Deep Learning for Super-Resolution DOA Estimation in Massive MIMO Systems

Hongji Huang*, *Student Member*, *IEEE*, Guan Gui*‡, *Senior Member*, *IEEE*, Hikmet Sari*, *Fellow*, *IEEE*, Fumiyuki Adachi†‡, *Life Fellow*, *IEEE*

*Nanjing University of Posts and Telecommunications, Nanjing 210003, China.

†Research Organization of Electrical Communication (ROEC), Tohoku University, Sendai 980-8577, Japan.

‡E-mails: guiguan@njupt.edu.cn, adachi@ecei.tohoku.ac.jp

***Abstract*—The requirement of the increasing capacity of the communication networks promotes the massive multiple input multiple output (MIMO), which has attracted a lot of attention among academic and industry communities. Due to the inherent sparsity features of channel structure in uplink massive MIMO systems, conventional methods often bring about high computational complexity and also fail to make full use of the structural information. In order to solve this problem, this paper proposes a novel deep learning (DL) based super-resolution direction of arrivals (DOA) estimation method. Specifically, it is realized with the aids of the well-designed deep neural network (DNN). Then we employ the DNN to carry out offline learning and online deployment procedures. This learning mechanism can learn the features of the wireless channel and the spacial structures efficiently. Finally, simulation results are provided to show that the proposed DL based scheme can achieve better performance in terms of the DOA estimation compared with conventional methods.**

***Index Terms*—Massive multiple input multiple output (MIMO), deep learning, DOA estimation, learning policy.**

I. Introduction

The anticipated 1000-fold explosive data traffic growth by 2020 will lead to great challenges in the design of fifth generation communication systems (5G) [1]. In order to meet this challenging goal, many advanced techniques have been proposed, such as massive multiple input multiple output (MIMO) [2], non-orthogonal multiple access (NOMA) [3], [4], and Millimeter-wave (mmWave) communication [5]. Massive MIMO, as an emerging technique in the field of array signal processing has given rise to a great interest among the research community in the recent years.

The channel state information (CSI) is directly concerned with the performance of the massive MIMO system, implying that channel estimation is a key issue in such a large-scale uplink system. A unified transmission framework for multiuser time division duplex (TDD)/frequency division duplex (FDD) massive MIMO systems was explored in [6]. In [7], for leveraging sparsity statistics of the mmwave beamspace channel, the authors proposed a support detection-based channel estimation strategy with low pilot overhead to estimate the sparse beamspace channel. In addition, hybrid analog and digital (HAD) structure with two phase alignment (PA) method [8], angle domain channel tracking approach [9], and priori aided channel tracking scheme [10] were reported for actualizing direction of arrivals (DOA) estimation.

Although research efforts in the above direction aim to address super-resolution DOA estimation issue, tremendous computational complexity induced by non-linear optimization exists in previous researches. These works are unable to fully exploit the structural features in the uplink massive MIMO system. Also, traditional channel estimation methods have difficulties in tracking the time-varying channel. Meanwhile, since previous studies assume the channel sparsity patterns as unknown, nonlinear reconstruction procedures are unavoidable. Hence, it is necessary to develop an alternative technique to realize DOA estimation instead of simply optimizing traditional methods. Recently, the novel deep learning (DL) [11] concept (i.e., a typical algorithm of the machine learning) was proposed to handle big data problem and overcome nonlinear operation challenge. Until now, some works that incorporate DL into communication have been presented, e.g., channel coding, MIMO, and heterogeneous network traffic control [12]–[14].

To overcome the fundamental limitations, we investigate the DOA estimation optimization strategy in the uplink massive MIMO system by integrating the DL into multi-antennas system. The main contributions of this paper are summarized as follows.

