Correspondence

Intelligent and Reliable Deep Learning LSTM Neural Networks-Based OFDM-DCSK Demodulation Design

Lin Zhang, *Senior Member, IEEE*, Haotian Zhang, Yuan Jiang, *Member, IEEE*, and Zhiqiang Wu, *Senior Member, IEEE*

***Abstract*—Chaos communications have widely been applied to provide secure, and anti-jamming transmissions by exploiting the irregular chaotic behavior. However, the real-valued chaotic sequences imposed on the information induce interferences to the user data, thereby leading to reliability performance degradations. To address this issue, in this paper, we propose to utilize the intelligent, and feature extraction capability of the deep neural network (DNN) to learn the transmission patterns to demodulate the received signals. In our design, we propose to construct the long short-term memory (LSTM) unit-aided intelligent DNN-based deep learning (DL) demodulator for orthogonal frequency division multiplexing-aided differential chaos shift keying (OFDM-DCSK) systems. After learning, and extracting features of information-bearing chaotic transmissions at the training stage, the received signals can be recovered efficiently, and reliably at the deployment stage. Thanks to the recursive LSTM-aided DL design, correlations between information-bearing chaotic modulated signals can be exploited to enhance reliability performances. Simulation results demonstrate with the proposed DL demodulation design, the intelligent OFDM-DCSK system can achieve more reliable performances over additive white Gaussian noise (AWGN) channel, and fading channels compared with benchmark systems.**

***Index Terms*—Bit error rate, deep learning (DL), long short-term memory (LSTM) network, orthogonal frequency division multiplexing-aided differential chaos shift keying (OFDM-DCSK), reliability.**

I. Introduction

Chaos-based communications are capable of resisting jamming to enhance the security performances for wireless systems. By exploiting the irregular non-linear chaotic behavior to hide the determinism, chaotic sequences have been applied to modulate user data in practical systems such as ultra-wide-band (UWB) communication systems and power line communication systems [1], [2].

Chaotic modulation methods are usually classified into two types, i.e., coherent and non-coherent schemes. In contrast, non-coherent chaotic systems have attracted more research interests since no complicated chaotic synchronization circuits are required. Differential chaos shift keying (DCSK) [3] is one of the non-coherent modulation methods that can achieve better reliability performances than the counterpart correlation delay shift keying (CDSK) [4] scheme. In order to improve the spectral efficiency and remove the delay line of DCSK systems, the orthogonal frequency division multiplexing-aided DCSK (OFDM-DCSK) scheme has been proposed [5].

Although chaotic systems provide the advantages of enhanced security and anti-jamming capabilities, the real-valued chaotic sequences used for the information-bearing induce interferences to signals, which bring the issue of the reliability degradations [6]. In order to enhance the reliability performances, several research works have been performed to improve the transceiver structures. [7], [8] proposed to reduce the noise at the receiver via allowing the same reference signal shared by multiple information bits or transmitting multiple reference chaotic signals. Similarly, [9] presented a scheme to allocate more subcarriers for reference signals to utilize the diversity gain reducing the bit error rate (BER), but at the cost of lower energy and spectral efficiencies. In our recent research work [10], we proposed to demodulate the reference chaotic signals iteratively, which does not require to modify the transmitter structure but the complexity is relatively high.

Different from traditional demodulation methods utilizing time or frequency resources to improve reliability performances, in this paper, we propose to exploit the feature extraction capability of deep neural network (DNN) for reliability enhancements. Thanks to the powerful ability of pattern recognition from data sets, DNN has been employed in various communication applications to improve the reliability, flexibility and efficiency. For example, in the physical layer, DNN has been applied to implement demodulation [11], channel estimation [12], modulation recognition [13] and end-to-end communication [14], etc. Furthermore, in higher layers, DNN has been applied in the routing decision making for software defined networks [15], [16] and the resource allocation [17] in cognitive radio networks. However, few research works have been done to extract the deterministic feature hidden behind the chaotic modulated signals.

