IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 21 February 2024, accepted 13 March 2024, date of publication 21 March 2024, date of current version 28 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3380203

RESEARCH ARTICLE

Design of Deep Learning-Based 1-bit Transceiver With Oversampling and Faster-Than-Nyquist (FTN) Signaling

METASEBIA DEMISSU GEMEDA[1], MIN SIG HAN[1], AMEHA TSEGAYE ABEBE[2], AND CHUNG G. KANG[1], (Senior Member, IEEE)
[1]Electrical Engineering Department, Korea University, Seoul 02841, South Korea
[2]Advanced Communication Research Center, Samsung Research, Seoul 06765, South Korea

Corresponding author: Chung G. Kang (ccgkang@korea.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) funded by the Korean Government [Ministry of Science and Information and Communication Technologies (MSIT)] under Grant 2020R1A2C100998413, and in part by the Institute of Information & Communications Technology Planning & Evaluation (IITP) funded by the Korean Government (MSIT) (Intelligent 6G Wireless Access System) under Grant 2021-0-00467.

ABSTRACT The adoption of one-bit analog-to-digital converters (ADCs) in wireless communication has become increasingly popular, owing to their potential to reduce power consumption. However, there is a decreased spectral efficiency (SE) associated with their use, as they are unable to take advantage of high-modulation-order signaling. Faster-than-Nyquist (FTN) signaling along with oversampling are promising means of improving SE at the expense of additional inter-symbol interference (ISI). This study establishes a common platform for design and evaluation by providing a general system model and problem formulation, as well as proposing an alternative solution framework for the one-bit quantization system. By systematically evaluating optimizable factors that impact performance, such as signaling rate, transmit sequences, and error correction blocks, this study proposes a deep learning (DL)-based architecture, specifically channel autoencoders, for end-to-end communication over a one-bit quantization channel. Various transceiver designs are proposed for performance comparison as well as to enable one-bit quantized communication at previously unattainable information rates through conventional means. Numerical results showcase the superiority of jointly optimizing all blocks in an additive white Gaussian noise (AWGN) channel. The DL-based scheme operationalizes Bit Error Rate (BER) performance at information rates scaling up to 80% of Shamai's limit. Beyond AWGN, the autoencoder-based transceiver design extends to make one-bit quantization viable over a challenging Rayleigh multipath fading channel. A detailed analysis compares a pilot-based scheme for one-bit quantization with a pilotless option, thereby revealing their SE and BER relationship. The pilot-based quantized scheme outperforms conventional fading channel transmission without quantization by up to 10 dB.

INDEX TERMS Auto-encoder, deep learning, one-bit-quantization, oversampling, faster-than-Nyquist.

I. INTRODUCTION

As the wireless communication industry strives for higher data rates through transmission in higher frequencies such as in the terahertz band, the power efficiency of current analog-to-digital converters (ADCs) is becoming a significant concern. Research has shown that power consumption can be reduced by decreasing the resolution of ADCs and instead, by relying on time domain resolution. As a result, there has been a growing interest in transmission schemes that use one-bit ADCs, which are even simpler because they remove all amplitude information. However, using one-bit ADCs leads to performance degradation as the system is unable to exploit high-modulation-order signaling, resulting in a

The associate editor coordinating the review of this manuscript and approving it for publication was Arun Prakash.

decrease in spectral efficiency (SE). To counteract this loss of information rate, techniques such as oversampling in the time domain at the receiver and faster-than-Nyquist (FTN) signaling at the transmitter can be implemented. Oversampling takes advantage of the high correlation between oversampled binary samples, making it possible to detect zero crossings in a one-bit ADC system. On the other hand, FTN increases the temporal resolution at the transmitter, enabling it to meet the demand for zero crossings on a finer grid. However, the combination of FTN and oversampling can result in high inter-symbol interference (ISI), which can make the recovery of message bits quite challenging. A series of operations, including one-bit quantization, oversampling, and FTN-based transmission, reformulated the design problem of basic components for wireless communication. In particular, the non-orthogonal properties in both transmission and sampling pose new challenges for error correction, receiver design, and transmit filter design. Furthermore, zero crossing-based detection redefines the achievable capacity in terms of the amount of oversampling, which is bounded by Shamai's limit [1]. It defines the maximum achievable SE when oversampling a received signal M_{Rx} times the Nyquist rate as $\log_2(M_{Rx} + 1)$ bits per Nyquist Interval (T). This limit only considers quantization noise, without any type of channel noise, and serves as an ultimate bound that can be achieved. When the signal-to-noise ratio (SNR) is sufficiently high, it is possible to achieve a SE that is comparable to Shamai's noiseless limit [2]. However, to achieve this theoretical limit, it is essential to strategically design the transmitter. To detect a high modulation-order signal using a zero crossing-based scheme, it is necessary to construct a sequence that can be easily distinguished even after undergoing one-bit quantization and oversampling. This sequence should be robust against both quantization noise and ISI. However, this often leads to expanded codewords and low information rates that fail to reach the upper bound. On the other hand, the combination of FTN and oversampling results in high ISI, which makes it more challenging to recover message bits. Any system that would consider transmission in frequencies above 100 GHz (sub-terahertz band) would need to take into account the associated power consumption of its components, and enabling spectrally efficient and robust signaling for systems with one-bit quantized ADC receivers can open the door toward power efficient communication in such ranges.

Different methods, including both conventional and deep learning (DL)-based ones, have been proposed to enhance the bit error rate (BER) or information rate performance in noisy channels [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19]. The study in [2] demonstrated that by using band-limited signals with a specific zero-crossing pattern, it is possible to encode input information to achieve information rates similar to Shamai's limit even in noisy channels. The works in [3], [4], and [5] employed run-length-limited (RLL) sequences to encode information into the distance between consecutive 1s for ideal one-bit detection, but the sequence detection requires a complicated detection algorithm with memory. Other modern approaches aimed at reducing energy consumption in ADCs include sub-Nyquist sampling methods, exemplified by the methodology introduced in [20]. The work in [11] attempted channel estimation and detection over fading low-bit quantized, oversampled channels. However, BER performance was limited when extreme of one-bit quantization is applied. Moreover, to make up for the information rate lost due to quantization, adoption of FTN signaling [21], [22], [23], [24], [25], [26] and other non-orthogonal waveform designs [26] that enhance SE should be considered. To actualize the concept, there have been several works that have explored implementation of FTN signaling. References [23] and [24] contributed to this discourse by focusing on the development of joint channel estimation and decoding methods for FTN signaling systems over fading channels, addressing challenges such as ISI and colored noise at the receiver side. A DL based receiver was designed in [25] to mitigate ISI from FTN signaling.

Additionally, various studies have used neural networks to improve the performance of one-bit quantization and FTN communication systems [12], [13], [14], [15], [16], [17]. By representing the communication channel with a regularized autoencoder, the authors in [11] showed that optimal codes can be found within a DL framework. As a result, [11] and [12] used a serial concatenation of existing forward error correcting (FEC) codes, such as Turbo and LDPC codes, with an autoencoder for a successful error correction in the presence of non-linearity introduced by a one-bit ADC. This was able to make one-bit quantization along with oversampling and FTN signalling operational over an acceptable BER range for modulation orders as high as 64 QAM for additive white Gaussian noise (AWGN) channel. Due to the low rate of sequence generation required for the one-bit ADC system, the SE of the system was not optimized to reach the limit defined by Shamai.

Previous research on one-bit quantization systems has made significant progress in addressing the challenges of this design. However, there is currently no comprehensive system in place to serve as a benchmark for measuring the improvements in SE. The objective of our work is to develop a general system model and problem formulation that can establish a common platform for design and evaluation. Additionally, we aim to propose an alternative solution framework. By formulating the problem in this way, we aim to systematically evaluate factors that impact performance, such as signaling rate, transmit sequences, and error correction block, which must be optimized for the 1-bit channel. In this paper, we propose a scheme that enables simple transmission with one-bit quantization at the receiver by adopting DL approaches. Specifically, we employ channel autoencoders for end-to-end communication over a one-bit quantization channel, which allows us to create various transceiver designs. The aims and contributions of this study can be summarized as follows:

- Investigate DL-based transceiver designs that eliminate the non-linearity of one-bit quantization and ISI while allowing high information rate communication with one-bit quantization in an AWGN channel.
- Learn optimal joint error-correction and modulation, as well as symbol-to-sequence mapping suited for the one-bit quantization channel. This produces codewords with zero-crossing patterns for recovery after one-bit quantization and simplifies the general problem formulation of the system, and
- Extend the DL-based transceiver designs to enable pilot-based and pilotless transmission over a fading channel for a one-bit quantization and oversampling receiver. By utilizing DL tools, such as convolutional layers and sequential learning, as well as expert-domain knowledge, it is shown that end-to-end communication with a high information rate over a fading channel with one-bit quantization and oversampling is possible, despite the challenges of quantized pilot reception and corresponding channel estimation.

The rest of the paper is organized as follows: In Section II, a system model and problem formulation are presented. Section III details the proposed DL-based transceiver designs for the AWGN channel. The DL-based designs for fading channels are discussed in Section IV. The simulation results are presented in Section V. Finally, Section VI concludes the study.

Notation: All italic letters such as x and X, are variables. Boldface uppercase letters, such as $\mathbf{X}$, represent matrices and boldface lowercase letters like $\mathbf{x}$ are vectors. $\mathbf{X}^T$ is the transpose of $\mathbf{X}$. Double-struck letters indicate the base of values for vectors such that $\mathbb{B}^d$, $\mathbb{C}^d$ and $\mathbb{R}^d$ indicate a binary, complex and real vector respectively, all with dimension d. Mapping is represented by symbol $\rightarrow$ while sets are indicated by braces $\{.\}$. Moreover, $|.|$ is used for absolute value functions and $\mathcal{O}(.)$ is used for big O notation for upper bound on the growth rate of a function.

