# Autoencoder-Based Enhanced Orthogonal Time Frequency Space Modulation

Yusuf Islam Tek, *Student Member, IEEE*, Ali Tugberk Dogukan, *Student Member, IEEE*, and Ertugrul Basar, *Fellow, IEEE*

***Abstract*— Orthogonal time frequency space (OTFS) is a novel waveform that provides a superior performance in doubly-dispersive channels. Since it spreads information symbols across the entire delay-Doppler plane, OTFS can achieve full diversity. However, reliability still needs to be improved in OTFS systems to meet the stringent demands of future communication systems. To address this issue, we propose an autoencoder (AE)-based enhanced OTFS (AEE-OTFS) modulation scheme. By training an AE under an additive white Gaussian noise (AWGN) channel, a feasible mapper and demapper are learned to improve the error performance and decrease the detection complexity of the OTFS system. The learned mapper is used to map incoming bits into high-dimensional symbols while the learned demapper recovers the information bits in the delay-Doppler domain. Additionally, we derive a theoretical upper bound for the frame error rate (FER). Simulation results confirm that AEE-OTFS outperforms conventional OTFS in terms of FER under perfect and imperfect channel conditions. AEE-OTFS also enjoys low decoding complexity in addition to its superior error performance.**

***Index Terms*— Orthogonal time frequency space (OTFS), time-varying channel, autoencoder, minimum Euclidian distance, AWGN, diversity.**

## I. Introduction

HIGH mobility communication is one of the desired features of enhanced mobile broadband (eMBB) and ultra-reliable and low latency communications (URLLC) use cases in the IMT-2020 standard determined by the International Telecommunication Union (ITU) for 5G and beyond communication systems [1]. Orthogonal frequency division multiplexing (OFDM) has been widely used in 4G, 5G, and many other wireless communication standards due to its superior performance in time-invariant frequency-selective channels. However, in the presence of high mobility, the wireless channel becomes time-varying. The performance of OFDM modulation deteriorates because of the inter-carrier interference (ICI) caused by Doppler spread in the time-varying channel. Recently, orthogonal time frequency space (OTFS) modulation has been proposed to bring a clever solution to this problem [2]. In the OTFS modulation, information symbols are multiplexed in the delay-Doppler domain and spread across the time-frequency domain. This process converts the doubly-dispersive channel into a virtual channel that can be considered non-fading. Hence, all information symbols experience almost the same channel gain, and OTFS can achieve full diversity [2], [3].

Recently, deep learning (DL) approaches have been used in various fields, such as speech/image recognition, medical diagnosis, natural language processing, etc. With the help of DL, data properties are learned automatically using computers with high computing power. A deep neural network (DNN) has the capability of successful feature extraction with its multi-layered structure. In general, DNNs process data in complicated ways using advanced mathematical modeling. Moreover, DNNs have been successfully implemented for wireless communication systems [4]. For instance, DNN-based techniques can be applied to the blocks of the transceiver pipeline, such as coding, channel estimation, equalization, modulation, and detection. For OTFS modulation, DNN-based techniques are presented for signal detection in [5] and [6]. Furthermore, in [7], the authors have proposed a clever DNN-based precoder technique that does not require instantaneous channel state information (CSI) to ensure good reliability.

The autoencoder (AE) presents the transmitter and the receiver of a communication system as its encoder and decoder, respectively [8]. Therefore, the transmitter and the receiver can be learned by training this AE-based structure. Whereas the encoder learns a representation for a given set of data samples, the decoder aims to reconstruct the original data samples from this representation. For OFDM modulation, in [9], an AE-based scheme is trained to learn the transmitter and receiver structures of the OFDM system. On the other hand, for OTFS modulation, the authors proposed a DNN-based AE architecture in which the encoder is trained to decrease the peak-to-average power ratio (PAPR) while the decoder is trained to reconstruct the original signal [10]. Since the whole OTFS frame is inputted into the AE in [10], the computational complexity increases significantly with frame size.

It has been shown that the OTFS system can provide full diversity, equal to the number of paths, in doubly-dispersive channels specifically for large frame sizes [3]. Although conventional OTFS provides remarkable performance in high mobility environments, there is still a need for the design of more sophisticated OTFS-based modulation schemes to meet the rigid requirements of the 6G standard [11], [12]. For instance, to improve the error performance of OTFS, one can design a scheme that maximizes the squared minimum Euclidean distance (SMED).

