Two-Stage LMMSE/DNN Receiver for High-Order Modulation

Zhaohui Huang, *Graduate Student Member, IEEE*, Dongxuan He, and Zhaocheng Wang, *Fellow, IEEE*

***Abstract*— High-order modulation has been seen as a promising technology to increase the spectral efficiency of wireless communications. Unfortunately, higher-order signals are susceptible to power amplifier (PA) nonlinearities and multi-path effects of the channel, leading to nonlinear distortion and severe inter-symbol interference. To guarantee the reliable detection of high-order signals, a two-stage receiver consisting of linear minimize mean squared error (LMMSE) equalizer and deep neural network (DNN) demodulator, namely TSLD receiver, is proposed, where LMMSE equalizer and DNN are deployed to eliminate inter-symbol interference and handle the amplifier nonlinearity, respectively. To facilitate the implementation of our proposed receiver, the specific pilot structure is designed, where low-order modulations are used for LMMSE equalizer and high-order modulations are used for DNN training. Simulation results show that our proposed TSLD receiver for high-order demodulation could recover the transmitted symbols effectively, even under serious multi-path and nonlinerity scenarios.**

***Index Terms*— High order modulation, deep learning, signal demodulation, nonlinear distortion.**

I. Introduction

BENEFITED from its high spectral efficiency and transmission capacity, high-order modulation has been regarded as a powerful technology, which is widely adopted in high-throughput satellite communications with limited bandwidth and regarded as a candidate uplink scheme in various wireless applications with high signal noise ratio (SNR) [1]. However, due to its sensitivity to signal quality, the reception performance of higher order modulation will degrade significantly in the face of low SNR and high nonlinearity scenarios [1]. The nonlinearity of power amplifier (PA) is related to the operating voltage, which will affect the energy efficiency. Specifically, the low working voltage in the load modulation region will lead to little nonlinearity but low energy efficiency, while the high working voltage will result in high energy efficiency but severe nonlinearity [2], [3]. The methodology to efficiently deal with the nonlinearity distortion while still reaping a high energy efficiency becomes an imperative challenge.

Manuscript received 15 February 2023; revised 19 April 2023 and 17 May 2023; accepted 27 May 2023. Date of publication 30 May 2023; date of current version 12 August 2023. This work was supported by the National Key R&D Program of China under Grant 2018YFB1801501. The associate editor coordinating the review of this letter and approving it for publication was A. Yadav. *(Corresponding authors: Zhaocheng Wang; Dongxuan He.)*

Zhaohui Huang and Zhaocheng Wang are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: hzh21@mails.tsinghua.edu.cn; zcwang@tsinghua.edu.cn).

Dongxuan He is with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: dongxuan_he@bit.edu.cn).

Digital Object Identifier 10.1109/LCOMM.2023.3281464

To reduce the multi-path effects, the linear minimize mean squared error (LMMSE) equalizer is usually adopted to estimate the signals in the metric of mean square error (MSE), which, however, cannot handle the nonlinearity of PA. The digital pre-distortion method could deal with the nonlinearity by adding a powerful pre-distortion device before the PA, but requires high-resolution radio frequency devices to get feedback of PA, which will consume much energy in transmitter [4]. Therefore, these traditional methods can not deal with the multi-path and nonlinearity effectively simultaneously.

Recently, with strong ability in fitting and scaling, deep learning has been used in multiple input multiple output (MIMO) detection, channel estimation, and coding [5]. Besides, there are also deep learning methods aiming at the modulation [6]. Although deep learning has great potential in fitting nonlinearity effects, it requires more pilot data, large training and has less adaptation to dynamic channel environment when compared with traditional methods, and may lead to overfitting or degeneration problem [7].

