Deep Learning-based Transceiver Design for Pilotless Communication over Fading Channel with one-bit ADC and Oversampling

Metasebia D. Gemeda, Min S. Han, Ameha T. Abebe, and Chung G. Kang
School of Electrical Engineering, Korea University
Seoul, Korea, Republic of
Email:{met4, als4585, ameha_tsegaye, ccgkang}@korea.ac.kr

***Abstract*—With the aim of addressing power consumption issues for terahertz band wireless communication, this work presents a deep learning-based solution for transceiver design with 1-bit quantization and oversampling at the receiver, and Faster-than-Nyquist transmission over fading channel. Specifically, by implementing the transceiver using a convolutional autoencoder, our work allows higher-order modulation transmission over one-bit fading channel without pilots. Transfer learning from previously trained blocks over simple noisy channel is used to minimize the probability of bit error and outperforms the convolutional autoencoder at low oversampling rates. The bit-error-rate gain offered at 20dB SNR by the transfer learning is seen to be as high as half of one order at low oversampling rates and to saturate as oversampling rate increases. Furthermore, by allowing explicit phase synchronization, the autoencoder-based transceiver with partial channel matching is able to approach unquantized performance with 4dB gap in Rayleigh fading environment.**

***Index Terms*—one-bit-quantization, oversampling, auto-encoder, deep learning**

I. Introduction

As the wireless communication industry reaches for higher data rate by transmission over terahertz band, the power efficiency of current analog to digital converters (ADCs) proves to be a problem. Moreover, it has been shown that power consumption can be reduced by limiting the resolution of the ADCs and relying on time domain resolution rather that amplitude resolution. Hence, transmission schemes utilizing one-bit ADC, which offers even better simplicity by removing all amplitude information, have been a subject of much interest. However, employing one-bit ADC results in performance degradation as the system is hindered from exploiting high-modulation-order signaling and it consequently results in reduced spectral efficiency (SE). In order to overcome this loss in information rate, oversampling in time domain can be implemented at the receiver along with Faster-than-Nyquist (FTN) signaling at the transmitter.

Oversampling takes advantage of high correlation between oversampled binary samples to allow zero crossing-based detection for the 1-bit ADC system. FTN increases the temporal resolution at the transmitter which allows to meet demand for zero crossings on a finer grid. Furthermore, the zero crossing-based detection redefines the achievable capacity in terms of the amount of oversampling. As shown in the work in [1], the achievable capacity for one-bit noiseless channel is bounded by Shamai's limit which 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).

A zero crossing-based scheme to detect a high-modulation-order signal requires the construction of a sequence which can be distinguished after one-bit quantization and oversampling, i.e., robust against quantization noise and inter-symbol interference (ISI). This results in expanded codewords and low information rates that fail to reach near the defined upper bound. On the other hand, FTN along with oversampling results in high ISI that will make the recovery of message bits challenging. Most of the conventional and deep learning (DL)-based one-bit transceivers have been proposed with the objective of improving either bit error rate (BER) or information rate performance in noisy and/or fading channels. In order to balance between robustness and higher information rates, a combinatorial optimization problem has been defined in section II. The work in [2] makes use of a conventional method like run-length-limited (RLL) sequences to encode information into the distance between consecutive 1's for ideal one-bit detection. However, the introduced channel memory for the long sequences require receivers with maximum-likelihood sequence detection to achieve predicted information rates.

One-bit transceiver designs using DL have been found to be beneficial in [4-6] by offering a less complex decoding as well an efficient way to reduce non-linearity introduced by the noisy or faded one-bit channel. They also allow joint optimization of the individual blocks of conventional communication systems where other attempts to jointly optimize components revealed intractable. In our previous work in [7], autoencoders (AEs) were used to generate encoder-decoder pairs to perform as the conventional of 1-bit transceivers over Additive White Gaussian Noise (AWGN) channel. The DL solution was able to achieve BER performance comparable to that of an end-to-end channel AE without the constraint of 1-bit quantization. The scheme also simplified the generalized problem for the l-bit channel while revealing a straightforward BER and SE trade-off. But as 1-bit quantization makes channel estimation and equalization challenging, transceiver designs over AWGN might not parallel real environment transmission. Quantization of the pilots results in the need for a large number of pilots in each channel coherence interval and therefore having a pilotless scheme is desirable. Therefore, our current work focuses on extending the DL based transceiver for pilotless transmission over fading channel.