In this letter, we propose a scheme called *AE-based enhanced OTFS (AEE-OTFS)* that models the mapper and learned demapper of an OTFS system as the encoder and decoder of an AE, respectively. Here, the AE is trained offline in an AWGN channel so that the encoder of AE provides a set of $n$-dimensional ($n$-dim) symbols that maximize SMED between them. Furthermore, the decoder of AE enables

Manuscript received 7 June 2023; revised 11 July 2023; accepted 17 August 2023. Date of publication 22 August 2023; date of current version 11 October 2023. This work is supported by TUBITAK under Grant Number 121C254. The associate editor coordinating the review of this letter and approving it for publication was L. Dai. *(Corresponding author: Ertugrul Basar.)*

The authors are with the Communications Research and Innovation Laboratory (CoreLab), Department of Electrical and Electronics Engineering, Koç University, Sariyer, 34450 İstanbul, Turkey (e-mail: ytek21@ku.edu.tr; adogukan18@ku.edu.tr; ebasar@ku.edu.tr).

Digital Object Identifier 10.1109/LCOMM.2023.3307423

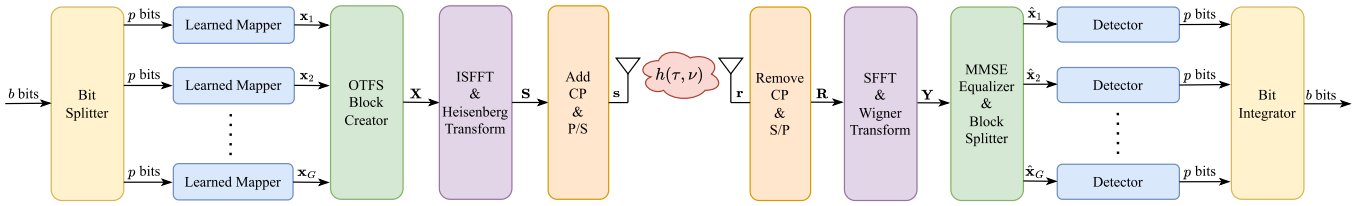


Fig. 1. Transceiver scheme of AEE-OTFS system.

low-complexity detection of $n$-dim symbols of the equalized OTFS signal on the receiver side. Note that the modulator and demodulator blocks can be learned using the OTFS framework under a doubly-dispersive channel rather than an AWGN channel; however, this causes the AE structure to be more complex and increases training time. In contrast to [10], we train our AE-based scheme over the AWGN channel, which does not include the doubly-dispersive channel and signal-processing blocks of the OTFS transceiver in the training process. Note that the learned mapper (LM) and demapper (LD) blocks, trained in the AWGN channel, are employed in the real-time OTFS system that experiences a doubly-dispersive channel. Furthermore, our only focus is improving the error performance, while PAPR reduction is considered in [10]. In contrast to [10], the proposed scheme does not exploit the whole OTFS frame to train the AE, which enables a less complex AE structure.

The rest of this letter is summarized as follows. In Section II, we describe the transceiver scheme of the proposed method and the AE-based mapper and demapper. Our performance analysis is presented in Section III. Computer simulation results are discussed in Section IV, and finally, we conclude this work in Section V.

*Notations:* $a$, $\mathbf{a}$, and $\mathbf{A}$ stand for a scalar, a vector, and a matrix, respectively, and $a_i$ denotes $i$th element of $\mathbf{a}$. $\mathbf{I}_N$ is the $N \times N$ identity matrix. $\mathbf{A} = \text{diag}([a_0, \ldots, a_M])$ denotes a diagonal matrix with diagonal elements $[a_0, \ldots, a_M]$. $\delta(\cdot)$ and $Q(\cdot)$ refer to the Dirac delta function and the Gaussian tail function, respectively. The superscripts $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ state the transpose and the Hermitian transpose operators, respectively. The operators $\otimes$, $\|\cdot\|$, $|\cdot|$, and $E\{\cdot\}$ represent the Kronecker product, Euclidean norm, cardinality of a set, and the expectation, respectively. The determinant and the rank of a matrix are denoted by $\det(\cdot)$ and $\text{rank}(\cdot)$, respectively. The zero mean circular symmetric complex Gaussian distribution with variance $\sigma^2$ is denoted by $\mathcal{CN}(0, \sigma^2)$. Finally, $\mathbf{F}_N$ refers to the $N$ point normalized DFT matrix.

## II. System Model

In this section, first, we describe the transceiver scheme of the proposed AEE-OTFS system, then, we explain the details of the proposed AE-based NN.

### A. Transmission of AEE-OTFS

We consider an AEE-OTFS transmission system with $N$ symbols and $M$ subcarriers, where its block diagram is given in Fig. 1. In this OTFS-based scheme, $\Delta f$ and $T = 1/\Delta f$ represent subcarrier spacing and symbol duration, respectively. A total number of $b$ bits enter the system. These bits are then split into $G$ groups, each including $p = b/G = \log_2(|\mathcal{X}|)$ bits, where $\mathcal{X}$ is the set of all possible $n$-dim symbols. For the $g$th group, $p$ bits are mapped into an $n$-dim symbol, $\mathbf{x}_g \in \mathbb{C}^{n\times 1}$, $g = 1, 2, \cdots, G$, where $\mathbf{x}_g \in \mathcal{X}$. The steps of this mapper are discussed in the sequel. All selected $n$-dim symbols are concatenated as $\mathbf{x} = [\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}, \cdots, \mathbf{x}_G^{\mathrm{T}}]^{\mathrm{T}}$ and $\mathbf{x}$ is provided to the OTFS block creator and overall delay-Doppler domain OTFS frame is obtained as $\mathbf{X} \in \mathbb{C}^{M\times N}$. In the next step, $\mathbf{X}$ is converted to the time-frequency domain by using inverse symplectic finite Fourier transform (ISFFT), and after, using the Heisenberg transform with pulse-shaping waveform $g_{\text{tx}}(t)$, time-domain transmit signal is obtained as

