Hindawi
Mobile Information Systems
Volume 2022, Article ID 7301273, 6 pages
https://doi.org/10.1155/2022/7301273

Research Article

Wideband mmWave MIMO: One Prior-Aided LAMP-Based Network Combining with Residual Learning for Beamspace Channel Estimation

Chunhua Zhu,[1,2,3] **Qinwen Ji**,[1,2,3] **and Xinying Guo**[1,2,3]

[1]*The Key Laboratory of Grain Information Processing and Control, Zhengzhou 450001, China*
[2]*The Henan Key Laboratory of Grain Photoelectric Detection and Control, Henan University of Technology, Zhengzhou 450001, China*
[3]*The College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, China*

Correspondence should be addressed to Chunhua Zhu; zhuchunhua@haut.edu.cn

Received 6 July 2022; Revised 27 October 2022; Accepted 16 November 2022; Published 28 November 2022

Academic Editor: Quanzhong Li

Existing channel estimation schemes for wideband systems generally estimate the channel matrix by pre-estimating the common support, which has a limited validity owing to the effect of beam squint. It is worth noting that the learned approximate message passing (LAMP) network need not pre-estimate the support and has obtained a relatively reliable estimation quality in the narrowband systems. In order to degrade the performance penalty caused by inaccurate estimation of the support in the wideband millimeter-wave MIMO systems, a prior-aided Gaussian mixture LAMP (GM-LAMP) network combining with residual learning is presented. Specifically, the multicarrier Gaussian mixture threshold shrinkage function is constructed for the GM-LAMP network, which can directly estimate the wideband beamspace channel while avoiding pre-estimating the support; then, considering the impact of channel noise and the coarse estimation error by the GM-LAMP, a residual network (ResNet) is designed to improve the estimation performance. Simulation results validate the efficiency of the proposed network (referred to as GLAMP-ResNet) with the lower computational complexity compared with the existing schemes.

1. Introduction

Beamspace channel estimation is essential for wideband millimeter-wave (mmWave) Multiple-Input Multiple-Output (MIMO) [1, 2] systems with lens antenna array to select a small number of dominant beams with larger path gains [3] and reduce the number of radio-frequency (RF) chains [4] for the lower hardware cost and power consumption. However, the effect of beam squint (Beam squint is used to imply that the support (i.e., the index set of non-zero elements in a sparse vector) of wideband beamspace channel tends to be frequency-dependent) [5] brings new challenges for the beamspace channel estimation, especially under the wider wideband and the larger number of antennas [6].

By exploiting the sparsity of beamspace channel, the beamspace channel estimation can be formulated as a sparse signal recovery problem [7]. Existing channel estimation schemes for wideband systems generally estimate the channel matrix under the common support assumption in frequency domain [8–10], which has a limited validity owing to the effect of beam squint and inaccurate estimation of frequency-dependent support. Specifically, a simultaneous orthogonal matching pursuit (SOMP)-based method is presented in [8]. Therein, the channel estimation is expressed as a multiple measurement vector problem with a common support (i.e., the channel support at different frequencies is assumed to be the same), which can be solved by the SOMP algorithm. In 2017, the orthogonal matching pursuit (OMP)-based algorithm [9] was proposed. It first estimates the support of the wideband beamspace channel at some frequencies independently by the OMP algorithm. Then, it combines them into the common support at all

frequencies. Unfortunately, the common support assumption in [8, 9] has limited validity in the practical wideband mmWave MIMO systems. To circumvent this problem, the successive support detection (SSD) algorithm [10] is proposed, where each support is uniquely determined by its spatial direction at the carrier frequency. Yet, it assumes a fixed sparsity for different frequencies, which will discard more nonzero elements of the components with greater sparsity.

Notedly, the studies in [8–10] adopt the scheme of pre-estimating support to estimate the wideband beamspace channel. However, therein, different frequencies share a common support and identical sparsity, which is not completely consistent with the actual channel. It is worth noting that the learned approximate message passing (LAMP) network need not pre-estimate the support and has obtained a relatively reliable estimation quality in the narrowband systems [11, 12]. If we can construct the threshold shrinkage function (TSF) of the LAMP network for wideband systems, the possibility of achieving satisfactory estimation accuracy will increase.

