ and ξ is a regularization parameter.

Problems (4) and (5) share the same solution when the sensing matrix $\boldsymbol{\Psi}$ has the restricted isometry property (RIP) [9]. Problem (5) is a typical sparse recovery problem. One of the commonly used approaches to solve (5) is the Iterative Shrinkage and Thresholding Algorithm (ISTA) [12], which recovers the sparse $\boldsymbol{\alpha}$ by iterating the following recursive equation: $\boldsymbol{\alpha}^{(t+1)} = h_{(\eta)}\left((\mathbf{I} - \frac{1}{\lambda_{\max}}\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi})\boldsymbol{\alpha}^{(t)} + \frac{1}{\lambda_{\max}}\boldsymbol{\Psi}^{\mathrm{T}}\mathbf{y}\right)$, where $h_{(\eta)}$ is the element-wise soft shrinkage function, given by $[h_{(\eta)}]_g = \mathrm{sign}([\boldsymbol{\alpha}]_{\mathrm{g}})(|[\boldsymbol{\alpha}]_{\mathrm{g}}| - \eta)_+$, where $[\cdot]_g$ is the g-th elements of vector; $\mathrm{sign}(\cdot)$ returns the sign of a scalar; $(\cdot)_+$ means $\max(\cdot, 0)$; η is a constant threshold. $\lambda_{\max}$ is the maximum eigenvalue of $\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi}$. $\mathbf{I}$ is the identity matrix.

The recovering accuracy of ISTA is heavily influenced by η and iteration time, which are hard to select for each $\boldsymbol{\alpha}$. To achieve higher estimation accuracy and efficiency for (5), we apply LISTA, a neural network version of the ISTA, which overcomes the disadvantages of ISTA by fixing the iteration time and learning η from the data [13]. The key idea of LISTA is to unfold each iteration of ISTA to a neural network layer. The input and output relationship of the l-th layer is formulated as

$$\boldsymbol{\alpha}^{(l+1)} = h_{(\hat{\eta}^{(l)})}\left(\mathbf{V}_a^{(l)}\boldsymbol{\alpha}^{(l)} + \mathbf{V}_b^{(l)}\mathbf{y}\right), \tag{6}$$

where $\mathbf{V}_a$, $\mathbf{V}_b$, and $\hat{\eta}$ are learnable parameters for representing $\mathbf{I} - \frac{1}{\lambda_{\max}}\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi}$, $\frac{1}{\lambda_{\max}}\boldsymbol{\Psi}^{\mathrm{T}}$ and η, respectively. We denote $\boldsymbol{\alpha}^{(L)}$ as the output feature of LISTA network, which is also the solution of (5). Then, the near-field channel $\mathbf{h}$ can be obtained from (2) accordingly.

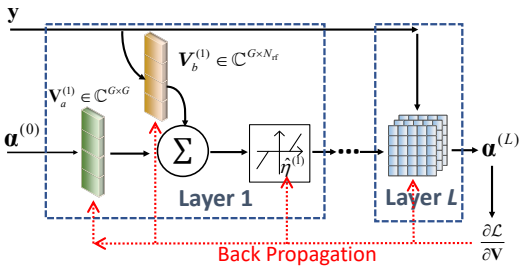

Fig. 2: The block diagram of LISTA algorithm

As a data-driven method, the parameters of LISTA are updated by back-propagating the gradient of the loss function $\mathcal{L} = \sum_{s=1}^{S} \left\|\boldsymbol{\alpha}_s^* - \boldsymbol{\alpha}^{(L)}\right\|_2$, on the data set $\mathbb{D} = \{[\mathbf{y}_s, \boldsymbol{\alpha}_s^*] | \mathbf{y}_s = \boldsymbol{\Psi}\boldsymbol{\alpha}_s^* + \mathbf{n}, s = 1, ..., S\}$, where $\boldsymbol{\alpha}_s^*$ is the accurate channel gain; $\mathbf{y}_s$ as the corresponding received signal of $\boldsymbol{\alpha}_s^*$; S is the total number of samples. Fig. 2 illustrates the neural network of LISTA with L layer, in which we represent the forward-solving process with dark lines and the back propagation with red dashed lines.

B. SDL-LISTA-based Near Field Channel Estimation

The above Section III-A provides a LISTA-based approach to study the new near-field channel estimation problem. It is worth noting that in Section III-A, the sparsifying dictionary $\mathbf{A}$ is constructed by simple spatial gridding. For the task of near-field channel estimation, the fixed spatial gridding $\mathbf{A}$ brings several challenges, which will be detailed later.