$$\mathbf{S} = \mathbf{G}_{\text{tx}}\mathbf{F}_M^{\mathrm{H}}(\mathbf{F}_M\mathbf{X}\mathbf{F}_N^{\mathrm{H}}) = \mathbf{G}_{\text{tx}}\mathbf{X}\mathbf{F}_N^{\mathrm{H}}, \tag{1}$$

where $\mathbf{G}_{\text{tx}} = \text{diag}[g_{\text{tx}}(0), g_{\text{tx}}(T/M), \ldots, g_{\text{tx}}((M-1)T/M)] \in \mathbb{C}^{M\times M}$. For transmission, a column-wise vectorization is applied to $\mathbf{S} \in \mathbb{C}^{M\times N}$, and the transmit vector $\mathbf{s} \in \mathbb{C}^{MN\times 1}$ is transmitted over the time-varying channel. Before transmission, a CP is added to $\mathbf{s}$ in order to prevent inter-symbol interference (ISI). The time-varying channel with response is represented by $h(\tau, \nu) = \sum_{i=1}^{P} h_i \delta(\tau - \tau_i)\delta(\nu - \nu_i)$, where $P$ is the number of propagation paths, $h_i$, $\tau_i$ and $\nu_i$ stand for the complex channel gain, delay shift, and Doppler shift for the $i$th path, respectively.

At the receiver, after discarding CP, the received time domain signal can be expressed as $\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{w}$, where $\mathbf{w} \in \mathbb{C}^{MN\times 1}$ is a vector of zero mean AWGN samples with $\mathcal{CN}(0, N_0\mathbf{I}_{MN})$, $N_0$ is the noise variance, and $\mathbf{H} \in \mathbb{C}^{MN\times MN}$ is the time domain channel matrix given by

$$\mathbf{H} = \sum_{i=1}^{P} h_i \mathbf{\Pi}^{l_i} \mathbf{\Delta}^{k_i+\kappa_i}, \tag{2}$$

where $\mathbf{\Pi} = \text{circ}\{[0, 1, \ldots, 0]^{\mathrm{T}}\} \in \mathbb{R}^{MN\times MN}$ is the cyclic-shift matrix, $\mathbf{\Delta} = \text{diag}[1, e^{-2\pi j/MN}, \ldots, e^{-2\pi j(MN-1)/MN}] \in \mathbb{C}^{MN\times MN}$ represents the diagonal Doppler shift matrix. We define the fading parameters for the $i$th path as $h_i \sim \mathcal{CN}(0, 1/P)$, $\tau_i = \frac{l_i}{M\Delta f}$, $\nu_i = \frac{k_i+\kappa_i}{NT}$ where $l_i$, $k_i$ and $\kappa_i$ represent delay taps, Doppler taps, and fractional Doppler shift, respectively. We assume that $l_i$ and $k_i$ are integers and there is no fractional Doppler shift, $\kappa_i = 0$.

After the signal is received, $\mathbf{r}$ is de-vectorized and converted to a matrix $\mathbf{R}$ with the size of $M \times N$. Then, the delay-Doppler domain received signal $\mathbf{Y}$ is obtained by applying Wigner transform and SFFT to $\mathbf{R}$ and it can be expressed as:

$$\mathbf{Y} = \mathbf{F}_M^{\mathrm{H}}(\mathbf{F}_M\mathbf{G}_{\text{rx}}\mathbf{R})\mathbf{F}_N = \mathbf{G}_{\text{rx}}\mathbf{R}\mathbf{F}_N. \tag{3}$$

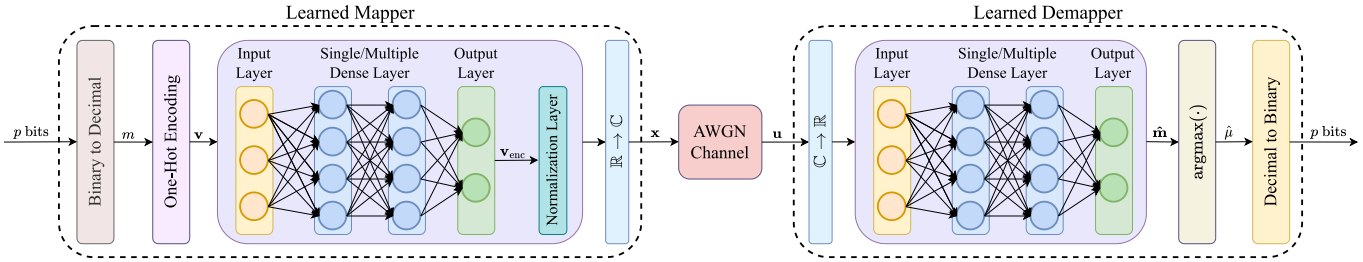


Fig. 2. The structure of the proposed AE.