In our design, we propose to replace the conventional correlator for demodulations [5], [6] with the DNN-aided deep learning (DL) demodulation module. To the best of our knowledge, this is the first work to apply DL to intelligently recover received signals for chaotic communications. With the aid of the long short-term memory (LSTM) network, we first construct the DNN architecture to learn the information transmission pattern at the training stage with OFDM-DCSK symbol samples. Subsequently, at the online deployment stage, after performing equalization and fast Fourier transform (FFT) operations, both the reference chaotic sequence and information-bearing chaotic modulated symbols act as the inputs to the LSTM-based DNN for information recovery.

Notably, the DL-aided OFDM-DCSK demodulators do not require to modify the transmitter structure, and it will not affect the security as well as energy efficiency. Besides, they can adapt to time changing

Manuscript received March 9, 2020; revised June 14, 2020, July 28, 2020, and September 1, 2020; accepted September 2, 2020. Date of publication September 7, 2020; date of current version January 22, 2021. This work was supported in part by the National Key Research, and Development Program of China under Grant 2018YFB1802300, in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2020A1515010703, and in part by the Key Research, and Development and Transformation Plan of Science and Technology Program for Tibet Autonomous Region under Grant XZ201901-GB-16. The review of this article was coordinated by Dr. B. Mao. *(Corresponding author: Yuan Jiang.)*

Lin Zhang is with the School of Electronics, and Information Technology, Sun Yat-sen University, Guangzhou 510006, China, and also with the Department of Science, and Technology of Guangdong Province, Southern Marine Science and Engineering Guangdong Laboratory, Zhuhai 519000, China (e-mail: isszl@mail.sysu.edu.cn).

Haotian Zhang is with the School of Electronics, and Information Technology, Sun Yat-sen University, Guangzhou 510006, China (e-mail: sam_3132@qq.com).

Yuan Jiang is with the Naihai Research Institute, Sun Yat-sen University, Guangzhou 510275, China (e-mail: jiangyuan3@mail.sysu.edu.cn).

Zhiqiang Wu is with the Department of Electrical Engineering, Tibet University, Lhasa 850000, China, and also with the Wright State University, Dayton, OH 45435 USA (e-mail: zhiqiang.wu@wright.edu).

Digital Object Identifier 10.1109/TVT.2020.3022043

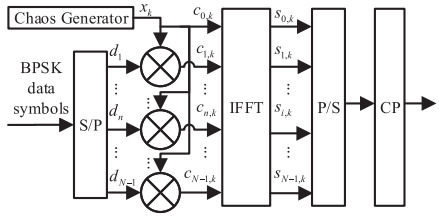

Fig. 1. The transmitter structure.

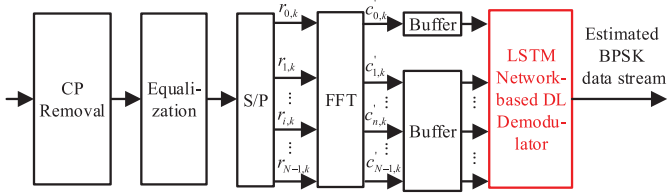

Fig. 2. The intelligent DNN-aided receiver structure.

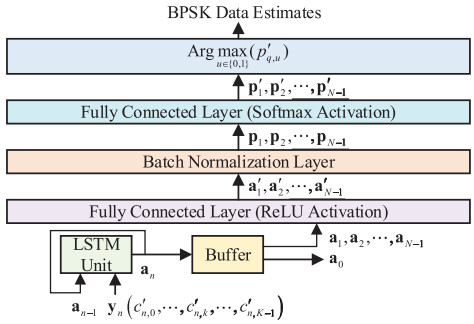

Fig. 3. The architecture of LSTM-based DNN for the DL demodulator.

conditions and extract information transmission patterns intelligently. Thus, the proposed DL demodulators could be easily integrated with existing OFDM-DCSK transmitters. Simulation results demonstrate that the proposed design achieves better BER performances than counterpart systems over additive white Gaussian noise (AWGN) channel and fading channels.

Briefly, the main contributions include: (1) we propose a novel intelligent DL-aided OFDM-DCSK demodulator with no requirement for modifying the transmitter structure; (2) we construct the LSTM based DNN module to learn and extract the information transmission patterns by exploiting the memory cells in LSTM units to formulate the correlation relationship between received chaotic signals; (3) the proposed design could achieve the intelligent information delivery with adaptability to time-changing channel conditions and enhanced reliability, while keeping satisfactory security and energy efficiency performances.

