International Journal of Electrical and Computer Engineering (IJECE)
Vol. 13, No. 4, August 2023, pp. 4169~4183
ISSN: 2088-8708, DOI: 10.11591/ijece.v13i4.pp4169-4183

Investigation of the performance of multi-input multi-output detectors based on deep learning in non-Gaussian environments

Mohammad Reza pourmir[1], Reza Monsefi[1], Ghosheh Abed Hodtani[2]
[1]Computer Department, Engineering Faculty, Ferdowsi University of Mashhad, Mashhad, Iran
[2]Electrical Engineering Department, Engineering Faculty Ferdowsi, University of Mashhad, Mashhad, Iran

Article Info

Article history:

Received Sep 25, 2022
Revised Dec 23, 2022
Accepted Jan 14, 2023

Keywords:

Deep learning
Information theory
Multi-output multi-input
Signal detection

ABSTRACT

The next generation of wireless cellular communication networks must be energy efficient, extremely reliable, and have low latency, leading to the necessity of using algorithms based on deep neural networks (DNN) which have better bit error rate (BER) or symbol error rate (SER) performance than traditional complex multi-antenna or multi-input multi-output (MIMO) detectors. This paper examines deep neural networks and deep iterative detectors such as OAMP-Net based on information theory criteria such as maximum correntropy criterion (MCC) for the implementation of MIMO detectors in non-Gaussian environments, and the results illustrate that the proposed method has better BER or SER performance.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Reza Monsefi
Computer Department, Engineering Faculty Ferdowsi, University of Mashhad
Mashhad, Iran
Email: monsefi@um.ac.ir

1. INTRODUCTION

The fifth generation (5G) mobile communication system has several advantages, especially with the use of multi-output multi-input (MIMO) receiver-transmitter, leading to lower error probability and or increased information throughput, however with orthogonal frequency division multiplexing (OFDM). MIMO systems face two challenges at the link layer: channel estimation and signal detection. The signal detection problem in the MIMO system involves retrieving the original messages sent by the sender on the receiver side when the received message is a noisy signal. The optimal technique for solving signal detection is called maximum likelihood estimation (MLE), but it is not scalable and therefore cannot be used for MIMO systems. In recent years, approximate message passing (AMP) techniques have yielded interesting results.

MIMO technology [1]–[3] is critical to modern wireless communication systems to assist the developing need for operational power [4]–[6]. In general, a maximum a posteriori detector (MAP) offers optimal detection performance but has exponential computational complexity that is impossible for large MIMO systems. Linear detectors [7], such as matched filter (MF), zero forcing (ZF) and linear minimum mean squared error (LMMSE), have low complexity but perform poorly compared to MAP detectors. On the other hand, iterative detection algorithms, for example, AMP [8], the sphere decoder (SD) [9], soft interference cancellation (SIC) [10], can achieve good results. They need complete knowledge of channel state information (CSI), which is prone to error if the system model does not conform to the actual transmission model or provides a flawed CSI and suffers from serious performance deterioration. Over the past decade, deep learning (DL) has revolutionized technology in many areas, such as computer vision and speech recognition.

Inspired by these successes, deep learning (DL) has recently been used in the design of communication systems, In [11]–[20] an overview is presented on various aspects of using deep learning in MIMO communications, for example, channel estimation and signal detection, resource management, power control, and end-to-end systems. Of all the DL-based applications, MIMO detection is one of the most fundamental. DL-based detectors can map the acquired signals into symbols sent from coaching records and operate better than common detection algorithms. In general, DL-based detectors can be divided into two categories: deep neural network (DNN) data-based DL detectors and model-based detectors based on a model of iterative detection algorithms.

Data-based DL detectors use deep neural network architectures to put in force signal detection [21]–[23]. These DNN architectures are model-independent and can retrieve sent symbols in a range of eventualities with excessive accuracy. However, such features can be realized at the cost of a giant variety of coaching parameters and examples. Model-based DL detectors are primarily based on a model of normal iterative detection algorithms, in which each layer of the network adds a single generation with some trainable variables [24]–[28]. The ensuing detectors perform better and converge faster than the original iterative detection algorithms [26]. However, contemporary model-based DL detectors are built on the assumption that the channel mannequin is linear, and CSI is available, which limits their use in complex environments.