Moreover, by combining (1) and (3), the vectorization of $\mathbf{Y}$ can be given as:

$$\mathbf{y} = (\mathbf{F}_N \otimes \mathbf{G}_{\text{rx}})\mathbf{H}(\mathbf{F}_N^{\text{H}} \otimes \mathbf{G}_{\text{tx}})\mathbf{x} + \tilde{\mathbf{w}} = \mathbf{H}_{\text{eff}}\mathbf{x} + \tilde{\mathbf{w}}, \quad (4)$$

where $\mathbf{H}_{\text{eff}} \in \mathbb{C}^{MN\times MN}$ is the effective delay-Doppler domain channel matrix and $\tilde{\mathbf{w}}$ is the noise in delay-Doppler domain. In case of rectangular waveforms (i.e., $\mathbf{G}_{\text{rx}} = \mathbf{I}_M$), $\tilde{\mathbf{w}}$ has the same statistical properties of $\mathbf{w}$. By exploiting the minimum mean square error (MMSE) equalizer, the delay-Doppler domain data symbols equalized are obtained as

$$\hat{\mathbf{x}} = \mathbf{H}_{\text{eff}}^{\text{H}}(\mathbf{H}_{\text{eff}}\mathbf{H}_{\text{eff}}^{\text{H}} + N_0\mathbf{I}_{MN})^{-1}\mathbf{y} = [\hat{\mathbf{x}}_1^{\text{T}}, \cdots, \hat{\mathbf{x}}_G^{\text{T}}]^{\text{T}}. \quad (5)$$

Finally, the $g$th equalized group $(\hat{\mathbf{x}}_g^{\text{T}})$ is input to the LD or maximum-likelihood (ML) detector to decode the information bits. The above process is performed for each group.

### *B. AE-Based Mapper and Demapper*

In this section, we present the proposed mapper and demapper for OTFS systems. In the AEE-OTFS method, we exploit the AE technique to model the mapper and demapper as two different DNNs. Here, the LM and LD are considered the encoder and decoder of an AE-based system, respectively. Since the OTFS system already provides a full diversity [3], to enhance the error performance of any OTFS system, one can increase the SMED, which is given by $d_{\min}^2 = \min_{\mathbf{x}_\phi \neq \mathbf{x}_\varphi} \|\mathbf{x}_\phi - \mathbf{x}_\varphi\|^2$, where $\mathbf{x}_\zeta$ is the $\zeta$th $n$-dim symbol, $\zeta = 1, \cdots, |\mathcal{X}|$. Here, our proposed AE-based system is trained under an AWGN channel and learns a set of $n$-dim symbols ($\mathcal{X}$) that maximizes the SMED.

As seen from Fig. 2, firstly, a total number of $p$ bits are converted into a decimal number $m$, $m \in \{m_1, \cdots, m_{|\mathcal{X}|}\}$. Then, $m$ is encoded as a one-hot vector $\mathbf{v} \in \mathbb{R}^{|\mathcal{X}|\times 1}$ whose only single element is one where the others are all zeros. Then, $\mathbf{v}$ is passed through single or multiple dense layers to obtain a transmit vector $\mathbf{v}_{\text{enc}} \in \mathbb{R}^{2n\times 1}$. This encoding process can be expressed as a function $\mathbf{v}_{\text{enc}} = f_{\Omega_{\text{enc}}}(\mathbf{v})$, where $\Omega_{\text{enc}}$ is the parameter set of dense layers of encoder NN. Then, we utilize a normalization layer to satisfy energy constraints, i.e., $\|\mathbf{v}_{\text{enc}}\|^2 = n$. Moreover, $\mathbf{v}_{\text{enc}}$ is converted into a complex vector $\mathbf{x} \in \mathbb{C}^{n\times 1}$, where the first and second halves of $\mathbf{v}_{\text{enc}}$ represent the real and imaginary parts of $\mathbf{x}$, respectively. After that, $\mathbf{x}$ is transmitted over an AWGN channel and $\mathbf{u} = \mathbf{x} + \bar{\mathbf{w}}$ is obtained, where $\bar{\mathbf{w}}$ is the vector of AWGN samples whose elements follow the distribution $\mathcal{CN}(0, \bar{N}_0)$ and $\bar{N}_0$ is the noise variance. The training signal-to-noise ratio (SNR) is $\rho = 1/\bar{N}_0$. At the demapper side, $\mathbf{u}$ is first converted into a real-valued vector with dimensions $2n \times 1$ by separating the real and imaginary parts and then concatenating them. After that, it is passed through single or multiple dense layers to obtain $\mathbf{v}_{\text{dec}} \in \mathbb{R}^{|\mathcal{X}|\times 1}$. This process can be represented as a function $\mathbf{v}_{\text{dec}} = f_{\Omega_{\text{dec}}}(\mathbf{u})$, $\Omega_{\text{dec}}$ is the parameter set of dense layers of decoder NN. The softmax function is used for the output layer of decoder NN to output a probability vector $\hat{\mathbf{m}} = [\hat{m}_1, \hat{m}_2, \cdots, \hat{m}_{|\mathcal{X}|}]^{\text{T}}$, where the $\mu$th element of $\hat{\mathbf{m}}$ denotes the probability of message $m_\mu$ being transmitted. The index of the transmitted message is determined as $\hat{\mu} = \underset{\mu\in\{1,\cdots,|\mathcal{X}|\}}{\text{argmax}} (\hat{m}_\mu)$, and $\hat{\mu}$ is converted into bits.

