Channel Estimation Method Based on Transformer in High Dynamic Environment

1st Zhuolin Chen
College of Electronic Science and Technology
National University of Defense Technology
Changsha, China
chenzhuolin1996@163.com

2nd Fanglin Gu
College of Electronic Science and Technology
National University of Defense Technology
Changsha, China
gu.fanglin@nudt.edu.cn

3rd Rui Jiang
School of Computer Science
Central South University
Changsha, China
184612340@csu.edu.cn

***Abstract*—As the relative movement speeds of the communication parties increase, the Doppler frequency offset gradually increases, and the speed of channel state information(CSI) change also increases, which limits the performance of traditional channel estimation algorithms. To solve the above problems, we propose a channel estimation structure based on Transformer. Convolutional Neural Network (CNN) is used to extract the feature vectors of channel response and Transformer is used for channel estimation. Utilize Transformer's deep learning capabilities to better track channel variation characteristics in highly dynamic environments. By simulation, we get results that compared with traditional channel estimation methods, the performance of the proposed channel estimation method is improved significantly under high dynamic environment.**

***Index Terms*—OFDM; fast time-varying channel; channel estimation; deep learning**

I. INTRODUCTION

The concept of Orthogonal Frequency Division Multiplexing (OFDM) can be traced back to the 1960s. In the literature [1], Chang first proposed to divide a high speed serial input data into multiple parallel low speed data, and modulated them to many orthogonal carriers respectively, so as to eliminate inter-symbol interference by extending the transmission time cycle of parallel data. In the field of wireless communications, the demand for high-speed data transmission is growing rapidly. OFDM systems are usually used in slow fading channels to provide high-speed data transmission, that is, the channel remains unchanged within one OFDM symbol period [2], [3], and the subcarriers need to be strictly orthogonal to resist the frequency selective fading of the channels. The literature [4] mentioned the necessity of high reliability for modern life, and the problem of channel estimation under high dynamic environment has gradually become the focus of people.

In a high dynamic environment, the fast movement of the receiving end will cause produce significant Doppler frequency shift [5], which causes the subcarrier frequency shift in the OFDM system, resulting in the destruction of the orthogonality between the subcarriers, resulting in ICI, making the system bit error rate deterioration. OFDM has the disadvantage of being very sensitive to fast time-varying channel fading and Doppler frequency shift in high mobility systems [6]. In a highly dynamic environment, the perfect assumption of channel state information in low-mobility systems is invalid now. The fast time varying fading of the channel makes the estimation error of traditional channel estimation methods unavoidable.

This work is supported by the National Natural Science Foundation of China under Grant 61931020.

According to the judgment condition of whether the prior information is used in the channel estimation algorithm, it is divided into two types : estimation algorithm based on the reference signal and blind estimation algorithm. Blind estimation algorithms usually require the channel to remain stable over multiple symbols, so this is not suitable for high dynamic scenarios. In an OFDM system under high dynamic environment, due to the inherent characteristics of fast time-varying channels, reference signal-based estimation algorithms are usually used to obtain better estimation performance. Literature [7]–[9] propose a channel estimation method based on a linear interpolation channel model. The main idea is to assume that the time-domain channel follows a linear variation within a single OFDM symbol or multiple OFDM symbol periods. In literature [8], Mostofi and Cox proposed two channel estimation methods based on a piecewise linear channel model : one uses CP of OFDM symbols to obtain additional channel information, and the other uses three consecutive OFDM Correlation of channels between symbols, and proved that the system performance can be effectively improved under the environment of high delay. Reference [9] considers the joint channel estimation in time-frequency domain on the basis of reference [8], which further improves the system performance. However, as the speed of the communication parties increases, the Doppler frequency offset will also increase accordingly, and the channel changes within a single OFDM symbol will become more severe, which makes the modeling error value caused by the linear interpolation channel model is becoming larger and larger, directly leads to the decrease of the channel estimation accuracy.

