Technical Report: Improved Low-Resolution Quantized SIMO Detection via Deep Learning

Saichao Liu, Shengbo Chen, Cong Shen, and Roshan Kumar

***Abstract*—The signal detection problem in a single input multiple output (SIMO) system with low-resolution analog-to-digital converters (ADCs) is studied. The optimal minimum mean square error (MMSE) detector is difficult to compute due to the nonlinearity introduced by ADCs, while linear MMSE (LMMSE) suffers from performance degradation. Motivated by the advances in deep learning, a convolutional neural network detector (CNND) is proposed, leveraging the highly successful TextCNN. Numerical results suggest that the performance of CNND outperforms LMMSE in the majority of the cases, and has less runtime. We also show that the CNND is robust to the inaccurate channel coefficients.**

I. Introduction

Massive multiple-input multiple-output (MIMO) systems are a key component in 5G [1]. In practice, each receive RF chain is typically equipped with a low-resolution analog-to-digital converter (ADC), whose precision is limited due to cost and hardware considerations in massive MIMO [2], [3], [4]. This nonlinearity brings significant challenge to the receiver design, especially the detector [5]. The optimal minimum mean square error (MMSE) detector for massive MIMO with quantized output typically does not have close-form expressions, and even for some special cases when it does, it is complex and difficult to implement [6]. Linear MMSE (LMMSE), on the other hand, has explicit analytical expressions and is implementation-friendly, but its performance can be sub-optimal. Designing an efficient detector for massive MIMO systems with quantized output is an active research area.

In recent years, many researchers have attempted to apply machine learning to signal detection with low-resolution ADCs. The authors in [7] use a feedforward neural network based approach to equalize signals. Besides, they design a new activation function and add an unsupervised loss to improve the generalization ability of the neural network. In [8], the authors propose three machine learning based algorithms for blind MIMO detection with low-resolution ADCs.

In this work, we are inspired by the success of deep learning in applications such as computer vision [9], speech recognition [10], and natural language processing (NLP) [11]. We propose a convolutional neural network detector (CNND) for a single-input multiple-output (SIMO) system with low-resolution quantized output. A more refined CNN design, called text convolutional neural network (TextCNN), is adopted in our detector design. TextCNN is highly successful in NLP [12]. which has a similar two-dimensional data structure as our problem. In addition, the convolution operation can efficiently extract the correlation features, which matches our model (the same signal is received at different receive antennas). Two fully-connected layers follow the convolution layer, which can improve the learning ability. This model is trained offline to minimize the normalized mean-squared error (NMSE) and the trained model is used as an online detector. Simulation results show that the performance and runtime of CNND is better than LMMSE in majority of the cases. In addition, the performance loss of CNND under inaccurate channel coefficients is small compared to the case with perfect channel coefficients.

The rest of this paper is organized as follows. Section II describes system model and Section III presents the proposed CNND. Simulation results are provided in Section IV. Finally, Section V concludes the paper.

S. Liu and S. Chen are with School of Computer and Information Engineering, Henan University. C. Shen is with the Department of Electrical and Computer Engineering, University of Virginia. R. Kumar is with the Department of Electronic and Information Technology, Miami College of Henan University.

II. System model

The SIMO system is similar to [6] and [13], with a single transmit antenna and N receive antennas as shown in Fig. 1. We use $h_i \sim \mathcal{CN}(0,1)$ to denote the channel coefficient between the transmit antenna and the i-th receive antenna, $\forall i = 1, \cdots, N$, which is assumed to be fixed throughout one coherence time period and thus can be perfectly estimated at the receiver. Therefore, before quantizing the signal by ADCs, the received signal of the i-th antenna can be expressed as

$$r_i = h_i x + z_i, \tag{1}$$

where $x \triangleq x^R + jx^I \sim \mathcal{CN}(0, \varepsilon_s)$ is the transmitted signal with ε_s denoting the variance of x, and $z_i \triangleq z_i^R + jz_i^I$ is the independent Gaussian noise with distribution $\mathcal{CN}(0,1)$.