This study proposes a scheme to allow pilotless transmission over fading channel with one-bit quantization and oversampling receiver by adopting DL methodologies. Specifically, it utilizes convolutional AEs for end-to-end communication with the main objective of learning the reconstruction task for faded and quantized transmission. Furthermore, the work aims to learn optimal joint error-correction and modulation as well sequence construction suited for the one-bit quantization and oversampling scenario for the best BER performance. The jointly optimized block of neural network aims at producing codewords with zero-crossing patterns that would be ideal for recovery after one-bit quantization as well as particular channel effects. Moreover, the parameters of the otherwise compartmentalized blocks can be represented with a single variable which allows us to not worry about many optimizable variables in system design. On the receiver side, the end-to-end training employed along with correlation property of convolutional layers have allowed joint optimization of channel estimation, equalization, and decoding blocks.

The proposed DL solution enables one-bit quantization along with oversampling and FTN channel to be operational for modulation orders as high as 64-QAM over fading channel without employing pilots. The use of increased oversampling rates is found to boost the performance of our proposed one-bit pilotless channel convolutional AE into an acceptable BER range. Furthermore, this work portrays the constellation BER gain offered by using Transfer Learning (TL) to improve initialization of training. Finally, a phase-synchronized version of AE with some pilots for partial channel estimation is discussed.

The rest of the paper is organized as follows. Section II presents a system model and problem formulation while the Proposed DL-based model is detailed in Section III. The simulation results are presented in Section IV and the paper concludes in Section V.

II. System Model and Problem Formulation

Fig. 1 presents a comprehensive system model that includes existing implementations performing 1-bit ADC and oversampling at the receiver and FTN signaling at the transmitter [2-6]. Let $\mathbf{b} \in \mathbb{B}^s$ denote information bits to be transmitted, which is then channel-encoded using with code rate $r = s/k$ whose output is given as $\mathbf{c} \in \mathbb{B}^k$. Afterwards, the coded bits $\mathbf{c}$ is modulated with an order of M (bits/symbol) to produce complex-modulated symbol $\mathbf{d} \in \mathbb{C}^{k/\log_2 M}$. In this system model, the transmitted signal has to encode its information in the distance between zero crossings, allowing good detection after the quantization at the receiver. This sequence construction has a symbol-to-sequence mapping rate of $\Delta \in (0, 1]$ that governs transmission rates. The mapping can be expressed as $\mathbb{C}^{k/\log_2 M} \to \mathbb{C}^N$ on the transmitter side such that $\Delta = k/N\log_2 M$ and it outputs the vector $\mathbf{x} = \{x_1, ..., x_N\}$. Finally, the DAC block transmits M_{Tx} symbols per one Nyquist interval T in an FTN manner by utilizing a pulse shaping filter $g(t)$ like a root-raised cosine (RRC) filter. The matched filter $g(-t)$ on the receiver side will filter the now noisy signal before temporal oversampling and 1-bit quantization is performed. Therefore, there will be the same ISI filtering on both in-phase and quadrature branches of the sequence of coded symbols.