II. SYSTEM MODEL AND PROBLEM FORMULATION

Fig. 1 presents a comprehensive system model that includes existing implementations performing one-bit ADC and oversampling at the receiver and FTN signaling at the transmitter [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13]. Let $\mathbf{b} \in \mathbb{B}^s$ denote information bits to be transmitted, which is then channel-encoded using code rate $r = s/k$ whose output is given as $\mathbf{c} \in \mathbb{B}^k$. Afterward, the coded bits $\mathbf{c}$ are modulated with an order of M (bits/symbol) to produce a complex modulated symbol $\mathbf{d} \in \mathbb{C}^{k/\log_2 M}$. In this system model, the transmitted signal is required to encode its information in the distance between zero crossings to ensure reliable detection after quantization at the receiver. This sequence construction has a symbol-to-sequence mapping rate of $\Delta \in (0, 1]$ that governs transmission rates. The

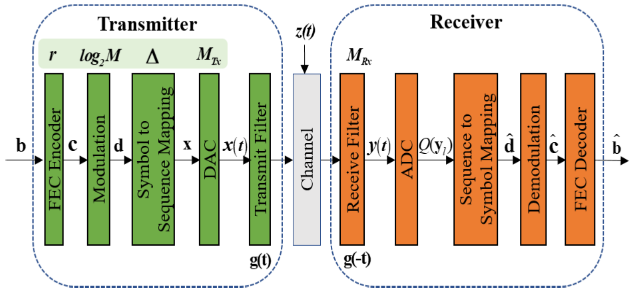


FIGURE 1. Comprehensive system model for one-bit quantization and oversampling communication schemes.

mapping can be expressed as $\mathbb{C}^{k/\log_2 M} \rightarrow \mathbb{C}^N$ such that $\Delta = k/(N \log_2 M)$ and it outputs the vector $x = [x_1\, x_2 \cdots x_N]$ where $x_n \in \{0, 1\}, \forall n$. Finally, the DAC block transmits M_{Tx} symbols per one Nyquist interval T in an FTN manner by utilizing a pulse shaping filter $g(t)$, particularly, a root-raised cosine (RRC) filter [21], where $M_{Tx} \in [1, \infty)$, which is referred to as the FTN signaling factor. Note that FTN enhances the temporal resolution of the transmitter by generating zero crossings on a more refined grid and therefore increases the throughput. However, the accelerated signaling rate induces ISI as a tradeoff, causing adjacent symbols to interfere with each other [22]. The resulting signal is represented as $x(t) = \sum_{n=1}^{N} g\left(t - \frac{nT}{M_{Tx}}\right) x_n$.

Consider the impulse response of the multipath fading channel with P paths, $h(t) = \sum_{i=1}^{P} \alpha_i u_i(t)\delta(t - \tau_i)$, where each $u_i(t)$ is a slowly time-varying, zero-mean, unit-variance Gaussian random process, and τ_i and α_i are the delay and root-mean-square value of the magnitude of the i-th path, respectively. Then, the received signal for the system can be represented as

$$y(t) = \left(\sum_{i=1}^{P} \sum_{n=1}^{N} \alpha_i u_i(t) g\left(t - \frac{nT}{M_{Tx}} - \tau_i\right) x_n + z(t) \right) * g(-t) \quad (1)$$

where $z(t)$ represents AWGN process, which is circularly symmetric, i.e., all real and imaginary samples are i.i.d. with $\mathcal{CN}\left(0, \sigma_h{}^2\right)$. Here, the channel noise and fading have standard deviations of σ_z and σ_h, respectively. We assume a slowly varying random process that does not change with the pulse duration; therefore, we can drop the time index on $u_i(t)$. A matched filter $g(-t)$ on the receiver side will filter the channel-distorted signal before oversampling and one-bit quantization is performed. There will be the same matched filtering on both in-phase and quadrature branches of the transmitted sequence of coded symbols. The ADC will temporally oversample this received signal with a rate of M_{Rx} with respect to the Nyquist rate, where the l-th sample can be represented as

$$y_l = \sum_{i=1}^{p}\sum_{n=1}^{N}\alpha_i u_i \bar{g}\left(\frac{lT}{M_{Rx}} - \frac{nT}{M_{Tx}}\right)x_n + z_l, \quad (2)$$

After receive filtering, the combined transmit and receive filter can be expressed as $g'(t) = g(t) * g(-t)$ and we can define $\bar{g} = g'(t-\tau_i)$ as the effective filter. It is the combined channel filter that captures the effect of the transmitter filter, its matched filter, and ISI due to oversampling and FTN. Additionally, $z_l = z\left(\frac{lT}{M_{Rx}}\right) * g\left(\frac{lT}{M_{Rx}}\right)$ is the filtered and oversampled noise. The oversampled real and imaginary signals are then one-bit quantized by the ADC such that

$$\begin{aligned} Q\left(Re(y_l)\right) &= sgn\left(Re(y_l)\right) \\ Q\left(Im\left(y_l\right)\right) &= sgn\left(Im\left(y_l\right)\right) \end{aligned} \quad (3)$$

where $sgn(x) = 1$ if $x \geq 0$ and $sgn(x) = -1$, otherwise. The quantized signal is demodulated and decoded to produce reconstructed message bits $\widehat{b}$. Since the one-bit ADC can only differentiate between two levels in both real and imaginary dimensions, the maximum modulation order that can be implemented is up to $M = 2$. This consequently results in a reduced SE and temporal oversampling can be applied to allow information rates up to Shamai's defined limit in [1]. The achieved transmission rate in this FTN transmitter for one-bit ADC can be given as $r \cdot \Delta \cdot \log_2 M \cdot M_{Tx}(\text{bits}/T)$, as depicted in Fig. 1. On a receiver side, the oversampling factor M_{Rx} determines the achievable rate of the scheme as given by $2\log_2(M_{Rx}+1)(\text{bits}/T)$ [1]. This leads to the following constraint:

$$r \cdot \Delta \cdot \log_2 M \cdot M_{Tx}(\text{bits}/T) \leq 2\log_2(M_{Rx}+1) \quad (4)$$

Therefore, our first objective is to design a system that can approach Shamai's limit for a given M_{Rx} by developing a detailed transceiver structure and optimizing the values of M_{Tx}, r, M, and Δ. In the course of formulating the design problem, the SNR γ, with its threshold $\gamma^{(th)}$ to satisfy a target BER, may constrain SE maximization. The BER is a function of M_{Tx}, r, M, δ, and M_{Rx} as errors are incurred by quantization and ISI, in addition to the channel noise and fading. All the design parameters should be optimized to achieve the highest possible SE subject to an upper bound on a required BER for reliable communication. Furthermore, practical hardware limits of one-bit ADC should be considered by setting ${M_{Rx}}^{(th)}$ and ${M_{Tx}}^{(th)}$ as design constraints. Subsequently, the following problem formulation comprehensively summarizes the practical design objectives of one-bit ADC as given in [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], and [13]:

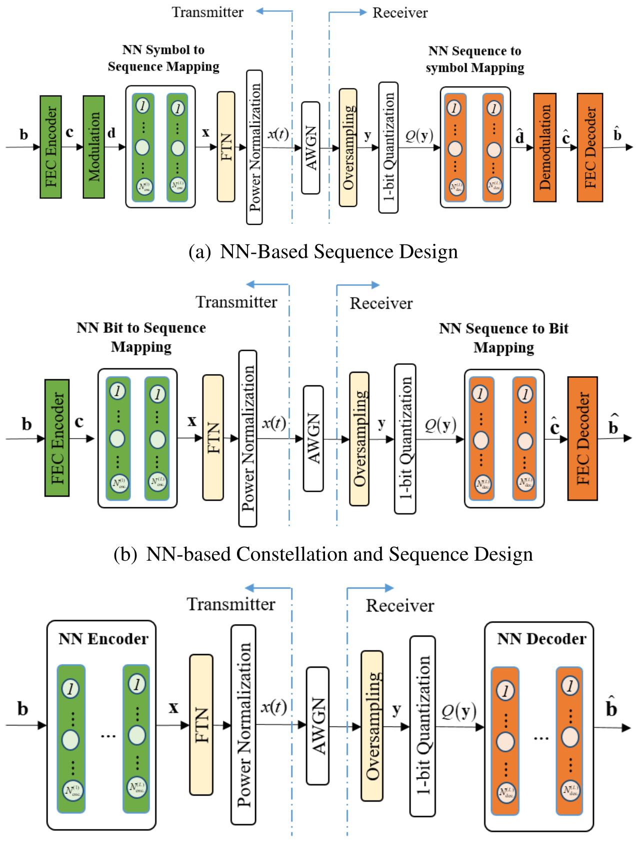
$$\begin{aligned} (r^*, \Delta^*, M^*, M_{Tx}^*) &= \underset{(r,\Delta,M,M_{Tx})}{\arg\max}\ \left(r \cdot \Delta \cdot \log_2 M \cdot M_{Tx}\right) \\ s.t.\quad & r \cdot \Delta \cdot \log_2 M \cdot M_{Tx} \leq 2\log_2(M_{Rx}+1) \\ & M_{Tx} \leq M_{Tx}^{(th)}, M_{Rx} \leq M_{Rx}^{(th)}, \\ & \gamma(r, \Delta, M, M_{Tx}, M_{Rx}, \sigma_z, \sigma_h) < \gamma^{(th)} \end{aligned} \quad (5)$$

The design approach in this problem formulation recognizes the challenge of meeting the many constraints and limits of one-bit ADC systems. It organizes the design efforts of various existing state-of-the-art schemes for one-bit quantization channel [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13]. However, the non-explicit representation of BER with respect to SNR and other parameters makes it a non-trivial optimization problem.

III. DL-BASED DESIGN FOR ONE-BIT QUANTIZATION AND OVERSAMPLING TRANSCEIVER IN AWGN CHANNEL

To solve the problem formulated above, we must identify optimal parameters for the transmitter given a particular oversampling rate. But in order to reach for this optimization, we have to assume that the blocks of the comprehensive system model for one-bit quantization that is displayed in Fig. 1 are working optimally. Although error-correcting and modulating blocks have undergone significant refinement over the years to achieve exceptional performance in conventional communication systems, the symbol-to-sequence mapping block specific to one-bit quantization systems has yet to be optimized. To clarify, one of the biggest challenges in trying to reach Shamai's limit for one-bit ADC systems is the construction of a sequence that is mapped to modulation symbols in order to allow the detection of high-modulation order signaling through a zero-crossing-based scheme. The sequence has to be distinguished after one-bit quantization and oversampling, i.e., robust against quantization noise and ISI. This results in expanded codewords and low information rates that fail to reach near the defined upper bound and hinder the optimization given in (5). Specifically, while the values for error correcting rate and modulation can be adjusted accordingly to increase information rate, conventional methods of mapping symbol-to-sequence fail at producing compact codewords with low SE without compromising error tolerance. Conventional sequence design relies on methods like RLL sequences or using other hand-crafted techniques. Studies in [3], [4], and [5] use RLL sequences to encode data by leveraging the gaps between consecutive 1s, in order to achieve one-bit detection. However, as the modulation order increases, such as in higher-level quadrature amplitude modulation (QAM) schemes, the complexity of RLL encoding and decoding also increases as it was originally designed for simpler binary modulation. This results in intricate algorithms and increased computational and memory demands. Similarly, hand-crafting pulses to avoid receiver ambiguity is only feasible for low-order modulation schemes. As modulation order grows, the complexity of sequence design becomes too intricate for a manual approach, as demonstrated in [2]. Utilizing a data-driven approach with DL can be an effective means of simplifying or solving this problem by optimizing one or more transceiver parameters. In [34], they have demonstrated that machine learning techniques can be utilized to automatically uncover the relationship between the problem representation and the optimal solution, by leveraging training data, instead of attempting to convert the problem into a computationally efficient or convex form. The work delves into the mitigation of model-driven optimization challenges in scenarios such as Reflective Intelligent Surface and MIMO channel estimation, primarily focusing on the receiver side of the transceiver. Moreover, a series of related studies [27], [28], [29], [30], [31], [32], [33], [34] leverage neural networks (NNs) to enhance the functionality of individual blocks and further on, show how the NN can be used to streamline and optimize the overall communication system from bit-to-bit.