Every receive antenna is equipped with two M-bit ADCs which are applied separately to the real and imaginary components of the received signal. Thus the quantized outputs of the ADCs are

$$\begin{aligned} y_i^R &= \mathcal{Q}(r_i^R) = \mathcal{Q}(h_i^R x^R - h_i^I x^I + z_i^R), \\ y_i^I &= \mathcal{Q}(r_i^I) = \mathcal{Q}(h_i^R x^I + h_i^I x^R + z_i^I), \end{aligned} \tag{2}$$

where h_i^R and h_i^I are the real and imaginary parts of the i-th channel coefficient. $\mathcal{Q}$ is the M-bit symbolic quantizer which maps r_i^R and r_i^I to $y_i^R \triangleq [y_{i,1}^R, \cdots, y_{i,M}^R]$ and $y_i^I \triangleq [y_{i,1}^I, \cdots, y_{i,M}^I]$, where all elements in the vectors are binary symbols.

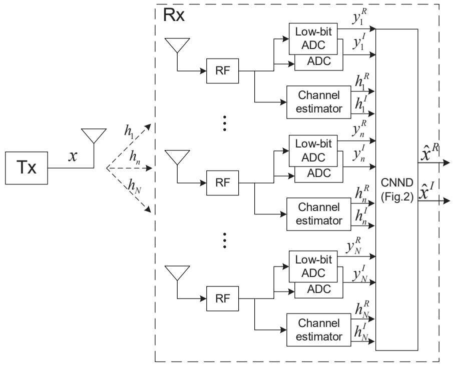

Fig. 1: System architecture

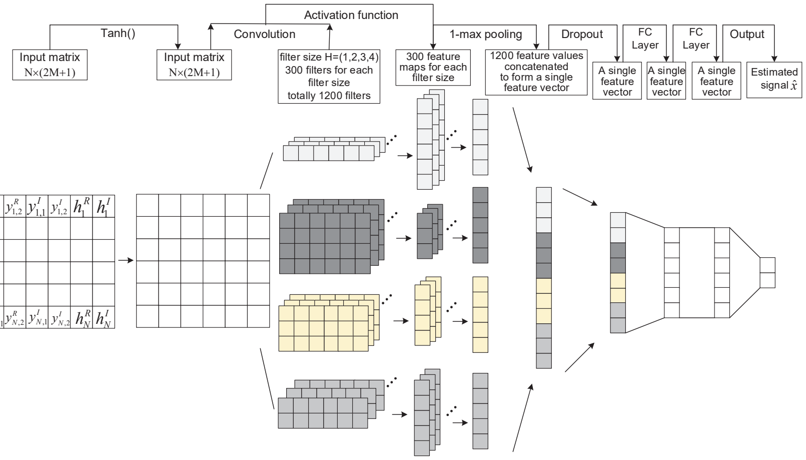

Fig. 2: The structure of CNND under 2-bit ADC

III. CNND DESIGN

In this section, we first present the network structure of the convolutional neural network detector (CNND), and the neural network training procedure. The conventional LMMSE detector is then described.

A. CNND architecture

The structure of the CNND is depicted in Fig. 2, which is inspired by the convolutional neural network for text [12], called TextCNN. The TextCNN is usually adopted for sentence classification tasks in NLP. A convolution operation can capture the local correlation between the word embedding vectors in context. In the system model of this paper, although the signals propagating through different channels experience mutually independent channel realizations, they are actually from the same transmitted signal. It means that they share some common features. Therefore, it is reasonable to carry out the signal detection with quantized output via TextCNN.

As the data analyzed by neural networks is always real-valued, we use both the real and imaginary parts of the quantized signal as input to the CNND. The quantization at the receiver, especially when M is small, leads to significant loss of information. To compensate for such loss and provide the CNND sufficient information to perform inference, the real and imaginary parts of channel coefficients $h_i, \forall i = 1, \cdots, N$, are also added as inputs to the network.

We stack the quantized output and channel coefficients vertically to create a matrix $\mathbf{X} \in \mathbb{R}^{N \times 2(M+1)}$. So we have

$$\mathbf{X} = [\mathbf{c}_1;\ \mathbf{c}_2; \cdots;\ \mathbf{c}_i; \cdots;\ \mathbf{c}_N]^\top,$$
$$\mathbf{c}_i = [y_{i,1}^R, \cdots, y_{i,M}^R, y_{i,1}^I, \cdots, y_{i,M}^I, h_i^R, h_i^I],$$

where $\mathbf{c}_i \in \mathbb{R}^{2(M+1)}$ is the vector of the real and imaginary parts of the quantized received signal and the channel coefficients at the i-th receive antenna. To satisfy the [-1, 1] range constraint of word embedding matrix of CNND, the input data $\mathbf{X}$ needs to be transformed with the tanh function.

