

This is a repository copy of *Deep Learning Based Underwater Acoustic OFDM Communications*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/145146/>

Version: Accepted Version

Article:

Zhang, Youwen, Li, Junxuan, Zakharov, Yuriy orcid.org/0000-0002-2193-4334 et al. (2 more authors) (2019) Deep Learning Based Underwater Acoustic OFDM Communications. *Applied Acoustics*. pp. 53-58. ISSN 0003-682X

<https://doi.org/10.1016/j.apacoust.2019.04.023>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Deep Learning Based Underwater Acoustic OFDM Communications

Youwen Zhang^{a,b,c}, Junxuan Li^c, Yuriy Zakharov^d, Xiang Li^{a,b,c,*}, Jianghui Li^e

^a*Acoustic Science and Technology Laboratory,
Harbin Engineering University, Harbin 150001, China*

^b*Key Laboratory of Marine Information Acquisition and Security,
Harbin Engineering University, Ministry of Industry and Information Technology, Harbin
150001, China*

^c*College of Underwater Acoustic Engineering,
Harbin Engineering University, Harbin 150001, China*

^d*Department of Electronic Engineering,
University of York, York, YO10 5DD, UK*

^e*Institute of Sound and Vibration Research,
University of Southampton, UK*

Abstract

In this paper, we present a deep learning based underwater acoustic (UWA) orthogonal frequency-division multiplexing (OFDM) communication system. Unlike the traditional receiver for UWA OFDM communication system that performs explicitly channel estimation and equalization for the detection of transmitted symbols, the deep learning based UWA OFDM communication receiver interpreted as a deep neural network (DNN) can recover the transmitted symbols directly after sufficient training. The estimation of transmitted symbols in the DNN based receiver is achieved in two stages: 1) training stage, when labeled data such as known transmitted data and signal received in the unknown channel are used to train the DNN, and 2) test stage, where the DNN receiver recovers transmitted symbols given the received signal. To demonstrate the performance of the deep learning based UWA OFDM communications, we generate a large number of labeled and unlabeled data by using an acoustic propagation model with a measured sound speed profile to train and test the DNN receiver. The performance of the deep learning based UWA OFDM communications is

*Corresponding author
Email address: lixianghrbeu1982@163.com (Xiang Li)

evaluated under various system parameters, such as the cyclic prefix length, number of pilot symbols, and others. Simulation results demonstrate that the deep learning based receiver offers consistent improvement in performance compared to the traditional UWA OFDM receiver.

Keywords: acoustic propagation model, channel estimation and equalization, DNN, OFDM, underwater acoustic communication.

1. Introduction

Underwater acoustic (UWA) channel poses a significant challenge for reliable communications due to its significant multipath spread and rapid time variation due to Doppler effects[1]. Orthogonal frequency-division multiplexing (OFDM) is an attractive scheme for UWA communications because of its capability of dealing with long multipath spread of UWA channels without resorting to complicated time-domain equalization techniques used in single carrier systems[2, 3, 4, 5, 6]. Many experiments have demonstrated operation of OFDM in UWA communications. The performance of OFDM systems mainly depends on the accuracy of estimation of the channel state information (CSI) which is used to recover the transmitted data at the receiver. Note that the long cyclic prefix (CP) required for the accurate channel estimation reduces the spectral efficiency of the OFDM communication system. A large number of multipath taps required for accurate channel estimation, also leads to reduced spectral efficiency[2, 3, 4, 5, 6].

In recent years, machine learning has been considered as an effective way to solve complex problems such as object detection and recognition, voice recognition, and object tracking[7, 8, 9]. In particular, deep learning has significantly improved performance in field such as computer vision and natural language processing[10]. The general deep learning model consists of multiple layers with multiple neurons. Through the training process, the model can adjust the weights and bias values within each layer, then the layer output is applied to the input of the next layer. In supervised deep learning, a loss function or

cost function that maps an event or values of one or more variables onto a real
25 number intuitively representing some “cost” associated with the event is used
for parameter estimation in a deep neural network (DNN)[11, 12]. We minimize
the loss function to obtain the optimized weights and bias of a DNN by the
supervision data. The estimated weights and bias can be used in the online
task[11]. Deep learning methods are a core part in many emerging applications
30 in communications, e.g. smart cities or the internet of things[13]. Some ini-
tial research works demonstrate the successful application of deep learning in
various communication applications[13, 14]. The learning based communication
systems can be competitive with respect to the current state of the art receivers
based on channel estimation, equalization and detection. Different from the
35 idea of conventional receiver, the deep learning based communication systems
regard the transmitter, channel, and receiver as one DNN[15, 16].