In recent years, deep learning has been gradually used for the research of wireless communication system, and researchers at home and abroad have made some preliminary explorations. Some preliminary progress has been made in the study of wireless transmission technology and deep learning, including channel estimation [10], [11], signal detection [12], [13], CSI feedback and reconstruction [14], [15], and channel coding [16], [17]. For channel estimation in high dynamic environment, deep learning has not been fully studied. Therefore, in view of the insufficient estimation performance of traditional channel estimation methods under high dynamic environment, we propose a channel estimation algorithm based on Convolutional Neural Network (CNN) and Transformer, which uses CNN to extract data feature vectors and Transformer to estimate channel state information, and predicts the CSI at the symbol position of data through the CSI at the pilot position.

II. System Model

At present, OFDM is a relatively effective transmission technology, which is a kind of multi-carrier modulation. The channel divided into several orthogonal sub-channels, converts high-speed data signals into parallel low-speed sub-data streams, and modulates to each sub-channel for transmission, so its spectrum utilization ratio is higher than that of other multi-carrier systems. The lowpass equivalent subcarrier set in the OFDM system can be denoted as:

$$\left\{1, e^{j\Delta\omega t}, \ldots, e^{jk\Delta\omega t}, \ldots, e^{j(K-1)\Delta\omega t}\right\}, 0 < t < T \quad (1)$$

where K is the number of sub-carriers, T is the symbol period, $\Delta\omega = 2\pi/T$ represents the lowpass equivalent angular frequency of the subcarrier. The orthogonal characteristics between subcarriers can be expressed as:

$$\int_0^T e^{jm\Delta\omega t} e^{-jk\Delta\omega t} dt = \delta_{mk} = \begin{cases} 0; m \neq k \\ T; m = k \end{cases} \quad (2)$$

It is defined that $X_n(k)$ represents the signal transmitted on the k^{th} subcarrier in the n^{th} symbol. We assumed that the number of symbols in the OFDM system is N and the number of subcarriers is K, that is, $n = 1, 2..., N$ and $k = 1, 2, ..., K$. $X = [X_n, X_n, \ldots, X_n]^T$ indicates the signal to be sent within a symbol. IFFT is carried out at the sending side, and the lowpass equivalent signal modulated by the OFDM system is:

$$x(t) = \sum_{k=0}^{K} x_k e^{jk\Delta\omega t}, 0 < t < T \quad (3)$$

Add cyclic prefix to the signal, then through the channel, remove the cyclic prefix at the receiving end, carry out the Fourier transform, the received frequency signal domain is expressed as:

$$Y_n = H_n X_n + W_n, \quad (4)$$

where $X_n = [X_n(1), X_n(2), \ldots, X_n(K)]^T$ is transmission symbol, $Y_n = [Y_n(1), Y_n(2), \ldots, Y_n(K)]^T$ is receive symbol, H_n is frequency domain channel matrix, W_n is additive white Gaussian noise (AWGN).

When the communication speed is higher and the mobile speed is faster, the fast-changing characteristics of the channel are more obvious, and the channel parameters will change significantly within an OFDM symbol cycle. As shown in the Fig.1 and Fig.2, we simulated and analyzed the change of time-varying channel in an OFDM symbol. The simulation

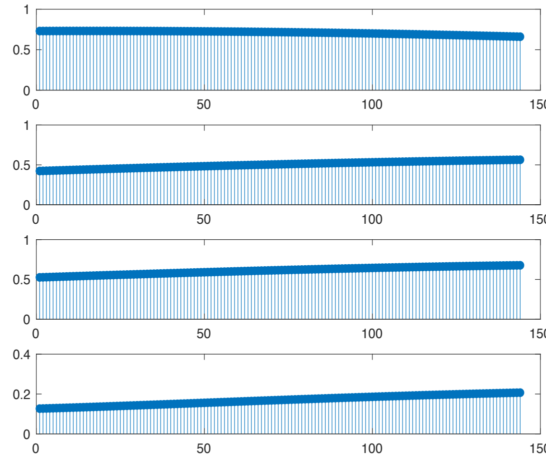

Fig. 1. Change of channel gain within an OFDM symbol when the normalized Doppler frequency offset is 0.01.