A channel autoencoder (AE) with the characteristics of presenting a higher dimensional representation of its input is reminiscent of a simple communication system [12], [27]. It can, therefore, be employed to template an encoder-decoder pair that will interpret the functions of one or more of the transmitter-receiver block pairs. AEs can be implemented for the one-bit quantized channel communication system since they can learn to communicate over any channel, even for which no information-theoretically optimal scheme is known [27]. By training to minimize communication errors, the AE can learn to design features fitted for one-bit quantization environment. Therefore, it is logical to utilize NN structures like AEs to solve the issue of the symbol-to-sequence block which was first proposed in the works [12], [13] by concatenating the error-correcting block and AE for one-bit quantization channel. The biggest obstacle in training NNs over one-bit-quantized channel is that one-bit quantization impedes gradient-based learning for layers preceding quantization. This is because quantization produces undefined derivatives at zero values and derivatives that are equal to zero at values different from zero. In order to address this challenge, [12] utilizes a two-step training policy in which the decoder parameters are first trained, while encoder parameters remain randomly initialized due to the one-bit quantization constraint. Afterward, encoder parameters are trained based on the previously trained and frozen decoder parameters. This resulted in a sub-optimally trained AE that failed to reach good performance for higher-order modulation. The work in [13] also failed to have generalization capability. Therefore we propose a new method of training as well as different architectures with the aim of making one-bit quantized reception possible at high information rate and good error performance using DL. Since AEs can be trained to perform various tasks, we will first begin by addressing the challenge of symbol-to-sequence mapping. And go on to explore possible NN-based enhancements to the conventional blocks of the system model for one-bit quantization and oversampling transceivers. Our investigation extends the roles of NN encoder and decoder to handle modulation, demodulation, and sequence design suitable for zero-crossing detection and de-mapping. Finally, we contemplate an end-to-end communication scenario, adapting the AE for seamless integration by consolidating the functionalities of all transceiver blocks.



(a) NN-Based Sequence Design

(b) NN-based Constellation and Sequence Design

(c) NN-based Error Correction, Constellation, and Sequence Design

FIGURE 2. A system model for the proposed autoencoder-based transceivers with one-bit quantization and oversampling.

A. NN-BASED SEQUENCE DESIGN

The task of the sequence design block is to produce codewords meeting the SE requirements necessary to approach Shamai's limit while simultaneously preserving resilience against ISI and channel-induced effects during zero-crossing detection. A communication system with a sequence design block enhanced by a NN would involve the transmission of the information bit sequence which is channel-coded and modulated to form a block of complex symbols $\mathbf{d} \in \mathbb{C}^{k/\log_2 M}$. The blocks of coded symbols are reshaped into smaller blocks before being input into the AE, in order to lower complexity. Moreover, the real and imaginary parts of the coded symbols are concatenated to be more suitable to be used as data for training and testing. The encoder portion of the AE will then receive the complex symbols as in Fig. 2(a), where it will be trained to optimize a codeword ideal for zero-crossing detection under one-bit quantization and oversampling scenarios. After the information is encoded into the distances between zero crossings, it can be recovered at the receiver even after one-bit quantization. This potential transceiver design is proposed with the objective of allowing simple transmission with one-bit quantization at the receiver. The structure is composed of a deep fully-connected $L^{(f)}$-layered auto-encoder and $L^{(g)}$-layered auto-decoder with

$N_L^{(f)}$ and $N_L^{(g)}$ nodes respectively. The NN-based symbol-to-sequence mapper performs the embedding $f_A : \mathbb{C}^{k/\log_2 M} \rightarrow \mathbb{R}^N$, where f is the encoder and this results in an N-dimensional signal $x = [x_1 \ldots x_N]$ from the designated transmitter. Non-trainable Lambda layers are introduced between the auto-encoder and auto-decoder to enable the adaptation of the signal to the specific channel characteristics it will encounter. These Lambda layers introduce FTN, power normalization, channel and oversampling as well as single-bit quantization, as seen in Fig. 2(a)-(c). The effects of FTN and oversampling on the transmitted vector as expressed in (2) are incorporated into an ISI filtering layer. The non-orthogonal transmission can be represented by the effect on the received vector, $y = [y_1\, y_2 \ldots y_N]$, as follows:

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{z}^{(\mathbf{g})} \tag{6}$$

where $\mathbf{G}$ is an $N \times N$ Toeplitz Gram matrix that represents the ISI and $\mathbf{z}^{(\mathbf{g})}$ denotes the filtered and oversampled noise vector. The overall channel distortion from all four non-trainable layers can be represented as

$$\lambda(\mathbf{x}) = Q\left(\mathbf{G}\mathbf{x}+\mathbf{z}^{(\mathbf{g})}\right) \tag{7}$$

where $Q(\cdot)$ quantizes each element of the received and oversampled vector. The transceiver is designed by the principle of minimizing the reconstruction error of the output from the input message vector, which is in parallel to the design goals of minimizing transmission error. After the signal undergoes distortion from the particular channel, the receiver decoder block that will decode receive the now quantized and noisy encoded data sequence and attempt to recover the original submitted coded symbols by using the transformation $g_A : \mathbb{R}^N \rightarrow \mathbb{C}^{k/\log_2 M}$. During training, the autoencoder will learn the encoder and decoder parameters for the two hidden-layer functions, f_A and g_A, such that the difference between the input and the predicted vectors is minimized by some loss function $L(\mathbf{d}, \widehat{\mathbf{d}})$ where $\widehat{\mathbf{d}} = g_A(\lambda(f_A(\mathbf{d})))$. Given the parameters of θ_{f_A} and θ_{g_A} for the encoder and decoder, respectively, the hidden-layer function must be determined to minimize a transmission error probability as follows:

$$\left(f_A{}^*, g_A{}^*\right) = \underset{(f,\, g)}{\arg\min}\, L\left(d, \widehat{d} \,\middle|\, \theta_{f_A}, \theta_{g_A}\right) \tag{8}$$

The NN encoder will project the input symbol into a higher dimensional codeword such that it will be robust against both the AWGN channel noise and quantization noise as well as ISI. In contrast to the sub-optimal two-step training policy in [12] to combat the gradient issue one-bit quantization layer, we adopt a simpler training policy. This involves introducing a hyperparameter during the training phase only. A hyperparameter ε is incorporated in the one-bit quantization layer of the training phase, such that the quantization function is adjusted to $Q(y_l) = \frac{y_l}{|y_l|+\varepsilon}$, where $\varepsilon \in [0, 1)$. This modification helps mitigate the hard quantization effect of squeezing all the weights and parameters into binary values and thereby avoids the occurrence of undefined values. Furthermore, it allows simple training without compromising performance as shown in Section V. During training, the neural network operates in a soft quantization environment, facilitated by the introduced hyperparameter. However, testing is conducted in a hard one-bit quantization environment. It is important to note that the hyperparameter requires careful tuning for optimal performance. Despite the addition of hyperparameters, the possibility of vanishing gradients remains a concern, particularly with the occurrence of the sigmoid activation effect at the soft-quantization layer. To address these challenges, various techniques are employed, such as utilizing ReLU activation and random weight initialization in the hidden layers. Additionally, the width and depth of the layers are optimized to balance the number of parameters required for robust encoding and decoding while avoiding the vanishing gradient problem.

B. NN-BASED CONSTELLATION AND SEQUENCE DESIGN

Another approach to improving communication over a one-bit quantized channel using an AE involves training the AE to perform both the design of the zero-crossing sequence and the modulation task. This approach enhances the capabilities of the AE beyond just designing the optimal codeword and expands its ability to encode information in a way that is better suited for one-bit quantization. As seen in Fig 2(b), the outputs of the error correcting block are input into the auto-encoder, which is responsible for the encoding function $f_B : \mathbb{B}^k \rightarrow \mathbb{R}^N$. The Lambda layers are maintained for the scheme as well in order to assimilate the effects of FTN, channel, oversampling, and quantization. The decoder will have the task of de-mapping and demodulation such that $g_B : \mathbb{R}^N \rightarrow \mathbb{B}^k$. If the AE is trained well, then the use of the NN in the architecture allows the optimization of the function of the two different conventional blocks in the transceiver. As seen in the seminar of [27], AEs can be trained to learn (produce) constellations suited for specific channels and have superior performance. In this scheme, the constellation design is optimized jointly with the sequence design. The conventional works [3], [4], [5] that utilized RLL have tried to design the transmit signal by employing zero-crossing modulation (ZXM) which is a modulation matched to a receiver employing one-bit quantization and temporal oversampling, only tractable for low modulation orders. The use of the NN-based constellation and sequence design allows communication at a higher information rate that would have been curbed by the lower modulation levels. The hyperparameter ε is still in use for the one-bit quantization layer. Then, the AE is similarly trained to the transceiver with an NN sequence design block.

C. NN-BASED ERROR-CORRECTION, CONSTELLATION, AND SEQUENCE DESIGN

The third transceiver design that is proposed makes use of a black-box AE that is expected to perform the tasks of all the blocks of the comprehensive system model described in Fig. 1. While training to map the input vector to an embedding vector, the AE learns ideal encoding weights that would replace the functions of modulation, channel coding, and symbol-to-sequence mapping. This simplified communication system using an AE end-to-end, as shown in Fig. 2(c), would involve the transmission of a bit sequence $\mathbf{b}$ such that $\mathbf{b} \in \mathbb{B}^s$ is transmitted making N uses of channel. The transmitter performs embedding $f_C : \mathbb{B}^s \rightarrow \mathbb{R}^N$ to output the N-dimensional signal $\mathbf{x}$. The signal will undergo distortion from the particular channel in the form of Lambda layers before it encounters the receiver block where transformation $g_C : \mathbb{R}^N \rightarrow \mathbb{B}^s$ allowed the recovery of the original information bits. By learning the optimal encoder-decoder pair, our approach achieves accurate signal recovery after one-bit quantization for high information rates. We optimize the various blocks within the conventional system design concurrently, maintaining the use of the hyperparameter ε in the one-bit quantization layer. The transceiver minimizes reconstruction error from input to output (end-to-end), aligning with the goal of minimizing transmission error. The NN encoder ensures robustness against AWGN channel noise, quantization noise, and ISI, learning an ideal bit-to-sequence mapping during training. Addressing the highly ISI-afflicted channel, a complex decoding block is replaced by an NN-based decoder. The decoder handles what would have been the task of de-mapping, demodulating, and decoding, recovering original message bits from the received one-bit-quantized codewords. Leveraging methodologies from [27], these NN blocks as transmitter and receiver enable low-complexity one-bit quantization, providing a framework for transceiver performance optimization.