In the convolutional layer, a filter $\mathbf{W} \in \mathbb{R}^{H \times 2(M+1)}$ is applied to perform convolution operation, where H is the height of the filter. It can be expressed as $a_i = g(\mathbf{W} \cdot tanh(c_{i:i+H-1}) + b)$, where $b \in \mathbb{R}$ is a bias term and $g(\cdot)$ is the activation function, rectified linear unit (ReLU). All the convolution results are concatenated to form a feature map $\boldsymbol{a} \triangleq [a_1, a_2, \cdots, a_{N-H+1}]$. It is noted that we may apply a bunch of filters with different H values. In the example shown in Fig. 2, there are 300 filters for each $H \in \{1, 2, 3, 4\}$. A max-over-time pooling operation is then applied on each feature map $\boldsymbol{a}$ to produce a feature value. All the feature values processed by pooling layer are concatenated to form a feature vector. After applying dropout for reducing overfitting, the vector is followed by two fully connected layers so as to improve the generalization ability of the CNND. Finally, the second fully connected layer connects output layer with two neurons to generate the detection output.

B. Neural network training

Both training and validation data are generated from the statistical SIMO channel model described in Section II. It is worth mentioning that the transmitted signals also serve as the labeled (ground-truth) data.

The back propagation algorithm is adopted to train the CNND, and NMSE is used to measure the performance of prediction, which is mathematically expressed as

$$\mathcal{J}_{\text{NMSE}} = \frac{\mathcal{J}_{\text{MSE}}}{\text{Var}(x)} = \frac{\frac{1}{T}\sum_{i=1}^{T} |x_i - \hat{x}_i|^2}{\frac{1}{T-1}\sum_{i=1}^{T} |x_i - \mu|^2}, \tag{3}$$

where

$$\mathcal{J}_{\text{MSE}} = \frac{1}{T}\sum_{i=1}^{T} |x_i - \hat{x}_i|^2 \tag{4}$$

$$\mu = \frac{1}{T}\sum_{i=1}^{T} x_i. \tag{5}$$

$\hat{x}_i$ is the detected signal of the i-th test sample, while μ and Var(x) denote the expectation and variance of transmitted signal, respectively. T is the total number of test samples. Note that Eqn. 4 is also the loss function for the training.

C. LMMSE detector

The LMMSE detector aims to minimize the mean square error (MSE) between the actual and detected signals, which can be expressed as

$$\hat{y} = \mathbf{R}_{yx}^{H}\,\mathbf{R}_{yy}^{-1} y. \tag{6}$$

where $\mathbf{R}_{yx} = \mathbb{E}[yx^*]$, $\mathbf{R}_{yy} = \mathbb{E}[yy^H]$. x^* denotes the conjugate of x and y^H denotes the conjugate transpose of y. For $M = 1$, the closed-form expressions of $\mathbf{R}_{yx}$ and $\mathbf{R}_{yy}$ are as follows [13]:

$$(\mathbf{R}_{yx})_n = h_n\,\varepsilon_s \sqrt{\frac{4}{\pi(|h_n|^2\,\varepsilon_s + 1)}}, \tag{7}$$

$$(\mathbf{R}_{yy})_{n,m} =$$

$$\begin{cases} 2, & n = m \\ \frac{4}{\pi}\left[\arcsin\left(\frac{(h_n\,h_m^*)^R\,\varepsilon_s}{\sqrt{|h_n|^2\varepsilon_s+1}\sqrt{|h_m|^2\varepsilon_s+1}}\right)\right. & \\ \left.+i\cdot\arcsin\left(\frac{(h_n\,h_m^*)^I\,\varepsilon_s}{\sqrt{|h_n|^2\varepsilon_s+1}\sqrt{|h_m|^2\varepsilon_s+1}}\right)\right], & n \neq m. \end{cases} \tag{8}$$

The case for general M is more involved but can be similarly handled.

IV. SIMULATION RESULTS

Majority of the computation tasks are carried out on two RTX 2080Ti GPUs and Xeon E5-2680 CPUs. The neural network is built using *Pytorch* with Python 3.6.7. The dataset setting is summarized in Table I.