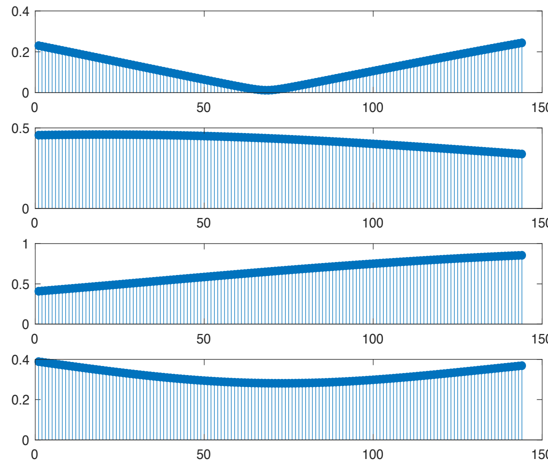

Fig. 2. Change of channel gain within an OFDM symbol when the normalized Doppler frequency offset is 0.2.

conditions are as follows: the symbol number is 100, the subcarrier number is 128, the sampling frequency is 2MHz, and the modulation mode is QPSK. The Fig.1 and Fig.2 show the variation of a tap component of the channel impulse response in the OFDM symbol at the normalized Doppler frequency of 0.01 and 0.2 respectively. It can be seen from the Fig.1 and Fig.2, due to the Doppler effect, the channel impulse response has obvious changes in the OFDM symbol. Also can be seen from the figures: if the normalized Doppler frequency is 0.01, the transformation of the channel impulse response in an OFDM symbol (block) approximately satisfies the linear characteristic, while if the normalized Doppler frequency is 0.2, the change of the channel impulse response in an OFDM symbol (block) no longer satisfies the linear rule. This is consistent with the viewpoints in the literature [3], [18]. The variation characteristics of the channel impulse response within an OFDM symbol (block) are critical to the research of channel estimation algorithms.

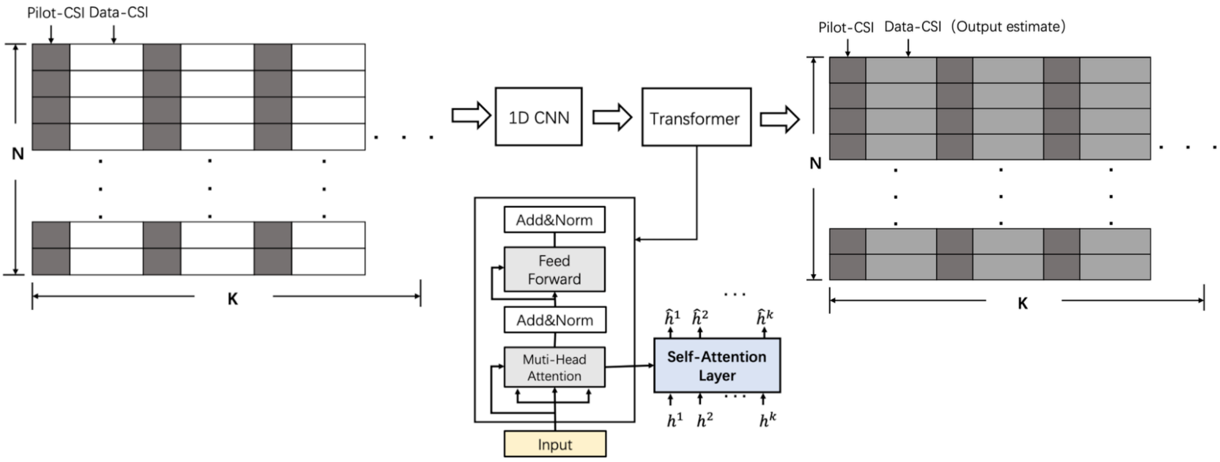

Fig. 3. Transformer-Estimation structurem.