In this paper, a prior-aided Gaussian mixture LAMP (GM-LAMP)-based network combining with the residual network (ResNet) is presented, referred to as the GLAMP-ResNet. The contributions of this paper can be summarized as follows:

(1) The GM-LAMP network is proposed to obtain a preliminary estimation of the wideband beamspace channel. In particular, the multicarrier Gaussian mixture TSF (mcGM-TSF) is constructed in the GM-LAMP network. Different from the channel estimation methods in [8–10], the mcGM-TSF need not pre-estimate the support, which can degrade the performance penalty caused by inaccurate estimation of the support in the wideband millimeter-wave MIMO systems. Furthermore, the mcGM-TSF exploits the prior information that beamspace channel elements follow the Gaussian mixture distribution, which can improve the estimation performance compared with the traditional TSF of the LAMP network.

(2) The ResNet is designed to refine the preliminary estimation of the GM-LAMP network. Specifically, the ResNet contains several residual blocks and each consists of three convolutional layers, and the basic idea of ResNet is to learn the mapping from the output of GM-LAMP to the noise. After training, the ResNet can further reduce the impact of channel noise and refine the coarse estimation from the GM-LAMP network.

(3) Simulation results verify that the proposed GLAMP-ResNet can achieve better channel estimation performance by normalized mean squared error (NMSE), which has the comparable computational complexity, compared with the traditional LAMP network.

2. System Model

This paper considers an uplink wideband mmWave MIMO system with M subcarriers, as shown in Figure 1, where the base station (BS) employs a lens antenna array with N antennas and N_{RF} RF chains to serve K single-antenna users. To characterize the dispersive wideband mmWave MIMO channel, the Saleh–Valenzuela multipath channel model [13, 14] is adopted. The spatial channel vector $\mathbf{h}_m$ of a certain user at subcarrier m can be presented as

$$\mathbf{h}_m = \sqrt{\frac{N}{L}} \sum_{l=1}^{L} A_l e^{-j2\pi\tau_l f_m} \boldsymbol{\psi}\left(\eta_{l,m}\right), m = 1, 2, \cdots M, \tag{1}$$

where L is the number of paths, A_l and τ_l are the complex gain and the time delay of the l-th path, respectively, and $\eta_{l,m}$ is the spatial direction at subcarrier m that can be defined as $\eta_{l,m} = f_m/cd \sin\alpha_l$, where $f_m = f_c + f_s/M\,(m - 1 - M - 1/2)$ is the frequency of subcarrier m with f_c and f_s presenting the carrier frequency and bandwidth, respectively, c is the speed of light, α_l is the physical direction, d is the antenna spacing, and $\psi(\eta_{l,m}) = 1/\sqrt{N}[e^{-j2\pi b\eta_{l,m}}]_{b\in\mathbb{j}(N)}$ denotes the array steering vector, where $\mathbb{j}(N) = \{j - N + 1/2, j = 1, 2, \cdots, N\}$ [15].

The wideband mmWave MIMO system as Figure 1 employs a lens antenna array to convert spatial channel into beamspace channel, thereby this mmWave system is also often referred to as the beamspace MIMO system. The role of lens antenna array can be simulated by a spatial discrete Fourier transform matrix $\mathbf{U} \in \mathbb{j}^{N\times N}$, whose elements denote the array steering vectors of N predefined spatial directions. Specifically, the $\mathbf{U}$ can be expressed as $\mathbf{U} = [\psi(\overline{\eta}_1), \psi(\overline{\eta}_2), \cdots, \psi(\overline{\eta}_n), \cdots, \psi(\overline{\eta}_N)]^H$, where $\overline{\eta}_n = 1/N\,(n - N + 1/2)$, $n = 1, 2, \cdots N$ are the N spatial directions. Then, the beamspace channel vector $\widetilde{\mathbf{h}}_m$ of a certain user at subcarrier m can be expressed as $\widetilde{\mathbf{h}}_m = \mathbf{U}\mathbf{h}_m$.

In the uplink channel estimation, the orthogonal pilot transmission strategy is used and a certain user is considered without loss of generality. After Q instants of pilot transmission, the received pilot vector at BS after combining, CP removal, and M point FFT can be presented as

$$\mathbf{z}_m = \mathbf{W}\widetilde{\mathbf{h}}_m s + \mathbf{W}\mathbf{n}_m = \mathbf{W}\widetilde{\mathbf{h}}_m s + \mathbf{n}_m^{eff}, \tag{2}$$

where s is the transmitted pilot symbol, $\mathbf{n}_m \mathbb{j}\mathbb{j}(0, \sigma^2\mathbf{I}_N)$ is the noise vector with σ^2 representing the noise power, $\mathbf{n}_m^{eff}$ is the effective noise vector and $\mathbf{W}$ of size $P \times N$ ($P = QN_{RF}$) is the combining matrix, whose elements can be randomly selected from the set $1/\sqrt{P}\{-1, +1\}$ with equal probability if the adaptive selection network is realized by low-cost 1 bit phase shifters [16]. Combining all frequencies with assuming $s = 1$ [10], the overall measurement matrix can be obtained as