TABLE I: Dataset setting

Type	Training Set	Validation Set	Test Set
Size	1,000,000	50,000	50,000

In the SIMO system simulation, the SNR range is set as [0, 2.5, 5, 7.5, 10, 12.5, 15, 17.5, 20] dB. The system complexity varies with the number of receive antennas, and we simulate N = {8, 32, 64, 128} receive antennas.

For simplicity, the same amount of filters are used in the construction of the CNND for different filter sizes. Also, in order to reduce the training complexity, the hyperparameter tuning is limited to the following: the size of filters H, the number of filters (denoted by N_{filter}), the number of neurons in the fully connected layer (denoted by N_{neuron}) and the number of epochs (denoted by N_{epoch}). The chosen values are listed in Table II.

TABLE II: Hyperparameters for CNND

N	8,32,64			
Name	H	N_{filter}	N_{neuron}	N_{epoch}
Size	{1,2,3,4}	300	500	20
N	128			
Name	H	N_{filter}	N_{neuron}	N_{epoch}
Size	{1,2,3,4}	400	500	20

A. Performance Evaluation

For different numbers of receive antennas, we compare the NMSE performance of the CNND and LMMSE detectors under different SNRs and quantization levels M. All simulation results demonstrate that both detectors get lower NMSE as M increases.

Fig. 3 reports the performance with 8 receive antennas. When $M = 1$ or $M = 2$, the performance of CNND is better than LMMSE. However, when $M = 3$, CNND becomes a little worse than LMMSE except for SNR = 20dB. The reason is that when the number of receive antennas is small, the input matrix to the CNND is also small, and the convolution layer cannot extract sufficient features.

Fig. 4 shows the performance with 32 antennas, where performance is improved compared to the 8-antenna case for both detectors. Also, we notice that the gap between CNND and LMMSE becomes larger compared to the previous case. When $M = 3$, the CNND outperforms LMMSE when SNR $\geq$ 7.5dB, which is due to the larger input matrix to CNND. Similar trend can be observed for the 64 and 128 antenna cases, as shown in Fig. 5 and 6, respectively.

One observation is that the NMSE performance may get worse for both detectors as the SNR increases for $M = 1$, which is also observed in the previous literature [14], [15]. This situation is due to the coarse quantization of 1-bit ADCs, which leads to magnitude information loss. To alleviate such bad influence, dithering is proposed to solve the problem [13], which is outside of the scope of this paper and will be pursued in our future work.

B. Robustness to Inaccurate Channel Coefficients

In the previous simulations, we assume perfect knowledge of channel coefficients h in the testing. However, in practice, the actual channel coefficients used in the test stage might be different from those used for training the network. For example, channel coefficients may be inaccurate due to estimation error. In this simulation, we simulate the system with 32 receive antennas when $M = 3$ and the CNND is trained with perfect channel coefficients. The model is tested at SNR = [10dB, 20dB] with different percentages of channel errors. In Fig. 7, for the percentage of channel error = [20%, 40%, 60%, 80%, 90%], we can see that as the error percentage becomes larger, the performance loss increases accordingly. However, the performance loss is relatively small compared to the gain brought about by the CNND. For instance, when the error percentage is 20%, the NMSE performance has 0.1dB loss at SNR = 10dB. Hence, we conclude that the performance of CNND is robust under inaccurate channel coefficients.

C. Complexity Analysis

In this subsection, we compare the computational complexity of the CNND and LMMSE by considering two metrics: the number of floating point operations (FLOPs) and the runtime. Under the 32 receive antennas, according to [16], the FLOPs for CNND and LMMSE per sample point are 2458600 and 85330 for $M = 1$ case, respectively. However, the runtime of CNND in this case is 0.011851s by using two