For a P- path channel with impulse response of $h(t) = \sum_{i=1}^{P} \alpha_i u_i(t)\delta(t - \tau_i)$ 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 i-th path, the received signal at the 1-bit ADC can be represented as

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

where $z(t)$ represents AWGN and $*$ is the convolution operator. We assume a slowly varying random process that does not change with the duration of pulse, so we can drop time index on $u_i(t)$. The ADC will temporally oversample this received signal with a rate of M_{Rx} with respect to Nyquist rate, where the l-th sample can be represented as

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

where $\bar{g} = g * g^H$ is the combined channel filter which captures the effect of the transmitter filter, its matched receiver filter and ISI due to oversampling and FTN. Additionally, $z_l = \tilde{z}\left(\frac{lT}{M_{Tx}}\right)$ where $\tilde{z} = z * h^H$. The oversampled real and imaginary signals are then one-bit quantized by the ADC such that

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

where $sgn(x) = 1$ if $x \geq 0$ and $sgn\left(\mathrm{Re}(y_l)\right) = -1$ otherwise. The quantized signal is demodulated and decoded to reconstruct message bits $\widehat{b}$. Since the 1-bit ADC can only differentiate two levels in both real and imaginary dimensions, the maximum modulation order that can be implemented at the transmitter 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 $r\Delta M_{Tx}\log_2 M(bits/T)$, as depicted in Fig. 1. On a receiver side, the oversampling factor M_{Rx} determines the achievable rate of the scheme as given by $2\log_2(M_{Rx} + 1)(\mathrm{bits}/T)$ [1]. This leads to the following constraint:

$$r\Delta M_{Tx}\log_2 M(bits/T_N) \leq 2\log_2(M_{Rx} + 1) \tag{4}$$

Therefore, our main objective is to design a system that can approach the 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

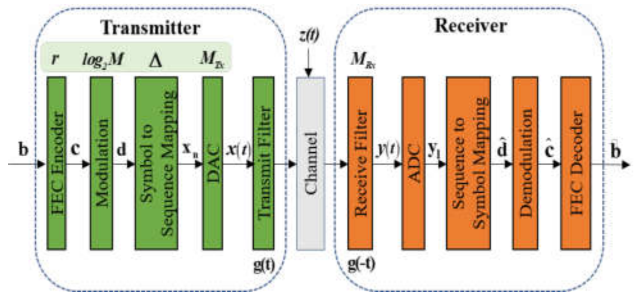


Fig. 1: Comprehensive system model for one-bit quantization and oversampling communication schemes

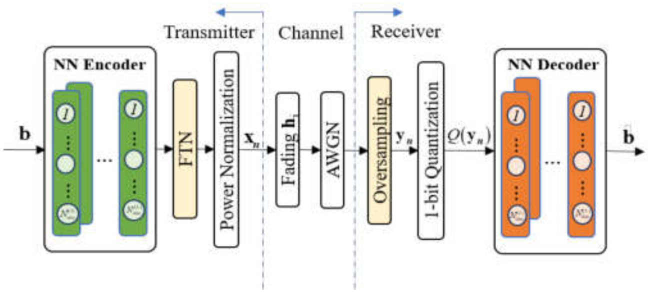


Fig. 2: A system model for the proposed autoencoder-based transceiver with one-bit quantization and oversampling

problem, the SNR γ , with its threshold $\gamma^{(th)}$ to satisfy a target BER, may constrain the SE maximization.

The BER is a function of M_{Tx}, r, M, Δ, and M_{Rx} as error is incurred by quantization and ISI, in addition to AWGN with the standard deviation σ_z. The variables should be optimized between the need for higher SE and 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. Consequently, the following problem formulation comprehensively summarizes the practical design objectives of 1-bit ADC given as in [2-6]:

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

The common design objective for one-bit quantized transceiver can be summarized by the above formulation. However, the non-explicit representation of BER with respect to SNR and other parameters makes the problem a non-trivial optimization problem. A design offering an alternative option for optimization is presented in the following section and is examined for a particular fading channel.

III. End-to-end DL-based Solution for One-bit Quantization and Oversampling transceiver

A. One-bit Channel Autoencoder

A *channel AE* can learn to communicate over any channel, even for which no information-theoretically optimal scheme is known. Therefore, the channel AE for end-to-end communication from the seminal work in [8] can be adapted for the one-bit quantization and oversampling scheme, as shown in Fig. 2, by fitting the comprehensive system described in Fig. 1. By learning the optimal encoder-decoder pair, we can accurately recover signals after one-bit quantization for higher modulation transmission while jointly optimizing the different blocks that make up the conventional system design. The use of these neural network blocks as transmitter and receiver will make one-bit quantization operational at low complexity, serving as a framework for optimizing the transceiver performance.