Despite their great success, data-based DL detectors are considered a black box for receiving signals, and only experimental evaluation is available to demonstrate their performance. Understanding the internal mechanism of model-based MIMO detection and providing an overall design guide is essential. In fact, there is a lot of literature on the analysis of internal DNN mechanisms. Pioneering work in [29], [30] has shown that any continuous function in a compact set can be approximated with any accuracy by a DNN with a sigmoid activation function. Recently, studies [31], [32] proved that DNNs with rectified linear units (ReLU) can also be used as a large family of approximation functions. In addition, model-based channel estimation has been proven to converge to the least squares error estimator (MMSE) through growing the size of the coaching set at [33]. However, MIMO detection is a classification problem, and the model-based channel estimation evaluation in Hu *et al.* [33] cannot be generalized directly to model-based MIMO detection. In this paper, we analyze the performance of model-based MIMO detection. Our contributions are as follows: This paper describes signal detection and its problems by choosing non-Gaussian noise models, focusing on the signal detection algorithm in the MIMO system based on deep learning, and using criteria based on information theory. The purpose of this paper is to investigate the techniques of detection in the presence of additive non-Gaussian noise on different models of non-Gaussian noise and quadrature amplitude modulation (QAM) modulation that can be used to solve the problem of MIMO detection in the presence of non-Gaussian noise.

We show that model-based DL detectors perform better in non-Gaussian environments and based on the use of loss functions based on information theory criteria in different scenarios but require sufficient training examples to converge. Therefore, we seek to create a reliable diagnostic method with a small set of educational data. Due to the specificity of the noise model, the performance of the DL detector is largely determined by the basic iterative detection algorithm, which is usually better than the DL detectors based on the MSE criterion in non-Gaussian noise environments. The simulation results confirm our analytical results and show the effectiveness of model-based MIMO detection for non-Gaussian noises.

2. PRELIMINARIES

In the following, we briefly survey the MIMO system, QAM Modulation, signal detection in MIMO system and maximum correntropy criterion (MCC). What we are stating in this section is a brief description of the above definitions and non-Gaussian noise models, which are useful for expressing the problem of signal detection in a signal detection and the proposed algorithm.

2.1. Impulse noise model

Although noise is commonly modeled as Gaussian noise for classical wireless communication channels, in this paper, we use the Gaussian mixture model to model impulse noise. Impulse noise is very different from Gaussian noise because it is correlated and the samples are based on a distribution of the Gaussian mixture. As a non-Gaussian model, the stable distribution of alpha has attracted much attention due to its generality for modeling heavy tail and impact noise, which is widely observed in many communication channels. function density for statistical analysis no unfortunately, there is no probability density function (PDF) for the stable symmetric alpha (SαS) to approximate the PDF of the SαS distribution, we use the Gaussian mixed model, and in order to obtain a suitable approximation, we use the kernel density estimation (KDE) method based on the Gaussian kernel. in the form of an impulse noise and its approximation Figure 1.

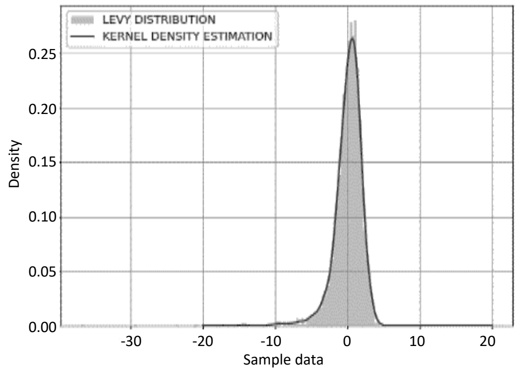


Figure 1. approximate impulse noise

2.2. MIMO system

As an advantage, MIMO technology improves link reliability and spectral performance. Because these advantages are necessary, we need efficient channel estimation methods and signal detection algorithms that balance performance and complexity. Demonstrated through a mathematical approach. Each transmitter antenna i sends a message s_i that follows different paths with different channel characteristics to reach the receiver antennas. In this article, we use the linear model of the MIMO system as (1).

$$y = Hx + n \tag{1}$$

where $H \in \mathcal{C}^{N_r \times N_t}$ and $y \in \mathcal{C}^{N_r}$ and $c \in \mathcal{C}^{N_r}$ and $x \in \mathcal{A}^{N_t}$ and $\mathcal{A} \in \mathcal{C}$ are a discrete alphabet. The y is the received vector and x is transmitted signal vector, whereas n represents the additive noise. values that x can be assumed are defined by an alphabet called constellation. The channel matrix H can consider different structures called channel models. this paper focuses on the independent and distributed Gaussian channel model (IID).