Motivated by the research works on deep learning based radio communica-
tion, in this paper, we introduce the deep learning into the UWA communica-
tions. In particular, we present a deep learning based UWA communications
40 with the OFDM modulation. Our contributions are as follows:

- 1) By representing the receiver as a DNN, we present a deep learning based
UWA OFDM communication system. The UWA communication systems
are modeled using one input layer, a number of hidden layers, and one out-
put layer. Without explicit channel estimation and equalization used in
45 conventional UWA communications, the deep learning UWA communica-
tion systems can directly recover the transmitted symbols after a sufficient
training stage.
- 2) We propose to generate the labeled data for training and the unlabeled
data for test which are based on the channel impulse response (CIRs)
50 generated by the ray tracing software Bellhop with measured sound speed
profile.
- 3) To evaluate the performance of proposed deep learning based UWA OFD-
M communication systems, several numerical experiments are conducted

with changing system parameters, including the cyclic prefix length, number of pilot symbols, and others. Numerical results show that the deep learning based UWA OFDM communication system outperforms conventional UWA OFDM communication system with the least squares (LS) channel estimation.

The remainder of this paper is organized as follows. In Section II, the conventional baseband UWA OFDM communication system is reviewed, and then the deep learning based UWA OFDM communication system is presented in detail. Simulation results are presented in Section III. Finally, conclusions are drawn in Section IV.

Notation: Matrices and vectors are denoted by boldface uppercase and lowercase letters, respectively. $(\cdot)^T$ denotes the transposition. \otimes denotes cyclic convolution. $\mathbf{0}_N$ denotes a zero vector of length N .

2. System model for deep learning based UWA OFDM communications

2.1. Review of a conventional UWA OFDM communication system

Fig. 1 depicts a conventional baseband UWA OFDM system with N subcarriers for single-input single-output (SISO) configuration. For convenience and without loss of generality, we just consider one OFDM symbol in following derivation. On the transmitter side, the binary information bit vector \mathbf{b} is first grouped and mapped into an information symbols vector \mathbf{s} according to the specified modulation mode such as BPSK, QPSK, etc. After that, a vector of data and pilot symbols $\bar{\mathbf{x}}(k)$ is obtained from the vector \mathbf{s} and pilot symbols, where k denotes the index of subcarriers in the frequency domain. A N -point inverse discrete Fourier transform (IDFT) transforms the frequency-domain data, $\bar{\mathbf{x}}(k)$, into the time domain signal $\mathbf{x}(n)$ given by

$$\mathbf{x}(n) = \text{IDFT} [\bar{\mathbf{x}}(k)] = \sum_{k=0}^{N-1} \bar{\mathbf{x}}(k) e^{j(2\pi kn/N)}, \quad n = 0, 1, 2, \dots, N-1, \quad (1)$$

70 where n denotes the time index[3, 5]. Following the IDFT block, a CP of length of N_{CP} is inserted to mitigate the inter-symbol interference (ISI), and its length should be larger than the maximum delay spread of the channel \mathcal{K} . Then the transmitted signal $\mathbf{x}_{CP}(n)$ will pass through the UWA channel with additive noise.

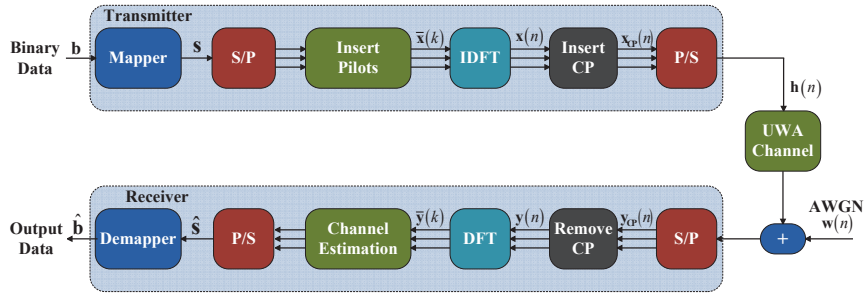


Figure 1: Block diagram of conventional baseband OFDM system in a multipath channel.