Lemma 1 (Theorem 1.11 of [9]): To accurately recover $\boldsymbol{\alpha}$ from $\mathbf{y}$ in (5), $\boldsymbol{\Psi}$ should satisfy the coherence guarantees, i.e., $\nu(\boldsymbol{\Psi}) < \frac{1}{2Q-1}$, where $\nu(\boldsymbol{\Psi})$ is the coherence of a matrix in Definition 1.5 of Literature [9].

Lemma 2: The spatial gridding dictionary $\mathbf{A}$ defined in (2) may not support $\boldsymbol{\Psi}$ to guarantee Lemma 1.

Proof: Please refer to Appendix A. ∎

Lemma 1 and 2 indicate that the spatial gridding dictionary $\mathbf{A}$ may degrade the estimation accuracy of near-field channels. Meanwhile, the size of $\mathbf{A}$ is much larger than that of the far-field spatial gridding dictionary, resulting in serious storage and computational burden.

To address these challenges, we propose a novel method, SDL-LISTA, that learns a sparsifying dictionary from data to achieve high channel estimation accuracy with reduced computational complexity by utilizing fewer atoms. The sparsifying dictionary optimization problem is formulated as

$$\min_{\mathbf{A}} \ \mathbb{E}\left\{\|\mathbf{h}^* - \mathbf{A}\mathcal{N}(\mathbf{y}_s | \mathbf{V}_a^*, \mathbf{V}_b^*, \hat{\eta}^*)\|_2\right\}, \tag{7}$$

where $\mathcal{N}(\mathbf{y}_s | \mathbf{V}_a, \mathbf{V}_b, \hat{\eta})$ represents the LISTA neural network. $\mathbf{h}^*$ represents the noise-free channel, wherein the incident angle and distance of channel path follow a distribution [14]. $\mathbf{V}_a^*$ and $\mathbf{V}_b^*$ and $\hat{\eta}^*$ are well-trained parameters under a given $\mathbf{A}$.

Directly solving (7) with alternating iteration is challenging because it needs to train the LISTA whenever $\mathbf{A}$ is updated.

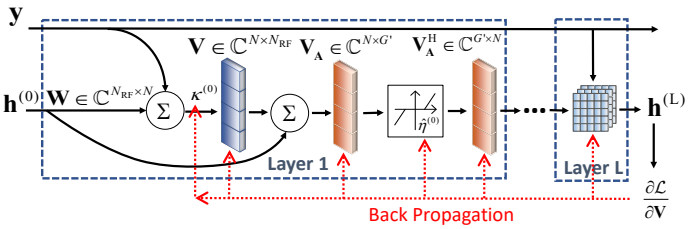

Fig. 3: The block diagram of SDL-LISTA

We turn this problem into a model-based learning problem that trains a desired $\mathbf{A}$ and the parameters of LISTA jointly. To realize this, we map $\mathbf{A}$ as a learnable parameter $\mathbf{V}_{\mathbf{A}} \in \mathbb{C}^{N \times G'}$ and construct it as a sparsifying dictionary learning layer, where G' is the number of atoms. Then, we embed this layer into the LISTA and propose the SDL-LISTA, which can be formulated as

$$\mathbf{h}^{(l+1)} = \mathbf{V}_{\mathbf{A}}^{\mathrm{H}} \cdot h_{(\hat{\eta}^{(l)})} \left(\mathbf{V}_{\mathbf{A}} \left(\mathbf{h}^{(l)} - \kappa^{(l)} \mathbf{V} (\mathbf{W}\mathbf{h}^{(l)} - \mathbf{y}) \right) \right). \quad (8)$$