$$\mathbf{Z} = \mathbf{W}\widetilde{\mathbf{H}} + \mathbf{N}^{eff}, \tag{3}$$

where $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2 \mathbb{j} \cdots \mathbb{j}\mathbf{z}_M]$ is the measurement matrix, $\widetilde{\mathbf{H}} = [\widetilde{\mathbf{h}}_1, \widetilde{\mathbf{h}}_2, \cdots, \widetilde{\mathbf{h}}_M]$ is the beamspace channel of a certain user, and $\mathbf{N}^{eff} = [\mathbf{n}_1^{eff}, \mathbf{n}_2^{eff}, \cdots, \mathbf{n}_M^{eff}]$ is the effective noise.

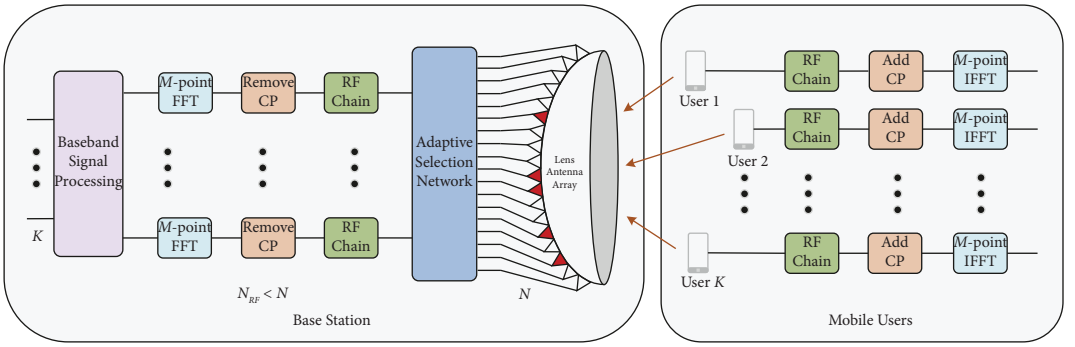

Figure 1: Architecture of the wideband mmWave MIMO system relying on lens antenna array.

3. Wideband Beamspace Channel Estimation

3.1. Pre-Estimation Based on GM-LAMP. In the proposed GM-LAMP network as shown in Figure 2, the overall measurement matrix $\mathbf{Z}$ will be inputted to the AMP network for training layer parameters under the Gaussian mixture channel prior assumption (Considering the previous contribution of modeling beamspace channel with the Gaussian mixture distribution [17]). For the T-layers GM-LAMP network, the t-th layer channel pre-estimation is as follows:

$$\widehat{\widetilde{\mathbf{H}}}'_{t+1} = \boldsymbol{\xi}_{gm}\left(\mathbf{r}_t; \boldsymbol{\theta}_t; \sigma_t^2\right), \tag{4}$$

where $\mathbf{r}_t = \widehat{\widetilde{\mathbf{H}}}'_t + \mathbf{A}_t\mathbf{u}_t$ is one of the layer parameters to be trained, $\mathbf{u}_t = \mathbf{Z} - \mathbf{W}\widehat{\widetilde{\mathbf{H}}}'_t + a_t\mathbf{u}_{t-1} + b_t\mathbf{u}^*_{t-1}$ for $t = 0, 1, \cdots, T-1$ with first-layer inputs $\widehat{\widetilde{\mathbf{H}}}'_0 = 0$, $\mathbf{u}_{-1} = \mathbf{0}$, $a_0 = 0$ and $b_0 = 0$. The $a_t\mathbf{u}_{t-1}$ and $b_t\mathbf{u}^*_{t-1}$ are called Onsager correction [18], which can be obtained by calculating the derivative of ξ_{gm} with respect to $\mathbf{r}_t$. It should be noted that $\mathbf{r}_t \in \mathbb{j}^{N\times K}$ contains the channel data of K users and each user has one carrier in narrowband mmWave systems [11]. However, a single user has multicarriers in wideband mmWave systems, so the dimension of $\mathbf{r}_t$ in (4) is $N \times M$ (i.e., $\mathbf{r}_t$ has the channel characteristics of M subcarriers). In this paper, the mcGM-TSF ξ_{gm} is constructed to make it suitable for wideband mmWave systems. The ξ_{gm} in (4) can be presented by