To train the AE model offline, a set of randomly generated bits and noise vectors are used. We employ the categorical cross-entropy loss function and adaptive moment estimation (Adam) optimizer. Once the training stage of AE is performed, the learned mapper and demapper can be implemented in the AEE-OTFS scheme as in Fig. 1. As seen in Fig. 2, the learned mapper takes $p$ bits as input and outputs the $n$-dim symbol, whereas the learned demapper takes the complex symbol vector as input and outputs $p$ bits.

## III. Performance Analysis

In this section, for the AEE-OTFS, we obtain an upper bound on the frame error rate (FER) by using ML.

According to (4), under the assumption of rectangular pulses, the received signal can be expressed as

$$\begin{aligned}\mathbf{y} &= \sum_{i=1}^{P} h_i(\mathbf{F}_N \otimes \mathbf{I}_M)\mathbf{\Pi}^{l_i}\mathbf{\Delta}^{k_i}(\mathbf{F}_N^{\text{H}} \otimes \mathbf{I}_M)\mathbf{x} + \tilde{\mathbf{w}} \\ &= \mathbf{\Phi}(\mathbf{x})\mathbf{h} + \tilde{\mathbf{w}}, \end{aligned} \quad (6)$$

where $\mathbf{h} = [h_1, h_2, \ldots, h_P]^{\text{T}} \in \mathbb{C}^{P\times 1}$ is the channel gain vector, and $\mathbf{\Phi}(\mathbf{x}) \in \mathbb{C}^{MN\times P}$ is a concatenated matrix given by $\mathbf{\Phi}(\mathbf{x}) = [\mathbf{\Psi}_1\mathbf{x}, \mathbf{\Psi}_2\mathbf{x}, \ldots, \mathbf{\Psi}_P\mathbf{x}]$ and $\mathbf{\Psi}_i \triangleq (\mathbf{F}_N \otimes \mathbf{I}_M)\mathbf{\Pi}^{l_i}\mathbf{\Delta}^{k_i}(\mathbf{F}_N^{\text{H}} \otimes \mathbf{I}_M)$.

We assume that perfect CSI is available at the receiver. When $\mathbf{x}$ is transmitted and erroneously detected $\hat{\mathbf{x}}$, the corresponding conditional PEP can be expressed as

$$\begin{aligned} P(\mathbf{x} \to \hat{\mathbf{x}}|\mathbf{h}) &= Q\left(\sqrt{\frac{\|(\mathbf{\Phi}(\hat{\mathbf{x}}) - \mathbf{\Phi}(\mathbf{x})\mathbf{h})\|^2}{2N_0}}\right) \\ &= Q\left(\sqrt{\frac{\Theta}{2N_0}}\right), \end{aligned} \quad (7)$$

where $\Theta = \mathbf{h}^{\text{H}}\mathbf{\Gamma}(\bar{\mathbf{x}})\mathbf{h}$, $\mathbf{\Gamma}(\bar{\mathbf{x}}) = \mathbf{\Phi}(\bar{\mathbf{x}})^{\text{H}} - \mathbf{\Phi}(\bar{\mathbf{x}}), \bar{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}$. By exploiting the approximation of $Q(x)$ [13], we can express the unconditional PEP (UPEP) as

$$P(\mathbf{x} \to \hat{\mathbf{x}}) \cong E_{\mathbf{h}}\{(1/12)e^{-\frac{\Theta}{4N_0}} + (1/4)e^{\frac{\Theta}{3N_0}}\}. \quad (8)$$

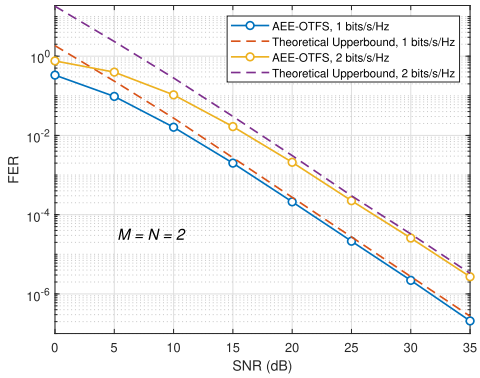


Fig. 3. Average FER performance of AEE-OTFS scheme for SE values of 1 and 2 bits/s/Hz and $\rho = 5$ dB.