We assume that the receiver is ideally synchronized in time and frequency, then the received signal is given by

$$\mathbf{y}_{CP}(n) = \mathbf{H}(n)\mathbf{x}_{CP}(n) + \mathbf{w}(n), \quad (2)$$

where $\mathbf{H}(n)$ is a circulant matrix that is stacked by the zero-padded channel impulse response $\mathbf{h}'(n) = [\mathbf{h}(n)^T, \mathbf{0}_{N+N_{CP}-\mathcal{K}}^T]^T$, $\mathbf{w}(n)$ is additive white Gaussian noise (AWGN) with zero mean and variance σ_n^2 . After removing the CP and performing DFT, the received equivalent frequency domain signal is given by

$$\bar{\mathbf{y}}(k) = \bar{\mathbf{H}}(k)\bar{\mathbf{x}}(k) + \bar{\mathbf{w}}(k), \quad (3)$$

75 where vector $\bar{\mathbf{y}}(k)$, $\bar{\mathbf{x}}(k)$, $\bar{\mathbf{H}}(k)$, and $\bar{\mathbf{w}}(k)$ are the DFT of vector $\mathbf{y}(n)$, $\mathbf{x}(n)$, $\mathbf{H}(n)$, and $\mathbf{w}(n)$, respectively.

Following the DFT block, the pilot signals are extracted and used to estimate the channel impulse response with a channel estimation scheme such as the LS estimator. The equivalent frequency domain channel estimate $\hat{\mathbf{h}}(k)$ of $\mathbf{h}(n)$ is written as follows:

$$\hat{\mathbf{h}}_{LS}(k) = [\bar{\mathbf{y}}(p_1)/\bar{\mathbf{x}}(p_1), \dots, \bar{\mathbf{y}}(p_p)/\bar{\mathbf{x}}(p_p), \dots, \bar{\mathbf{y}}(p_P)/\bar{\mathbf{x}}(p_P)]^T, \quad (4)$$

where p_p is the index of the p -th pilot symbol, P is the number of pilots. Then, the interpolation is performed to estimate the channel coefficients at the data symbol subcarriers.

80 *2.2. Deep learning based receiver for UWA OFDM communications*

2.2.1. Review of the deep neural networks

Feed-forward neural network (FNN) is an artificial neural network (ANN)[11, 16]. Data flow passes through the network in one direction, from the input layer to hidden layers, then to the output layer. The input layer consists of multiple
 85 neurons, each of them representing one feature. Generally speaking, DNN are deeper version of ANN with multiple hidden layers in order to improve the ability in representation or recognition[11, 16]. The structure of a DNN model with Q layers is shown in Fig. 2, where the layer 1 is called as the input layer, the layer Q is the output layer, and other layers are called as the hidden layers.

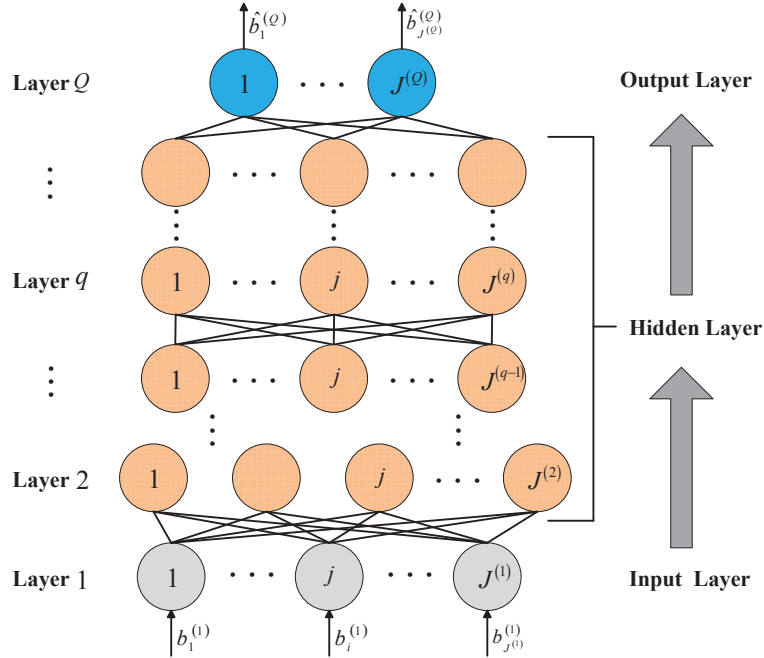


Figure 2: An example of DNN structure.