$$\boldsymbol{\xi}_{gm}\left(\mathbf{r}_t; \boldsymbol{\theta}_t; \sigma_t^2\right) = \frac{\sum_{g=0}^{G} p_{t,g}\widetilde{\mu}_{t,g}\left(\mathbf{r}_t\right)\mathbb{j}\mathbb{j}\left(\mathbf{r}_t; \mu_{t,g}, \sigma_t^2 + \sigma_{t,g}^2\right)}{\sum_{g=0}^{G} p_{t,g}\mathbb{j}\mathbb{j}\left(\mathbf{r}_t; \mu_{t,g}, \sigma_t^2 + \sigma_{t,g}^2\right)}, \tag{5}$$

where σ_t^2 is the noise variance in the t-th layer, G is the number of Gaussian variables following to the Gaussian mixture distribution, $p_{t,g}$, $\mu_{t,g}$,and $\sigma_{t,g}^2$, respectively, represent the probability, mean, and variance of the g-th Gaussian variable in the t-th layer; the set of previous distribution parameters $\theta_t = \left\{p_{t,g}, \mu_{t,g}, \sigma_{t,g}^2\right\}$ is a trainable matrix, generally called the shrinkage parameters, $\widetilde{\mu}_{t,g}\left(\mathbf{r}_t\right)$ and $\mathbb{j}\mathbb{j}\left(\mathbf{r}_t; \mu_{t,g}, \sigma_t^2 + \sigma_{t,g}^2\right)$ can be expressed as

$$\begin{aligned}\widetilde{\mu}_{t,g}\left(\mathbf{r}_t\right) &= \frac{\sigma_t^2\mu_{t,g} + \sigma_{t,g}^2\mathbf{r}_t}{\sigma_t^2 + \sigma_{t,g}^2} \\ \mathbb{j}\mathbb{j}\left(\mathbf{r}_t; \mu_{t,g}, \sigma_t^2 + \sigma_{t,g}^2\right) &= \frac{1}{\pi\sigma_{t,g}^2}e^{-\left(\mathbf{r}_t - \mu_{t,g}\right)*\left(\mathbf{r}_t - \mu_{t,g}\right)/\sigma_{t,g}^2},\end{aligned} \tag{6}$$

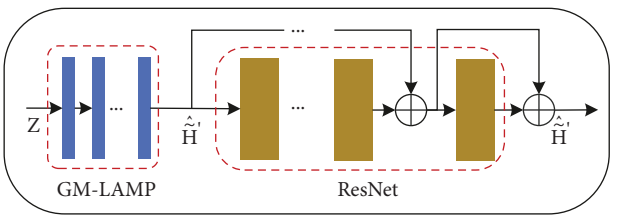

Figure 2: Network architecture of the GLAMP-ResNet.

where $\mathbb{j}\mathbb{j}\left(\mathbf{r}_t; \mu_{t,g}, \sigma_t^2 + \sigma_{t,g}^2\right)$ denotes the probability density function of the g-th Gaussian component in the t-th layer.

By constructing the mcGM-TSF $\xi_{g,m}$ as (5) in the GM-LAMP network, the pre-estimation of wideband beamspace channel can be obtained by (4). After T iterations, the final output of the GM-LAMP network can be written as $\widehat{\widetilde{\mathbf{H}}}' = \widehat{\widetilde{\mathbf{H}}}'_T$.

3.2. Residual Optimization Based on ResNet. Considering the impact of channel noise and the coarse estimation error by the GM-LAMP network, the ResNet is introduced to narrow the gap between the pre-estimated channel $\widehat{\widetilde{\mathbf{H}}}'$ and the label value $\widehat{\widetilde{\mathbf{H}}}$. Compared with some traditional convolutional neural networks, e.g., AlexNet, VGG, and GoogleNet, the ResNet has achieved promising performance by introducing fast links that directly pass the data flow to later layers, which can ease the training of the neural network. As a result, signal attenuation caused by multiple stacked nonlinear transformations can be effectively avoided and faster training speed can be achieved. Therefore, this paper introduces the ResNet as an important module to further improve the estimation performance. The proposed ResNet is composed of multilevel residual blocks, where each block consists of convolutional layers and activation function layers as shown in Figure 3.

In the offline training phase, $\widehat{\widetilde{\mathbf{H}}}'$ and $\widehat{\widetilde{\mathbf{H}}}$ are the input and label of the ResNet, respectively, and the corresponding residual optimization problem can be modelled as

$$\min L\left(\boldsymbol{\Psi}\right) = \sum_{n=1}^{N}\sum_{m=1}^{M}\left\|\widehat{\widetilde{\mathbf{H}}}_{n,m} - \widetilde{\mathbf{H}}_{n,m}\right\|_2^2, \tag{7}$$

where Ψ is the optimization variables, also the trainable parameters, and $\widetilde{\mathbf{H}}$ is the output of the ResNet.