III. Channel Estimation Method Based on Transformer

Because the transformation of the channel in a period of OFDM symbol no longer satisfies the linear feature, the traditional estimation method is no longer applicable. In this paper, a high dynamic channel estimation method based on deep learning is proposed, and a Convolutional Neural Network (CNN) feature vector and Transformer are used to track channel changes to obtain better channel estimates.

A. Pilot Position CSI Acquisition

We use an estimation algorithm based on the reference signal, insert pilot information at the sending end, and then obtain channel state information from the pilot at the receiving end based on the known pilot information. According to different pilot arrangements, the pilot structure can be divided into three types, namely block pilot, comb pilot and lattice pilot. In a highly dynamic environment, the channel has time-varying and frequency-selective characteristics. In order to better track the channel, this paper uses a block pilot structure. The channel estimation method adopts the DFT channel estimation method, which is a transform domain channel estimation method, which utilizes the characteristic that the time domain impulse response energy is concentrated on a few paths to reduce the noise of LS estimation.

B. Transformer-Estimation Structure

The Transformer-Estimation structure is shown in the Fig.3, which is divided into two parts: offline training and online prediction. In the offline training stage, a large amount of channel data is used to train the learning network. In the online learning stage, the CSI at the pilot position is taken as the network input, the CSI of data position is set to zero, and the output is a CSI matrix with the same dimension as the input. In 2017, Vaswani [19] proposed a sequence-encoding model Transformer based on self-attention mechanism, which includes two parts: encoder and decoder. We use the encoder part of it, which contains a six-layer superimposed Transformer-block. The Transformer-block structure is shown in the Fig.3. It mainly includes two parts: Muti-Head Attention and Feed Forward. Muti-Head Attention is self-attention structure.

1) 1D CNN is mainly used to extract features. A CNN network is composed of multiple convolution filters, each of which processes a group of data and uses sliding windows to carry out convolution summation. The input is the CSI matrix, $H = [h_1, h_2, \ldots, h_N] \in^{N\times K}$. The CSI of pilot position is obtained by DFT estimation method, and the CSI of data position is set to zero. The output dimension is unchanged, $H' = [h'_1, h'_2, \ldots, h'_N] \in^{N\times k}$. Set W as the convolution filter, then the transformation formula of CNN is:

$$h' = f(W * h + \varepsilon), \tag{5}$$

Where $\mathcal{E}$ is the bias and f is the activation function. ReLU function is selected in this paper.

2) Transformer structure is used for channel estimation, where the core structure is Self-Attention. The input is$H' = [h'_1, h'_2, \ldots, h'_N] \in^{N\times K}$, the Fig.4 is part of an OFDM symbol size CSI data processing procedure.

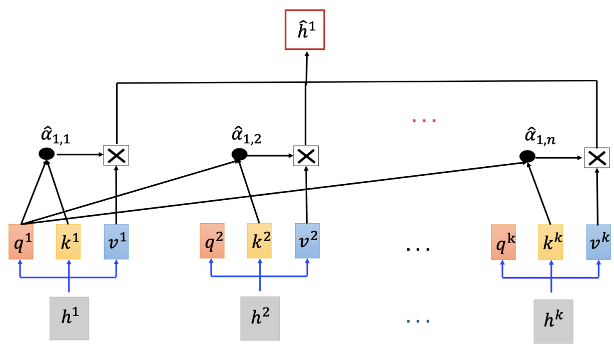

Fig. 4. data processing in self-attention.

The first step is to generate three vectors from h at each position, namely q, k, and v. They are as follows:

q: query (to match others)

$$q^i = W^q a^i, \tag{6}$$

k: key (to be matched)

$$k^i = W^k a^i, \tag{7}$$

v: information to be extracted

$$v^i = W^v a^i. \tag{8}$$

The second step is to take q^1 do attention to each key k, then get $\hat{\alpha}_{1,i}$:

$$\hat{\alpha}_{1,i} = q^1 \cdot k^i / \sqrt{d}, \tag{9}$$

where d is the dimension of q and k.