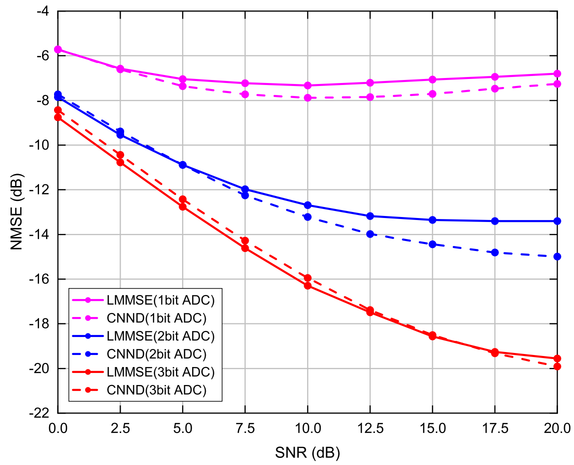

Fig. 3: Performance comparison of CNND with LMMSE with 8 receive antennas

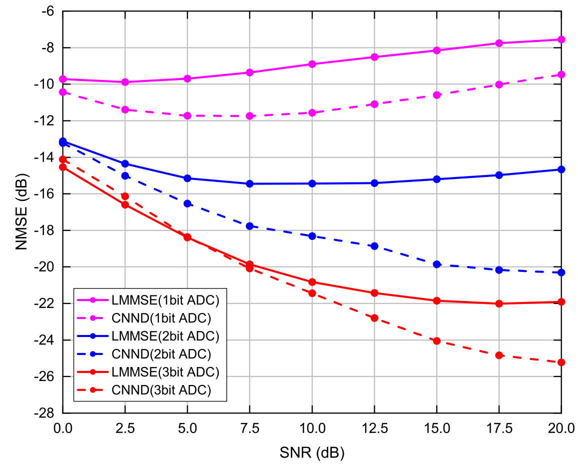

Fig. 4: Performance comparison of CNND with LMMSE with 32 receive antennas

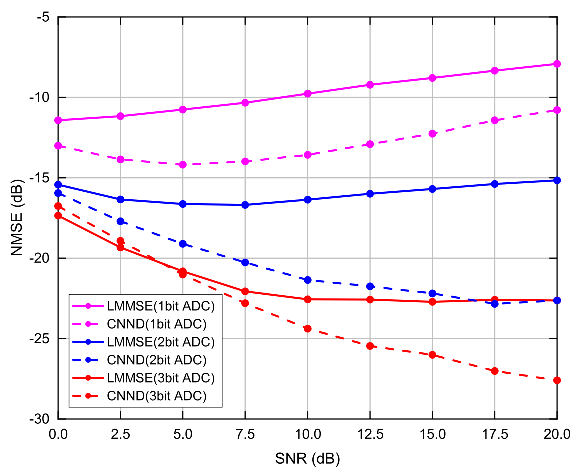

Fig. 5: Performance comparison of CNND with LMMSE with 64 receive antennas

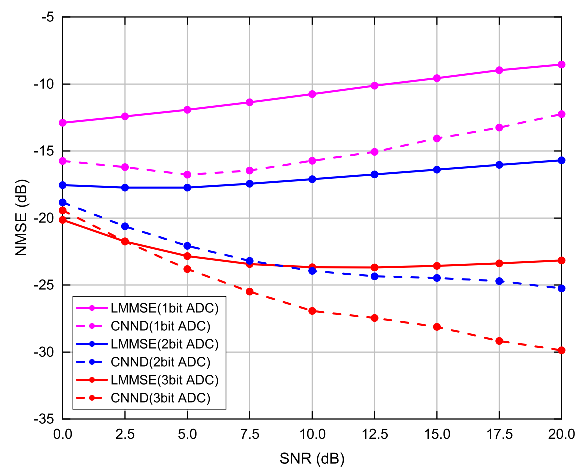

Fig. 6: Performance comparison of CNND with LMMSE with 128 receive antennas

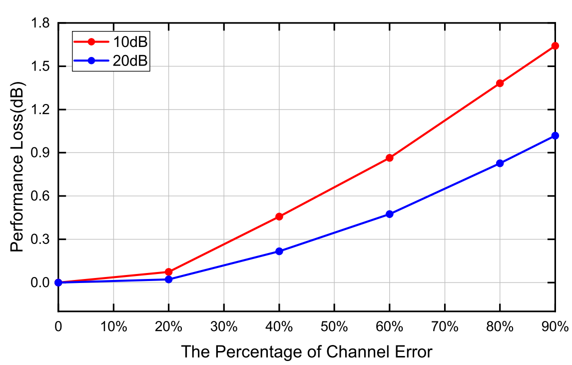

Fig. 7: Performance Loss of CNND (3-bit ADC) with 32 receive antennas when measured channel coefficients have different percentages of errors

RTX 2080ti GPUs, while LMMSE takes 0.188134s on two Xeon E5-2680 CPUs. The reason is that GPU allows lots of parallel computations, while CPU has limitations. Therefore, we conclude that CNND has more computational FLOPs but less runtime.