If perfectly trained, the encoder part of our proposed AE jointly optimizes channel coding, modulation, and sequence generation by end-to-end training. As a result, we introduce a new parameter κ, which represents the ratio of the input bits to the transmitted sequence after error correction, modulation, and generation of sequence suited for one-bit quantization. κ can be defined as the bit-to-sequence mapping rate. It corresponds to the rate of the proposed AE that is given as

$$\kappa = r \cdot \Delta \cdot \log_2 M \tag{9}$$

Accordingly, the SE and problem formulation in (5) now can be restated for the proposed scheme as follows:

$$\begin{aligned}
&(\kappa^*, M_{Tx}^*) = \underset{(\kappa, M_{Tx})}{\arg\max}\ (\kappa M_{Tx}) \\
&s.t. \kappa \cdot M_{Tx} \leq 2\log_2(M_{Rx} + 1) \\
&M_{Tx} \leq M_{Tx}^{(th)}, M_{Rx} \leq M_{Rx}^{(th)}, \\
&\gamma(\kappa, M_{Tx}, M_{Rx}) < \gamma^{(th)}
\end{aligned} \tag{10}$$

The parameter described in equation (9) indicates the efficient signal-packing capability of the autoencoder in the presence of high ISI caused by FTN with one-bit quantization and oversampling. This capability enables the AE to learn optimal parameters for the block that is represented by the encoder. Note that it significantly simplifies the problem formulation. By our AE design, we can observe an inherent trade-off between transmitting SE and BER. In a given operating SNR region, the optimized SE can be heuristically found by manipulating the value of bit-to-sequence mapping rate κ for our proposed AE transceiver. The simplification of the problem formulation using κ has facilitated trade-off analysis by providing a simple parameter, which in turn enables heuristic optimization.

The new transceiver cannot help with the explicit solution to the above optimization problem as the bit error rate as a function of κ, M_{Tx}, and M_{Rx} is unknown. However, the particular values for κ, M_{Tx}, and M_{Rx} can be found using the new non-complex model for one-bit quantization channel. Compared to conventional schemes for one-bit quantization and oversampling which require huge block-by-block optimization, however, the AE structure for the one-bit channel allows end-to-end optimization through backpropagation. Such a simple optimization process allows us the flexibility in optimizing system parameters described in our problem formulation (3) by trial and error.

Finding transmitting schemes in the noisy channel that can achieve SE near the defined limit within an acceptable BER range is still an open problem. The limiting factor in the previous conventional systems was the symbol-to-sequence mapping and its resulting loss in information rate. Since the limit is not bound by a specific transmitting scheme, our proposed one-bit quantized AWGN channel autoencoder can open the door to observing how much of Shamai's limit can be achieved. Thanks to the straightforward end-to-end autoencoder design, we may be able to optimize the one-bit quantization system model across a range of performance parameters with ease. As a result, selecting the optimal value of κ for a given BER performance is a simple task. We demonstrate the efficacy of our new simplified problem formulation by simulating selected cases in Section V.

IV. DL-BASED DESIGN FOR ONE-BIT TRANSCEIVER WITH FTN AND OVERSAMPLING IN FADING CHANNEL

Although the overall structure of the autoencoder transceiver design has been explored in many previous works [27], [28], [29], [30], [31], [32], [33], [34], our proposed training and testing method has facilitated the extension to one-bit quantization. Additionally, in this section, we investigate the necessary adjustments to the end-to-end transceiver architecture to enable one-bit quantization within fading channel parameters. While the AWGN model holds significant value in the research community and has driven numerous advancements in communication theory, a design that accommodates a fading channel environment offers a closer approximation to reality and presents a genuine challenge to the system's capabilities. However, to deploy one-bit quantization effectively in a fading channel scenario, especially at information rates competing with high-order modulation, we must introduce new functional blocks beyond the transceiver depicted in Fig. 1—specifically, pilot generation and channel estimation. These blocks need optimization for seamless integration into the system. Furthermore, expanding the DL solution initially designed for AWGN to a real-world environment demands consideration of the impact of one-bit quantization on performance. The challenge arises as received pilots undergo quantization before reaching the channel estimation block, necessitating a large number of pilots. Thus, a critical investigation is required to understand how a DL-based transceiver can adeptly communicate in scenarios involving *faded* and *quantized* signals without a dire SE sacrifice.

The receiver in such a scenario would have to estimate channel from quantized signals which poses the challenge of finding the correlation of the quantized signal to the encountered channel. While no traditional methods have endeavored to make one-bit quantization operational in fading channel for SE competitive to high order modulation, the work in [15] has explored the integration of OFDM with one-bit quantized reception through supervised deep learning models where the scheme relies on the quantized pilots to estimate the channel, and subsequently communicates this information to the transmitter for signal precoding. It is worth noting that this approach was only implemented for QPSK modulation. By utilizing an end-to-end NN-based scheme, we propose two different transceiver designs for one-bit quantization in fading channel. In an autoencoder-based communication system for a standard transceiver without one-bit quantization, pilots can be used to facilitate channel estimation and improve the system's performance. On the other hand, a pilotless system can make use of superimposed pilots in the autoencoder-generated transmitted codewords to reduce the overhead of pilot signaling and increase the data rate. The decision to employ pilots or adopt a pilotless approach hinges on the tradeoff between channel estimation accuracy and data rate. However, the introduction of one-bit quantization in transceivers complicates this trade-off, as channel estimation becomes challenging when the arriving pilots at the receiver are quantized. The work in [15] has derived an expression to demonstrate that a fading channel would be estimated perfectly with one-bit ADCs if there was a very large number of pilots. The number of pilots required will certainly influence how much SE would be sacrificed to achieve a certain BER performance expected for such a system. Consequently, in this work, we compare a pilot-based end-to-end NN scheme with one that operates without pilots, examining the implications for SE and BER performance.

A. PILOT-BASED TRANSCEIVER DESIGN

The pilot-based autoencoder transceiver that is being proposed in this section is portrayed in Fig. 3(a) where the communication system is designed specifically for one-bit quantization in fading channel using a sequential learning method and expert knowledge. In order to extend our results in AWGN channel to fading channel, several different manipulations for robustness are made to the end-to-end AE transceiver design, since the distortion between the encoder and decoder is more severe due to channel fading as well quantization. The fully connected layers in the previous AE are exchanged with one-dimensional convolutional layers, with the aim of taking advantage of their correlation property. By helping resolve the low correlation issue between quantized signal and channel, the convolutional layers can help boost reconstruction capability of the AE. Another modification made to the end-to-end AE is the incorporation of expert domain knowledge during the training process, specifically in the context of fading channels. By incorporating signaling and detection manipulation used during conventional communication over a wireless fading channel, we can improve the reconstruction ability of the NN [27]. Hence, we propose the separation of the decoder into two NN-based blocks to reflect the traditional roles of estimation and detection. As can be seen there, the system consists of an NN-based encoder, decoder, and channel estimation block as well as a traditional pilot sequence generation block.

After being transmitted through the channel, the pilot sequence $\mathbf{x_p}$ undergoes quantization, leading to the vector $Q(\mathbf{y_p}) = Q(\mathbf{hGx_p}+\mathbf{z}^{(\mathbf{g})})$ where $\mathbf{h}$ is the L-tap channel vector, $\mathbf{G}$ represents the ISI, and $\mathbf{z}^{(\mathbf{g})}$ denotes the oversampled noise. Furthermore, a sequential training method is applied to the design, where the estimation block is trained separately from the other blocks in the shaded region and frozen before being end-to-end training of the system, as shown in Fig. 3(a). During the end-to-end training, the encoder maps input $\mathbf{s}$ to the encoded data vector $\mathbf{x_d}$, which is equivalent to learning to take the data bits and encode them to be robust against random fading channel and quantization errors, as well as ISI due to FTN and oversampling. The encoded vector is then altered by the non-trainable Lambda layers portraying channel effects to construct the received signal $Q(\mathbf{y_d}) = Q(\mathbf{hGx_d}+\mathbf{z}^{(\mathbf{g})})$. The received signal is then passed through the estimator block, which is already trained to recognize the channel by comparing the received quantized and faded pilot sequence to the original pilot sequence. During its training, the estimator is exposed to a dataset of randomly faded and quantized pilot sequences labeled by the fading channel where a mean-squared logarithmic error (MSLE) loss function $L\left(\mathbf{h}, \widehat{\mathbf{h}}\right) = \frac{1}{B}\sum_{i=0}^{B}\left\{\log\left(h_i+1\right) - \log\left(\widehat{h}_i+1\right)\right\}^2$ is used where h_i and $\widehat{h}_i$ are the elements of the channel vector $\mathbf{h}$ and its estimate $\widehat{\mathbf{h}}$ and where B is the designated batch size. MSLE is chosen as a loss function because the particular regression task of estimating the channel has a target value distribution that is quite dispersed and the natural logarithm before calculating the minimum square error has the effect of relaxing the punishing effect of large differences in predicted values. Here, once the estimator block has been trained to reach the loss value for peak estimation capability, its weights are made non-trainable. The estimated channel is then used to aid the decoder in detecting the transmitted signal and reconstructing the original bits.