1) To the best of our knowledge, we first consider a framework that incorporates the DL technique into the multiple antennas uplink systems for spacial characteristics detection. Specifically, deep neural network (DNN) is introduced to extract the features of the communication model, in which different layers can process specific activation functions and realize corresponding mapping relationship.
2) In our work, we propose a high-resolution DOA estimation scheme in sparse channel scenario. In our framework, after obtaining real-time CSI and spacial structure information through offline and online training, these samples are used to train the DNN for DOA estimation. Extensive simulation results and comparison have verified the efficiency and robustness of the proposed DOA estimation scheme in the uplink massive MIMO system.

978-1-5386-6358-5/18/$31.00

II. System Model

We consider a typical uplink massive MIMO system, where one base station (BS) with a uniform linear array (ULA) of N_t antennas and D single-antenna users are designed. Here, the BS is assumed to have no information on the users. Furthermore, we introduce the classical narrowband ray-based channel model [15], [16], and the uplink model of user k can be given as

$$\mathbf{h}_k = g_k \mathbf{a}_t(\theta_k), \tag{1}$$

where g_k is denoted as the complex gain of the k-th user, while θ_k is noted as the physical DOA at the k-th user of massive MIMO. Also, the steering vector $\mathbf{a}_t(\theta_k)$ is defined as the array response at the BS. For a ULA, $\mathbf{a}_t(\theta_k) \in \mathbb{C}^{N_t \times 1}$ can be expressed as

$$\mathbf{a}_t(\theta_k) = \frac{1}{\sqrt{N_t}}[1, e^{-j2\pi\frac{d}{\lambda}\sin\theta_k}, \cdots, e^{-j2\pi\frac{d}{\lambda}(N_t-1)\sin\theta_k}]^T, \tag{2}$$

Here, d represents the antenna spacing, while λ is defined as the wavelength of the carrier frequency. Meanwhile, the uplink channel matrix is $\mathbf{A} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_D] \in \mathbb{C}^{N_t \times D}$.

Then, assuming the transmitted signal vector as $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$, the received signals at the BS can be formulated as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{3}$$

where $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_D)$ is additive white Gaussian noise (AWGN) with zero mean and variance σ^2. Concretely, the transmitted signal vector $\mathbf{x}$ is transmitted to the fixed channel matrix $\mathbf{A}$ in all direction, and corresponding received signal vector $\mathbf{y}$ is obtained.

From Eq. (1), it is observed that the DOA information is a vital feature to model the channel matrix in such a uplink massive MIMO system. In traditional schemes, the uplink training leads to unavoidable pilot contamination induced by the large number of the antennas. In contrast, the proposed DL based approach leverages the sparsity features of the multiple antennas system to achieve super-resolution DOA estimation based on the physical DOA.

III. Deep Learning Based Super-Resolution DOA Estimation

In this section, for the sake of realizing high-resolution DOA estimation in the uplink massive MIMO systems, we provide a DL based method which integrates the state-to-the-art DL into the uplink systems. In the past few years, many researches have been devoted to DOA estimation, and a lot of distinguished methods have been proposed, such as estimation of signal parameters via rotational invariance technique (ESPRIT) [17], [18] and multiple signal classification (MUSIC) [19]. Lately, paper [20] argued that eigen-decomposition method is one of the core steps in these subspace based strategies. Unfortunately, these methods are constrained in such a large-scale antennas system. In the proposed DL based framework, according to the well-known universal approximation theorem [21], a feed-forward network with multilayer perception is capable of approximating continuous functions on compact subsets of $\mathbf{R}^n$, suggesting that the performance of the DOA estimation can be facilitated with the aids of the strong recognition and mapping relations of the DL.

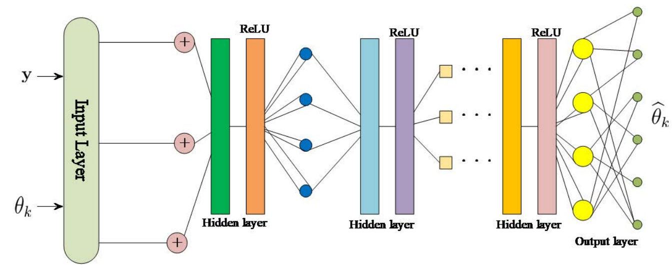

Fig. 1. DNN architecture in the proposed scheme.