Finally, consider the CSI at all positions, and multiply $\hat{h}^1$ by v to get the first outputh $\hat{\alpha}_{1,i}$:

$$\hat{h}^1 = \sum_i \hat{\alpha}_{1,i} v^i. \tag{10}$$

So forth, to get the output of each positions, denoted as $\hat{h}_1 = \left[\hat{h}^1, \hat{h}^2, \ldots, \hat{h}^k\right]$.

The matrix calculation in Self-attention, as shown in Fig 5. Expressed as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \tag{11}$$

Each symbol h in $H' \in^{N \times k}$ is subjected to the above process to finally obtain the estimated value $\widehat{H} \in^{N \times k}$.

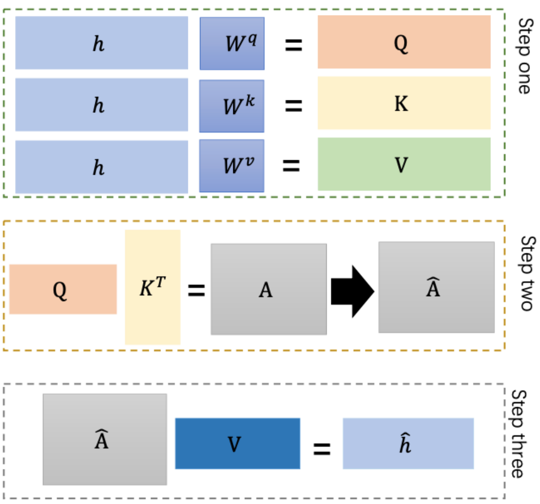

Fig. 5. Self-attention matrix calculation process.

C. Model Training

Let the transformation expression and all parameters of the entire Transformer network be $f_{est}(\cdot)$ and Θ_{est} respectively, so the CSI estimated by the Transformer network can be denoted as $\widehat{H} = f_{est}(H, \Theta_{est})$. We use the Adaptive torque estimation(Adam) algorithm to update the Transformer network parameter set. The loss function of this network is the Mean Squared Error (MSE), and the predicted loss of our model is:

$$L(\Theta_{est}) = \frac{1}{M} \sum_{i=1}^{M} \left(f_{est}(H_i, \Theta_{est}) - H_i^*\right)^2, \tag{12}$$

where, H_i^* is the supervision data, and M is the total number of samples in the training sample set.

IV. Simulation Evaluation

In this section, the Transformer channel estimation method is compared with the traditional frequency-domain channel estimation method to compare the performance of different estimation methods in different channel environments. The OFDM system parameters are shown in the table below. 3GPP-Extended Typical Urban model is selected for the channel model. In the Transformer network, the size of our training set, check set and test set are respectively 20000, 6000 and 3000.

TABLE I. Simulation Parameters

Parameters	Value
Bandwidth	3MHz
Number of subcarrier	400
Number of symbols	100
FFT size	512
Length of CP	32
Modulation	QPSK
Channel model	3GPP-Extended Typical Urban model

A. Estimated Mean Square Error

The Fig.6 compares the mean square error performance of frequency domain channel estimation methods LS estimation, DFT estimation, and the Transformer-based channel estimation method proposed in this paper in the presence of different normalized Doppler frequency offsets.

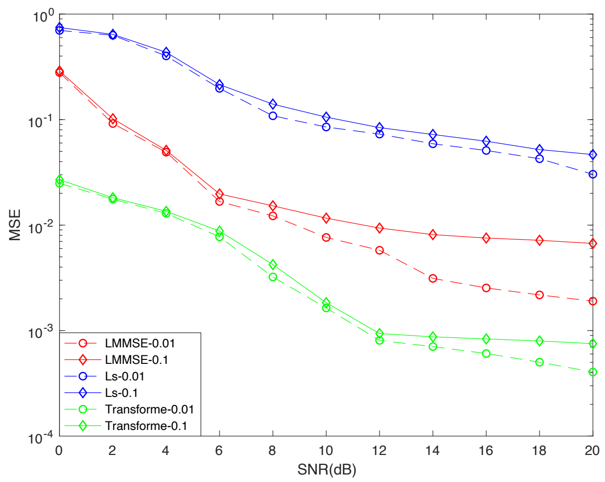

Fig. 6. MSE comparison of three estimation methods.