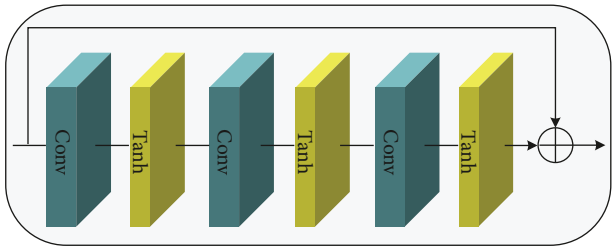

Figure 3: The structure of each residual block.

3.3. Structure of the Proposed GLAMP-ResNet. In the proposed GLAMP-ResNet scheme as shown in Figure 2, the number of Gaussian variables in the shrinkage function $\xi_{g,m}$ is set as $G = 4$, and the mean and variance of all Gaussian variables are both set as 0 considering the sparsity of the wideband beamspace channel. The convolution kernels of 7×7, 5×5, 3×3 and the features mapping of 64, 32, 1 are adapted, respectively, to the first, second, and third convolution layer of each residual block (The values of the hyper-parameters, such as the size of the convolution kernels and the number of feature maps are mainly chosen according to the scale and dimension of the dataset; furthermore, the trade-off between estimation accuracy and computational complexity also needs to be considered). In addition, the number of residual blocks is equal to 3 (considering an attractive trade-off between the computational complexity and system performance). Then, the output of the GLAMP-ResNet can be written as

$$\widetilde{\widehat{\mathbf{H}}} = \mathbb{j}_{\text{ResNet}}\left(\mathbb{j}_{\text{GM\_LAMP}}\left(\mathbf{A}_t, \boldsymbol{\theta}_t\right), \boldsymbol{\Psi}\right), \tag{8}$$

where $\mathbb{j}_{\text{ResNet}}(\cdot)$ and $\mathbb{j}_{\text{GM\_LAMP}}(\cdot)$ are the mapping function in the GM-LAMP network and ResNet, respectively.

4. Simulation Results

In our simulations, a lens antenna array with $N = 256$ and $N_{RF} = 8$ is used. The number of users and instants of pilot transmission is set as $K = 8$ and $Q = 16$, respectively. The channel parameters are set as follows [19, 20]: (1) $L = 3$; (2) $A_l \mathbb{j}\mathbb{j}\mathbb{j}(0, 1)$ for $l = 1, 2, 3$; (3) $\alpha_l \mathbb{j}\mathbb{j}\mathbb{j}(-\pi/2, \pi/2)$ for $l = 1, 2, 3$.

In order to better determine the number of GM-LAMP network layers (T) and the number of residual blocks (F), the NMSE performance comparison experiments have been conducted. The comparison results are shown in Figures 4 and 5, respectively. Relying on the classic AMP network, the number of T can be determined through the NMSE performance of the AMP network under different layers. It can be seen from Figure 4 that the NMSE performance can hardly be improved with increasing layers (T) when T is equal to 8, thereby the optimal number of T is determined as 8. Similarly, the number of residual blocks (F) can be determined by testing the NMSE performance under different number of residual blocks. It can be seen from Figure 5 that the improvement of NMSE performance reaches one limitation when the number of F is increased to 3, thereby the number of F is set to 3.

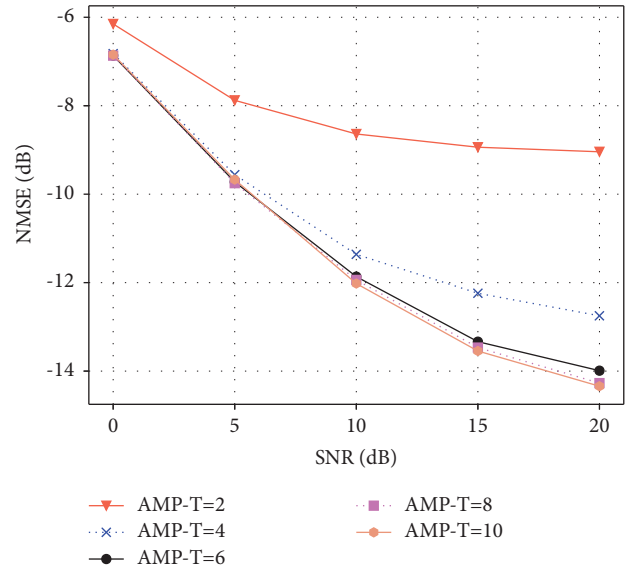

Figure 4: NMSE performance of AMP network under different layers (T).