D. Impact of Scalar and Symbolic Quantizer to CNND

In this subsection, we illustrate the influence of scalar and symbolic quantizer to the performance of CNND. In Fig. 8 $\sim$ Fig. 11, the performance of CNND with scalar quantizer is worse than CNND with symbolic quantizer. Although the performance of CNND with scalar quantizer is better than LMMSE under 1-bit ADC in all cases and CNND with scalar quantizer outperforms LMMSE under 2-bit ADC when the number of receive antennas = [32, 64, 128], the performance all experiences degradation as SNR increases. This is because the received signal quantized by scalar quantizer has small number of elements, leading to the small size of input matrix. Then the convolutional layer can not extract enough features, which makes the learning ability of CNND limited. Instead, the symbolic quantizer converts the received signal into binary symbols. The convolutional layer can extract enough features and CNND achieves the excellent performance. Therefore, the scalar quantizer is not suitable for application of CNND.

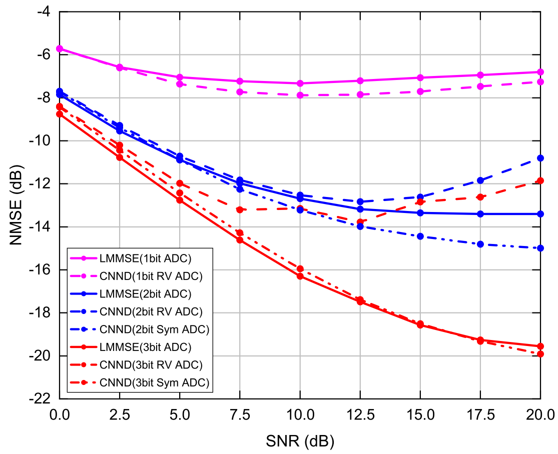

Fig. 8: Performance comparison of CNND with LMMSE with 8 receive antennas(include scalar quantization)

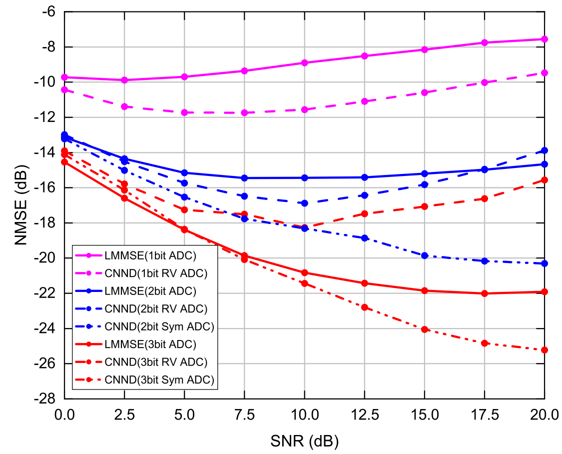

Fig. 9: Performance comparison of CNND with LMMSE with 32 receive antennas(include scalar quantization)

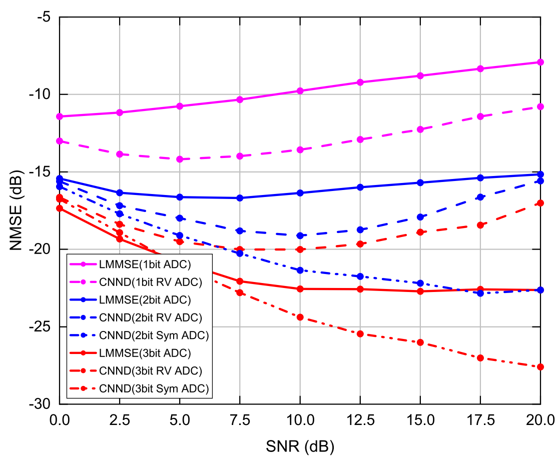

Fig. 10: Performance comparison of CNND with LMMSE with 64 receive antennas(include scalar quantization)