2.3. QAM modulation

One of the techniques used in communication systems to transmit signals is modulation. Modulation is an operation that is applied to a periodic waveform called a carrier signal to change its phase and/or amplitude and/or frequency to transmit information. The modulation technique used in this paper is QA), which changes the amplitude and phase of the carrier signal. This involves generating a signal in which two carriers with the same frequency are shifted in the 90° phase (they are square or orthogonal modulated and combined). On the receiver side, the signal can be split thanks to the orthogonal property. A base signal can only send 0 or 1 because it can show only two positions. Thanks to QAM, different points that are different in phase and amplitude can be increased. QAM points in a square grid with distance horizontal and vertical are equal, which is called a constellation diagram. Since digital communications use binary data, the number of points that make up a constellation is usually 2. The most common forms of QAM are QAM-4, QAM-16, QAM-64, QAM-256. In QAM-M, the points along each axis have values equal to:

$$\pm(\sqrt{M} - 1)d/2$$

where M is the power of two and d is the minimum distance between two different points in It is a constellation. In this paper, we evaluate the performance of different detector models in the presence of non-Gaussian noise and QAM modulation of the transmitted signal.

2.4. MIMO detection

Consider a MIMO system. The detection problem can be defined as retrieving the transmitted signal $x \in \mathcal{R}^{N_t}$ from a noise system known by the following relation, in which the channel matrix $H \in \mathcal{R}^{N_r \times N_t}$ and the noise vector $n \in \mathcal{R}^{N_r}$ are unknown.

$$y = Hx + n \in \mathcal{R}^{N_r} \tag{3}$$

This is known as "standard linear regression" or "the linear inverse problem" in signal processing literature. Multiple symbols can be identified separately or together. In the joint recognition to distinguish a symbol, the characteristics of other symbols must also be considered, while in a separate recognition, each symbol is recognized independently. Typically, shared diagnostics perform better than separate diagnostics despite their higher complexity.

The performance-optimized MIMO detector is the maximum probability (ML) detector, but its complexity increases exponentially with the number of transmitters. This includes a comprehensive search of all symbols for each user equipment (UE). Therefore, it is necessary to find algorithms with the best performance/complexity exchange, such as ZF and LMMSE, which have less computational complexity but worse performance. Promising detectors with excellent performance and reduced complexity detective detectors are based on AMP algorithms. The AMP-based detector approximates the Prior distribution using Taylor expansion and the central limit theorem on a dense operational graph and can achieve optimal Bayes performance in massive MIMO systems with channel matrices with independent elements. The same Gaussian distributions work well. In the following, the detection methods studied in this article are briefly discussed.

2.4.1. Iterative framework

One of the methods that can be followed to solve the MIMO detection problem is to repeat the transmitted signal. These algorithms are based on the number of T iterations. They include two stages of linear estimator and non-linear estimator as (4).

$$general\ iteration \begin{Bmatrix} z_t = \hat{x}_t + A_t\,(y - H\hat{x}_t\,) + b_t \\ \hat{x}_{t+1} = \eta_t(z_t) \end{Bmatrix} \tag{4}$$

The first step is as input $\hat{x}_t$ it calculates the current estimate of the transmitted signal x, the channel matrix H and the received signal y ، z_t which is a linear conversion. The second step is instead a nonlinear delimiter applied to z_t to generate a new estimate o $\hat{x}_{t+1}$ from x, which is used for the first step of the next iteration. The denoiser $\eta_t(\cdot)$ can be any non-linear function, but usually it applies the same thresholding function to each element, a common choice for the denoising function is the minimizer of $\mathbb{E}[\parallel \hat{x} - x \parallel_2 |z_t]$ that is given by $\eta_t(z_t) = \mathbb{E}[x|z_t]$. The purpose of each iteration is to improve the estimate $\hat{x}_t$ from x over the previous iterations Figure 2.