The neural network encoder will project the input symbol into a higher dimensional codeword such that it will be robust against the channel fading, channel noise, and quantization noise as well as ISI. During the training phase, the encoder will learn an ideal modulation scheme and error correcting code that are suited to the channel environment with one-bit quantization, oversampling, and FTN. Importantly, it learns the symbol-to-sequence mapping which encodes the information into the distances between zero crossings and can be recovered after channel effects and one-bit quantization at the receiver. While the new transceiver cannot help with the explicit solution to the optimization problem in (5) as the BER as a function of Δ, M_{Tx} and M_{Rx} is not known in any explicit form, the joint optimization of the first three blocks that make up the transmitter will introduce better performance and furthermore, a new way to perceive the optimization problem. Compared to conventional schemes for one-bit quantization and oversampling which require a huge block-by-block optimization, however, the AE structure for the one-bit channel, which was first discussed in our previous work in [7], allows end-to-end optimization through back-propagation. Such a simple optimization process allows us the flexibility in determining the system parameters involved in our problem formulation (3). We expect that the encoder part of our proposed AE jointly optimizes channel coding, modulation, and sequence generation by end-to-end training. As a result, a parameter κ is introduced, which is referred to as the information-to-transmit-sequence mapping rate as defined in [7]. It represents the mapping of the input information bits to transmitted sequence after error correction, modulation, and generation of sequence suited for one-bit quantization. It corresponds to the rate of the proposed AE that is given as

$$\kappa = r\Delta\log_2 M \tag{6}$$

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

$$\begin{aligned}(\kappa^*, M_{Tx}^*) = &\underset{(\kappa,M_{Tx})}{\arg\max} \,(\kappa M_{Tx})\\ \text{s.t.} \quad & \kappa M_{Tx} \leq 2\log_2(M_{Rx}+1)\\ & M_{Tx} \leq M_{Tx}^{(th)}, M_{Rx} \leq M_{Rx}^{(th)},\\ & \gamma(\kappa, M_{Tx}, M_{Rx}, \sigma_z, \sigma_h) < \gamma^{(th)}\end{aligned} \tag{7}$$

The parameter κ in (6) implies that the AE's efficient signal packing capability in the presence of high ISI due to FTN with one-bit quantization and oversampling allows the learning of optimal parameters for the blocks that is represented by the encoder. As shown in [7], the channel AE allows clear inherent trade-off between transmitting SE and BER for AWGN channel by simplifying the general problem. But extending the

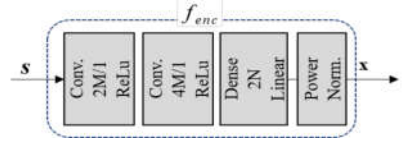
DL solution to real environment would be difficult since both received signal and pilots will be quantized before estimation, which results in the need for a very large number of pilots. Therefore, it is necessary to investigate the capability of DL-based transceiver to effectively communicate faded and quantized signals. The receiver in this scenario would have to estimate channel from quantized signals which creates the challenge of finding the correlation of the quantized signal to encountered channel. This arduous task is performed using supervised deep learning models in [5-6] where the schemes rely on the quantized pilots at receiver to estimate channel and then feedback is sent to transmitter for precoding the signal.

In this work, we tackle on transmission over fading channel with one-bit quantization at the receiver with two different levels of pilot-free transceiver designs. The first scheme attempts to perform this transmission without any explicit pilots by making use of only the proposed end-to-end training method and convolutional layers. The pilots for full channel estimation are only implicitly present and are learnt by the encoder and decoder through back propagation. The AE has convolutional layers to resolve the low correlation issue between quantized signal and channel. Here, matched filtering is fully taken care of by the neural network. Furthermore, oversampling on the receiver side is critical in increasing the dimension reduced by quantization. As transmission is pilotless, channel estimation and equalization and corresponding performance on fading channel is dependent of the 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 rates. Thus, utilizing the oversampling and FTN in the DL-based transceiver is shown to improve BER performance in the numerical results in Section IV.