To combine the advantages of both traditional and learning methods, a two-stage receiver consisting of LMMSE equalizer and DNN demodulator (TSLD) is proposed in this letter. Specifically, the received signals are firstly processed by the LMMSE equalizer, which aims at eliminating the multi-path effects, and then handled by a DNN, which aims at eliminating the nonlinearity distortion and obtaining the demodulated signals. Compared with traditional LMMSE method, it handles the nonlinearity of PA more efficiently. Compared with traditional deep learning method, our proposed two-stage method utilizes the knowledge of multi-path effect and relieves the fitting burden of neural network. Simulation results are provided to validate the high-performance of our proposed TSLD receiver.

II. System Model

In this work, a single-input single-output (SISO) equivalent baseband model is considered. The transmitted symbol $s(t) \in \left\{1, 2, \cdots, 2^N\right\}$ is modulated to a complex-value baseband symbol $x(t) = x_I(t) + \mathrm{j}\, x_Q(t), t = 0, 1, 2, 3, \cdots$, which is converted to a carrier signal by up-converter and amplified by PA. Hereby, N is the modulation order, $x_I(t)$ and $x_Q(t)$ denote the in-phase (I) and quadrature (Q) branches at time t, respectively. In order to improve the power efficiency, PA is assumed to operate in the nonlinear region, which makes the signal suffer from strong amplitude compression and phase rotation. The distortion of PA is formulated by the odd memoryless polynomial model, given by

$$z(t)=\sum_{k=1}^{K}\alpha_{2k-1}(t)x(t)|x(t)|^{2(k-1)}, \tag{1}$$

where $z(t)$ is the transmitted signal, $\alpha_{2k-1}(t)$ are the model parameters and $|\cdot|$ represents the absolute value of a complex scalar [8].

The multi-path channel impulse response $h(t)$ can be modeled as

$$h(t)=h_0e^{j\theta_0}\delta(t)+\sum_{k=1}^{M}(\gamma h_0 e^{-b\tau_k+j\theta_k})\delta(t-\tau_k), \tag{2}$$

where $\delta(\cdot)$ is the Dirac delta function, h_0 and θ_0 represent the amplitude and phase of the first path, respectively, M is the number of paths, γ describes the decrease of multi-path components, θ_k is the phase of the k-th multi-path component and b is the decay exponent [9].

The received signal can be expressed as

$$r(t)=h(t)*z(t)+n(t), \tag{3}$$

where $n(t)$ is the Guassian noise, and $*$ represents the convolution operation.

Due to the coupling of amplifier nonlinearity and multi-path effect, the received signal $r(t)$ is hard to be demodulated directly. To solve this problem, as shown in Fig. 1, a TSLD receiver consisting of LMMSE equalizer and DNN demodulator is proposed, where the LMMSE equalizer focus on eliminating multi-path effect while the DNN focus on counteracting the nonlinearity caused by PA. Subsequently, Viterbi-decision or trellis-based demodulator will be utilized to obtain $\hat{s}(t)$ which denotes the estimate of transmitted symbols $s(t)$.

III. Proposed Two-Stage Receiver

The architecture of the proposed TSLD receiver will be illustrated in this section. To train the LMMSE equalizer and DNN respectively, a two-stage training method, which includes channel estimation phase and DNN training phase, is required. The input of LMMSE equalizer is defined as $r(t)$ and the output is defined as $\hat{z}(t)$, i.e. the estimate of $z(t)$. The input of DNN, which is also the output of LMMSE equalizer, is $\hat{z}(t)$ while the output, as the estimate of $x(t)$, is denoted by $\hat{x}(t)$.

A. LMMSE Equalizer

To reduce the negative effective of multi-path, MSE, i.e. $\mathbb{E}[(\hat{z}(t)-z(t))^2]$, is selected as the metric of channel estimation, where $\mathbb{E}[\cdot]$ denotes the expectation. The LMMSE output $\hat{z}(t)$ is a linear convolution of the input signal $r(t)$:

$$\hat{z}(t)=c(t)*r(t)=\sum_{i=-L}^{L}c(i)r(t-i), \tag{4}$$

where $2L+1$ is the order of equalizer and $c(t)$ is a complex scalar as the coefficient of the LMMSE equalizer for $t=-L,-L+1,\cdots,L$.