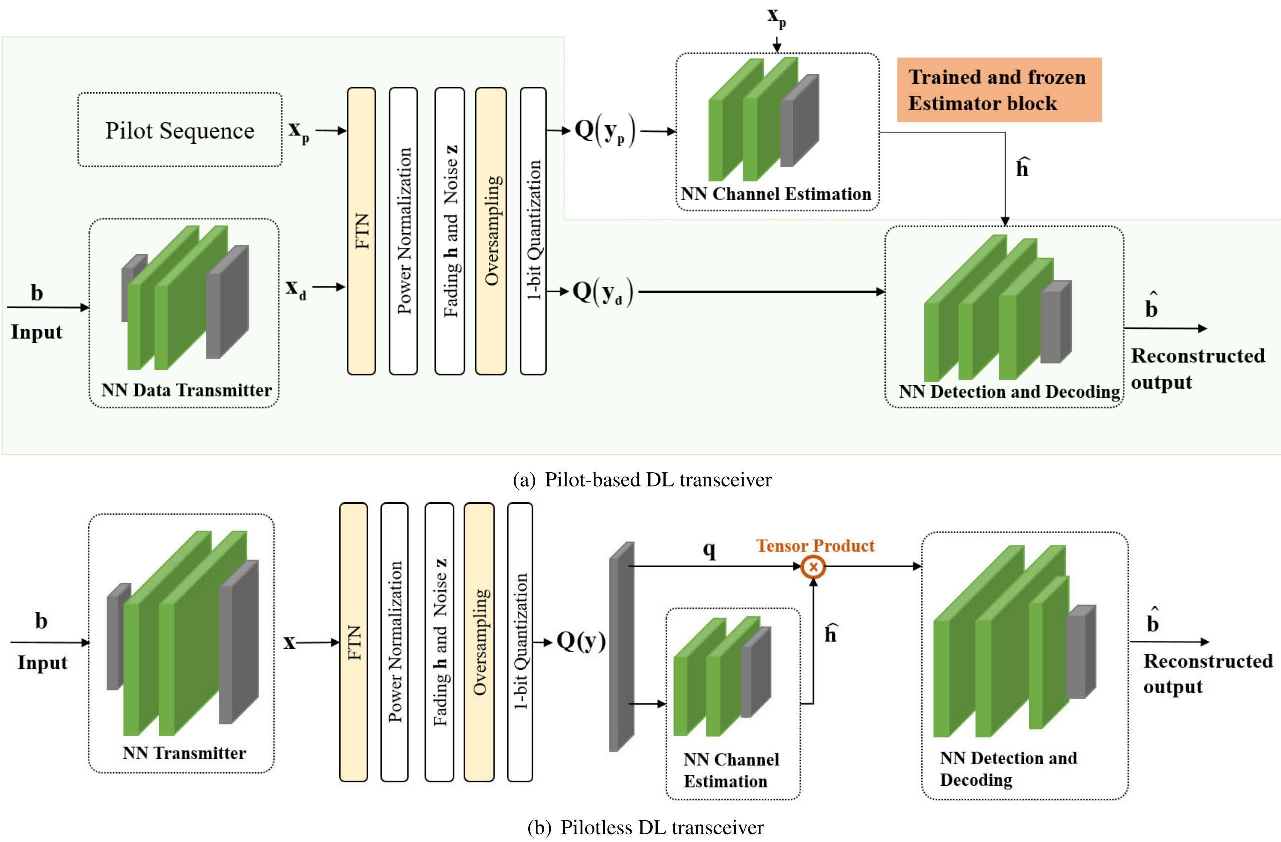


FIGURE 3. **Proposed transceiver designs for fading channel.**

Overall, the pilot-based autoencoder transceiver is a powerful communication system that can effectively deal with the challenges posed by one-bit quantization, channel fading, and ISI. The system leverages an NN-based estimator to accurately estimate the channel, which is then utilized to assist the decoder in effectively decoding the received signal.

B. PILOTLESS TRANSCEIVER DESIGN

Our second proposed scheme aims to achieve transmission without the use of explicit pilots, relying solely on our end-to-end training approach and convolutional layers. In this scheme, the full channel estimation pilots are only implicitly present and are learned by the encoder and decoder through the process of backpropagation. The learned constellation during a training process implicitly reflects the channel conditions and is leveraged by the trained decoder to detect and reconstruct the transmitted signal. Here, matched filtering is fully taken care of by the neural network. Furthermore, a high oversampling rate on the receiver side is critical so as to increase the dimension reduced by quantization. Since the proposed transmission scheme is pilotless, the accuracy of channel estimation and equalization, as well as the resulting performance on fading channels, depends on the degree of oversampling, which controls the dimension of the received vector. The robustness of the neural network-based transceiver allows the scheme to disregard the high ISI that would be encountered by high oversampling and FTN rates, which was only done for AWGN channel in [25].

Here, we also make use of convolutional layers as well as expert domain knowledge in order to make pilot-less communication possible over one-bit quantized fading channel. In our proposed scheme, as shown in Fig. 3(b), during a process of training on a fading channel, the autoencoder portion aims to leverage conventional processing techniques by dividing the block into an NN-based channel estimator and an NN-based detection and decoding block. The convolutional layers of the channel estimation block are used to obtain comprehensive channel characteristics that can be given to the detection block. With the lack of explicit pilots, there is still a need to infer the channel information for good data reconstruction. The lack of explicit pilot symbols and the dependence on the received vector $\mathbf{y}$ to observe channel characteristics means that the necessary correlation for data recovery cannot be fully utilized at every node of the convolutional layer, especially as the signal will be quantized before processing. Hence, there is a need to boost the channel information observed by the decoder.

Inspired by the work in [28] which makes use of bilinear product to augment pilotless end-to-end communication for MIMO, we incorporate a tensor product between our

NN-based channel estimation and detection blocks. The tensor product function, which is a multi-linear algebra tool, uses the outer product of two vectors to connect the output of a channel estimation module to the input of a detector by means of creating a new vector that combines the channel estimates with the received signal. The tensor product of the feature vectors that are typically representing different information results in a vector of outer productions for each position of the received signal. The bilinear features of the resulting tensor capture the interactions between these different pockets of information for extracting higher-level features that can be harnessed for more robust pilotless detection. For a transmitted vector $\mathbf{x}$, when the channel information from the NN estimation block, represented by a vector $\widehat{\mathbf{h}} \in \mathbb{R}^{E\times 1}$, is multiplied at every position by the received signal $\mathbf{q} = Q(\mathbf{y}) = Q(\mathbf{hGx}+\mathbf{z}^{(\mathbf{g})}) \in \mathbb{R}^{1\times N}$, the resulting tensor product is a new vector $\widehat{\mathbf{h}} \otimes \mathbf{q}$, that contains all possible products of the elements of $\widehat{\mathbf{h}}$ and $\mathbf{q}$. The tensor product is a bilinear map from the Cartesian product $\widehat{\mathbf{h}} \times \mathbf{q}$ to the tensor product space $\widehat{\mathbf{h}} \otimes \mathbf{q}$, which is defined as

$$\widehat{h} \otimes q := \widehat{h} q^T = \begin{pmatrix} \widehat{h}_1 q_1 & \widehat{h}_1 q_2 & \dots & \widehat{h}_1 q_N \\ \widehat{h}_2 q_1 & \widehat{h}_2 q_2 & \dots & \widehat{h}_2 q_N \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{h}_E q_1 & \widehat{h}_E q_2 & \dots & \widehat{h}_E q_N \end{pmatrix} \tag{11}$$

where $\widehat{h}_i$ and q_j are the elements of the vectors $\widehat{\mathbf{h}}$ and $\mathbf{q}$, respectively. Using a matrix-vector duality, the matrix can be reshaped into a vector of size $1 \times NE$. Here, the basis vectors, $\widehat{\mathbf{h}}_i$ and $\mathbf{q}_j$, whose weighted sums make up the vectors, $\widehat{\mathbf{h}}$ and $\mathbf{q}$, respectively, result in a new basis for $\widehat{\mathbf{h}} \otimes \mathbf{q}$ which is the set of all vectors of the form $\widehat{\mathbf{h}}_i \otimes \mathbf{q}_j$ for $i = 1, 2, \dots, E$ and $j = 1, 2, \dots, N$. This new vector can then be used as input to the detector. The tensor product magnifies all possible products of the received signal, which provides additional information that can be utilized by the convolutional layers of the detector block. By incorporating this additional information, the convolutional layers are better able to make the necessary correlations to recover the quantized and faded signals. Here, due to the absence of explicit pilot data and the reliance solely on the received vector $\mathbf{y}$ to observe channel characteristics, the correlation required for data recovery cannot be fully exploited at each node of the convolutional layers. Therefore, the tensor product operation effectively enhances the capability of the convolutional layers to exploit the correlation in the received signal, even in the absence of explicit pilot data. This leads to improved channel estimation and data recovery performance, which is particularly beneficial in scenarios where pilot symbols are not available or are limited.

V. SIMULATION RESULTS

A. COMPARISON OF NN-BASED DESIGNS OVER AWGN CHANNEL

Three transceiver designs detailed for AWGN channel are implemented and compared on similar simulation settings for fair comparison. The common architecture used for the AE in each scheme is presented in Table 1, where u represents the blocks of information that are processed by the AE. All AE-based transceivers were implemented using Keras with TensorFlow as the backend. Training was performed using Adam optimizer with a learning rate of 0.001 and categorical cross-entropy loss function. Batch normalization and ReLU activation are applied in the hidden layers of the encoder and decoder for all scenarios. Additionally, the AE employs one-hot encoding to represent input information and there is a

TABLE 1. Parameters for AWGN transceivers.

	Type of Layer	Output Size
Encoder	Input	2^u
	FC + ReLU	u/κ
	FC + Linear	u/κ
Decoder	FC + ReLU	$2 * u/\kappa$
	FC + ReLU	$2 * u/\kappa$
	FC + ReLU	$2 * u/\kappa$
	FC + SoftMax	u/κ

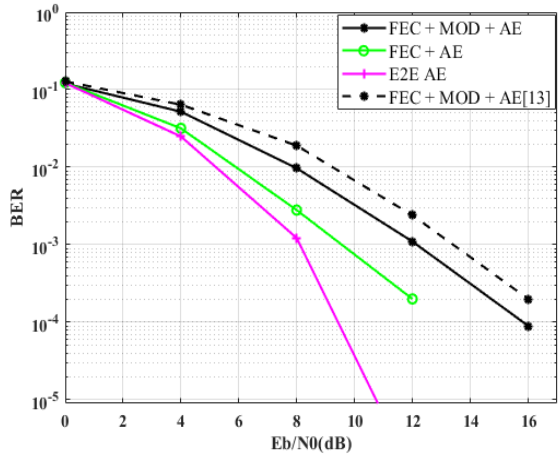


FIGURE 4. BER comparison over AWGN channel for proposed transceiver designs at $M_{Tx} = M_{Rx} = 10$.