According to the simulation results, the LS estimation method shows poor MSE performance, and the main reason for its limited estimation accuracy is that the time domain correlation coefficient is a time varying parameter in the non-stationary channel environment, and the rule of channel variation does not meet the linear hypothesis. If the normalized Doppler frequency offset is 0.01, the Signal-Noise Ratio (SNR) gain of LMMSE is around 6dB higher than LS method. At this time, Our proposed channel estimation method is about 6dB higher than LMMSE. The performance of our proposed estimation method is improved significantly, mainly because the learning network can track the variation characteristics of the channel, so it can better adapt to the variation of the channel.

If the normalized Doppler frequency offset is 0.1, the MSE of the three methods increases compared with that of 0.01, and there is not much difference between the MSE performance curve of the LS method and that of the normalized Doppler frequency offset is 0.01, indicating that LS combined with linear interpolation method has poor estimation performance under the circumstance of larger or smaller normalized Doppler frequency offset. At this point, the MSE performance curve of LMMSE has an obvious upward trend when compared with the normalized Doppler frequency offset of 0.01, indicating that the LMMSE estimation method is no longer appropriate under the high dynamic environment. For Transformer estimation method, although the MSE is larger than that of the normalized Doppler frequency offset of 0.01, it is still smaller than the SNR gain of the traditional frequency domain channel estimation method in any case. To sum up, Our proposed method has the best MSE performance, indicating that our channel estimation network can well learn the channel variation characteristics under high dynamic environment.

B. Bit Error Rate

Bit error rate is a measure of the effect of different channel estimation methods on the overall performance of the system. The Fig.7 below compares the BER of LS estimation, LMMSE estimation and the Transformer channel estimation under different normalized Doppler frequency offset environments.

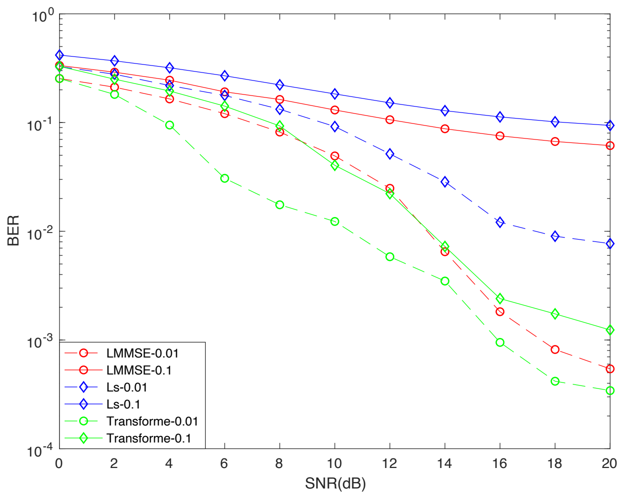

Fig. 7. BER comparison of three estimation methods.

As can be seen from the Fig.7, if the normalized Doppler frequency offset is 0.01, the BER performance of each estimation method is better than that if the normalized Doppler frequency offset is 0.1, indicating that the rapid change of channel under high dynamic environment has an impact on the performance of the channel estimation method. If the normalized Doppler frequency offset is 0.01, with the increase of the SNR, the difference of the bit error rate of the LS and LMMSE estimation methods becomes larger and larger, indicating that LMMSE estimation method still has good performance when the normalized Doppler frequency offset is small, but LS is not good. The reason is that due to the existence of Doppler frequency offset, the change rule of channel no longer meets the linear hypothesis. If the normalized Doppler frequency offset is 0.1, the BER performance of LS and LMMSE are both poor, and the performance of the channel estimation method recommended in this paper is significantly better than other estimation methods, which shows that the traditional channel estimation method is no longer in use under a highly dynamic environment, while the channel estimation method based on deep-learning can still show excellent BER performance.