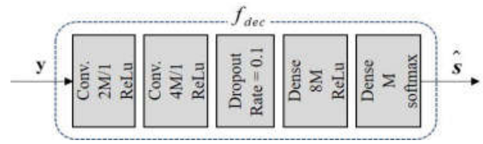
The second proposed scheme makes use of partial channel matching by using separate entity for phase synchronization but allowing the AE-based transceiver to recover the signal after the channel fading in amplitude, ISI and quantization. The target transmission channel in terahertz band is populated with large propagation losses and phase offsets. The prevailing cause of demodulation performance loss for this range of transmission is the phase offset, which cannot be adjusted once the signal has been quantized [9]. Consequently, phase offset needs to be corrected before sampling. In this second scheme, an explicit phase estimation and compensation is assumed to have been handled beforehand. Henceforth, the end-to-to-end training as well as the convolutional layers implemented in the second AE transceiver are aided by the partial matched filtering for the signal reconstruction task.

B. Training of One-bit Channel Autoencoder

For both of the proposed transceiver, while training to map the input vector to an embedding vector, the AE learns ideal modulation, channel coding, and symbol-to-sequence mapping. The straightforward end-to-end communication with AE incorporates the channel characteristics by the use of a non-trainable layer between the auto-encoder and auto-decoder.



(a) Encoder architecture



(b) Decoder architecture

Fig. 3: Network Architecture for proposed scheme (Convolutional layer parameters are indicated as "kernel size/filters")

This lambda layer introduces FTN, power normalization, fading channel, and oversampling as well single-bit quantization as seen in Fig. 2. The effects of FTN and oversampling on the transmitted vector as expressed in (2) is incorporated into an ISI filtering layer. For the first scheme, the non-orthogonal transmission can be represented by the effect on received vector as follows:

$$\mathbf{y} = \mathbf{HGx} + \mathbf{z}^{(\mathbf{g})} \tag{8}$$

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

$$\lambda(\mathbf{x}) = Q\left(\mathbf{HGx}+\mathbf{z}^{(\mathbf{g})}\right) \tag{9}$$

where $Q(\cdot)$ quantizes each elements of the received and oversampled vector. Here, channel matching is not performed as in the conventional system but rather rely on the robust encoding and decoding of the AE. 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 design goals of minimizing a transmission error. On the other hand, the second transceiver design with partial channel matching due to phase synchronization has the task of reconstructing a signal with the following overall channel distortion:

$$\lambda(\mathbf{x}) = Q\left(\widetilde{\mathbf{H}}\mathbf{Gx}+\mathbf{z}^{(\mathbf{g})}\right) \tag{10}$$

where $\widetilde{\mathbf{H}}$ is Toeplitz of the N-tap channel amplitude formed from vector $\mathbf{hh}^{\mathbf{H}}$ representing channel with matched phase. Here, $\mathbf{h}^{\mathbf{H}}$ is the Hermitian (conjugate) of channel vector. In the course of training, both AEs will learn the encoder and decoder parameters for the two hidden-layer functions, $f_{enc}(\cdot)$ and $f_{dec}(\cdot)$, such that the difference between the input and the predicted vectors is minimized by some loss function, denoted as $L(\mathbf{b}, \widehat{\mathbf{b}})$ where $\widehat{\mathbf{b}} = g(\lambda(f(\mathbf{b})))$. Given the parameters of θ_{enc} and θ_{dec} for the encoder and decoder, respectively, the hidden-layer function must be determined to minimize a transmission error probability as follows:

$$(f_{enc}^{*}, f_{dec}^{*}) = \underset{(f_{enc}, f_{dec})}{\arg\min}\ L\left(\mathbf{b}, \widehat{\mathbf{b}}\,|\theta_{enc}, \theta_{dec}\right) \tag{11}$$