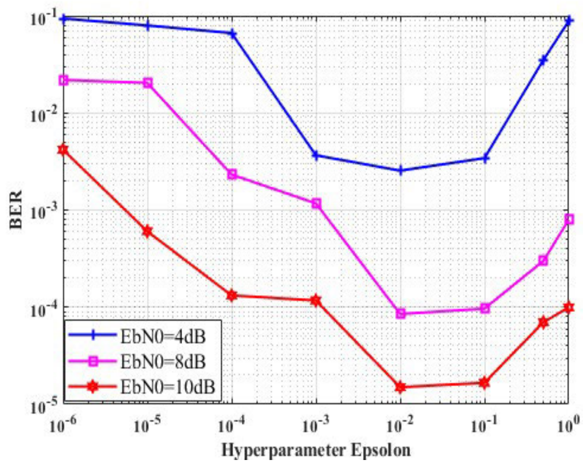
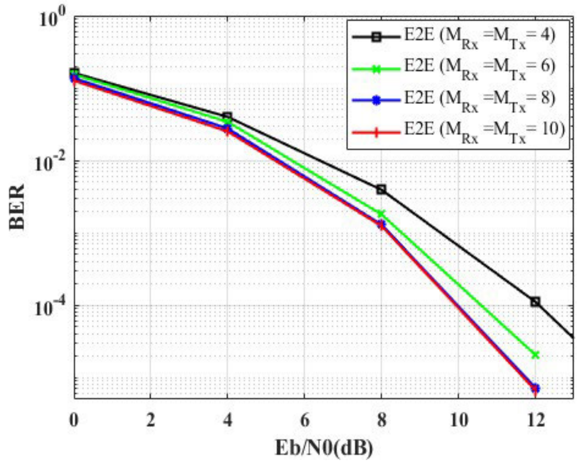
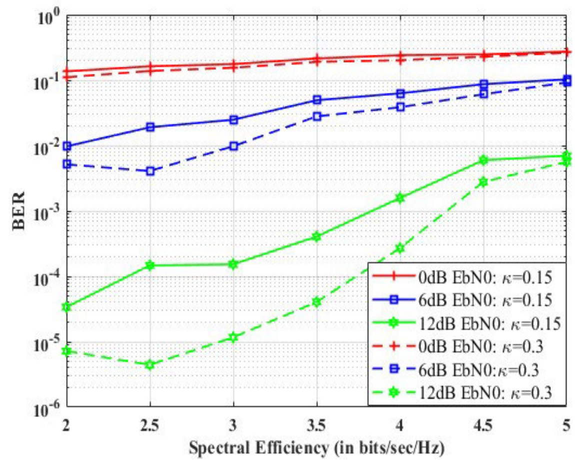


FIGURE 5. Hyperparameter Tuning for proposed DL-transceiver.



(a) Effect of increasing oversampling and FTN rate



(b) Effect of only increasing FTN rate

FIGURE 6. **Effect of oversampling and FTN rate on BER performance of end-to-end transceivers over AWGN channel.**

Softmax activation function at the decoder. The training SNR is tuned to various parameters, such as power normalization, designated equivalent modulation scheme, and the amount of ISI observed. In our simulations, we considered the offline training of the AEs with the proposed layout. Subsequently, our numerical results were obtained by Monte-Carlo simulation of the trained autoencoder with a test dataset. The performance of the proposed encoding schemes in terms of SE and BER is evaluated for a fixed oversampling rate of $M_{Tx} = M_{Rx} = 10$. In the first scheme, LDPC is used for error correcting with a block length of 640 bits and a rate of $\frac{1}{2}$. Furthermore, the modulation block in this case is set to 16-QAM and the modulated symbols are fed to the AE in blocks of $u = 16$. When we evaluate the computational complexity of the proposed systems, drawing a direct comparison between the end-to-end approach and the conventional communication system proves challenging due to the complexity being influenced by numerous parameters inherent to the specific architecture employed. If we take the complexity of the transmitter and receiver as the maximum of the complexity for each of the individual blocks, then the complexity for the transmitter in the first transceiver is $\max\left\{\mathcal{O}(K\cdot V),\ \mathcal{O}(\log_2 4),\ \sum_{L^{(f)}=1}^{2}\mathcal{O}(u/\kappa)\right\}$ for the LDPC encoding, modulation, and NN encoder respectively, where K is the length of the LDPC codeword and V is the number of non-zero elements in the parity check matrix. Likewise, the receiver will have complexity $\max\left\{\sum_{L^{(g)}=1}^{4}\mathcal{O}(u/\kappa),\ \mathcal{O}(\log_2 4), \mathcal{O}(K\cdot I\cdot D)\right\}$ where I is the number of decoding iterations and D is the degree of the variable nodes. Therefore, it is logical to assume that the end-to-end transceiver will have equal or lower complexity than its two counterparts. The second proposed transceiver also uses a similar LDPC for error correction, concatenated with NN-designed constellation and sequence. Here, coded bits are fed into the AE in blocks of $u = 16$.

Fig. 4 compares the BER of all three schemes at SE = 3 bits/sec/HZ. This result shows that the end-to-end AE has the best performance, outperforming the first and second schemes by 5 and 2 dBs at BER of 10^{-4}, respectively. Furthermore, enabling the NN to jointly learn constellation and sequence mapping has been shown to provide a 3dB gain at a BER of 10^{-4}. This implies that joint optimization of all blocks has better performance in one-bit AWGN channel than the separate blocks of the conventional schemes with some enhanced NN blocks. Furthermore, all 3 transceivers outperform the work done in [13] at the given SE. The first NN transceiver with concatenated LDPC and modulation block differs in training methodology from the work in [13] by the use of hyperparameter for circumventing the vanishing gradient effect of one-bit quantization, and also by using cross-entropy loss function as opposed to mean-squared-error (MSE). Furthermore, the generalizability of all the AEs is evident in their capability to accurately reconstruct examples beyond the training dataset, which [13] was unable to achieve. Hyperparameter tuning is performed for the optimal ε value used in the training for all the transceivers. If the value of the hyperparameter is excessively high, the AE might not perform well in the hard quantization testing environment as it is dissimilar to its soft quantization training. If ε is too small, the AE does not converge as fast to the global optima because of vanishing gradients. The tuning for the end-to-end AE is shown in Fig. 5 by comparing BER performance in the testing phase where $\varepsilon = 10^{-2}$ is found to be the best-performing hyperparameter for the observed cases.

Fig 6 investigates how different time-packing factors impact performance. By maintaining the SE goal as a percentage of Shamai's limit, we assess the end-to-end AE's performance at oversampling and FTN rates $M_{Rx} = M_{Tx} = 4, 6, 8$, and 10 in Fig. 6(a). Increasing the

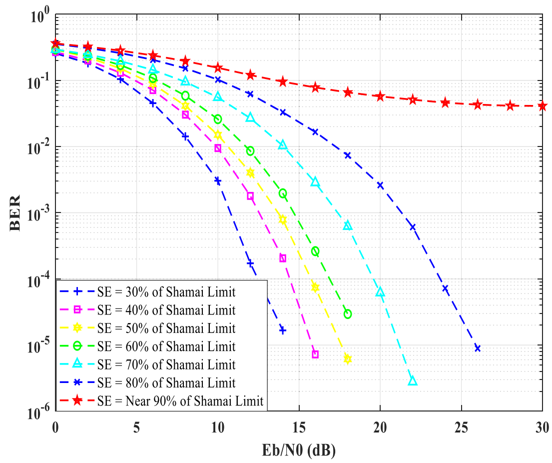


FIGURE 7. Performance of end-to-end AE at different Spectral Efficiency (SE-BER trade-off).

oversampling rate means increasing the effective resolution of the received signal, allowing for more accurate detection of the transmitted symbol. Therefore, by escalating the transceiver oversampling rate from $M_{Rx} = M_{Tx} = 4$ to 8, gains in BER are shown. However, at $M_{Rx} = M_{Tx} = 10$, the gain saturates. Beyond this threshold, further time packing and oversampling do not offer extra information to enhance detection accuracy, yielding no significant performance gains and therefore, we can determine the optimal rates for a given SE goal. In Fig. 6(b), we explore how our scheme can be used to achieve SE close to Shamai's limit by utilizing FTN paradigm and the simpler optimization offered by the use the AE. By fixing the oversampling rate at $M_{Rx} = 8$, we can increase FTN rate M_{Tx} of the system beyond the information rate needed to capture the information for oversampling. This will allow us to reach for higher and higher SE. However, the corresponding ISI will have to be mitigated to some degree by the AE if good SE performance is required without compromising the BER. Hence, by increasing the SE from 2 to 5 bits/sec/Hz through only increasing our FTN rate, we can observe in Fig. 6(b) that the AE allows us to somehow maintain the BER performance while increasing information rates as much as 3.5 bits/sec/Hz (55% of Shamai's limit for the given oversampling rate). After that, the BER performance starts to deteriorate for both bit-to-sequence mapping rates (κ). Furthermore, we can observe that the BER for $\kappa = 0.3$ outperforms $\kappa = 0.15$ since the transceiver with the latter required more FTN packing to reach the same spectral efficiency and the incurring ISI compromises BER performance.

For a given FTN and oversampling rate of $M_{Tx} = M_{Rx} = 10$, by setting a particular SE target, we can design the parameters of the autoencoder in terms of κ and compare its BER performance which can be observed in Fig. 7. The performance at each level of SE is depicted, suggesting that the proposed scheme has operational BER performance at information rates as high as 80% of Shamai's limit. The different slopes of the relationship allow the selection of optimal SE for designated reliable communication requirements. To illustrate, for a system with defined 10^{-4} BER for reliable communication and a minimum Eb/N0 of 24 dB, then the best possible SE is at 80% of Shamai's limit which is found to be 5.6 bits/sec/Hz.

Fig. 8 compares the proposed DL solution with the conventional transmission for one-bit quantization similar to the work in [4]. The conventional system utilizes Finite-state-machine (FSM) based RLL sequences to generate zero crossing transmit signals with a soft-input soft-output decoding using BCJR, and employs LDPC error correcting. By setting oversampling rate at $M_{Rx} = 9$, we compare the performance of the 2 one-bit transceivers by varying FTN rates and SE. The comparison shows that the conventional and the proposed DL-transceiver can reach this low SE goal of 2 bits/sec/Hz with low oversampling nearly equally. But at higher oversampling ($M_{Tx} = 6$), DL is able to deal with ISI much better and outperforms the conventional by about 2 dBs at 10^{-4} BER. The gap widens as the SE goal increases to 2bits/sec/Hz, where even at lower oversampling of $M_{Tx} = 3$, the conventional system cannot generate compact and robust sequences for one-bit quantization as the DL. At low SNR, the threshold effect of LDPC, where a certain SNR value is needed for reliable communication creates lower BER performance in comparison the DL transceiver which is not conventionally coded. The effect of quantization on an end-to-end DL scheme is shown in Fig 9, which portrays the performance of the AE from the seminal work [27] as well as the same AE when one-bit quantization layer before decoding and also when FTN transmission is considered. As can be observed, the AE from [27] cannot deal with the non-linearity of one-bit quantized layer. Furthermore, for the sake of fair comparison in terms of SE, when FTN with rate $M_{Tx} = 3$ is applied, a larger codeword is used for transmission and the performance is shown to be improve upon the original AE. Here, for the proposed AE, its training using soft quantization allow to training over the non-linear quantization layer. Therefore, the proposed AE with a quantized layer approaches the non-quantized AE from [27] with only 3dB difference at BER of 10^{-5}.