Define $\boldsymbol{r}^T=[r(t+L),r(t+L-1),\cdots,r(t-L)]$ as the vector of inputs to the equalizer and $\boldsymbol{c}^T=[c(-L),c(-L+1),\cdots,c(L)]$ as the vector of coefficients of LMMSE equalizer. Then we have $\hat{z}(t)=\boldsymbol{r}^T\boldsymbol{c}$, and $\mathbb{E}[(\hat{z}(t)-z(t))^2]=\mathbb{E}[\boldsymbol{c}^T\boldsymbol{r}\boldsymbol{r}^H\boldsymbol{c}^*-2\Re\{\boldsymbol{r}^H\boldsymbol{c}^*z(t)\}+|z(t)|^2]$, where $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ denote the transpose, conjugate and conjugate transpose, respectively.

By differentiating the MSE with respect to $\boldsymbol{c}$ and setting the result to zeros, the optimal tap weights are given by:

$$\boldsymbol{c}_{opt}=(\mathbb{E}[\boldsymbol{r}^*\boldsymbol{r}^T])^{-1}\mathbb{E}[\boldsymbol{r}^*z(t)], \tag{5}$$

where $\boldsymbol{c}_{opt}=[c(-L)_{opt},c(-L+1)_{opt},\cdots,c(L)_{opt}]^T$.

By taking z-transforms, we obtain

$$\mathcal{Z}(c(t)_{opt})=\frac{1}{\mathcal{Z}(h(t))+1/SNR}, \tag{6}$$

where $\mathcal{Z}(\cdot)$ represents z-transformation and $SNR=\mathbb{E}[z(t)^*z(t)]/\mathbb{E}(n(t)^*n(t))$ [10].

LMMSE equalizer only focuses on eliminating the inter-symbol interference. However, the nonlinearity of PA is a kind of distortion of each symbol rather than inter-symbol interference, which could not be handled by LMMSE equalizer efficiently. Therefore, the performance of LMMSE is limited when the PA nonlinearity becomes dominant.

B. DNN Demodulator

Once the LMMSE equalizer is well-trained, the inter-symbol interference could be compensated. Then, DNN is trained to eliminate the nonlinear distortion caused by power amplifier.

1) DNN Structure: A feedforward network consisting of multiple fully-connected layers is utilized to demodulate the distorted signals. Since the input and output of DNN should be real-valued, the real and imaginary part of the compensated signal are input into DNN. At time t, the input and output of DNN can be expressed as two-dimension vectors $[\Re(\hat{z}(t)),\Im(\hat{z}(t))]^T$ and $[\Re(\hat{x}(t)),\Im(\hat{x}(t))]^T$, respectively, where $\Re(\cdot)$ and $\Im(\cdot)$ represent taking the real and imaginary part of a complex-valued scalar, respectively.

The DNN consists of L hidden layers to obtain strong fitting ability. Besides, the number of hidden layers should not be too large to avoid overfitting and degeneration, or too small to avoid underfitting. The output of a hidden layer is the input of the subsequent layer. The output of i-th hidden layer $\boldsymbol{y}_i$ could be expressed by

$$\boldsymbol{y}_i=f(\boldsymbol{W}_i*\boldsymbol{y}_{i-1}+\boldsymbol{b}_i), \tag{7}$$

where $\boldsymbol{W}_i$ and $\boldsymbol{b}_i$ are the weight matrix and bias of the i-th hidden layer, respectively, and $f(\cdot)$ is the activation function, which is nonlinear and usually adopted by *ReLU* or *Sigmoid*.