In the proposed schemes, one-hot encoding is employed to represent input bits b_s as a one-hot 2^s-dimensional vector and there is SoftMax activation function at the decoder. The network design will be the same for the two proposed schemes which is a combination of convolutional and dense layers for both encoder and decoder as shown in Fig 3. Furthermore, batch normalization and Rectified Linear Unit (ReLu) activation are applied in the hidden layers of the encoder and decoder. Non-differentiable lambda layers are applied between encoder and decoder layers. Moreover, a loss function of categorical cross-entropy $L(\mathbf{b}, \widehat{\mathbf{b}}) = -\sum_{i=1}^{s} b_i \log \widehat{b}_i$, is used, where b_i is the i-th element of $\mathbf{b}$. The training of the AEs described above poses a challenge as one-bit quantization impedes gradient-based training due to resulting undefined derivatives at zero values and derivatives equal to zero at values different from zero. In our AE training, we make use a hyperparameter ε in the one-bit quantization layer of training phase such that the quantization function is adjusted to soft quantization function $\tilde{Q}(y_l) = y_l/(|y_l| + \varepsilon)$, where $\varepsilon \in [0, 1)$. This hyperparameter helps with avoiding a hard-decoding effect of squeezing all the weights and parameters into binary values. This means that while the neural network is being trained in the soft quantization environment, it will be tested in the hard one-bit quantization environment. The hyperparameter ε must be tuned for optimal performance. After the addition of the hyperparameter, the vanishing gradient is still a possibility since the sigmoid activation effect might occur at the soft-quantization layer. Thus, to combat these effects, some methods like using *ReLu* activation as well random weight initialization are applied in the hidden layers. Furthermore, the breadth and width of layers are balanced to have enough parameters for robust encoding and decoding while employing a smaller number of parameters to avoid the vanishing gradient. Training was performed using the Adam optimizer with a learning rate of 0.001. The AE-based transceiver was implemented using *Keras* with *TensorFlow* as backend.

C. Transfer Learning (TL) - AWGN to Rayleigh Channel

Transfer learning attempts to improve on traditional machine learning by transferring knowledge learned in one or more source tasks and using it to improve learning in a related target task. The effectiveness of any transfer method depends on the source task and how it is related to the target. If the relationship is strong and the transfer method can take advantage of it, the performance in the target task can significantly improve through transfer learning [10]. In the case of the proposed first scheme of AE transceiver, it has the task of reconstructing transmitted bits after two counts of distortion in the form of fading channel, one-bit quantization and ISI without the aid of any pilots. Therefore, the transmitter has to learn the right type of robustness for all counts. While the end-to-end learning should be enough for all properties to be learnt by the trained models, training eventually is imperfect due to lack of enough datasets and possible mismatches associated with hyperparameter tuning. Henceforth, it would be advantageous to use method such as TL to improve the learning ability of the AE by transferring a useful prior knowledge in another training to this particular scenario.

In this work, we propose training the architecture described above to minimize transmission error over a simple channel like AWGN and then freezing the first two encoder layers followed by fine-tuning of the AE to a fading channel such as Rayleigh fading. The AWGN-trained encoder will transmit a vector such that

$$\mathbf{x} = f_{enc}^{(AWGN)}(\mathbf{s}\,;\,\omega_0, \beta_0, \omega_1, \beta_1,\, \omega_2, \beta_2) \tag{12}$$

where (ω_i, β_i) are the weights and biases of the i-th layer of encoder. After the AE has been trained over Rayleigh fading channel, the weights and biases for the decoder as well as the last layer of the encoder will be fine-tuned to the new channel characteristics such that the transmitted vector can be described by

$$\mathbf{x}' = f_{enc}^{(Rayleigh)}\left(f_{enc}^{(AWGN)}(\mathbf{s}\,;\,\omega_0, \beta_0, \omega_1, \beta_1\,), {\omega'}_2, {\beta'}_2\right) \tag{13}$$

where ω_i', β_i' represent the encoder's fine-tuned weights and biases. This allows various parts of the transmitter to learn the specific properties for zero crossing and precoding sequences for one-bit quantization and fading, respectively. By employing the TL over the proposed AE scheme from a similarly constructed AE with a related task, we can initialize weights in our neural network for the best performance. The gain observed due to using TL is discussed in the next section.