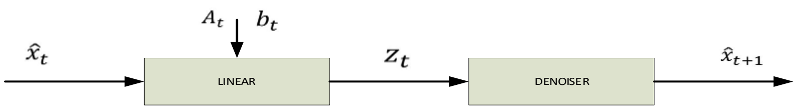


Figure 1. A block of the iterative detector that consists of a linear transformation and a denoising section in each block

2.4.2. Approximate message passing

a. AMP

Another way to almost solve the problem is to detect MIMO through belief propagation (BP). BP requires a few update messages, which is $\mathcal{O}(N_r N_t)$ for each iteration, which is not possible for large MIMO systems. To meet this limitation, Tan *et al.* [34] provided an AMP to solve the MIMO detection problem in the Gaussian scenario with less complexity. In fact, AMP uses $\mathcal{O}(N_r{+}N_t)$ messages for each iteration. The AMP algorithm performs the following steps:

$$AMP \begin{Bmatrix} z_t = \hat{x}_t + H^H\,(y - H\hat{x}_t\,) + b_t \\ b_t = \alpha_t\,(H^H(y - H\hat{x}_{t-1})\,) + b_{t-1} \\ \hat{x}_{t+1} = \eta_t(z_t\,;\sigma_t) \end{Bmatrix} \tag{5}$$

AMP is an iterative algorithm that uses $A_t = H^H$ and a b_t term that is called Onsanger correction term. Both σ_t and α_t can be computed using signal to noise ratio (SNR) and system parameters such as the dimension of the system or the constellation.

b. OAMP

A type of AMP that alleviates the Gaussian channel assumption. The orthogonal AMP (OAMP) works for single fixed channel matrices. OAMP is an optimal estimator in terms of MSE with excellent convergence properties. In [27], [35], the principle of OAMP is to divide the probability $p\left(x\middle|y;\widehat{H}\right)$ into a series of probabilities $p\left(x_i\middle|y;\widehat{H}\right)$ repeatedly. The OAMP detector can be written as the following algorithm:

$$OAMP\begin{Bmatrix} z_t = \hat{x}_t + \gamma_t\, H^H({v_t}^2\, HH^H + \sigma^2 I)^{-1}(y - H\hat{x}_t) \\ \hat{x}_{t+1} = \eta_t\,(z_t\,; {\sigma_t}^2) \end{Bmatrix} \tag{6}$$

where it γ_t normalizes coefficient and ${v_t}^2$ can be derived from system dimensions and SNR and it is proportional to the average noise power at the output of the denoiser at iteration t.

2.4.3. Deep learning

Deep learning is a subset of machine learning that tries to use a set of data consisting of pairs of attributes, tags $\{(y^{(d)}, x^{(d)})\}_{d=1}^{D}$ where D represents the number of pairs. Slowly learn some parameters: An artificial neural network (ANN) recognizes an unknown $\hat{x}$ tag associated with new data y recognizes the network and uses it in many processing layers, where each layer is converted by one It is composed of linear and non-linear components.

There are several benefits to using deep learning in MIMO diagnostics. First, deep learning can significantly increase the convergence rate compared to traditional iterative algorithms. Second, DL methods can reduce the average recovery error compared to duplicate because they do not require problem modeling but learn a mapping directly from input to output. The following DL methods are briefly reviewed in this paper.

a. DetNet

Recently, research on MIMO detection has been conducted on machine learning and deep learning approaches. Samuel *et al.* [25], [36] proposed detection network (DetNet), a deep learning network that performs well in the Gaussian scenario for small MIMO systems. it is a multi-layer neural network and its architecture follows the following steps:

$$DETNET\begin{Bmatrix} q_t = \hat{x}_{t-1} - {\theta_t}^{(1)}\, H^H\, y + {\theta_2}^{(2)}\, H^H\, H\hat{x}_{t-1} \\ u_t = [{\theta_t}^{(3)} q_t + {\theta_t}^{(4)} v_{t-1} + {\theta_t}^{(5)}]_+ \\ v_t = {\theta_t}^{(6)} u_t + {\theta_t}^{(7)} \\ \hat{x}_{t+1} = {\theta_t}^{(8)} u_t + {\theta_t}^{(9)} \end{Bmatrix} \tag{7}$$

where $[x]_+ = max(x, 0)$ is an elemental function called the ReLU activation function. DetNet performance can be promising, but architecture has two main issues [37]. First, it is difficult to adapt the network to spatially correlated channels or higher modulation schemes. Second, it does not use the known features of duplicate methods and therefore leads to unnecessary complexity. At the receiver of a wireless channel, a method known as successive interference cancellation (SIC) is utilized to decode two or more packets that are received simultaneously (in a normal system, packets arriving at the same time cause an interference). By first decoding the stronger signal, taking it out of the combined signal, and then decoding the difference into a lesser signal, the SIC is produced [38]. Therefore, with the combination of SIC and DetNet detectors, we can have better detection, and the results presented in this article show that good results have been obtained by using criteria based on information theory in the DetNet detector and combining it with SIC in non-Gaussian noises.