2) DNN Training: Let $\boldsymbol{W}=[\boldsymbol{W}_1,\boldsymbol{W}_2,\cdots,\boldsymbol{W}_L,\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_L]$ be the parameters that need to be updated in the training process of DNN. To get the accurate estimation of $x(t)$, the mean square error between $x(t)$ and $\hat{x}(t)$ could be used as the loss function to train the DNN, which is defined as

$$J(\boldsymbol{W})=\mathbb{E}[(\hat{x}(t)-x(t))^2]. \tag{8}$$

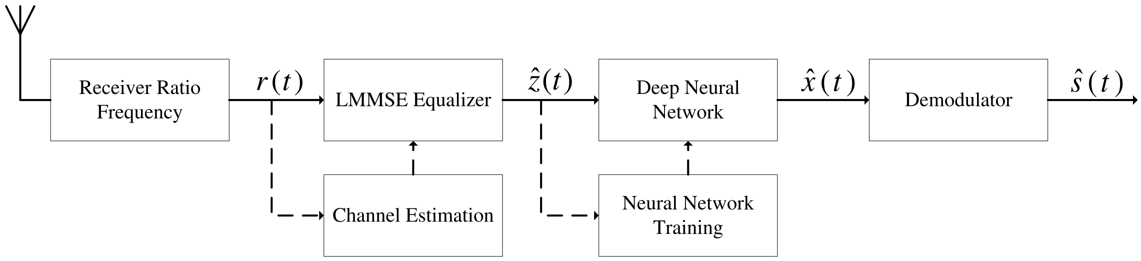

Fig. 1. Illustration of the proposed TSLD receiver.

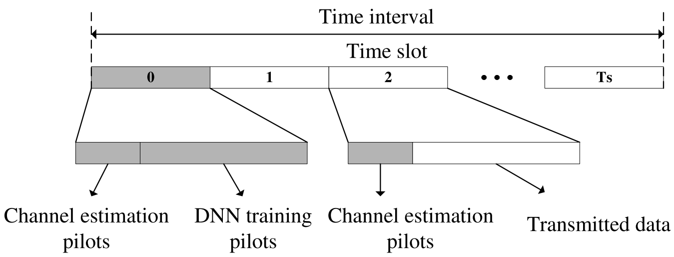

Fig. 2. Pilot structure.

Adam [11] or gradient descent method could be used to train DNN by updating parameters $\boldsymbol{W}$ until the loss function $J(\boldsymbol{W})$ is smaller than a predefined threshold or the number of iteration is satisfied. Once the DNN is well-trained, it could be adopted to eliminate the nonlinearity in the data transmission phase.

IV. Pilot Structure

Since the variation of amplifier nonlinearity (which are mainly determined by the characteristics of hardware and temperature) is much slower than the change of wireless channel, a hierarchical pilot structure is proposed as shown in Fig. 2. Specifically, within a fixed time interval, the characteristic of PA distortion is assumed to be unchanged. In addition, the time interval is divided into $1+T_s$ time slots, where each time slot is much shorter than the coherent time of wireless channel. Among them, time slot 0 consists of all training symbols to estimate the channel and train DNN, while other time slots only focus on channel estimation and data transmission.

Specifically, time slot 0 contains two parts, namely, channel estimation pilots and DNN training pilots. It is worth noting that channel estimation (or equivalently, training of LMMSE equalizer) must be executed before DNN training due to the fact that the inter symbol interference will deteriorate the demodulation ability of DNN. Other time slots consist of two parts, namely, channel estimation pilots and transmitted data, respectively. Due to fast channel variations, channel state information needs to be re-estimated in each time slot first.

Furthermore, the channel estimation pilots are made of low-order modulation signals such as binary phase shift keying (BPSK) or quadrature phase shift keying (QPSK) in order to estimate channel more accurately and energy-efficiently. At the same time, the DNN training pilots are made of high-order modulation signals to acquire the behavior of PA.