II. DL OFDM-DCSK Transceiver

In this section, we will present the DL-aided OFDM-DCSK transceiver structure. As shown in Fig. 1, at the transmitter which has the same structure as that given in [10], user data bits are firstly modulated by the binary phase shift keying (BPSK) scheme, while the chaos generator uses the antisymmetric cubic map to generate the chaotic sequence with more complex dynamic properties [18], which can be expressed as $x_{k+1} = 4x_k^3 - 3x_k$ with $0 \le k \le K-1$, where $-1 < x_k < 1$ denotes the kth chip, K is the length, and the initial value x_0 is uniformly distributed between 0 and 1.

Then, after the serial to parallel (S/P) conversion, the nth ($0 \le n \le N-1$) BPSK symbol is modulated by the kth chaotic chip as

$$c_{n,k} = d_n x_k, \tag{1}$$

where d_1 to d_{N-1} represent the BPSK user data symbols, $d_0 = 1$, thus $c_{0,k} = x_k$, which means only the reference chaotic symbols are delivered to the receiver.

Subsequently, the inverse fast Fourier transform (IFFT) operations are performed over N available subcarriers, and the resultant ith ($0 \le i \le N-1$) OFDM symbol in the kth chip time slot, which is denoted by $s_{i,k}$, is expressed as

$$s_{i,k} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} c_{n,k} e^{j\frac{2\pi n}{N} i}, \tag{2}$$

where j is the imaginary unit with $j^2 = -1$. After the parallel to serial (P/S) conversion and adding the cyclic prefix (CP), signals are transmitted over the channels.

At the receiver, we propose an intelligent LSTM network based DNN-aided demodulator to replace the traditional correlated demodulator. Without loss of generality, we assume perfect channel state information is known. As shown in Fig. 2, after removing the CP, equalization, and S/P conversions, fast Fourier transform (FFT) operations are performed. The resultant nth information-bearing chaotic modulated symbol in the kth chip slot, which is denoted by $c'_{n,k}$, could be retrieved from the received symbol $r_{i,k}$ as

$$c'_{n,k} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} r_{i,k} e^{-j\frac{2\pi i}{N} n}. \tag{3}$$

Next, $c'_{n,k}$ acts as the inputs to the intelligent LSTM demodulator to recover the BPSK data based on the parameter set learned from the training. More details about the intelligent demodulator design will be described in detail as follows.

III. LSTM-Based Deep Learning Demodulator

In this section, we propose an LSTM-based DNN, then we present the training procedure and analyze the complexity.

A. Architecture of LSTM-Based DL Demodulator

Figure 3 illustrates the proposed architecture of an LSTM-based DNN. At the bottom layer, the LSTM unit performs nonlinear operations recursively on inputs and to capture the temporal dynamic behavior of the input data. The resultant output vector is fed back as the input vector and buffered to input to the next layer.

To be more explicit, at the first time slot, set the initial value $\mathbf{a}_{-1} = \{0, \ldots, 0, \ldots, 0\}$. Then $\mathbf{a}_{-1}$ is combined with $\mathbf{y}_0 = \{c'_{0,0}, \ldots, c'_{0,k}, \ldots, c'_{0,K-1}\}$ to act as the input vectors to generate the vector $\mathbf{a}_0$. Subsequently, $\mathbf{a}_0$ is fed back and combined with $\mathbf{y}_1$ to produce $\mathbf{a}_1$. Similar procedure is carried out with $\mathbf{a}_{n-1}$ and $\mathbf{y}_n = \{c'_{n,0}, \ldots, c'_{n,k}, \ldots, c'_{n,K-1}\}$ till all the received data are processed and input to DNN. Notably, since $\mathbf{a}_0$ only contains the received reference chaotic sequence and no user data is delivered via $\mathbf{a}_0$, it is no need to input $\mathbf{a}_0$ to the next layer, thus the remaining $\mathbf{a}_1, \ldots, \mathbf{a}_{N-1}$ will act as inputs for further processing. Besides, the input data in the complex domain are divided into the real part and the imaginary part, both of which are concatenated into one vector before processed by the DNN.