Under the assumption $\mathbf{h} \sim \mathcal{CN}(0, \frac{1}{P}\mathbf{I}_P)$, the UPEP can be written as

$$P(\mathbf{x} \to \hat{\mathbf{x}}) = \frac{1/12}{\det(\frac{1}{P}\mathbf{I}_P + \alpha\mathbf{\Gamma}(\bar{\mathbf{x}}))} + \frac{1/4}{\det(\frac{1}{P}\mathbf{I}_P + \beta\mathbf{\Gamma}(\bar{\mathbf{x}}))} \tag{9}$$

where $\alpha = 1/4N_0$ and $\beta = 1/3N_0$. As $\mathbf{\Gamma}(\bar{\mathbf{x}})$ is a Hermitian matrix, the $i$th nonzero eigenvalue of $\mathbf{\Gamma}(\bar{\mathbf{x}})$ can be described as $\lambda_i, i \in \{1, \ldots, r\}$, where $r = \text{rank}(\mathbf{\Gamma}(\bar{\mathbf{x}}))$. As mentioned in [14], the coding gain depends on $\lambda_i$ values and there is a relation between $\lambda_i$ and the SMED of $\bar{\mathbf{x}}$ as $\sum_{i=1}^{r} \lambda_i = P d_{\min}^2$. Using the above analysis, we can express the average FER of the AEE-OTFS as follows

$$P_e \approx \frac{1}{|\mathcal{X}|} \sum_{\mathbf{x}} \sum_{\hat{\mathbf{x}} \neq \mathbf{x}} P(\mathbf{x} \to \hat{\mathbf{x}}). \tag{10}$$

## IV. Simulation Results and Comparisons

In this section, the FER performance of the AEE-OTFS scheme is evaluated via computer simulations. In all simulations, we define SNR as $\gamma = 1/N_0$ and assume that CSI is available at the receiver and the carrier frequency is $f_c = 4$ GHz. Each delay path in all simulations except theoretical results has a single Doppler shift generated with Jakes' model $\nu_i = \nu_{\max} \cos(\theta_i)$, where $\nu_{\max}$ corresponds to the maximum Doppler shift depending on the receiver speed and $\theta_i$ is uniformly distributed over $[-\pi, \pi]$. The receiver speed is taken as 506 km/h, which corresponds to $\nu_{\max} = 1875$ kHz. For the encoder and decoder NNs of the proposed AE, we consider one input, one hidden, and one output layer. For training, hyperparameters are selected such that SMED is maximized by exploiting a manual search method. The number of hidden units is selected as 128. Epoch, batch size, and learning rate are selected as $10^3$, 512, and 0.0001, respectively, and varying $\rho$ values have been exploited. ML detector is used for conventional OTFS system while LD is used for AEE-OTFS.

In Fig. 3, we compare computer simulation results for SE values of 1 and 2 bits/s/Hz with the upper bound FER curves obtained by (10). In these simulations, an ML detector is used and the selected system parameters are $M = N = 2$, $n = 2$, subcarrier spacing $\Delta f = 3.75$ kHz, and $P = 2$ with delay-Doppler profile $(l_i, k_i) = [(0,0)(1,1)]$, $i = 1, 2$. As seen from Fig. 3, it is shown that theoretical upper bounds and computer simulation results coincide with each other in the high SNR region.

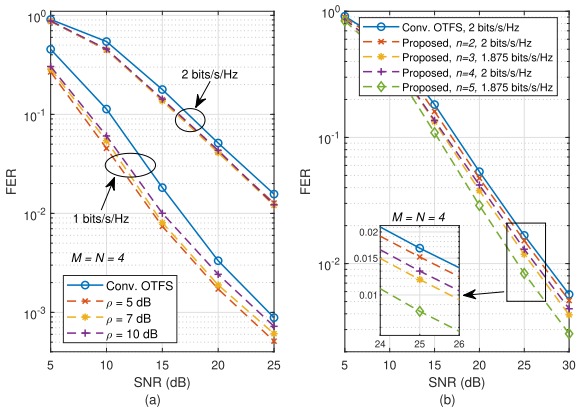


Fig. 4. FER performance comparison of AEE-OTFS scheme using LD for (a) varying $\rho$ with $n = 4$ and (b) varying $n$ with $\rho = 5$ dB.