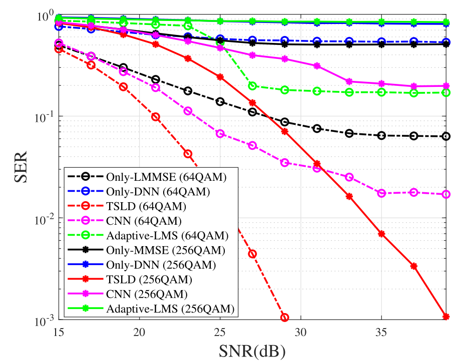

Fig. 3. SER performance versus SNR with different receivers for 64/256QAM.

V. Simulation Results and Discussions

Simulation results are presented to verify the effectiveness of our proposed TSLD methodology. The order of nonlinearity of PA is $2K-1=5$ with $\alpha_1(t)=1.8206+j0.2987$, $\alpha_3(t)=-1.2729+j0.3561$ and $\alpha_5(t)=0.3231-j0.1578$ [2]. As for the multi-path channel, the number of received path M is uniformly sampled from 3 to 6, τ_k is independently uniformly sampled from 1 to 15 and $b=1/2, \gamma=1/2$. The phases θ_k of all multi-path components independently follow uniform distribution.

The receiver schemes are detailed as follows:

1) Only-LMMSE: Only frequency-domain LMMSE equalization is executed to eliminate the multi-path effect.
2) Only-DNN: Only feedforward neural network is utilized to deal with multi-path and nonlinearity [12], whose input is a few taps of received signals and output is the recovered signals.
3) CNN: A convolutional neural network is utilized to recover the signals [13].
4) Adaptive-LMS: A linear least mean square (LMS) adaptive equalizer is executed to recover the signals, where the parameters are updated during data transmission [10].
5) TSLD: Our proposed two-stage receiver consisting of LMMSE equalizer and DNN demodulator.

The neural network used in TSLD has 6 hidden layers and the neurons in each layer are (16, 128, 128, 128, 128, 16), while the DNN in TSLD has 7 hidden layers and the neurons in each layer are (128, 200, 200, 200, 200, 128, 64). CNN consists of 5 convolutional layers and a fully-connected layer, where the output channels and kernel seizes

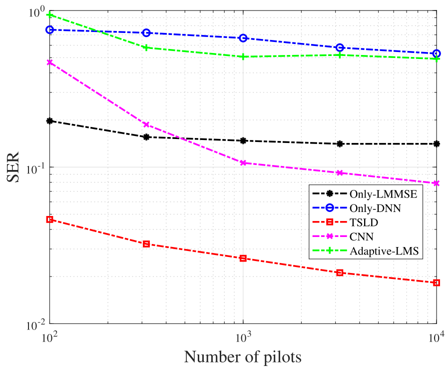

Fig. 4. SER performance versus number of pilots with different receivers for 64QAM.

of convolutional layers are (10, 20, 40, 10, 5) and (3, 3, 3, 1, 3), respectively. For all five schemes, 10^4 uniformly and randomly generated symbols are adopted to train the receiver. In TSLD, 5000 symbols are utilized to train LMMSE and other 5000 symbols are utilized to train DNN in TSLD. In Fig. 3, the symbol error rate (SER) versus SNR, defined as E_s/N_0, for different schemes with various high order modulations, including 64QAM and 256QAM, where E_s is the average power of the symbol and N_0 is the power of the Guassian noise.

When 64QAM is transmitted, it is evident that the proposed TSLD performs much better than all other benchmarks when the SNR is higher than 20dB. For example, at the SER level of 10^{-1}, TSLD has around 6dB gain in SNR compared with Only-LMMSE. Besides, compared to our proposed TSLD, all other schemes show high SER floors, demonstrating that they cannot handle nonlinearity and multi-path effect simultaneously. Particularly, for 256QAM, Adaptive-LMS and Only-DNN approach the BER floor at very low SNR. In addition, with higher modulation order, the performances of all five schemes degrade due to possible overlapping in the constellation points, but TSLD still has the best performance.