B. IMPLEMENTATION OVER FADING CHANNEL

The AE transceivers for fading channel were trained in a similar fashion to the end-to-end AE for AWGN channel, with Monte-Carlo simulation to obtain our numerical results. The simulations are performed over Rayleigh multipath fading channel with $L = 3$ channel taps using the AE architectures shown in Table 2. In Fig. 10, we can observe the performance of the two proposed AE-based transceivers in the given settings along with some performance references at same SE. One of the baselines is the BER of a 16-QAM modulation signal with LDPC error correction over a similarly tapped Rayleigh Fading channel with unquantized reception and minimum mean-squared error (MMSE) estimation. In this

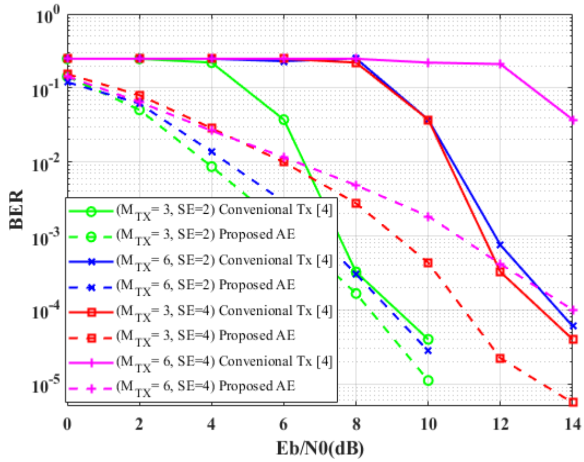


FIGURE 8. Performance comparison with conventional system.

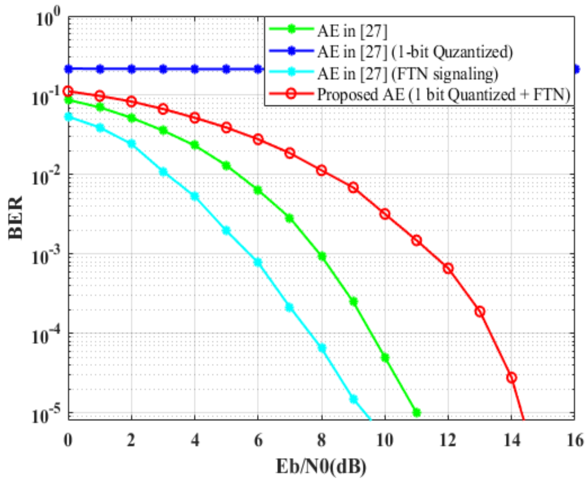


FIGURE 9. Effect of Quantization and FTN.

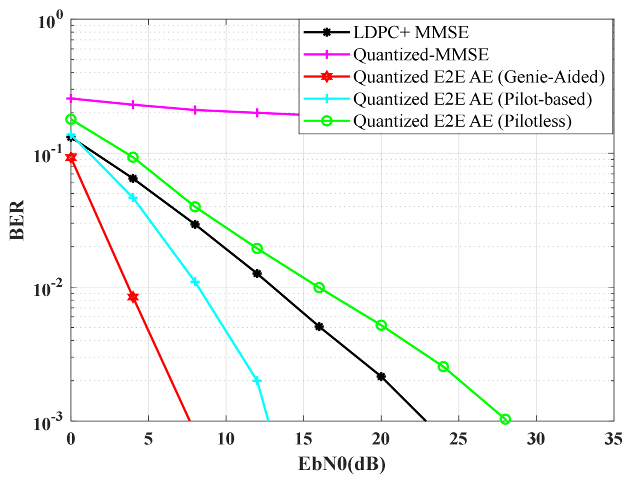


FIGURE 10. Performance comparison of pilotless vs. pilot-based transceivers over quantized Rayleigh fading channel.

context, it is assumed that pilots are sent before each block transmission in order to determine the characteristics of the channel. Moreover, the performance of the conventional

TABLE 2. Parameters for fading channel transceivers.

	Type of Layer	Kernel Size	Output Size
Pilot-based Transceiver			
Encoder	Input		2^u
	Conv1D + ReLU	3	$6 * u/\kappa$
	Conv1D + ReLU	3	u/κ
	FC + Linear		u/κ
Decoder	Conv1D + ReLU	3	$6 * u/\kappa$
	Conv1D + ReLU	3	$3 * u/\kappa$
	FC + ReLU		$3 * u/\kappa$
	FC + ReLU		$3 * u/\kappa$
	FC + SoftMax		2^u
Pilotless Transceiver			
Encoder	Input		2^u
	Conv1D + ReLU	3	$6 * u/\kappa$
	Conv1D + ReLU	3	u/κ
	FC + Linear		u/κ
Decoder	Conv1D + ReLU	3	$6 * u/\kappa$
	Conv1D + ReLU	3	$3 * u/\kappa$
	FC + SoftMax		2^u
Estimator	Conv1D + LeakyReLU	5	256
	Conv1D + LeakyReLU	5	128
	FC + ReLU		u/κ
	FC + Sigmoid		u/κ

estimation and detection scheme when using 1-bit ADC's is investigated and we can observe that the reconstruction using conventional means is quite unmanageable. Additionally, we are employing a genie-aided scheme as a comparison lower bound, in which perfect channel state information is furnished to our end-to-end AE transceiver. This approach eliminates the necessity for a channel estimation block in the two proposed schemes designed for a fading channel. In this genie-aided scheme, the AE is able to recover the quantized and faded signal quite well once it is provided with full channel coefficients and its performance can serve as the goal for our pilot-based and pilotless schemes.

The pilot-based scheme uses the trained and frozen estimator to capture channel information and provide it to the AE transceiver. In order for the estimator block to estimate the channel, at least one pilot symbol must be inserted for a certain number of symbols, which we denote by $\zeta = n_p/n_d$ where n_p represents the number of pilot symbols required to estimate a block of n_d data symbols. For $\zeta = 0.2$, it has been demonstrated that the pilot-based design outperforms the non-quantized LDPC-MMSE scheme by up to 10dB at a BER of 10^{-3}. On the other hand, the pilotless scheme does not perform as well as the pilot-based scheme with a 5dB disadvantage to the non-quantized LDPC-MMSE scheme at a BER of 10^{-3}. For the BER comparison of the schemes, we are disregarding the SE lost due to pilot transmission. The pilotless scheme allows high-modulation transmission for one-bit quantization in fading channel without any explicit pilots, which could offer the SE advantage when we consider the SE lost due to pilots.

As shown in Fig. 11, when the pilotless and the pilot-based transceivers are compared in the SE vs. BER graph by taking into account the SE lost due to pilots in transmission, we can

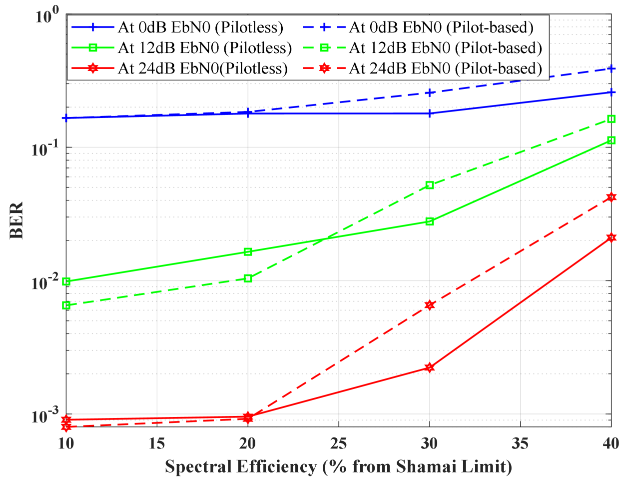


FIGURE 11. SE vs. BER for pilotless and pilot-based transceivers over Rayleigh fading channel.

see that the pilotless transceiver offers a BER advantage on average at $Eb/N0 = 0, 12$ and 24 dB. The amount of SE that is sacrificed due to the steep number of pilots needed for one-bit communication in fading channel makes pilotless transmission an attractive option. We can also observe that at very low SE, both schemes have similar BER performance, and therefore, either can be used in such requirement scenarios.

The ultimate design goal for such systems is to balance the tradeoff between SE and accuracy, taking into consideration the specific requirements and limitations of the application scenario. For example, in a low-bandwidth communication system, achieving high SE might be the priority, whereas in a mission-critical communication system, high accuracy might be more important. One possible practical application of pilot-based and pilotless one-bit communication systems is in the development of low-power, high-speed wireless communication systems for Internet of Things (IoT) devices and next-generation networks, where power efficiency and SE are crucial due to the limited power availability and bandwidth constraints. By carefully balancing the tradeoff between SE and accuracy, both pilot-based and pilotless one-bit communication systems can help enhance the overall performance and reliability of such systems.

VI. CONCLUSION

This work addresses the formidable challenge posed by the shift towards terahertz band communication, where power efficiency becomes a paramount concern. The implementation of one-bit quantization necessitates a re-imagination of traditional transmission and detection processes, leading to a critical limitation–the inability to achieve information rates sufficient to reach Shamai's limit. The intricate interplay of symbol-to-sequence mapping and zero-crossing detection becomes a stumbling block in conventional systems. Our approach stands out as a unique exploration into identifying the most effective transceiver design for this reimagined transmission paradigm involving one-bit quantized reception as well as oversampling in the time domain and faster-than-Nyquist (FTN) signaling.

By meticulously comparing conventional methodologies with previous DL-based approaches, we discern a specific path that allows us to approach Shamai's limit while maintaining robust error performance. Crucially, our work not only replaces the conventional sequence design for zero-crossing detection with a DL scheme but also quantifies the advantages DL can offer in this context. We address the limitations of prior DL-based works by enhancing BER performance at similar spectral efficiency (SE) levels, unraveling factors that contribute to increase in information rate, and achieving generalization in our proposed DL transceiver. By adopting the DL-based transceiver designs that were implemented for AWGN channels, the study achieves the feat of allowing the high information rate communication in fading channel with one-bit quantization receivers. Pilot-based and pilotless schemes are proposed to take on the challenge of fading channel estimation with quantized pilot sequences. The simulation results demonstrated that the pilotless autoencoder-based transmission is capable of offsetting the information rate lost due to the high amount of pilots needed for channel estimation after quantization and fading. Moreover, our work paves the way for power-efficient, high-frequency communication by allowing one-bit quantization to operate without substantial SE or BER performance loss. In essence, this research not only pushes the boundaries of current communication systems but also underscores the transformative potential of DL in shaping the landscape of wireless communication in the terahertz era.