Next, the $N-1$ information-bearing output vectors from the LSTM unit are processed by the network constituted by two fully connected (FC) layers and a batch normalization (BN) layer which is used to mitigate the vanishing and exploding gradient problem and to accelerate

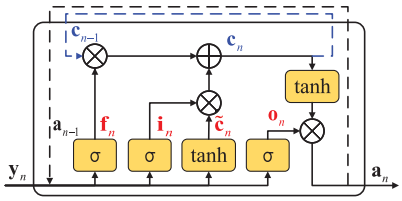

Fig. 4. The structure of an LSTM unit.

the convergence. In addition, FC layers are employed to perform further nonlinear operations and calculate the probability of each BPSK symbol carried by the information-bearing chaotic sequences. The output probability vector $\mathbf{p}'_q\,(1 \le q \le N-1)$ contains two elements $p'_{q,0}$ and $p'_{q,1}$, which respectively represents the probability that the transmitted BPSK symbol is $+1$ or -1. Then by identifying the index of the largest element of $\mathbf{p}'_q$, the proposed DNN will output the BPSK data estimates from $\mathbf{y}_q$ $(1 \le q \le N-1)$. Notably, when $q = 0$, $\mathbf{y}_0$ is the reference chaotic sequence, hence no BPSK symbol is delivered.

More details about how LSTM unit, FC layers and BN layer work are provided as follows.

B. Operation Principles of DNN Layers

1) LSTM Unit: As shown in Fig. 4, at time n, with the input vectors $\mathbf{y}_n$ and $\mathbf{a}_{n-1}$ obtained at time $n-1$, the output vector $\mathbf{a}_n$ is calculated as follows [19]:

$$\mathbf{f}_n = sigmoid(\mathbf{W_f} \cdot [\mathbf{a}_{n-1}, \mathbf{y}_n] + \mathbf{b_f}), \tag{4a}$$

$$\tilde{\mathbf{c}}_n = tanh(\mathbf{W_c} \cdot [\mathbf{a}_{n-1}, \mathbf{y}_n] + \mathbf{b_c}), \tag{4b}$$

$$\mathbf{i}_n = sigmoid(\mathbf{W_i} \cdot [\mathbf{a}_{n-1}, \mathbf{y}_n] + \mathbf{b_i}), \tag{4c}$$

$$\mathbf{c}_n = \mathbf{f}_n \odot \mathbf{c}_{n-1} + \mathbf{i}_n \odot \tilde{\mathbf{c}}_n, \tag{4d}$$

$$\mathbf{o}_n = sigmoid(\mathbf{W_o} \cdot [\mathbf{a}_{n-1}, \mathbf{y}_n] + \mathbf{b_o}), \tag{4e}$$

$$\mathbf{a}_n = \mathbf{o}_n \odot tanh(\mathbf{c}_n), \tag{4f}$$

where $\mathbf{f}_n$ is Forget Gate controls what memory in the cell $\mathbf{c}$ should be forgotten at time n, $sigmoid(\cdot)$ and $tanh(\cdot)$ are the sigmoid and tanh activation function, respectively, $\mathbf{W_f}$, $\mathbf{b_f}$, $\mathbf{W_c}$, $\mathbf{b_c}$, $\mathbf{W_i}$, $\mathbf{b_i}$, $\mathbf{W_o}$, $\mathbf{b_o}$ are all parameters that are learnt and obtained through the neural network training, $[\mathbf{a}_{n-1}, \mathbf{y}_n]$ is the vector constituted by $\mathbf{a}_{n-1}$ and $\mathbf{y}_n$, and $\odot$ means the element-wise product operation. Meanwhile, $\tilde{\mathbf{c}}_n$ is the optional update value generated from the inputs at time n. In addition, $\mathbf{i}_n$ represents Input Gate, which controls what values in $\tilde{\mathbf{c}}_n$ should be used to calculate $\mathbf{c}_n$.