The effect of the number of pilots on SER performance is demonstrated in Fig. 4, where SNR is set as 25 dB and 64QAM is adopted. It is evident that the performance of TSLD obtains the optimal SER performance regardless of the number of pilots, which improves with the increase of pilots. Meanwhile, the performance of Only-LMMSE does not change with different number of pilots, since it could obtain the optimal parameters with limited number of pilots. The performance of CNN degrades rapidly when the number of pilots is less than 10^3, due to the overfitting problem introduced by insufficient training data.

The detailed comparison concerning the computational complexity, training overhead, and training complexity is presented in Table I. Specifically, N_{tap}, N_{DNN}, and N_{CNN} represent the number of taps in traditional methods, total neurons in DNN and CNN, respectively. N_c and N_n are the number of pilots for channel estimation and nonlinearity estimation, respectively. Usually, the neurons of networks are much larger than the number of taps in traditional methods, i.e. $N_{DNN}, N_{CNN} \gg N_{tap}$, and the pilots for channel estimation could be much less than the pilots for nonlinearity estimation, i.e. $N_c \ll N_n$. It can be observed from Table I and Fig. 3, Adaptive-LMS possesses the lowest computational complexity, training overhead and training complexity but has the worst SER performance. Our proposed TSLD outperforms CNN and DNN in SER performance, training overhead and training complexity. Compared to Only-LMMSE, TSLD offers a significant performance improvement at an acceptable cost in terms of computational complexity, training overhead and training complexity. Besides, it is apparent that the training complexity and training overhead of TSLD method are close to Only-LMMSE method when T_s becomes large.

The performances of Only-LMMSE, Only-DNN and TSLD schemes with different nonlinearity levels is demonstrated in Fig. 5. The nonlinearity levels are measured by PA voltages, which are normalized by the linear maximum voltage of PA. It is evident that performances of both TSLD and Only-LMMSE degrade with the increasing of PA voltage, due to more severe nonlinearity caused by higher PA volt-

TABLE I

PERFORMANCE COMPARISON OF DIFFERENT SCHEMES

Schemes	Computational complexity	Training overhead	Training complexity
Only-LMMSE	N_{tap}^3	N_c	N_{tap}^3
Only-DNN	N_{DNN}	N_n	N_{DNN}^2
TSLD	$N_{tap}^3 + N_{DNN}$	$N_c + \frac{N_n}{T_s}$	$N_{tap}^3 + \frac{N_{DNN}^2}{T_s}$
CNN	N_{CNN}	N_n	N_{CNN}^2
Adaptive-LMS	N_{tap}	N_c	N_{tap}

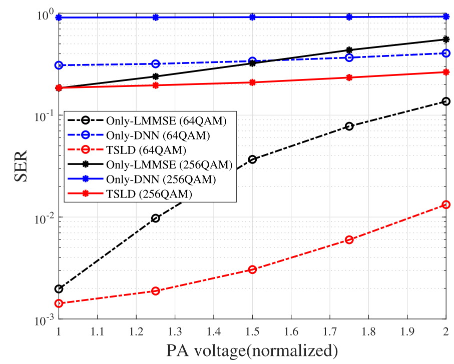

Fig. 5. SER performance versus PA voltage with different receivers for 64/256QAM.

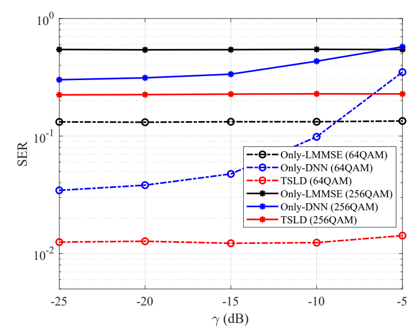

Fig. 6. SER performance versus attenuation coefficient γ with different receivers for 64/256QAM.

age. In addition, TSLD outperforms Only-LMMSE along the whole PA voltage range, which is even more significant with the increase of PA voltage. Furthermore, SER gap between Only-LMMSE and TSLD is larger with higher modulation order.