V. Conclusion

In this paper, we propose a channel estimation method based on deep learning to solve the problems of high dynamic environment, such as the increase of doppler frequency offset, rapid channel change and limited estimation performance of traditional channel estimation algorithms. CNN is used to extract channel response feature vectors, and the Transformer has strong deep learning ability to track channel changes, so as to better adapt to the channel changes in high dynamic environment. The self-attention mechanism in Transformer enables the network to continuously learn more CSI locations to obtain better channel estimation results. The optimization of the computational complexity of the system will be the focus of later work.The simulation shows that the proposed method has higher estimation accuracy and robustness than the traditional channel estimation method.

Acknowledgment

This work is supported by the National Natural Science Foundation of China under Grant 61931020.

References

[1] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell System Technical Journal*, vol. 45, no. 10, pp. 1775–1796, 1966.

[2] None, "Robust ofdm receivers for dispersive time-varying channels: Equalization and channel acquisition," *IEEE Transactions on Communications*, vol. 52, no. 3, pp. 518–518, 2004.

[3] J. P. M. G. Linnartz and A. Gorokhov, "New equalization approach for ofdm over dispersive and rapidly time varying channel," vol. 2, pp. 1375–1379, 2000.

[4] M. Chen, X. Wei, J. Chen, L. Wang, and L. Zhou, "Integration and provision for city public service in smart city cloud union: Architecture and analysis," *IEEE Wireless Communications*, vol. 27, no. 2, pp. 148–154, 2020.

[5] L. Liu, C. Tao, J. Qiu, H. Chen, L. Yu, W. Dong, and Y. Yuan, "Position-based modeling for wireless channel on high-speed railway under a viaduct at 2.35 ghz," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 834–845, 2012.

[6] J. Wu and P. Fan, "A survey on high mobility wireless communications: Challenges, opportunities and solutions," *IEEE Access*, vol. 4, pp. 450–476, 2016.
[7] S. Sung and D. J. Brady, "Spectral spatial equalization for ofdm in time-varying frequency-selective multipath channels," pp. 434–438, 2000.
[8] Y. Mostofi and D. C. Cox, "Ici mitigation for pilot-aided ofdm mobile systems," *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 765–774, 2005.
[9] K. Kwak, S. Lee, H. Min, S. Choi, and D. Hong, "New ofdm channel estimation with dual-ici cancellation in highly mobile channel," *IEEE Transactions on Wireless Communications*, vol. 9, no. 10, pp. 3155–3165, 2010.
[10] 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.
[11] D. Neumann, T. Wiese, and W. Utschick, "Learning the mmse channel estimator," *IEEE Transactions on Signal Processing*, vol. 66, no. 11, pp. 2905–2917, 2018.
[12] 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.
[13] H. He, C. Wen, S. Jin, and G. Y. Li, "A model-driven deep learning network for mimo detection," pp. 584–588, 2018.
[14] T. J. Oshea, T. Erpek, and T. C. Clancy, "Deep learning based mimo communications.," *arXiv: Information Theory*, 2017.
[15] T. Wang, C. Wen, S. Jin, and G. Y. Li, "Deep learning-based csi feedback approach for time-varying massive mimo channels," *IEEE Wireless Communications Letters*, vol. 8, no. 2, pp. 416–419, 2019.
[16] T. Gruber, S. Cammerer, J. Hoydis, and S. T. Brink, "On deep learning-based channel decoding," pp. 1–6, 2017.
[17] S. Cammerer, T. Gruber, J. Hoydis, and S. T. Brink, "Scaling deep learning-based decoding of polar codes via partitioning," pp. 1–6, 2017.
[18] A. Gorokhov and J. . Linnartz, "Robust ofdm receivers for dispersive time-varying channels: equalization and channel acquisition," *IEEE Transactions on Communications*, vol. 52, no. 4, pp. 572–583, 2004.
[19] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," pp. 5998–6008, 2017.