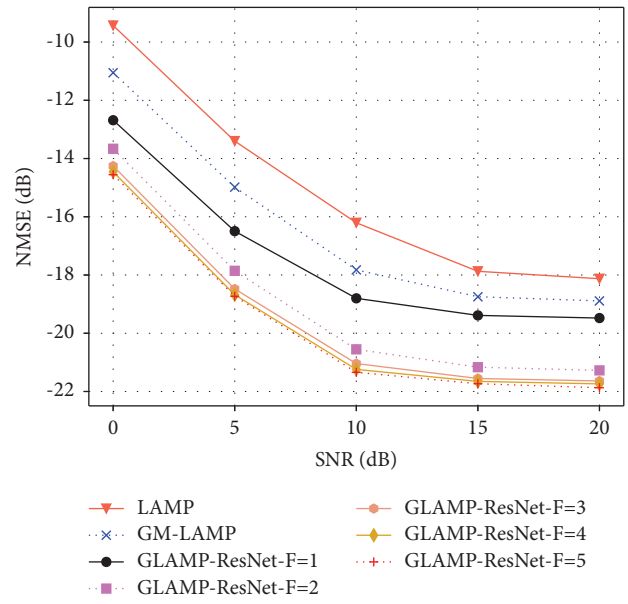

Figure 5: NMSE performance of GLAMP-ResNet under different number of residual blocks (F).

Figure 6 shows the NMSE performance of the proposed GLAMP-ResNet compared with the existing wideband beamspace channel estimation algorithms. It can be seen that the performance of the SOMP algorithm [8] and the OMP algorithm [9] are inferior, because they both assume that the support of sparse channel is the same in frequency domain. In the SSD algorithm [10] the support is differentiated by the spatial direction, thereby the corresponding performance is significantly improved. From another perspective, the AMP algorithm introduces the threshold shrinkage function to pass the approximate channel estimation among different layers of the AMP network;

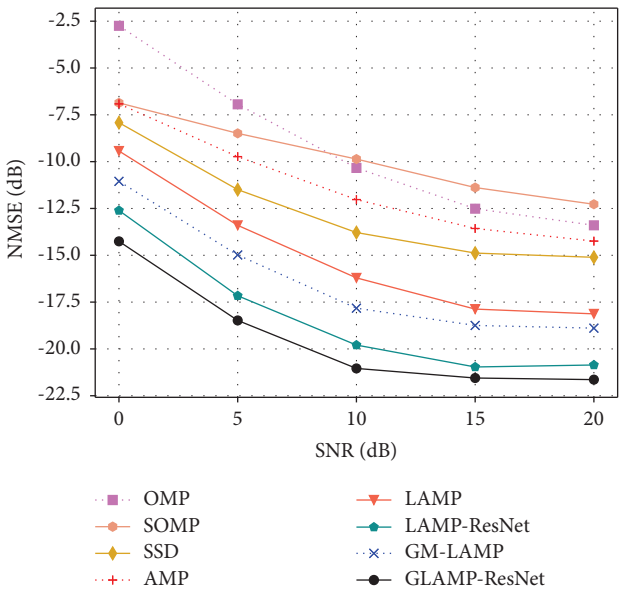

Figure 6: NMSE performance of several schemes.

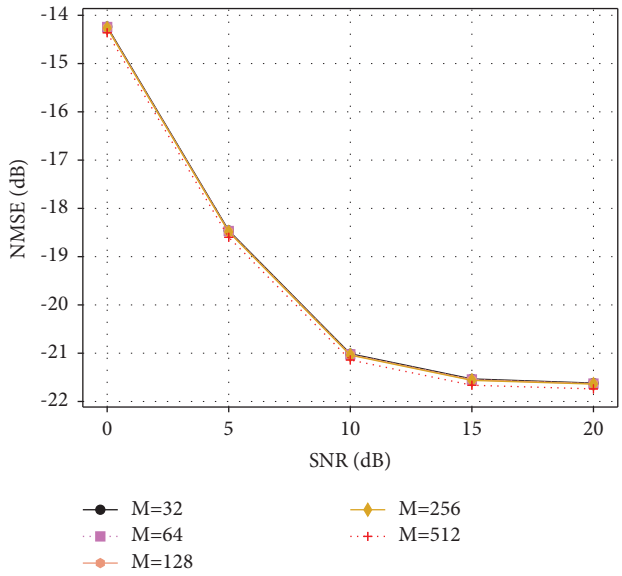

Figure 7: NMSE performance under different number of subcarriers (M).