Assume that the input layer (i.e. layer 1) has $J^{(1)}$ variables in vector $\mathbf{b}^{(1)}$. We rewrite $\mathbf{b}^{(1)}$ as $\mathbf{b}^{(1)} = [b_1^{(1)}, \dots, b_j^{(1)}, \dots, b_{J^{(1)}}^{(1)}]^T$ to associate it with input layer 1, then the j -th neuron's input of layer 2 is

$$a_j^{(2)} = \sum_{i=1}^{J^{(1)}} u_{ij}^{(1)} b_i^{(1)} + v_j^{(1)}, \quad j = 1, 2, \dots, J^{(2)}, \quad (5)$$

where $u_{ij}^{(1)}$ is called a weight between the i -th neuron of layer 1 and the j -th neuron of layer 2, $v_j^{(1)}$ is a bias of the j -th neuron in layer 2, $J^{(2)}$ is the number of neurons of layer 2. In each hidden layer, there is a non-linear activation function $f(\cdot)$ which transforms the linear combinations of inputs to non-linear combinations, the j -th neuron's output of layer 2 is thus given by:

$$b_j^{(2)} = f(a_j^{(2)}), \quad (6)$$

then $\mathbf{b}^{(2)} = [b_1^{(2)}, \dots, b_j^{(2)}, \dots, b_{J^{(2)}}^{(2)}]^T$ will be the next layer's input. Similarly, the j -th neuron's input $a_j^{(q)}$ and output $b_j^{(q)}$ of layer q are

$$a_j^{(q)} = \sum_{i=1}^{J^{(q-1)}} u_{ij}^{(q-1)} b_i^{(q-1)} + v_j^{(q-1)}, \quad j = 1, 2, \dots, J^{(q)}, \quad (7)$$

$$b_j^{(q)} = f(a_j^{(q)}), \quad (8)$$

where $J^{(q-1)}$ and $J^{(q)}$ are the number of neurons of layer $(q-1)$ and q , respectively.

Several activation functions can be chosen, i.e., the sigmoid function $f_S(n) = \frac{1}{1+e^{-n}}$, Rectified Linear Unit (ReLU) nonlinearity $f_R(n) = \max(0, n)$ [16], they are shown in Fig. 3(a) and Fig. 3(b), respectively. Hence, the final output of the DNN $\hat{\mathbf{b}}^{(Q)} = [\hat{b}_1^{(Q)}, \dots, \hat{b}_j^{(Q)}, \dots, \hat{b}_{J^{(Q)}}^{(Q)}]^T$ is a cascade non-linear transformation of input $\mathbf{b}^{(1)} = [b_1^{(1)}, \dots, b_j^{(1)}, \dots, b_{J^{(1)}}^{(1)}]^T$, expressed as:

$$\begin{aligned} \hat{\mathbf{b}}^{(Q)} &= f\left(\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(q)}, \dots, \mathbf{b}^{(Q-1)}; \mathbf{u}^{(1)}, \dots, \mathbf{u}^{(q)}, \dots, \mathbf{u}^{(Q-1)}; \right. \\ &\quad \left. \mathbf{v}^{(1)}, \dots, \mathbf{v}^{(q)}, \dots, \mathbf{v}^{(Q-1)}\right) \\ &= f^{(Q-1)}\left(f^{(Q-2)}\left(\dots f^{(1)}\left(\mathbf{b}^{(1)}\right)\right)\right), \end{aligned} \quad (9)$$

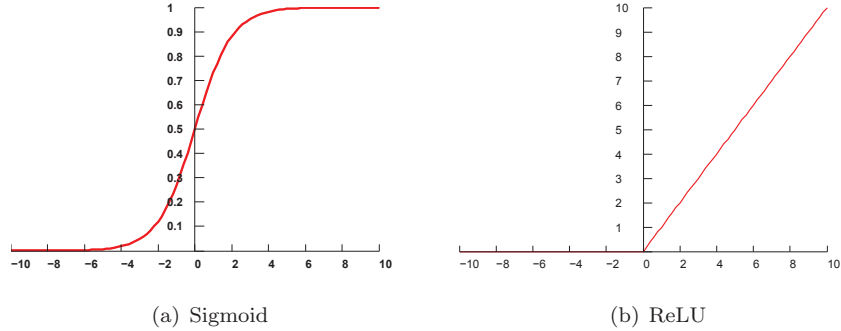


Figure 3: Activation functions for deep learning.

where $f^{(q)}(\cdot)$ denotes the activation function adopted by layer q . $\mathbf{u}^{(q)}$ and $\mathbf{v}^{(q)}$ denote the weights and bias of the layer q .