In future research, the practicality and generality of these transceivers for one-bit quantization can be tested by implementing over real-world channel environment and also by incorporating more advanced DL techniques, such as meta-learning and transformers, to improve the performance and robustness of the transceiver design. Lastly, it could also be useful to extend the current work to consider more complex scenarios, such as multi-user or multi-antenna systems. This would enable the exploration of the performance of the DL-based transceiver in more realistic and challenging wireless communication environments

REFERENCES

[1] S. Shamai, "Information rates by oversampling the sign of a bandlimited process," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1230–1236, Jul. 1994.

[2] R. Deng, J. Zhou, and W. Zhang, "Bandlimited communication with one-bit quantization and oversampling: Transceiver design and performance evaluation," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 845–862, Feb. 2021.

[3] S. Bender, M. Dörpinghaus, and G. Fettweis, "On the spectral efficiency of bandlimited 1-bit quantized AWGN channels with runlength-coding," *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2147–2151, Oct. 2020.

[4] P. Neuhaus, M. Dörpinghaus, H. Halbauer, S. Wesemann, M. Schlüter, F. Gast, and G. Fettweis, "Sub-THz wideband system employing 1-bit quantization and temporal oversampling," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–7.

[5] G. Fettweis, M. Dorpinghaus, S. Bender, L. Landau, P. Neuhaus, and M. Schluter, "Zero crossing modulation for communication with temporally oversampled 1-bit quantization," in *Proc. Asilomar Conf. Signals, Systems Computers*, Pacific Grove, CA, USA, Nov. 2019, pp. 207–214.

[6] L. Landau, M. Dorpinghaus, and G. Fettweis, "Communications employing 1-bit quantization and oversampling at the receiver: Faster-than-Nyquist signaling and sequence design," in *Proc. IEEE Int. Conf. Ubiquitous Wireless Broadband (ICUWB)*, Montreal, QC, Canada, Oct. 2015, pp. 1–5.

[7] D. Marcela V. Melo, L. T. N. Landau, and R. C. De Lamare, "State machine-based waveforms for channels with 1-bit quantization and oversampling with time-instance zero-crossing modulation," *IEEE Access*, vol. 12, pp. 21520–21529, 2024.

[8] P. Neuhaus, M. Dörpinghaus, H. Halbauer, V. Braun, and G. Fettweis, "On the spectral efficiency of oversampled 1-bit quantized systems for wideband LOS channels," in *Proc. IEEE 31st Annu. Int. Symp. Pers., Indoor Mobile Radio Commun.*, Aug. 2020, pp. 1–6.

[9] R. R. M. de Alencar and L. T. N. Landau, "Continuous phase modulation with faster-than-Nyquist signaling for channels with 1-bit quantization and oversampling at the receiver," in *Proc. 54th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2020, pp. 17–21.

[10] L. T. N. Landau, M. Dörpinghaus, R. C. de Lamare, and G. P. Fettweis, "Achievable rate with 1-bit quantization and oversampling using continuous phase modulation-based sequences," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 7080–7095, Oct. 2018.

[11] X. Cheng, B. Xia, K. Xu, and S. Li, "Bayesian channel estimation and data detection in oversampled OFDM receiver with low-resolution ADC," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5558–5571, Sep. 2021.

[12] E. Balevi and J. G. Andrews, "Autoencoder-based error correction coding for one-bit quantization," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3440–3451, Jun. 2020.

[13] E. Balevi and J. G. Andrews, "High rate communication over one-bit quantized channels via deep learning and LDPC codes," in *Proc. IEEE 21st Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Atlanta, GA, USA, May 2020, pp. 1–5.

[14] Y. Zhang, M. Alrabeiah, and A. Alkhateeb, "Deep learning for massive MIMO with 1-bit ADCs: When more antennas need fewer pilots," *IEEE Wireless Commun. Lett.*, vol. 9, no. 8, pp. 1273–1277, Aug. 2020.

[15] E. Balevi and J. G. Andrews, "One-bit OFDM receivers via deep learning," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4326–4336, Jun. 2019.

[16] D. He, Z. Wang, T. Q. S. Quek, S. Chen, and L. Hanzo, "Deep learning-assisted TeraHertz QPSK detection relying on single-bit quantization," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8175–8187, Dec. 2021.

[17] Y.-S. Jeon, N. Lee, and H. V. Poor, "Robust data detection for MIMO systems with one-bit ADCs: A reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1663–1676, Mar. 2020.

[18] M. D. Gemeda, M. S. Han, A. T. Abebe, and C. G. Kang, "Deep learning transceiver for terahertz band communication system with 1-bit ADC and oversampling," in *Proc. IEEE 19th Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2022, pp. 515–516.

[19] M. D. Gemeda, M. S. Han, A. T. Abebe, and C. G. Kang, "Deep learning-based transceiver design for pilotless communication over fading channel with one-bit ADC and oversampling," in *Proc. 27th Asia–Pacific Conf. Commun. (APCC)*, Oct. 2022, pp. 60–65.

[20] H. Naaman, S. Mulleti, and Y. C. Eldar, "FRI-TEM: Time encoding sampling of finite-rate-of-innovation signals," *IEEE Trans. Signal Process.*, vol. 70, pp. 2267–2279, 2022.

[21] A. D. Liveris and C. N. Georghiades, "Exploiting faster-than-Nyquist signaling," *IEEE Trans. Commun.*, vol. 51, no. 9, Sep. 2003, Art. no. 15021511.

[22] J. Fan, S. Guo, X. Zhou, Y. Ren, G. Y. Li, and X. Chen, "Faster-than-Nyquist signaling: An overview," *IEEE Access*, vol. 5, pp. 1925–1940, 2017.

[23] Q. Shi, N. Wu, X. Ma, and H. Wang, "Frequency-domain joint channel estimation and decoding for faster-than-Nyquist signaling," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 781–795, Feb. 2018.

[24] W. Yuan, N. Wu, Q. Guo, D. W. K. Ng, J. Yuan, and L. Hanzo, "Iterative joint channel estimation, user activity tracking, and data detection for FTN-NOMA systems supporting random access," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2963–2977, May 2020.

[25] P. Song, F. Gong, Q. Li, G. Li, and H. Ding, "Receiver design for faster-than-Nyquist signaling: Deep-learning-based architectures," *IEEE Access*, vol. 8, pp. 68866–68873, 2020.

[26] N. Wu, Y. Zhang, H. Li, T. Zhang, J. Liu, X. Hou, and W. Liu, "Variational inference-based iterative receiver for unified non-orthogonal waveform (uNOW)," *IEEE Trans. Veh. Technol.*, vol. 73, no. 2, pp. 2848–2853, Feb. 2024.

[27] 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.

[28] H. Ye, G. Ye Li, and B.-H. F. Juang, "Bilinear convolutional auto-encoder based pilot-free end-to-end communication systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[29] H. Ye, G. Y. Li, B. F. Juang, and K. Sivanesan, "Channel agnostic end-to-end learning based communication systems with conditional GAN," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 1–5.

[30] S. Dörner, S. Cammerer, J. Hoydis, and S. T. Brink, "Deep learning based communication over the air," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 132–143, Feb. 2018.

[31] F. A. Aoudia and J. Hoydis, "Model-free training of end-to-end communication systems," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2503–2516, Nov. 2019.

[32] V. Raj and S. Kalyani, "Backpropagating through the air: Deep learning at physical layer without channel models," *IEEE Commun. Lett.*, vol. 22, no. 11, pp. 2278–2281, Nov. 2018.

[33] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.

[34] W. Yu, F. Sohrabi, and T. Jiang, "Role of deep learning in wireless communications," *IEEE BITS Inf. Theory Mag.*, vol. 2, no. 2, pp. 56–72, Nov. 2022.

[35] E. Soria-Olivas, *Handbook Of Research On Machine Learning Applications and Trends: Algorithms, Methods and Techniques*. Pennsylvania, USA: IGI Publishing, 2009.

METASEBIA DEMISSU GEMEDA received the B.S. degree in electrical and computer engineering from Addis Ababa University, Addis Ababa, Ethiopia, in 2019. She is currently pursuing the Ph.D. degree in electrical and computer engineering with Korea University, South Korea. Her research interests include the application of machine learning algorithms for PHY layer system design and optimization for beyond-5G and 6G wireless networks.

MIN SIG HAN received the B.S. degree in electrical engineering from Korea University, in 2016, where he is currently pursuing the Ph.D. degree with the School of Electrical Engineering. His research interests include machine learning-based cross-layer design and optimization for 5G/6G wireless communication systems. He was a finalist in the Communication and Networks Division at the 27th Samsung Humantech Paper Award, in 2020. He has also received the Best Paper Awards in the 27th IEEE Asia–Pacific Conference on Communications, in 2022.

AMEHA TSEGAYE ABEBE received the B.S. degree in electrical and computer engineering from Addis Ababa University, in 2011, and the Ph.D. degree in electrical and computer engineering from Korea University, in 2020. Then, he joined Samsung Electronics, where he is currently working on the physical layer standardization of 5G NR. His research interests include PHY and MAC layer design in wireless communication networks, including multiple and random access schemes, multiantenna communication, and AI/ML for air-interface. He has also received a KU Qualcomm Innovation Award and the Bronze Award in the IEEE Seoul Section Student Paper Contest, in 2017. He was a recipient of the Best Paper Awards from Korea Institute of Information and Communication Sciences Conference, in Summer 2015 and Winter 2018.

CHUNG G. KANG (Senior Member, IEEE) received the B.S. degree in electrical engineering from the University of California at San Diego, San Diego, in 1987, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California at Irvine, Irvine, in 1989 and 1993, respectively. He is currently a Distinguished Professor in the field of mobile communication and wireless networks. He is also a Full Professor with Korea University, Seoul, South Korea. He specializes in areas, such as massive MIMO, grant-free non-orthogonal multiple access, dynamic spectrum access, mobile caching networks, and AI-based communication system design. As a recipient of Dr. Irwin Jacobs Award sponsored by Qualcomm Inc., for his contribution to research on mobile communication technologies, he has over 250 refereed publications in international journals and conference proceedings in the fields of mobile communication and wireless networks, while holding over 100 patents. He also coauthored a reference textbook on the MIMO-OFDM wireless system, titled *MIMO-OFDM Wireless Communication with MATLAB* (Wiley, 2010). He served as the Chair for 2.3GHz IMT-WiBro Project Group (PG702) in Telecommunications Technology Association (TTA), which is a standard development organization (SDO) in South Korea. He is a Senior Member of IEEE COMSOC, IT, and VT, a KICS Fellow, and a member of the National Academy of Engineering of Korea. He has also played significant leadership roles in various organizations, such as the Chair of KICS Mobile Communication Technical Activity Group, the President of KICS, and the Chair of Wireless Technology Committee in 5G Forum of Korea. He also serves as the Chair for the Steering Committee in the Satellite Communication (SatCom) Forum of Korea.