however, the parameters of different layers are fixed, which induces the limited performance; in contrast, the Learned AMP (LAMP) can optimize the layer parameters by deep learning, and the NMSE performance can be improved by about 4 dB, which has been verified in this paper; then, the channel prior information of beamspace channel elements is exploited as the GM-LAMP algorithm, or the LAMP channel pre-estimation is combined with residual learning as the LAMP-ResNet, which can bring 1 dB and 3 dB performance improvement compared with the LAMP network, respectively; inspired by them, the GLAMP-ResNet is proposed with the NMSE up to 7 dB compared with the SSD algorithm.

To further verify the validity and universality of the GLAMP-ResNet, this paper tests the NMSE performance of the trained GLAMP-ResNet under different number of subcarriers (M) as Figure 7. It can be seen from Figure 7 that the NMSE performance is relatively better under $M = 512$, because the number of subcarriers in the training set is $M = 512$. When the number of subcarriers is changed during online estimation, it is found that the NMSE performance does not change significantly, which indicates that the NMSE performance of the GLAMP-ResNet is hardly influenced by the number of subcarriers. Besides, Figure 8 also verifies the stability of the proposed method under different bandwidth.

For the computational complexity, a single convolutional layer in the ResNet needs the computation as $\mathbb{j}(S^2V^2J_{\text{in}}J_{\text{out}})$ [21, 22], where S and V are the spatial lengths of the filter and the output feature map, respectively, J_{in} and J_{out} are the number of input and output channels, respectively, and the calculation time of the AMP and GM-LAMP is comparable to $\mathbb{j}(\text{TMN})$; here, T, M, and N are the number of AMP network layers, subcarriers, and antennas, respectively. Since the product $\mathbb{j}(\text{TMN})$ is far larger than

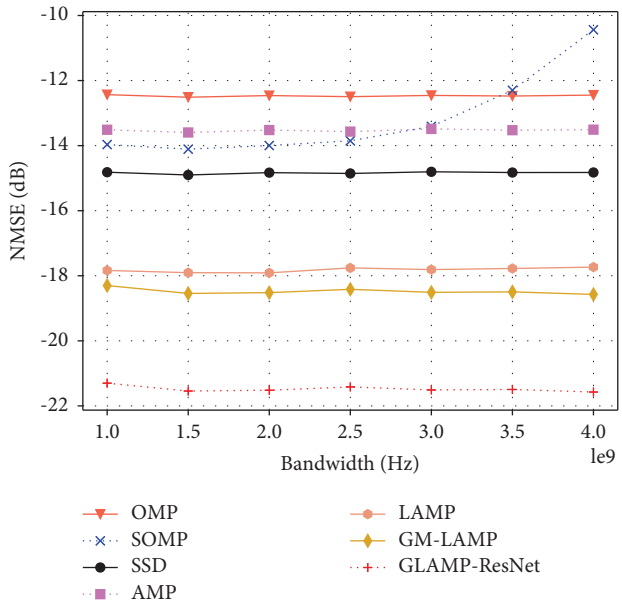

Figure 8: NMSE performance under different bandwidth with SNR = 15dB.

$\mathbb{j}(S^2V^2J_{\text{in}}J_{\text{out}})$ in mmWave MIMO systems, the computational complexity of the GLAMP-ResNet can be approximated as $\mathbb{j}(\text{TMN})$. By contrast, the complexity of both the OMP and SOMP algorithm can be presented by $\mathbb{j}(MPL^3\Omega^3) + \mathbb{j}(NMPL\Omega)$, and the complexity of the SSD algorithm is $\mathbb{j}(\text{NML}) + \mathbb{j}(MPL\Omega^2) + \mathbb{j}(MPL^2\Omega^2)$. Here, Ω (e.g., $\Omega = 4$) denotes the sparsity level of the beamspace channel. Thus, the proposed scheme can improve the system performance without significantly increasing the computational complexity.

5. Conclusion

This paper introduces one GLAMP-ResNet scheme for wideband beamspace channel estimation in mmWave MIMO systems, which mainly involves the following works: (1) constructing the multicarriers Gaussian mixture threshold shrinkage function for the GM-LAMP network to pre-estimate the wideband beamspace channel, not the support, thus the key challenge of channel estimation error from support estimation bias is avoided; (2) the strategy of pre-estimation combined with the second denoising can improve the final channel estimation accuracy without significantly increasing computation. Besides, the proposed channel estimation scheme can be applied in the other communication systems.