A. Deep Neural Network Architecture

The remarkable progress in DL has brought great advance in many promising areas, and the newly proposed DNN technique is regarded as a revolutionary technology which has been universally adopted among natural language processing (NLP), computer vision (CV), and automatize driving. These applications have corroborated that the performance of many information systems can be elevated dedicated by the DL. However, existing works have not focused on the DL aided multi-antennas systems, and we are wondering if the powerful DL can enhance its performance.

In order to spur the performance of artificial neural networks (ANN), DNN is developed and its learning capacity is promoted contributed by many hidden layers designed in the DNN. Specifically, multiple neurons exist in each hidden layer, and the output of the network is a weighted sum of these neurons processed by nonlinear functions. Generally speaking, we usually use the Sigmoid function and the Rectified linear unit (ReLU) function in the nonlinear operation, which can be written as $f_S(x) = \frac{1}{1+e^{-x}}$ and $f_R(x) = \max(0, x)$, respectively. We assume the output of the DNN as $\mathbf{o}$, and $\mathbf{v}$ represents the input data, we formulate

$$\mathbf{o} = f(\mathbf{v}, w) = f^{(n-1)}(f^{(n-2)}(\cdots f^1(\mathbf{v}))), \tag{4}$$

where n and w are denoted as the amount of layers of the DNN and the weights of the DNN, respectively.

As exhibited in Fig. 1, in our DNN framework, we define the length of each training sequence of the network as L, representing dimension of the input layer of the DNN. As a fully connected layer, it contains 256 neurons and acts as input of the transmitted signal vectors. Then, it conveys sparsity features to the first hidden layer with 300 neurons. To realize encoding, the second hidden layer is equipped with 256 neurons. In order to restrain overfitting, we design a dropout layer with retaining probability p. Afterwards, the next layer is designed as a noise layer with 200 neurons to corrupt the transmitted signals with the AWGN. Furthermore, the remaining hidden layer with 128 neurons is implemented as a decoder. In addition, the output layer is a linear layer, in which the estimated results of the DOA based on the uplink massive MIMO system can be obtained. Interestingly, the ReLU function is used as the activation function in the input layer and all the hidden layers, while we introduce the Sigmoid function in the output layer.

B. Learning Scheme

In the proposed DL based framework, we regard the uplink massive MIMO system as a mapping function. In order to learn the sparsity statistics and the channel characteristics of the developed uplink system, we derive a two-stages training policy. In the first stage, we propose an offline training method, which is adopted to capture the features of this multi-antennas system. Based on the system model, we can obtain the corresponding received signal $\mathbf{y}$ based on the fixed channel matrix $\mathbf{A}$ in different direction, which is collected as the training examples. Specifically, every time, once we transmit signal vector vias the channel through the massive MIMO system, the received signal $\mathbf{y}$ is acquired in a special direction. Hence, we can obtain the received signals in all direction. Synchronously, the physical DOA θ_k can be generated randomly to form a training dataset with the received signal vector $\mathbf{y}$, which is the training samples of the DNN. In the next stage, online learning mechanism is conducted based on the given channel model and the estimated DOA can be obtained without requiring iterations. To estimate the DOA information θ_k, a loss function based on the mean square error (MSE) concept is expressed as

$$\begin{aligned} \text{loss} &= \mathbb{E}\{\|\theta_k - \widehat{\theta}_k\|^2\} \\ &= \frac{1}{DM}\sum_{m=1}^{M}\sum_{k=1}^{D}\|\theta_k - \widehat{\theta}_k\|^2, \end{aligned} \tag{5}$$

Here, M is denoted as the number of examples, while $\widehat{\theta}_k$ is noted as the prediction. Thereafter, based on the loss function as Eq. (5), we present a novel DL based algorithm for DOA estimation, which is illustrated as **Algorithm 1**.