Then with (4d), $\mathbf{f}_n$, $\tilde{\mathbf{c}}_n$ and $\mathbf{i}_n$ are used to recursively generate $\mathbf{c}_n$. $\mathbf{o}_n$ is Output Gate which controls what values in $tanh(\mathbf{c}_n)$ should be the output result. At last, we could obtain the output $\mathbf{a}_n$ of the LSTM unit from $\mathbf{o}_n$ and $\mathbf{c}_n$. The operations presented above are repetitively carried out till all the input data in a specific period have been processed by the proposed DNN.

Notably, we can notice from Figs. 4 and (4a)–(4f) that the output of the LSTM unit is not only related to the current external input but also dependent on the parameters stored at the previous time slot, which matches the characteristics of OFDM-DCSK systems that the information-bearing chaotic sequences are correlated to each other due to the usage of the same reference chaotic sequence $[c_{0,0}, c_{0,1}, \ldots, c_{0,K-1}]$. Accordingly, with the aid of LSTM units, in DL-based chaotic demodulators, the correlations between information-bearing chaotic sequences delivered over different subcarriers could be utilized to enhance the BER performances.

2) FC Layer: By applying FC layers, the output of the LSTM unit can be further processed, and thus the capacity of the neural network can be improved. Note that our proposed DNN employs two FC layers as shown in Fig. 3. Let α denote the index of the FC layer, i.e., $\alpha = 1, 2$, then for input vector $\mathbf{z}_q$ of the αth FC layer, the output vector $\mathbf{z}'_q$ is calculated as $\mathbf{z}'_q = \sigma_\alpha(\mathbf{W}_\alpha \cdot \mathbf{z}_q + \mathbf{b}_\alpha)$, where $\mathbf{W}_\alpha$ and $\mathbf{b}_\alpha$ are parameters of the αth FC layer that are learnt and obtained through training, and $\sigma_\alpha(\cdot)$ is the activation function applied in the αth FC layer.

More explicitly, when $\alpha = 1$, $\mathbf{z}_q = \mathbf{a}_q$ and $\mathbf{a}'_q = \sigma_1(\mathbf{W}_1 \cdot \mathbf{a}_q + \mathbf{b}_1)$, while the rectified linear unit (ReLU) activation function $ReLU(a_{q,u}) = max(0, a_{q,u})$ is employed, where $a_{q,u}$ denotes the uth element of $\mathbf{a}_q$. Additionally, when $\alpha = 2$, we have $\mathbf{z}_q = \mathbf{p}_q$, $\mathbf{z}'_q = \mathbf{p}'_q$. Moreover, the softmax activation function $softmax(z_{q,u}) = \frac{e^{z_{q,u}}}{\sum_{z_{q,v} \in \mathbf{z}_q} e^{z_{q,v}}}$ is employed, which enables the last FC layer to learn to calculate the probability of each BPSK symbol, and to output the probability vector for further estimation of received symbols.

3) BN Layer: In this layer, the input vector is standardized, thus the output vector would have zero mean and unit variance, which is beneficial to increase the gradient when using back propagation algorithm during neural network training. For example, when the input vector has relatively smaller value, the gradient of the sigmoid activation function will be larger, while the standardization can help adjust the value of vectors in the area where the activation functions such as sigmoid, tanh and softmax have high gradients. As a result, the vanishing and exploding gradient problem can be mitigated and the convergence process can be accelerated.

More explicitly, as shown in Fig. 3, the input vector of the BN layer is $\mathbf{a}'_q$, then the output vector $\mathbf{p}_q$ can be calculated as follows [20]:

$$\mathbf{p}_q = \frac{\mathbf{a}'_q - mean(\mathbf{a}'_q)}{\sqrt{var(\mathbf{a}'_q) + \epsilon}}, \tag{5}$$

where $mean(\mathbf{a}'_q))$ and $var(\mathbf{a}'_q))$ are the mean and variance value of the input vector, respectively, which are estimated during the training and then directly used at the deployment stage. $\epsilon = 10^{-5}$ is an extremely small constant used to prevent the denominator from equalling to zero.

C. Training Procedures

At the offline training stage, DNN calculates the derivative with respect to variable parameters, which are updated by using the back propagation algorithm. Along with the updating, the differences between the samples and outputs will decrease gradually and converge to a specific threshold.