Data Availability

The simulation code and data used to support the findings of this study are available from the corresponding authors upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study is partly supported by the National Science Foundation of China: 61871176 and 61901159; the Applied Research Plan of Key Scientific Research Project for Henan Colleges and Universities: 22A510013.

References

[1] Y. Zhang, Y. Mu, Y. Liu, T. Zhang, and Y. Qian, "Deep learning-based beamspace channel estimation in mmwave massive mimo systems," *IEEE Wireless Communications Letters*, vol. 9, no. 12, pp. 2212–2215, 2020.

[2] L. You, Y. Huang, D. Zhang, Z. Chang, W. Wang, and X. Gao, "Energy efficiency optimization for multi-cell massive MIMO: centralized and distributed power allocation algorithms," *IEEE Transactions on Communications*, vol. 69, no. 8, pp. 5228–5242, 2021.

[3] C. Feng, W. Shen, and J. An, "Beam selection for wideband millimeter wave MIMO relying on lens antenna arrays," *IEEE Communications Letters*, vol. 23, no. 10, pp. 1875–1878, 2019.

[4] B. Ai, A. F. Molisch, M. Rupp, and Z.-D. Zhong, "5G key technologies for smart railways," *Proceedings of the IEEE*, vol. 108, no. 6, pp. 856–893, 2020.

[5] J. Brady and A. Sayeed, "Wideband communication with high-dimensional arrays: new results and transceiver architectures," in *Proceedings of the 2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 1042–1047, London UK, June 2015.

[6] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, and L. Hanzo, "Beamspace channel estimation for wideband millimeter-wave MIMO with lens antenna array," in *Proceedings of the 2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, Montreal, QC, Canada, May 2018.

[7] L. Dai, X. Gao, S. Han, I. Chih-Lin, and X. Wang, "Beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," in *Proceedings of the 2016 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1–6, Chengdu, China, July 2016.

[8] Z. Gao, C. Hu, L. Dai, and Z. Wang, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Communications Letters*, vol. 20, no. 6, pp. 1259–1262, June 2016.

[9] K. Venugopal, A. Alkhateeb, N. González Prelcic, 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.

[10] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, and L. Hanzo, "Wideband beamspace channel estimation for millimeter-wave MIMO systems relying on lens antenna arrays," *IEEE Transactions on Signal Processing*, vol. 67, no. 18, pp. 4809–4824, Sept. 2019.

[11] M. Borgerding, P. Schniter, and S. Rangan, "AMP-inspired deep networks for sparse linear inverse problems," *IEEE Transactions on Signal Processing*, vol. 65, no. 16, pp. 4293–4308, 2017.

[12] X. Wei, C. Hu, and L. Dai, "Deep learning for beamspace channel estimation in millimeter-wave massive MIMO systems," *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 182–193, Jan. 2021.

[13] J. Tao, C. Qi, and Y. Huang, "Regularized multipath matching pursuit for sparse channel estimation in millimeter wave massive MIMO system," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 169–172, Feb. 2019.

[14] L. Dai, B. Wang, M. Peng, and S. Chen, "Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 1, pp. 131–141, 2019.

[15] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, 2016.

[16] X. Gao, L. Dai, S. Han, and X. Wang, "Reliable beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6010–6021, 2017.

[17] C. Huang, L. Liu, C. Yuen, and S. Sun, "Iterative channel estimation using LSE and sparse message passing for mmWave MIMO systems," *IEEE Transactions on Signal Processing*, vol. 67, no. 1, pp. 245–259, 2019.

[18] T. Lin and Y. Zhu, "Beamforming design for large-scale antenna arrays using deep learning," *IEEE Wireless Communications Letters*, vol. 9, no. 1, pp. 103–107, 2020.

[19] L. Chen and X. Yuan, "Blind multiuser detection in massive MIMO channels with clustered sparsity," *IEEE Wireless Communications Letters*, vol. 8, no. 4, pp. 1052–1055, 2019.

[20] Y. Zeng and R. Zhang, "Millimeter wave MIMO with lens antenna array: a new path division multiplexing paradigm," *IEEE Transactions on Communications*, vol. 64, no. 4, pp. 1557–1571, 2016.

[21] Y. Wei, M. M. Zhao, M. Zhao, M. Lei, and Q. Yu, "An amp-based network with deep residual learning for mmwave beamspace channel estimation," *IEEE Wireless Communications Letters*, vol. 8, no. 4, pp. 1289–1292, 2019.

[22] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5353–5360, Boston, MA, USA, June 2015.