100 2.2.2. Deep learning based UWA OFDM communication system

In deep learning based OFDM system, the transmitter is the same as the conventional one, while the receiver is replaced by a deep learning model. There are two stages to obtain an effective deep learning model. In the offline training stage, we utilize received signals that are generated with various information sequences and under UWA channel conditions with certain statistical properties to train our model by reducing the difference between the prediction and supervision data, thus generating appropriate weights and bias of DNN. In the test stage, with the weights set $\mathbf{u} \triangleq \{\mathbf{u}^{(q)}\}_{q=1}^{Q-1}$ and bias set $\mathbf{v} \triangleq \{\mathbf{v}^{(q)}\}_{q=1}^{Q-1}$ obtained at the training stage, the deep learning model generates the received signal $\bar{\mathbf{y}}(k)$ in the frequency domain, see Fig. 1, and recovers the transmitted data without explicit estimation and equalization of the underwater acoustic channel[15]. The deep learning based OFDM system is shown in Fig. 4.

The model training process needs random data sequences and channel conditions. Here we use a channel model Bellhop to generate UWA channel statistics. Bellhop is a highly efficient ray tracing program, which can perform two-dimensional acoustic ray tracing with a given sound speed profile $c(z)$ or a given sound speed field $c(r, z)$. Output options include ray coordinates, travel time, amplitude, eigenrays, acoustic pressure or transmission loss[17]. These

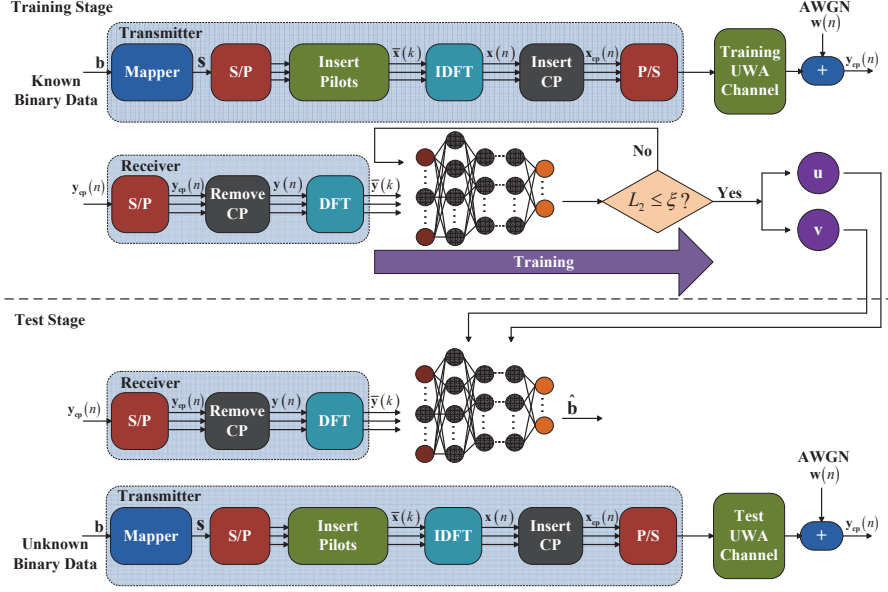


Figure 4: Deep learning based UWA OFDM communication system.

are utilized to generate channel conditions. By changing the position of transmitter or receiver different channel statistics are obtained. In each simulation, an OFDM symbol is formed with a random data sequence and a sequence of pilot symbols, and the pilot symbols need to be fixed during the training and deployment stages. The received OFDM signal is generated with the OFDM frames and the current channel obtained from Bellhop they go through. The training data include the received signal, the transmitted data sequence and the current channel impulse response which are generated using the Bellhop program. With these collected data, the training process aims to minimize the difference between the original transmitted data sequence and the output of the deep learning model. Here we choose the L_2 loss to define the difference as following:

$$L_2 = \frac{1}{N} \sum_{k=0}^{N-1} \left(\hat{\mathbf{b}}(k) - \mathbf{b}(k) \right)^2, \quad (10)$$

where $\hat{\mathbf{b}}(k)$ is the prediction and $\mathbf{b}(k)$ is the supervision data corresponding to Fig.4. In the training stage, the training process ends when the value of the loss

115 L_2 shown in (10) reaches a predefined threshold ξ . As a result, appropriate weights \mathbf{u} and bias \mathbf{v} for every layer of DNN are generated. These are then utilized in the test stage.