The performances of Only-LMMSE, Only-DNN and TSLD schemes with different attenuation coefficients γ are demonstrated in Fig. 6. It can be seen that the performances of TSLD and Only-LMMSE are stable with varying γ, while the performance of Only-DNN decreases with the increase of multi-path effect, which implies that the Only-DNN scheme could not deal with the strong multi-path effect effectively with limited training data.

VI. Conclusion

To improve the demodulation performance of high-order modulation in the presence of multi-path effect and amplifier nonlinearity, a two-stage receiver consisting of LMMSE equalizer and DNN demodulator is proposed firstly. In particular, LMMSE equalization is executed to eliminate the inter-symbol interference caused by the multi-path effect and DNN is then adopted to handle the nonlinearity. In addition, we proposed the corresponding pilot structure to facilitate the operation of our proposed TSLD receiver according to the time-variant channel characteristic and PA features, where low-order modulations are used for LMMSE equalizer and high-order modulations are used for DNN training. Simulation results have demonstrated that our proposed TSLD receiver outperforms its LMMSE-only and DNN-only counterparts.

For future research, the alternative methods could be explored to replace either LMMSE or DNN in our proposed two-stage architecture for obtaining better performance.

References

[1] R. Balasubramanian and M. Miller, "Pre-distortion techniques and bandwidth efficient modulation for military satellite communications," in *Proc. IEEE MILCOM. Mil. Commun. Conf.*, Nov. 2004, pp. 1028–1033.

[2] H. Lai and Y. Bar-Ness, "Minimum distortion power polynomial model (MDP-PM) of nonlinear power amplifiers and its application on analog predistorters," in *Proc. IEEE VTC*, vol. 3, Sep. 1999, pp. 1501–1505.

[3] Y. Chen, Y. Wang, J. Zhang, P. Zhang, and L. Hanzo, "Reconfigurable intelligent surface (RIS)-aided vehicular networks: Their protocols, resource allocation, and performance," *IEEE Veh. Technol. Mag.*, vol. 17, no. 2, pp. 26–36, Jun. 2022.

[4] H.-J. Wu et al., "A wideband digital pre-distortion platform with 100 MHz instantaneous bandwidth for LTE-advanced applications," in *Proc. Workshop Integr. Nonlinear Microw. Millimetre-Wave Circuits*, Sep. 2012, pp. 1–3.

[5] E. Nachmani, E. Marciano, L. Lugosch, W. J. Gross, D. Burshtein, and Y. Be'ery, "Deep learning methods for improved decoding of linear codes," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 119–131, Feb. 2018.

[6] Y. Fan, Y. Li, T. Chai, and D. Ding, "Satellite data transmission method for deep learning-based AutoEncoders," in *Proc. Int. Conf. Inf. Commun. Technol. Disaster Manag. (ICT-DM)*, Dec. 2021, pp. 38–42.

[7] H. Ye, L. Liang, G. Y. Li, and B. Juang, "Deep learning-based end-to-end wireless communication systems with conditional GANs as unknown channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3133–3143, May 2020.

[8] T. Schenk, *RF Imperfections in High-Rate Wireless Systems: Impact and Digital Compensation*. Berlin, Germany: Springer, 2008.

[9] J. Yuan, X. Yang, Z. Tian, S. Qiu, Q. Wu, and B. Wang, "A multipath delay model for indoor ultra-wideband systems," in *Proc. Int. Conf. Microw. Millim. Wave Technol. (ICMMT)*, vol. 5, May 2012, pp. 1–4.

[10] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[11] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. IEEE ICLR*, May 2015, pp. 1–15.

[12] L. Fang and L. Wu, "Deep learning detection method for signal demodulation in short range multipath channel," in *Proc. IEEE 2nd Int. Conf. Opto-Electronic Inf. Process. (ICOIP)*, Jul. 2017, pp. 16–20.

[13] X. Feng, Y. Wang, T. Qi, and Y. Chen, "The design of satellite equalizer based on CNN," in *Proc. IEEE 8th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC)*, May 2019, pp. 1647–1651.