IV. Numerical Results and Discussion

In our simulation, we consider an offline training of the AE with the proposed layout. Then our numerical results are obtained by Monte-Carlo simulation of the trained AE with test datasets. A BER performance of the proposed transmission scheme is evaluated for a fixed SE by varying different oversampling and FTN rates as well as κ values, under the constraint of one-bit ADC receivers over fading channel. In Fig 4, we can observe the performance of the pilotless AE-based transceiver over Rayleigh multipath channel with $L = 3$ channel taps at oversampling and FTN rates of $M_{Rx} = M_{Tx} = 4, 6,$ and 8. Their performance is lower-bounded by the BER of 16-QAM modulation with hamming code over similarly tapped Rayleigh Fading with unquantized reception and full channel-matched filter. At low SNR, the constellation gain and the error correcting capability are affected mainly by high noise. Consequently, in the low SNR region, the unquantized performance benchmark is outperformed by all the other AE transceivers, which are all utilizing oversampling. The oversampling adds a form of diversity gain in the low SNR region. However, one-bit quantization scheme is challenging for higher order modulation, especially over a fading channel. The AE transceiver with explicit phase synchronization has achieved the performance near to the lower bound of unquantized transmission with around 4dB gap at 10^{-4} BER requirement.

While the proposed pilotless solutions have performance much worse than the unquantized 16 QAM, the proposed method has resulted in a similar waterfall slope with nearly