3. Simulation results

3.1. Environment configurations

120 For the proposed DL-based receiver, the DNN has $Q = 5$ layers, in which $J^{(1)}, J^{(2)}, J^{(3)}, J^{(4)}$ and $J^{(5)}$ are set to 2048, 1024, 612, 128, and 32, respectively. The last layer utilizes the ReLu function, and other layers utilize the sigmoid function as the active function. The input number corresponds to the number of real parts and imaginary parts of 2 OFDM blocks with pilots and transmitted
 125 symbols. An OFDM block contains 512 sub-carriers and the CP length 128 is considered.

Fig. 5 depicts the layout of the simulation environment used in the Bellhop for generating the channel impulse response (CIR). The SSP measured in a sea experiment and transmission loss (TL) are depicted in Fig. 5. The depth of
 130 the sea is about 250 m. We consider a flat bottom with the parameters: sound speed is 1574 m/s, density is 1.268 g/cm^3 , and attenuation coefficient is 0.01875 dB/wavelength.

In order to train the DNN, we use the Bellhop to generate a large number of CIRs according to various combinations of the positions of the transmitter and
 135 receiver. The CIRs are further used to generate the received training signals (labeled signals) following the equation (2). Finally, the labeled signals are used for training the DNN. The amount of training set and test set ratio is 3:1.

The size of the training set has a significant impact on the accuracy of the DNN training. To investigate how the size of the training set impacts the DNN,
 140 we use two partition schemes for generating the labeled signals as shown in Fig. 6: one scheme is based on the dense grid for transmitter positions depicted in Fig. 6(a), the other is based on the coarse grid for transmitter positions depicted in Fig. 6(b). Specifically, the step in both range and depth between

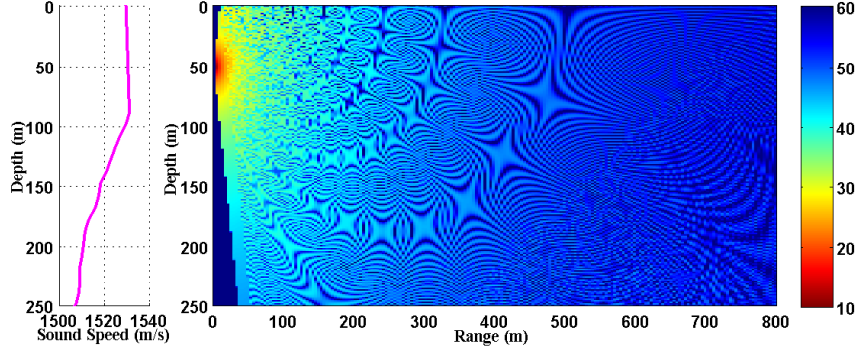


Figure 5: Measured SSP and predicted transmission loss with a source at 50 m depth. The transmission loss is shown in dB.

two transmitter positions is set to 5 m for the dense grid scheme. For the coarse
 145 grid scheme, the 30 m step in both range and depth is used. The minimum
 and maximum range between the transmitter and receiver are 700 m and 800
 m respectively. The depth step for the receiver positions is 10 m.