In order to assess the performance of the proposed DL based DOA estimation scheme, we introduce the well-known MSE principle to evaluate the estimation error, which is given by

$$\mathrm{MSE_d} = \frac{1}{M}\sum_{m=1}^{M}\|\theta_k - \widehat{\theta}_k\|^2, \tag{9}$$

C. Robustness Analysis of the DL based Scheme

Although the proposed DL based scheme for DOA estimation adapts to various channel conditions, it is still necessary to investigate the robustness of a DL based algorithm. We now consider the two typical cases to validate the robustness of the proposed scheme. For one thing, we test how the regularization item impacts on the performance of the DL based method. Also, extending our proposed method to various number of the hidden layers is important given the widely commercial use, but totally confused. To answer these questions, TABLE I and TABLE II present the results when adopting the hidden layers and regularization item of the DNN.

Algorithm 1 DNN based DOA estimation scheme in the uplink massive MIMO system.

Input: Environment simulator, DNN, received signal vector θ_k.
Output: $\widehat{\theta}_k$.

Initialize small constant τ with error threshold as $\tau^0 = 10^{-7}$, and gradient accumulation variable as $u = 0$. Also, initialize global learning rate η.
Start the environment simulator to generate wireless channel, and add random noise or distortion into the channel.
Process the uplink massive MIMO model, and obtain the corresponding received signal $\mathbf{y}$.
Combined with the received signal vector θ_k, we collect the training dataset Ψ.
while $\tau \geq \tau^0$ **:**
Sample a minibatch of M samples from the training dataset Ψ.
Construct the proposed DNN, and operate the network based on the selected data.
Update the output $\widehat{\mathbf{x}}$ of the DNN.
Compute gradient g as

$$g = 1/M \,\triangledown_{\theta_k} \sum_m \|\theta_k - \widehat{\theta}_k\|^2 \tag{6}$$

Accumulate squared gradient u, which is represented as

$$u = u + g \odot g \tag{7}$$

Calculate the update step of the estimated DOA $\widehat{\theta}_k$ of the k-th user as

$$\Delta\widehat{\theta}_k = -\frac{\eta}{\tau^0 + \sqrt{u}} \odot g \tag{8}$$

Update the output of the estimated DOA $\widehat{\theta}_k$: $\widehat{\theta}_k = \widehat{\theta}_k + \Delta\widehat{\theta}_k$.
end while
Obtain the estimated DOA $\widehat{\theta}_k$ after processing the DNN.
return: $\widehat{\theta}_k$.

TABLE I
TEST ACCURACY FOR HIDDEN LAYERS AND REGULARIZATION TERM.

Hidden Layers	Method	Test Accuracy(%)
2	DNN	98.122
	DNN + L^2 Regularization	98.763
3	DNN	99.001
	DNN + L^2 Regularization	99.003
4	DNN	99.569
	DNN + L^2 Regularization	99.603
5	DNN	99.722
	DNN + L^2 Regularization	99.751

TABLE II
OPERATION TIME FOR HIDDEN LAYERS AND REGULARIZATION TERM.

Hidden Layers	Method	Operation time(s)
2	DNN	0.0134
	DNN + L^2 Regularization	0.0134
3	DNN	0.0152
	DNN + L^2 Regularization	0.0168
4	DNN	0.0188
	DNN + L^2 Regularization	0.0189
5	DNN	0.0191
	DNN + L^2 Regularization	0.0193

As we can see from TABLE I, due to the regularization term can constrain the overfitting risk and optimize generalization performance, the proposed DL based scheme attains further improvement with the aids of the L^2 regularization terms method. We also see that the test accuracy is not been degraded when decreasing the number of hidden layers, inferring that the proposed DL based scheme is robust and efficient. Then, we now turn to evaluate the computational complexity of the DL based strategy, the operation time in the second stage (i.e., testing stage) is provided when changing the number of the hidden layers. We observe from TABLE II that the computational complexity of the proposed DL based scheme is low and this index is keeping stable with different hidden layers. Furthermore, it is clear that adding regularization term is not required extra operation time. Thus, the proposed DL based scheme for DOA estimation in the uplink massive MIMO systems is robust and high-efficiency.