In (8), we adopt another form of LISTA proposed in [15], in which $\mathbf{V} \in \mathbb{C}^{N \times N_{\mathrm{RF}}}$, $\hat{\eta}$ and κ are learnable parameters. $\mathbf{V}$ replaces the learnable parameters $\mathbf{V}_a$ and $\mathbf{V}_b$ used in (6), and it is shared among all layers. Meanwhile, it should be noticed that the sparsifying dictionary learning layer $\mathbf{V}_{\mathbf{A}}$ and $\mathbf{V}_{\mathbf{A}}^{\mathrm{H}}$ are embedded before and after the soft shrink function, respectively. $\mathbf{V}_{\mathbf{A}}$ transforms $\mathbf{h}$ into sparse form and $\mathbf{V}_{\mathbf{A}}^{\mathrm{H}}$ transforms it back. The block diagram of SDL-LISTA is illustrated in Fig. 3. Compared with LISTA, SDL-LISTA directly outputs the channel $\mathbf{h}$ instead of $\boldsymbol{\alpha}$. Thus, the neural network is trained using the data set $\mathbb{D}_{\mathrm{SDL}} = \{[\mathbf{y}_s, \mathbf{h}_s^*] | \mathbf{y}_s = \mathbf{W}\mathbf{h}_s^* + \mathbf{n}, s = 1, ..., S\}$. The loss function is given by $\mathcal{L}_{\mathrm{SDL}} = \sum_{s=1}^{S} \left\| \mathbf{h}_s^* - \mathbf{h}^{(L)} \right\|_2$. The SDL-LISTA-based near-field channel estimation approach is summarized in Algorithm 1.

Algorithm 1 SDL-LISTA for near-field channel estimation

1: Initialize the parameters $\mathbf{V}$, $\mathbf{V}_{\mathbf{A}}$, $\hat{\eta}$ and κ. Collect the data set $\mathbb{D}_{\mathrm{SDL}} = \{[\mathbf{y}_s, \mathbf{h}_s^*] | \mathbf{y}_s = \mathbf{W}\mathbf{h}_s^* + \mathbf{n}, s = 1, ..., S\}$.
2: **While**
3: Sample a batch of data from the data set $\mathbb{D}_{\mathrm{SDL}}$;
4: **Forward process**: Calculate the output of neural network $\mathcal{N}(\mathbf{y})$ and calculate the loss $\mathcal{L}_{\mathrm{SDL}}$;
5: **Backward process**: Update the parameters $\mathbf{V}$, $\mathbf{V}_{\mathbf{A}}$, $\hat{\eta}$ and κ using backpropagation.
6: **Until** $\mathcal{L}_{\mathrm{SDL}}$ converge.
7: **Output**: A well-trained neural network $\mathcal{N}_{\mathrm{SDL}}$.

SDL-LISTA generates a deep integration between the sparsifying dictionary and the distribution of the channel's incident direction. Due to the deep integration, the sparsifying dictionary biased in several directions provides a possibility to achieve higher accuracy with a much small dictionary. Another reason for higher accuracy is that the loss of SDL-LISTA is directly calculated by the channel instead of the sparse signal, which inherently avoids the sparse representation error to loss function compared with LISTA. Meanwhile, since we optimize the sparsifying dictionary at the training stage of LISTA, our algorithm will enjoy ultra-low computational complexity, which is definitely much less than traditional alternating minimization.

IV. Numberical Result

In this section, we provide numerical experiments to demonstrate the performance of our proposed near-field channel estimation approaches. In the following experiments, we consider a $N = 128$ uniform linear array with element space $\Delta d = \lambda/2$, which works at the carrier frequency 28 GHz (λ = 1.07 cm). The near-field distance under such a setting is around 87 m.

Our learning algorithm utilizes the Adam optimizer with the learning rate as $1e^{-4}$ in LISTA and SDL-LISTA. The batch size is 256. The initial values of $\mathbf{V}_a, \mathbf{V}_b$ for LISTA and $\mathbf{V}, \mathbf{V}_{\mathbf{A}}$ for SDL-LISTA are initialized with uniform distribution $\mathcal{U}(0, 1)$. The initial value of $\hat{\eta}$ is 10^{-4}.

The training, testing, and validation set, are generated with the mathematical simulation, which includes 256000, 2560, and 2560 samples, respectively. In the training and testing set, the signal-to-noise ratio (SNR) of receiving pilot is randomly chosen from 0 to 27 dB. The validation sets are evenly segmented into ten parts to measure the channel estimation accuracy when SNR is 0, 3,..., and 27 dB, respectively. The sample follows our channel model in (1), in which the incident angle $\phi = \sin\theta$ follows a Gaussian mixture distribution $\mathcal{N}(\boldsymbol{\mu}_\phi, \boldsymbol{\Sigma}_\phi)$, where $\boldsymbol{\mu}_\phi = [-0.6, -0.45, -0.2, 0.3, 0.6]$ and $\boldsymbol{\Sigma}_\phi = 0.15\mathbf{I}$. μ_q follows the uniform distribution $\mathcal{U}(2, 100)$ and d_q follows $\mathcal{U}(0, 100)$. Q is an integer randomly selected between 2 to 6. The label $\boldsymbol{\alpha}^*$ for LISTA is generated by assuming the signal is incident from the nearest grid.