More explicitly, let $\mathbf{d} = [d_1, d_2, \ldots, d_{N-1}]^T$ denote BPSK symbol samples, where $(\cdot)^T$ represents the transposition operation, and $\mathbf{y}_n = [c'_{n,0}, c'_{n,1}, \ldots, c'_{n,K-1}]^T$, $(0 \le n \le N-1)$ represent received chaotic sequence samples. After the training, we can formulate and establish the mapping from $\mathbf{y}_n$ to the estimates $\hat{\mathbf{d}}$ with the objective of minimizing the differences between $\mathbf{d}$ and $\hat{\mathbf{d}}$ based on the categorical cross entropy loss function.

D. Complexity Analysis

Finally, we analyze the computational complexity of the LSTM-based demodulator. Similar to the computational complexity analysis presented in [21] and [22], with considerations that the training is carried out offline, we only evaluate the complexity at the online deployment stage when the intelligent demodulator provides estimates for users.

Let L_{LSTM} denote the output dimension of the LSTM layer and L_{FC1} represent the output dimension of the first FC layer,

TABLE I
PARAMETER SETTINGS OF EACH LAYER IN THE PROPOSED DNN

Layer	Output dimensions			Number of Trainable Parameters		
	$K=8$	$K=16$	$K=32$	$K=8$	$K=16$	$K=32$
Input	16	32	64	0	0	0
LSTM	19	38	77	2812	10944	44044
1st FC	17	38	69	340	1482	5382
BN	17	38	69	0	0	0
2nd FC	2	2	2	36	78	140

the computational complexity will be $O(NKL_{LSTM} + NL_{LSTM}^2 + NL_{LSTM}L_{FC1})$ per OFDM-DCSK symbol, which is dependent on the value of L_{LSTM} and L_{FC1}. Since the demodulation performances are also related to these two parameters, there would be a trade-off between the complexity and reliability. For example, if the dimensions of each layer remain invariant, i.e., L_{LSTM} and L_{FC1} are constants, then the complexity would be $O(NK)$. In this case, the complexity of the proposed system is at the same level as that of conventional correlation demodulators, whose complexity is also $O(NK)$ [5]. However, if L_{LSTM} and L_{FC1} are set as being about $2K$ to achieve better reliability performances as shown in the following section, the complexity would become $O(NK^2)$, which is higher than conventional correlators.

IV. SIMULATIONS AND ANALYSIS

This section provides simulation results to validate our design. Table I presents the parameter settings of the DNN for simulations, wherein the output dimensions are determined by the dimension of input vectors with considerations of the capacity of the DNN. Then, the number of trainable parameters could be determined. As an example, when applying the PyTorch DL framework, let I and O respectively denote the input dimension and the output dimension, then the number of trainable parameters could be calculated as $4(I+O)O + 8O$ for the LSTM based bottom layer and $IO + O$ for FC layers. Notably, since the input layer only acts as an entrance for the data, while the BN layer only performs the standardization operation, thus no training is needed to learn the expected mappings in these two layers. Hence no trainable parameters are needed and the number of parameters are 0 for both the input layer and the BN layer.

When training DNN, the value of E_b/N_0 is set as 10 dB for AWGN channel and 20 dB for others, while for the Rician fading channel, the Rician factor is set as 1. In each epoch, 120 thousand BPSK symbols are generated randomly, modulated and pass through simulated AWGN or fading channels to provide the samples of training set, and the BPSK symbols themselves are used as the answers. In addition, the batchsize is set as 20, the categorical cross-entropy and the Adam algorithm [23] are employed as the loss function and the optimization algorithm respectively, and the learning rate is firstly set as 0.01 and gradually decreases until 10^{-5} if the training loss does not decrease in an epoch. Besides, the Intel Core i5-7300HQ is used for the computations.