IV. SIMULATION RESULTS AND ANALYSIS

In this section, numerical analysis for the DOA estimation of the proposed DL based schemes is presented. Here, we adopt the popular *Keras* to design this DNN based framework. In our simulation, we consider a typical massive MIMO system, in which the BS is equipped with $N_t = 128$ antennas and $D = 32$ users. Also, $d = \lambda/2$ and the wavelength λ of the carrier is selected as 5 mm. Moreover, the DOA $\theta_{k,i}$ is randomly distributed in the space $[-\pi/2, \pi/2]$. Additionally, the basic learning rate of the DNN is set as 0.024 and the learning rate decay is set as 0.96, while the batch size is 1200.

Fig. 2 shows the MSE performance of the DOA estimation against SNR of the proposed DL based DOA estimation scheme with different length L of the training sequences, in which $L = 4$ bits, $L = 8$ bits, $L = 16$ bits, and $L = 32$ bits are considered. It can be seen from Fig. 2 that the MSE of the DOA estimation is reducing with the increasing signal to noise ratio (SNR), and it becomes stable gradually until the SNR is large enough. Also, we can see from the simulation results that the MSE performance can be enhanced when adopting longer training sequences. This result is dedicated by the fact that longer length of training sequence can stir the convergence of the training stage of the DNN.

The performance comparison in terms of the MSE of the DOA estimation against SNR is shown in Fig. 3, where the number of θ consisted in a sample is set as 4, 8, 16, and 32, respectively. Here, the length of the training sequence is initial as 16 bits. It can be seen from Fig. 3 that the MSE performance in the case of 32 outperforms than that of other cases, which implies that adding more information of the physical DOA can elevate the performance of the DOA estimation based on the proposed scheme in the uplink massive MIMO system. In addition, it is pointed out that this advantage is based on the fact that more θ can accelerate the learning and representation procedures of the wireless channel statistics in the proposed DL based scheme.

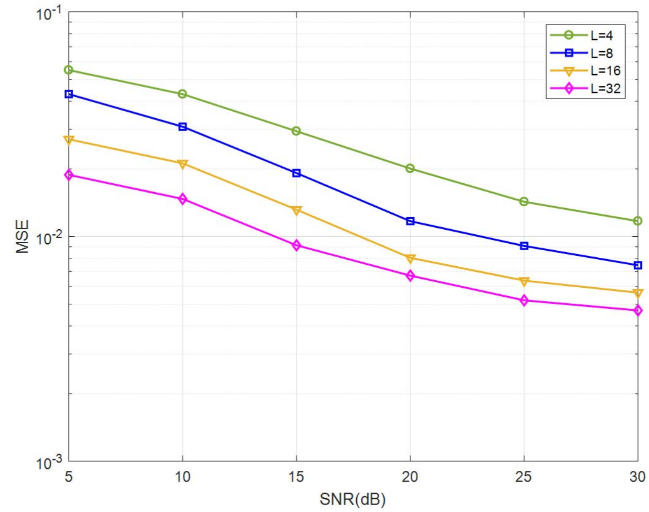

Fig. 2. MSE performance of the DOA estimation of the proposed DL based scheme when the length of training sequence is 4,8,16,and 32 (bits).