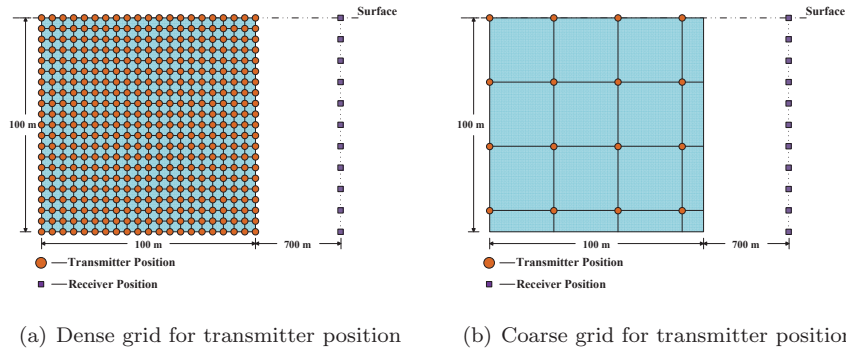


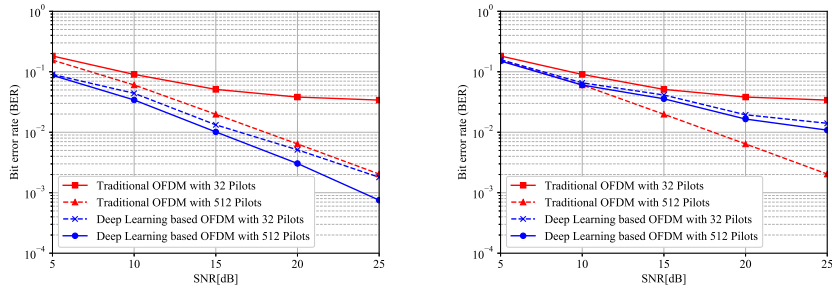
Figure 6: Grid size configurations of transmitter and receiver for different training schemes.

3.2. BER versus the number of pilots

We compare the deep learning based UWA OFDM system with the con-
 150 ventional UWA OFDM system with the LS channel estimation. Fig. 7 shows
 BER performance obtained with the two grid configuration schemes. As shown

in Fig. 7(a), with the dense training grid for transmitter position, the deep learning based UWA OFDM system outperforms the conventional UWA OFDM system. It is also seen that the deep learning based UWA OFDM system with 32 pilots can outperform the conventional UWA OFDM system with 512 pilots. For the conventional UWA OFDM system, the performance significantly depends on the accuracy of channel estimation. The performance gap between the conventional UWA OFDM system using 32 pilots and 512 pilots for channel estimation is large. Thus the conventional UWA OFDM system is very sensitive to the pilot number. By contrast, the deep learning based UWA OFDM system is robust to the pilot number.

Fig. 7(b) shows the performance of the deep learning based UWA OFDM system with coarse training grid for transmitter positions. When the training set is not enough, the performance of the deep learning based UWA OFDM system is limited in spite of how many pilot symbols are adopted. However the performance of the deep learning based UWA OFDM system still better than that of the traditional UWA OFDM system with 32 pilots, but it is inferior to that one with 512 pilots. Overall, the deep learning based UWA OFDM system is sensitive to the training size.

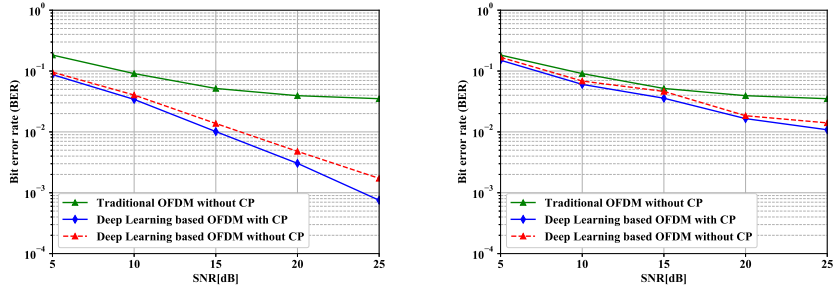


(a) Dense grid configuration for transmitter position (b) Coarse grid configuration for transmitter position

Figure 7: BER performance of deep learning based receiver and conventional receiver with different training grid configurations.

170 3.3. BER Performance for the OFDM system with or without CP

We consider how the performance of the deep learning based UWA OFDM system and traditional UWA OFDM system is affected by the CP. For fair comparison, we keep the 512 pilots unchanged for the two types of UWA OFDM systems. Under the dense grid for transmitter positions, without the CP, the traditional UWA OFDM system has a high error floor as shown in Fig. 8(a). However, for the deep learning based UWA OFDM system, the performance loss induced by the absence of the CP is small. With the coarse grid for the transmitter positions, as depicted in Fig. 8(b), the performance of the deep learning based UWA OFDM system is mainly limited by the size of the training set. Without the CP, the deep learning based UWA OFDM system still outperforms the traditional UWA OFDM system.



(a) Dense grid configuration for transmitter position (b) Coarse grid configuration for transmitter position

Figure 8: BER performance of deep learning based UWA OFDM and traditional UWA OFDM systems without CP.

Overall, as shown in Fig. 8, the CP has a limited impact to the performance of the deep learning based UWA OFDM system due to the capability of the DNN to learn the impact of the UWA channel during the training stage.