TABLE I
SMED Comparison of Conventional PSK and AEE-OTFS

| | PSK | AEE-OTFS | | | |
|---|---|---|---|---|---|
| | | $n = 2$ | $n = 3$ | $n = 4$ | $n = 5$ |
| 1 bits/s/Hz | 4 | 5.2277 | 5.6127 | 6.4352 | 7.0640 |
| 2 bits/s/Hz | 2 | 2.2550 | 2.3858 | 2.5767 | 2.6803 |

In Fig. 4(a), error performance results are presented for different training SNR $\rho$ values. The system parameters are specified as $M = N = 4$, $n = 4$, and $\Delta f = 3.75$ kHz. For the channel, we consider $P = 4$ taps with the maximum delay tap defined as $l_{\max} = 3$, and the delay profile is given as $l_i = [0, 1, 2, 3]$. Note that BPSK and QPSK modulations are implemented for the conventional OTFS system to achieve 1 and 2 bits/s/Hz SE values. As seen from Fig. 4(a), the optimum training SNR $\rho$ value can be selected as 5 dB. Additionally, at a FER value of $10^{-2}$, AEE-OTFS provides a 2.6 dB gain over the classical OTFS for a SE value of 1 bits/s/Hz.

In Fig. 4(b), the FER performance of AEE-OTFS is compared to conventional OTFS system for varying block sizes $n$. The simulation parameters are the same as those used in Fig. 4(a). Note that the SE value equals 1.875 bits/s/Hz for $n = 3$ and $n = 5$ since the frame size ($M \times N = 16$) is not an integer multiple of 3 and 5. Therefore, we apply zero-padding for $n = 3$ and $n = 5$. As seen from Fig. 4(b), the AEE-OTFS outperforms the conventional OTFS for all values of $n$. Moreover, AEE-OTFS with $n = 5$ performs better than other parameters. The superiority of the AEE-OTFS technique over OTFS in terms of error performance can be explained by the greater SMED values offered by the proposed method, as shown in Table I.

In Fig. 5, the FER performance of AEE-OTFS is compared to OFDM and OTFS systems for a larger frame size, $M = N = 32$. Extended Vehicular A (EVA) channel model [15] is used. The subcarrier spacing $\Delta f$ is 15 kHz. As seen from Fig. 5, at a FER value of $10^{-3}$, AEE-OTFS approximately provides 2.8 and 1.9 dB gains over classical OTFS for 1 and 2 bits/s/Hz, respectively. Similarly, LD provides the same FER performance as that of ML detector. To investigate the decoding complexity of AEE-OTFS and OTFS, we calculated runtime in milliseconds. As seen from Table II, although AEE-OTFS with ML detector is more complex than classical OTFS with ML, it provides better

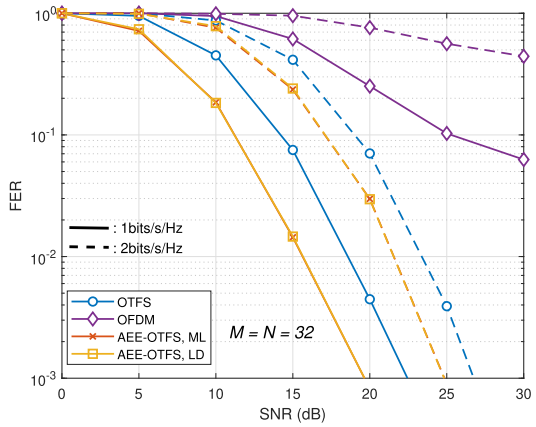


Fig. 5. FER performance comparison of AEE-OTFS scheme for $M = N = 32$ with $n = 4$, SE values of 1 and 2 bits/s/Hz, and $\rho = 5$ dB.

TABLE II
THE COMPLEXITY COMPARISON OF DIFFERENT DETECTORS FOR DIFFERENT FRAME SIZES AND BLOCK SIZES

| | OTFS, ML | AEE-OTFS, ML | | AEE-OTFS, LD | |
|---|---|---|---|---|---|
| | | $n = 2$ | $n = 4$ | $n = 2$ | $n = 4$ |
| $M = N = 16$ | 13.9 | 86.4 | 173.2 | 1.39 | 2.8 |
| $M = N = 32$ | 59.4 | 322.18 | 641.24 | 2.95 | 5.9 |

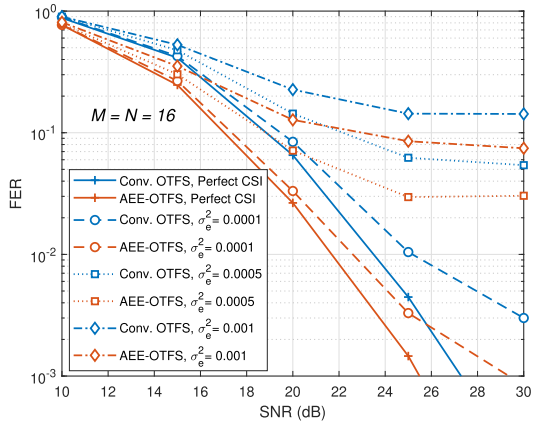


Fig. 6. FER performance comparison of AEE-OTFS scheme using LD with conventional OTFS for $M = N = 16$ with $n = 4$, SE value of 2 bits/s/Hz, $\rho = 5$ dB, and $\Delta f = 15$ kHz under imperfect CSI.