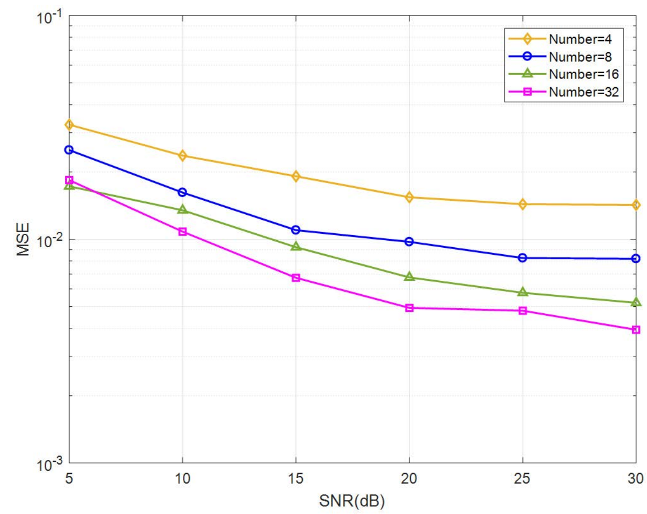

Fig. 3. MSE performance of the DOA estimation of the proposed DL based scheme when the number of θ is 4, 8, 16 and 32.

Fig. 4 exhibits the MSE performance of the DOA estimation against the SNR of the proposed DL based scheme, SBEM scheme [6], PA channel tracking scheme [10], ADMA user scheduling scheme [9], and MUSIC scheme [19], respectively. We can observe that the MSE performance can be improved with higher SNR among all the methods. Particularly, the proposed DL scheme has about one order of magnitude

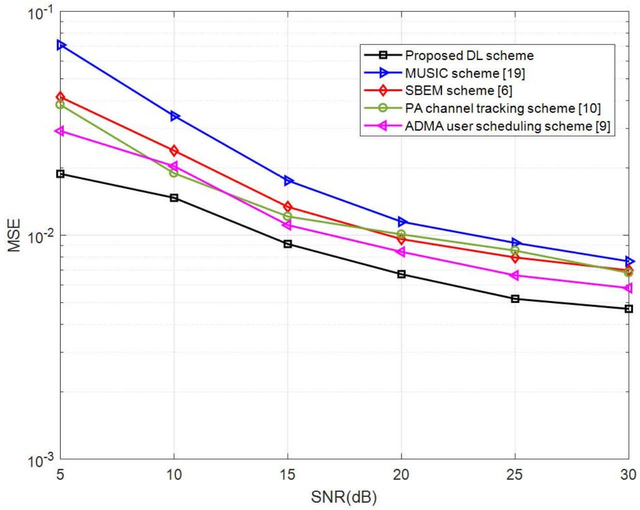

Fig. 4. Comparison of the MSE performance of the DOA estimation of the proposed DL based scheme, SBEM scheme, PA channel tracking scheme, ADMA user scheduling scheme, and MUSIC scheme.

reduction in terms of the MSE performance compared with the PA channel tracking method, which benefits from the DL based technique can realize end-to-end optimization whereas enormous channel power leakage and tremendous performance loss are induced by the selection of the single spatial support with the maximum amplitude for transmission of the method in [10]. Also, the MSE performance of the DOA estimation of the proposed DL based scheme outperforms that of the SBEM scheme, the ADMA user scheduling approach, and the MUSIC scheme, for the reason that the DOA can be obtained directly in angle domain with the aids of the powerful DL technique and the excellent generalization ability of the learning mechanism of the DNN. Additionally, the significant enhancement of the MSE performance of the DOA estimation is not required for high computational complexity in our proposed DL based scheme.

V. Conclusions

In this paper, we have proposed a DL based super-resolution DOA estimation in the uplink massive MIMO system. At first, we developed a typical massive MIMO system model. Then, we designed the DNN structure and mapping transformation of each layer of the DNN to learn the statistics of the channel model and capture the spatial features in angle domain. Particularly, the offline learning policy and online deployment method were provided and was incorporated into the proposed DNN to detect the wireless channel. In order to realize super-resolution DOA estimation, a DL based scheme was proposed in this paper. Simulation results have demonstrated that the proposed DL based scheme can achieve better MSE performance in terms of DOA estimation compared with previous methods. Importantly, we explored a new way to accelerate the development of conventional communication, showing that incorporating the DL into wireless communication scenarios is a feasible.