We measure the channel estimation accuracy with the normalized MSE (NMSE), given by $\mathrm{NMSE} = \sum_{n=1}^{N} \frac{\|h_n - h_n^*\|_2^2}{\|h_n^*\|_2^2}$.

In Fig. 4, we study the convergence behavior of SDL-LISTA for different layers, considering $\mathbf{V}_{\mathbf{A}}$ having $G' = 256$ atoms. Observations indicate that while the two-layer SDL-LISTA converges quickly, it falls short of fully recovering the channel. A minimum of 6 layers is required for a full recovery. In general, convergence can be achieved within 100-150 epochs.

In Fig. 5, we compare the achievable NMSE performance of our proposed model-based learning methods (LISTA and SDL-LISTA) with representative non-learning algorithms, including OMP [5] and Fast ISTA (FISTA) [12]. The number of iterations for OMP and FISTA is 10 and 100, respectively.

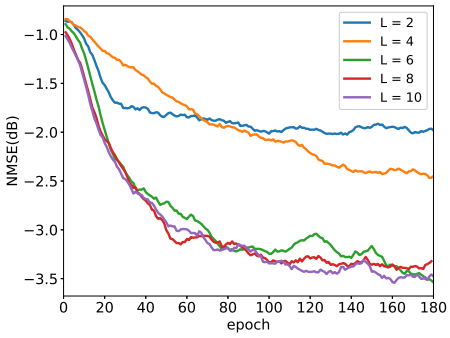

Fig. 4: The convergence behavior of SDL-LISTA

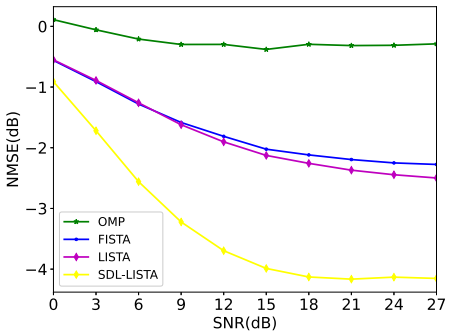

Fig. 5: The NMSE performance comparison

LISTA and SDL-LISTA use neural networks with ten layers. The OMP, FISTA, and LISTA use the sparsifying dictionary with $G = 2048$ atoms generated by spatial gridding. The grids' angles are obtained by evenly partitioning $\phi \in [-1, 1]$ into 256 grids, while the grids' distances are obtained by evenly partitioning $\frac{1}{\mu_q} \in [0, 0.5]$ into eight grids. From Fig. 5, it is observed that model-based algorithms outperform existing OMP and FISTA. For example, SDL-LISTA with $G = 256$ outperforms OMP and FISTA for 3 dB on average. Thereby, we can conclude that LISTA achieves better performance with less calculation compared with the existing methods. Moreover, thanks to the joint training of the sparsifying dictionary, SDL-LISTA can achieve better performance than LISTA with ten times atoms reduction. This implies that the SDL-LISTA further improves performance and reduces calculation complexity.

V. Conclusion

This letter studied the near-field channel estimation problem in XL-MIMO systems, and two model-based deep learning algorithms were proposed to estimate channel parameters efficiently. In particular, LISTA was first applied to the near-field channel estimation problem using the spatial gridding-based sparsifying dictionary. Then, SDL-LISTA, which embeds the sparsifying dictionary into LISTA neural network, was proposed further to enhance the near-field channel estimation accuracy with reduced complexity. Finally, simulation results were provided to verify the effectiveness of our proposed algorithms.

Appendix A
Proof of Lemma 2

First, the coherence of $\mathbf{\Psi}$ can be rewritten as $|\langle \boldsymbol{\psi}_t(\mu_t, \theta_t), \boldsymbol{\psi}_v(\mu_v, \theta_v)\rangle| = \left|\mathbf{a}_v(\mu_v, \theta_v)^{\mathrm{H}} \mathbf{W}^{\mathrm{H}} \mathbf{W} \mathbf{a}_t(\mu_t, \theta_t)\right|$. Without loss of generality, we assume $G = N$ and the coherence of $\mathbf{W}$ reached the Welch bound, which means $\mathbf{W}^{\mathrm{H}}\mathbf{W} = \mathbf{I}$ [9]. Then, we have