185 4. Conclusions

In this paper, a deep learning based UWA OFDM communication system, which treats the complicated UWA communication system as a DNN, is present-

ed. Unlike the traditional UWA communication, the deep learning based UWA communication can be trained to learn the complicated distortions induced by the UWA channel, and then recover the transmitted symbols directly from the received signal, subject to a sufficient training. The deep learning based UWA OFDM communication system has been trained and tested by using a ray tracing toolbox with an SSP measured in a real sea experiment. The performance of the receiver is evaluated for various training pilots, CP length, and the size of training set. Simulation results demonstrate that the deep learning based UWA OFDM communication is more robust to the training pilot number and the absence of CP than the traditional UWA OFDM receiver. Thus, the deep learning based UWA OFDM system offers a higher spectral efficiency.

Acknowledgment

The work of Y. Zhang and Hualin Lan were supported in part by the National Natural Science Foundation of China under Grant 61471138, Grant 51609052, Grant 61531012, and Grant 50909029, in part by the China Scholarship Council Funding, in part by the Program of International Science and Technology Cooperation under Grant 2013DFR20050, in part by the Defense Industrial Technology Development Program under Grant B2420132004, and in part by the Acoustic Science and Technology Laboratory in 2014. The work of Y. Zakharov is partly supported by the U.K. Engineering and Physical Sciences Research Council under Grant EP/P017975/1 and Grant EP/R003297/1. The work of Jianghui Li was partly supported by the European Unions Horizon 2020 research and innovation programme under the grant agreement number 654462 (STEMM-CCS).

References and links

- [1] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 84-89, Jan. 2009.

- [2] B. Li, S. Zhou, M. Stojanovic, L. Freitag, J. Huang, and P. Willett, "MIMO-OFDM over an underwater acoustic channel," in *Proc. IEEE OCEANS*, pp. 1-6, Sep. 2007.
- [3] M. Stojanovic, "OFDM for underwater acoustic communications: Adaptive synchronization and sparse channel estimation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, pp. 5288-5291, Apr. 4, 2008.
- [4] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts," *IEEE J. Ocean. Eng.*, vol. 33, no. 2, pp. 198-209, Apr. 2008.
- [5] Y. M. Aval and M. Stojanovic, "Differentially Coherent Multichannel Detection of Acoustic OFDM Signals," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 2, pp. 251-268, April 2015.
- [6] Y. V. Zakharov and A. K. Morozov, "OFDM transmission without guard interval in fast-varying underwater acoustic channels," *IEEE J. Ocean. Eng.*, vol. 40, no. 1, pp. 144-158, Jan. 2015.
- [7] R. M. Mehmood, R. Du, and H.-J. Lee, "Optimal feature selection and deep learning ensembles method for emotion recognition from human brain EEG sensors," *IEEE Access*, vol. 5, pp. 14797-14806, July 2017.
- [8] W. Liu et al., "An ensemble deep learning method for vehicle type classification on visual traffic surveillance sensors," *IEEE Access*, vol. 5, pp. 24417-24425, Oct. 2017.
- [9] X. L. Zhang and D. Wang, "A deep ensemble learning method for monaural speech separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 5, pp. 967-977, Mar. 2016.
- [10] Y. Cui, X. J. Jing, S. Sun, et al., "Deep learning based primary user classification in Cognitive Radios," *International Symposium on Communications and Information Technologies*, pp. 165-168, 2016.

- [11] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Massachusetts, USA: MIT Press, 2016.
- 245 [12] Wikipedia, “Loss function,” https://en.wikipedia.org/wiki/Loss_function.
- [13] P. J. Werbos, “Computational intelligence for the smart grid-history, challenges, and opportunities,” *IEEE Comput. Intell. Mag.*, vol. 6, no. 3, pp. 14-21, 2011.
- [14] C.-W. Tsai, C.-F. Lai, M.-C. Chiang, L. T. Yang, et al., “Data mining for internet of things: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 77-97, 2014.
- 250 [15] H. Ye and G. Y. Li, “Initial results on deep learning for joint channel equalization and decoding,” *Proc. IEEE 86th Annual Vehicular Technology Conference*, Toronto, Canada, Sept. 2017.
- [16] H. Ye, G. Y. Li, and B.-H. 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, Feb. 2018.
- 255 [17] M. B. Porter, “The Bellhop manual and users guide: Preliminary draft,” *Heat, Light, and Sound Research, Inc., La Jolla, CA, USA, Tech. Rep*, 2011.
- 260