References

[1] C. L. I, C. Rowell, S. Han, Z. Xu, G. Li and Z. Pan, “Toward green and soft: a 5G perspective,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 66-73, Feb. 2014.

[2] J. Hoydis, S. ten Brink, and M. Debbah, “Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160-171, Feb. 2013.

[3] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, “Systemlevel performance evaluation of downlink non-orthogonal multiple access (NOMA),” in *IEEE International Symposium Personal, Indoor and Mobile Radio Communications* (PIMRC), London, UK, Sep. 8-11, 2013, pp. 611–615.

[4] H. Huang, J. Xiong, J. Yang, G. Gui, and H. Sari, “Rate Region Analysis in a Full-Duplex-Aided Cooperative Nonorthogonal Multiple-Access System,” *IEEE Access*, vol. 5, pp. 17869-17880, Aug. 2017.

[5] 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 Trans. Commun.*, vol. 63, no. 9, pp. 3029-3056, Sep. 2015.

[6] H. Xie, F. Gao, S. Zhang, and S. Jin, “A Unified Transmission Strategy for TDD/FDD Massive MIMO Systems With Spatial Basis Expansion Model,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3170-3184, Apr. 2017.

[7] X. Gao, L. Dai, S. Han, C. L. I, and X. Wang, “Reliable Beamspace Channel Estimation for Millimeter-Wave Massive MIMO Systems with Lens Antenna Array,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6010-6021, Sep. 2017.

[8] F. Shu et al., “Low-Complexity and High-Resolution DOA Estimation for Hybrid Analog and Digital Massive MIMO Receive Array,” *IEEE Trans. Commun.*, vol. PP, no. 99, pp. 1-1, Feb. 2018.

[9] J. Zhao, F. Gao, W. Jia, S. Zhang, S. Jin, and H. Lin, “Angle Domain Hybrid Precoding and Channel Tracking for Millimeter Wave Massive MIMO Systems,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6868-6880, Oct. 2017.

[10] X. Gao, L. Dai, Y. Zhang, T. Xie, X. Dai, and Z. Wang, “Fast Channel Tracking for Terahertz Beamspace Massive MIMO Systems,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 5689-5696, Jul. 2017.

[11] Hinton, Geoffrey E., S. Osindero, and Y. W. The, “A Fast Learning Algorithm for Deep Belief Nets,” *Neural Computation*, vol.18, 2006.

[12] Tobias Gruber, Sebastian Cammerer, Jakob Hoydis, and Stephan ten Brink, “On Deep Learning-Based Channel Decoding,” *ArXiv preprint arXiv:1701.07738*, 2017.

[13] N. Kato et al., “The Deep Learning Vision for Heterogeneous Network Traffic Control: Proposal, Challenges, and Future Perspective,” *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 146–153, Jun. 2017.

[14] S. Shen, C. J. Chang, and L. C. Wang, “A cellular neural network and utility-based radio resource scheduler for multimedia CDMA communication systems,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 11, pp. 5508–5519, Nov. 2009.

[15] J. Brady, N. Behdad, and A. M. Sayeed, “Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements,” *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.

[16] P. V. Amadori and C. Masouros, “Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection,” *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2212–2223, Jun. 2015.

[17] A. Paulraj, R. Roy, and T. Kailath, “A subspace rotation approach to signal parameter estimation,” *Proc. IEEE*, vol. 74, no. 7, 1044–1046, Jul. 1986.

[18] R. Roy and T. Kailath, “Esprit-estimation of signal parameters via rotational invariance techniques,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[19] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[20] L. Cheng, Y.-C. Wu, J. Zhang, and L. Liu, “Subspace identification for DOA estimation in massive/full-dimension MIMO systems: Bad data mitigation and automatic source enumeration,” *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 5897–5909, Nov. 2015.

[21] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, 1989, vol. 2, no. 5, pp. 359-366, 1989.