$$\begin{aligned}\nu(\mathbf{\Psi}) &= \max_{1\le t<v\le G} \frac{|\langle \mathbf{a}_t, \mathbf{a}_v\rangle|}{\|\mathbf{a}_t\|_2 \|\mathbf{a}_v\|_2} \\ &\simeq \max_{1\le t<v\le G} \frac{1}{N}\left|\sum_{n=0}^{N-1} \exp\left(jkn\Delta d\left(\frac{1}{2}n\Delta d\Delta\mu - \Delta\phi\right)\right)\right| \end{aligned} \tag{9}$$

where $\mathbf{a}_t$ and $\mathbf{a}_v$ are any two column of $\mathbf{A}$; $\Delta\mu = \frac{1}{\mu_t} - \frac{1}{\mu_v}$ and $\Delta\phi = \sin\theta_t - \sin\theta_v$.

The maximum value of $\nu(\mathbf{\Psi})$ is normally obtained when t and v are adjacent grids. We denote the $\Delta\phi_0$ and $\Delta\mu_0$ as the grid interval. A common set of grid intervals is that $\Delta\phi_0 \le \frac{2}{N}$ and $\Delta\mu_0 \ge \frac{2\lambda}{D^2}$ [5], [6]. In such scenario, we can approximate the second-order function $\frac{1}{2}\Delta d\Delta\mu_0 n^2 - \Delta\phi_0 n$ with a first-order function $\left(\frac{1}{2}(N-1)\Delta d\Delta\mu_0 - \Delta\phi_0\right)n$. Then, we have

$$\nu(\mathbf{\Psi}) = \sum_{n=0}^{N-1} \exp\left(jk(N-1)\Delta d\left(\frac{1}{2}n\Delta d\Delta\mu_0 - \Delta\phi_0\right)\right) \tag{10}$$

It can be observed that $\nu(\mathbf{\Psi}) \simeq 1$ when $\frac{1}{2}(N-1)\Delta d\Delta\mu_0 \simeq \Delta\phi_0$, which is the case that two grids have one distance interval and an angle interval. Hence, $\nu(\mathbf{\Psi})$ with the spatial gridding dictionary $\mathbf{A}$ has big coherence and may not guarantee Lemma 1. The proof is completed.

References

[1] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a reality—what is next?: Five promising research directions for antenna arrays," *Digital Signal Processing*, vol. 94, pp. 3–20, 2019.

[2] H. Lu and Y. Zeng, "Communicating with extremely large-scale array/surface: Unified modelling and performance analysis," *IEEE Transactions on Wireless Communications*, 2021.

[3] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani, and Y. C. Eldar, "Beam focusing for near-field multiuser MIMO communications," *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 7476–7490, 2022.

[4] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani, and Y. C. Eldar, "Near-field wireless power transfer for 6g internet of everything mobile networks: Opportunities and challenges," *IEEE Communications Magazine*, vol. 60, no. 3, pp. 12–18, 2022.

[5] M. Cui and L. Dai, "Channel estimation for extremely large-scale MIMO: Far-field or near-field?" *IEEE Transactions on Communications*, vol. 70, no. 4, pp. 2663–2677, 2022.

[6] Y. Han, S. Jin, C.-K. Wen, and X. Ma, "Channel estimation for extremely large-scale massive MIMO systems," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 633–637, May 2020.

[7] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2370–2386, 2016.

[8] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.

[9] Y. C. Eldar and G. Kutyniok, *Compressed sensing: theory and applications*. Cambridge university press, 2012.

[10] F. Bohagen, P. Orten, and G. E. Oien, "Design of optimal high-rank line-of-sight MIMO channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1420–1425, 2007.

[11] J. and Wolper, "A mathematical introduction to compressive sensing," *Computing reviews*, vol. 55, no. 1, pp. 8–8, 2014.

[12] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm with application to wavelet-based image deblurring," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 693–696.

[13] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proceedings of the 27th international conference on international conference on machine learning*, 2010, pp. 399–406.

[14] J. Zhang, C. Pan, F. Pei, G. Liu, and X. Cheng, "Three-dimensional fading channel models: A survey of elevation angle research," *IEEE Communications Magazine*, vol. 52, no. 6, pp. 218–226, 2014.

[15] X. Chen, J. Liu, Z. Wang, and W. Yin, "Theoretical linear convergence of unfolded ISTA and its practical weights and thresholds," *arXiv:1808.10038*, Nov. 2018.