Figure 5 compares reliability performances, wherein the proposed intelligent demodulator and the benchmark OFDM-DCSK system [5] share the same transmitters, while the OFDM-BPSK system [24] does not use chaotic modulations and hence no chaotic interferences are induced. From Figs. 5(a) to 5(f), we could observe that the proposed system provides more reliable performances than the OFDM-DCSK system. Especially when transmitted over fading channels, more substantial performance gain can be attained thanks to the intelligent learning capability. By contrast, the OFDM-DCSK system has the worst reliability performances. Besides, even when compared with the OFDM-BPSK system [6], the proposed intelligent demodulator still can achieve better reliability performances at higher E_b/N_0 over fading

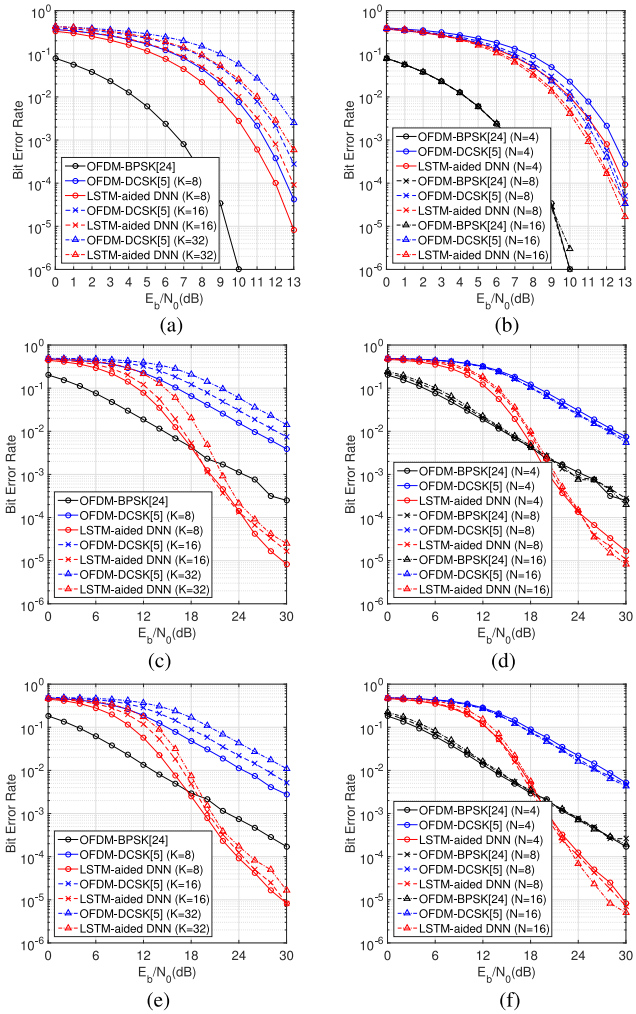

Fig. 5. BER performance comparisons over AWGN channel, Rayleigh fading channel and Rician fading channel when $N = 4, 8, 16$ and $K = 8, 16, 32$. (a) AWGN, $N = 4$. (b) AWGN, $K = 16$. (c) Rayleigh fading, $N = 4$. (d) Rayleigh fading, $K = 16$. (e) Rician fading, $N = 4$. (f) Rician fading, $K = 16$.

channels. It is worth mentioning that when no multiplicative channel gain is provided or E_b/N_0 is smaller, the performance gain becomes smaller since the DL demodulator could not extract information transmission features well and the learning capability becomes weak.

Furthermore, considering that K and N have impacts on system performances including the security, data rate, energy efficiency as well as the reliability [6], we investigate the performances of the proposed design with different K and N, and compare them with benchmarks. We could notice that similar to OFDM-DCSK systems, the proposed intelligent system achieves better BER performances with larger N or smaller K, thanks to relatively weaker interferences.

V. CONCLUSION

In this paper, we propose an LSTM-aided DNN to construct an intelligent demodulator for OFDM-DCSK systems for more reliable transmissions. In our design, we construct a recursive LSTM unit to formulate the correlations between chaotic modulated OFDM-DCSK signals. Then, we compose an LSTM-aided DNN by connecting two FC layers with one BN layer inserted between them. After the offline training, the proposed DL demodulator can utilize the formulated

parameter set and the network structure to retrieve the information. Simulation results over AWGN and fading channels demonstrate that the intelligent demodulator outperforms the benchmark OFDM-DCSK systems. Moreover, our design can achieve even better reliability performances than OFDM-BPSK systems at higher E_b/N_0 over fading channels. Therefore, our design can provide more reliable and adaptable services while retaining the high-security benefit brought by chaotic communications.