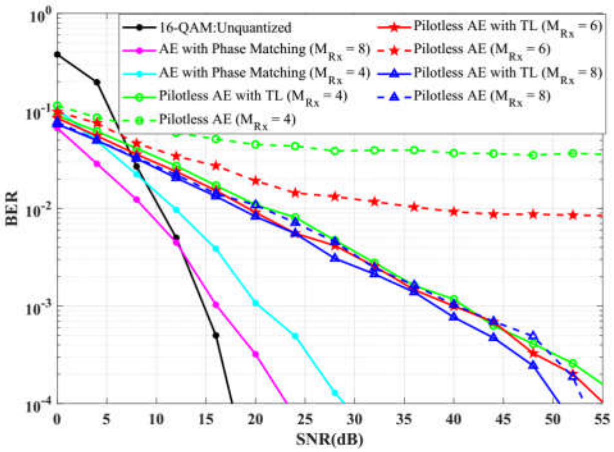


Fig. 4: BER performance with 16-QAM and Rayleigh fading channel

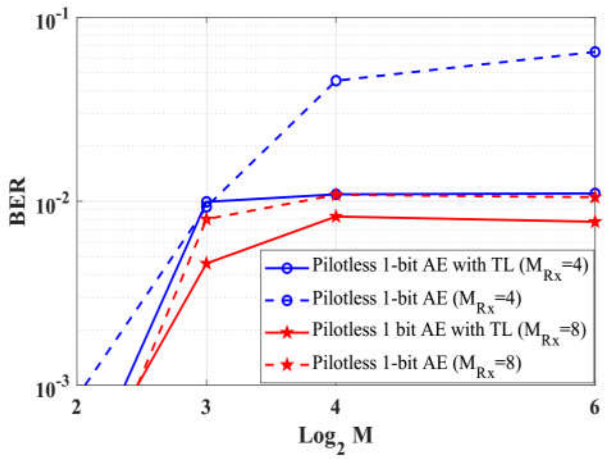


Fig. 5: BER gain due to transfer learning at SNR = 20dB

fixed SNR loss to the lower bound, which was not previously viable. At lower oversampling rate of $M_{Rx} = 4$ or 6, the AE-based transceiver is unable to reconstruct the faded and quantized signals. As the oversampling and FTN rate reaches to $M_{Tx} = M_{Rx} = 8$, the AE-based transceiver achieves the performance closest to the unquantized performance benchmark. Additionally, it can be observed that using TL over the proposed pilotless channel AE improves training and allows transmission with low oversampling rates to have BER performance close to that of higher oversampling rates. In fact, it is observed that all three cases for TL with $M_{Rx} = 4, 6$ and 8 achieve a similar performance, close to each other. The gain due to TL becomes minimal in higher oversampling rate due to the fact that the oversampling is able to capture the data distorted by one-bit quantization and the improved training offered by TL can no longer enhance the BER performance.

This constellation gain in BER in the higher SNR region is examined in the results in Fig 5. Here, at SNR = 20dB, the gain from TL is shown as the BER at different oversampling $M_{Rx} = 4$ is compared to that with $M_{Rx} = 8$ for different modulation levels. At 16-QAM and 64-QAM, we can see that TL offers a significant BER gain for the lower rate of $M_{Rx} = 4$, while showing litter performance gain at $M_{Rx} = 8$.

V. Conclusion

The current work aims at enhancing power efficiency in the terahertz communication systems by enabling high modulation transmission over fading channel with one-bit quantization and oversampling at the receiver and FTN at the transmitter. A deep learning (DL) solution is proposed to simplify and optimize the transceiver design. The implementation of the current solution has been shown to have made one-bit quantization and oversampling scheme operational enough to have an acceptable BER performance approaching to that of unquantized one in Rayleigh fading channel. Transfer learning is utilized to improve the BER performance at low oversampling rates. We can observe that while oversampling increases the input dimension for one-bit channel communication, it is the robust encoding offered by DL-based approaches that allow effective use of input dimension. As a future work, the current scheme can be extended to multiple antenna scenarios in order to explore new dimensions for further improvement.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2020R1A2C100998413).

References

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

[2] 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", *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020.

[3] R. Deng, J. Zhou, and W. Zhang, "Bandlimited Communication With One-Bit Quantization and Oversampling: Transceiver Design and Performance Evaluation," in *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 845-862, Feb. 2021, doi: 10.1109/TCOMM.2020.3037076.

[4] E. Balevi and J. G. Andrews, "High Rate Communication over One-Bit Quantized Channels via Deep Learning and LDPC Codes," *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2020, pp. 1-5, doi: 10.1109/SPAWC48557.2020.9154208.

[5] Y. Zhang, M. Alrabeiah, and A. Alkhateeb, "Deep Learning for Massive MIMO With 1-Bit ADCs: When More Antennas Need Fewer Pilots," in *IEEE Wireless Communications Letters*, vol. 9, no. 8, pp. 1273-1277, Aug. 2020, doi: 10.1109/LWC.2020.2987893.

[6] E. Balevi and J. G. Andrews, "One-Bit OFDM Receivers via Deep Learning," in *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4326-4336, June 2019, doi: 10.1109/TCOMM.2019.290381

[7] 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," *2022 IEEE 19th Annual Consumer Communications and Networking Conference (CCNC)*, 2022, pp. 515-516, doi: 10.1109/CCNC49033.2022.9700716.

[8] T. O'Shea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563-575, Dec. 2017, doi: 10.1109/TCCN.2017.2758370.

[9] D. He, Z. Wang, T. Q. S. Quek, S. Chen, and L. Hanzo, "Deep Learning-Assisted TeraHertz QPSK Detection Relying on Single-Bit Quantization," in *IEEE Transactions on Communications*, vol. 69, no. 12, pp. 8175-8187, Dec. 2021, doi: 10.1109/TCOMM.2021.3112216.

[10] E. Olivas, J. Guerrero, M. Sober, J. Benedito, and A. López, "Handbook Of Research On Machine Learning Applications and Trends: Algorithms, Methods and Techniques", IGI Publishing, Hershey, PA, 2009.