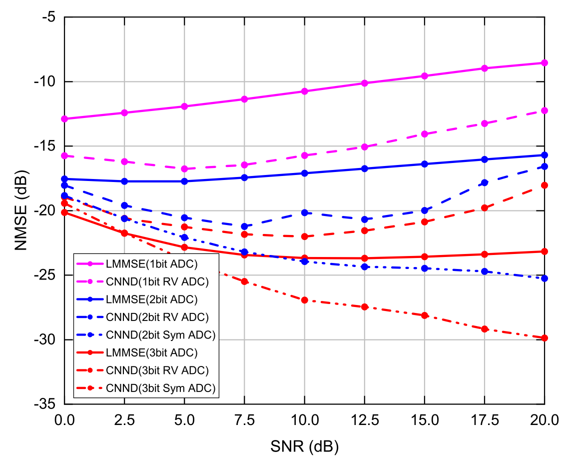

Fig. 11: Performance comparison of CNND with LMMSE with 128 receive antennas(include scalar quantization)

V. Conclusion

A neural network inspired detector has been proposed for a nonlinear SIMO system with low-resolution ADCs. The nonlinearity introduced by the quantization makes it difficult to implement the well-adopted MMSE detector, while the simple LMMSE suffers from performance degradation. A novel detector, CNND, leveraging the highly successful TextCNN outperforms LMMSE in most cases. Numerical simulation results show that the performance of CNND is better than that of LMMSE detector with 1-bit quantization output in all cases, and close or better than LMMSE for $M > 1$. We also show that the CNND is robust to the inaccurate channel coefficients and has less runtime. In the future work, we plan to explore how to generalize CNND to a full MIMO system.

References

[1] B. Panzner, W. Zirwas, S. Dierks, M. Lauridsen, P. Mogensen, K. Pajukoski, and D. Miao, "Deployment and implementation strategies for massive mimo in 5g," in *2014 IEEE Globecom Workshops (GC Wkshps)*, pp. 346–351, IEEE, 2014.

[2] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath, "Channel estimation in millimeter wave mimo systems with one-bit quantization," in *2014 48th Asilomar Conference on Signals, Systems and Computers*, pp. 957–961, IEEE, 2014.

[3] C. Risi, D. Persson, and E. G. Larsson, "Massive mimo with 1-bit adc," *arXiv preprint arXiv:1404.7736*, 2014.

[4] J. Zhang, L. Dai, S. Sun, and Z. Wang, "On the spectral efficiency of massive mimo systems with low-resolution adcs," *IEEE Communications Letters*, vol. 20, no. 5, pp. 842–845, 2016.

[5] J. Choi, J. Mo, and R. W. Heath, "Near maximum-likelihood detector and channel estimator for uplink multiuser massive mimo systems with one-bit adcs," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 2005–2018, 2016.

[6] T.-C. Zhang, C.-K. Wen, S. Jin, and T. Jiang, "Mixed-adc massive mimo detectors: Performance analysis and design optimization," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7738–7752, 2016.

[7] L. Chu, H. Li, and R. C. Qiu, "Lemo: Learn to equalize for mimo-ofdm systems with low-resolution adcs," *arXiv preprint arXiv:1905.06329*, 2019.

[8] Y.-S. Jeon, S.-N. Hong, and N. Lee, "Blind detection for mimo systems with low-resolution adcs using supervised learning," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2017.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[10] G. Hinton, L. Deng, D. Yu, G. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, B. Kingsbury, *et al.*, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal processing magazine*, vol. 29, 2012.

[11] I. Sutskever, O. Vinyals, and Q. Le, "Sequence to sequence learning with neural networks," *Advances in NIPS*, 2014.

[12] Y. Kim, "Convolutional neural networks for sentence classification," *arXiv preprint arXiv:1408.5882*, 2014.

[13] N. Liang and W. Zhang, "Mixed-adc massive mimo," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 983–997, 2016.

[14] S. Rao, A. L. Swindlehurst, and H. Pirzadeh, "Massive mimo channel estimation with 1-bit spatial sigma-delta adcs," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4484–4488, May 2019.

[15] I.-s. Kim and J. Choi, "Gradient pursuit-based channel estimation for mmwave massive mimo systems with one-bit adcs," in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–6, IEEE, 2019.

[16] P. Molchanov, S. Tyree, T. Karras, T. Aila, and J. Kautz, "Pruning convolutional neural networks for resource efficient inference," *arXiv preprint arXiv:1611.06440*, 2016.