REFERENCES

[1] Y. Fang, G. Han, P. Chen, F. C. M. Lau, G. Chen, and L. Wang, "A survey on DCSK-based communication systems and their application to UWB scenarios," *Commun. Surveys Tut.*, vol. 18, no. 3, pp. 1804–1837, Jul.–Sep. 2016.

[2] G. Kaddoum and N. Tadayon, "Differential chaos shift keying: A robust modulation scheme for power-line communications," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 1, pp. 31–35, Jan. 2017.

[3] G. Vizvari *et al.*, "Differential chaos shift keying: A robust coding for chaos communication," in *Proc. Nonlinear Dyn. Electron. Syst.*, 1996, pp. 87–92.

[4] M. Sushchik, L. S. Tsimring, and A. R. Volkovskii, "Performance analysis of correlation-based communication schemes utilizing chaos," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 47, no. 12, pp. 1684–1691, Dec. 2000.

[5] G. Kaddoum, "Design and performance analysis of a multiuser OFDM based differential chaos shift keying communication system," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 249–260, Jan. 2016.

[6] S. Li, Y. Zhao, and Z. Wu, "Design and analysis of an OFDM-based differential chaos shift keying communication system," *J. Commun.*, vol. 10, no. 3, pp. 199–205, Mar. 2015.

[7] G. Kolumbán, Z. Jákó, and M. P. Kennedy, "Enhanced versions of DCSK and FM-DCSK data transmission systems," in *Proc. IEEE Int. Symp. Circuits Syst. (Cat. No. 99CH36349)*, 1999, vol. 4, pp. 475–478.

[8] G. Kaddoum and E. Soujeri, "NR-DCSK: A noise reduction differential chaos shift keying system," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 7, pp. 648–652, Jul. 2016.

[9] H. Yang, G.-P. Jiang, W. K. Tang, G. Chen, and Y.-C. Lai, "Multi-carrier differential chaos shift keying system with subcarriers allocation for noise reduction," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 11, pp. 1733–1737, Nov. 2018.

[10] B. Chen, L. Zhang, and Z. Wu, "General iterative receiver design for enhanced reliability in multi-carrier differential chaos shift keying systems," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7824–7839, Nov. 2019.

[11] W. Hongmei *et al.*, "Deep learning for signal demodulation in physical layer wireless communications: Prototype platform, open dataset, and analytics," *IEEE Access*, vol. 7, pp. 30 792–30 801, 2019.

[12] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Deep learning-based channel estimation for beamspace mmWave massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 852–855, Oct. 2018.

[13] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074–4077, Apr. 2019.

[14] T. OShea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.

[15] B. Mao *et al.*,"A novel non-supervised deep-learning-based network traffic control method for software defined wireless networks," *IEEE Wireless Commun.*, vol. 25, no. 4, pp. 74–81, Aug. 2018.

[16] B. Mao, F. Tang, Z. M. Fadlullah, and N. Kato, "An intelligent route computation approach based on real-time deep learning strategy for software defined communication systems," *IEEE Trans. Emerg. Topics Comput.*, to be published, doi: 10.1109/TETC.2019.2899407.

[17] M. Liu, T. Song, J. Hu, J. Yang, and G. Gui, "Deep learning-inspired message passing algorithm for efficient resource allocation in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 641–653, Jan. 2019.

[18] Y. Zhang, Y. Xu, and Z. Wang, "Dynamical randomicity and predictive analysis in cubic chaotic system," *Nonlinear Dyn.*, vol. 61, no. 1, pp. 241–249, 2010.

[19] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.

[20] S. Loffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn.*, 2015, pp. 448–456.

[21] 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.

[22] G. de Veciana and A. Zakhor, "Neural net-based continuous phase modulation receivers," *IEEE Trans. Commun.*, vol. 40, no. 8, pp. 1396–1408, Aug. 1992.

[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Representations*, May 2015, pp. 1–9.

[24] Z. Du, J. Cheng, and N. C. Beaulieu, "Accurate error-rate performance analysis of OFDM on frequency-selective Nakagami-m fading channels," *IEEE Trans. Commun.*, vol. 54, no. 2, pp. 319–328, Feb. 2006.