error performance. Additionally, LD reduces the complexity significantly compared to the ML detector. Another important finding is that the complexity of AEE-OTFS with LD is considerably less than classical OTFS with ML detector. Also, note that for larger $n$ values, the number of $n$-dim symbols $|\mathcal{X}|$ increases. Consequently, decoding complexity also increases.

In Fig. 6, we investigate the error performance of AEE-OTFS under imperfect CSI and compare it with the conventional OTFS. As in [16], $\tilde{h}_i$ is exploited instead of $h_i$ to model the channel estimation error, where $\tilde{h}_i = h_i + \epsilon$, $i = 1, \cdots, P$, and $\epsilon \sim \mathcal{CN}(0, \sigma_e^2)$. As seen from Fig. 6, AEE-OTFS outperforms the classical OTFS regarding error performance for different values of $\sigma_e^2$. As a result, the proposed scheme is more robust to channel estimation errors compared to the classical OTFS.

## V. CONCLUSION

In this work, we propose an AE-based scheme to improve the error performance of OTFS by learning a set of $n$-dim symbols and maximizing the SMED between them. Using an AWGN channel instead of a doubly-dispersive channel provides a simple AE structure to be trained. The encoder and decoder of AE are used as the mapper and demapper blocks of a real-time OTFS system after training them under the AWGN channel. Moreover, a theoretical FER upper bound has been derived. Computer simulation results show that, in terms of the FER performance, AEE-OTFS performs better than the conventional OTFS system. In addition, LD significantly reduces the decoding complexity. We conclude that the proposed AEE-OTFS scheme can be a candidate for certain 6G use cases thanks to its outstanding error performance in the presence of high mobility and flexibility. Optimization of the proposed system in terms of bit error rate is considered future work.

## REFERENCES

[1] M. J. Marcus, "5G and 'IMT for 2020 and beyond' [spectrum policy and regulatory issues]," *IEEE Wireless Commun.*, vol. 22, no. 4, pp. 2–3, Aug. 2015.
[2] R. Hadani et al., "Orthogonal time frequency space modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–6.
[3] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Effective diversity of OTFS modulation," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 249–253, Feb. 2020.
[4] B. Ozpoyraz, A. T. Dogukan, Y. Gevez, U. Altun, and E. Basar, "Deep learning-aided 6G wireless networks: A comprehensive survey of revolutionary PHY architectures," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1749–1809, 2022.
[5] A. Naikoti and A. Chockalingam, "Low-complexity delay-Doppler symbol DNN for OTFS signal detection," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, Apr. 2021, pp. 1–6.
[6] X. Xu, M.-M. Zhao, M. Lei, and M.-J. Zhao, "A damped GAMP detection algorithm for OTFS system based on deep learning," in *Proc. IEEE 92nd Veh. Technol. Conf. (VTC-Fall)*, Nov. 2020, pp. 1–5.
[7] C. Liu, S. Li, W. Yuan, X. Liu, and D. W. K. Ng, "Predictive precoder design for OTFS-enabled URLLC: A deep learning approach," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 7, pp. 2245–2260, Jul. 2023.
[8] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cognit. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.
[9] A. Felix, S. Cammerer, S. Dörner, J. Hoydis, and S. Ten Brink, "OFDM-autoencoder for end-to-end learning of communications systems," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.
[10] M. Liu, M.-M. Zhao, M. Lei, and M.-J. Zhao, "Autoencoder based PAPR reduction for OTFS modulation," in *Proc. IEEE 94th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2021, pp. 1–5.
[11] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.
[12] Z. Wei et al., "Orthogonal time-frequency space modulation: A promising next-generation waveform," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 136–144, Aug. 2021.
[13] E. Basar, Ü. Aygölu, E. Panayirci, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5536–5549, Nov. 2013.
[14] S. Li, J. Yuan, W. Yuan, Z. Wei, B. Bai, and D. W. K. Ng, "Performance analysis of coded OTFS systems over high-mobility channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 6033–6048, Sep. 2021.
[15] ETSI, "Evolved universal terrestrial radio access (E-UTRA); base station (BS) radio transmission and reception," ETSI, Sophia Antipolis, France, Tech. Rep. TS 136 104, V14.3.0, 2021.
[16] P. Raviteja, K. T. Phan, Y. Hong, and E. Viterbo, "Interference cancellation and iterative detection for orthogonal time frequency space modulation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6501–6515, Oct. 2018.