b. OAMP-Net

OAMP-Net, an orthogonal AMP-based deep learning network with good performance in Gaussian channel models, was his idea in [26]. uses the best Gaussian denoiser possible as AMP. The network has a strong assumption that the system is modelled with unitarily-invariant matrices because it is based on OAMP. For each iteration, algorithms only add two parameters to the initial OAMP. The following is a description of OAMP-Net:

$$OAMPNET\begin{Bmatrix} z_t = \hat{x}_t + {\theta_t}^{(1)}\, H^H({v_t}^2\, HH^H + \sigma^2 I)^{-1}(y - H\hat{x}_t) \\ \hat{x}_{t+1} = \eta_t\,(z_t\,; {\sigma_t}^2) \end{Bmatrix} \tag{8}$$

c. OAMP-Net2

OAMP-Net2, a model-driven deep learning network based on OAMP that is similar to OAMP-Net but has more trainable parameters to adapt to various channel environments and take channel estimate error into account, was proposed by Zhang *et al.* in [26]. Unfolding the OAMP detector and adding some trainable variables results in the OAMP-Net2. In terms of channel correlation, SNR, modulation symbol, and MIMO configuration mismatches, OAMP-Net2 performs noticeably better than OAMP and is more resilient and MIMO configuration mismatches.

d. MMNet

Over the years, several machine-based or deep-learning detectors have been proposed that have yielded promising results in Gaussian channel models, but with reduced performance in real-world channel models with spatial correlation. Khani *et al.* [35] proposed the MMNet detector, a MIMO detection scheme based on deep learning and the theory of iterative soft threshold algorithms. Thanks to a new training algorithm that uses time and spectral correlation to accelerate training, MMNet outperforms real-channel approaches with the same or less computational complexity. MMNet adds the right degree of freedom to the iterative framework and balances the flexibility and complexity of the model. In the Gaussian channel, MMNet achieves the same performance as optimal detectors, with twice the complexity of other deep inclination approaches. It is also better than a classic linear layout like the MMSE detector. The advantage of MMNet is that the algorithm is taught online, and through this, it can be adapted to different channel models. There are two versions of the MMNet neural network, one for Gaussian channel matrices and the other for custom channels. For the Gaussian channel, the network has the following architecture:

$$MMNET \begin{Bmatrix} z_t = \hat{x}_t + {\theta_t}^{(1)}\, H^H(y - H\hat{x}_t) \\ \hat{x}_{t+1} = \eta_t\,(z_t\,;{\sigma_t}^2) \end{Bmatrix} \tag{9}$$

2.5. MCC

The square root of the differences between the corresponding elements of the two vectors is the Euclidean distance, or L2 distance. Very helpful: becoming infected [39]–[47]. The best model is found using MCC criteria by maximizing the correlation between the model's output and the variable target:

$$M^* = arg \max_{M\in\mathcal{M}} \mathcal{V}_\sigma(T، Y) = \mathbb{E}[G_\sigma(e)] \tag{10}$$

where M^* is the optimal model, Y is the model output and T is the target variable, and $G_\sigma(e)$ a Gaussian kernel function that is presented as (11):

$$G_\sigma(e) = \frac{1}{\sigma\sqrt{2\pi}} \exp\,(-\frac{e^2}{2\sigma^2}) \tag{11}$$

σ the kernel function's bandwidth and the error between the target variable and the output model variable are both expressed as $e = T - Y$. In machine learning and signal processing, the Gaussian kernel is a powerful measure since it is a local function of the error variable. In this paper, we are interested in implementing MCC and using it in a learning model for signal detection in a MIMO system. Note that the MCC loss function is non-convex, so its analysis is fundamentally different from the least squares method. Assume the following is the sample set.

$$\mathcal{D} = \{(x_i\, ، y_i)\}_{i=1}^N \tag{12}$$

In this case, the experimental form $\mathbb{V}_\sigma(f)$ is as (13).

$$\mathbb{V}_\sigma(f) = \frac{1}{N}\sum_{i=1}^N G\left(\frac{(f(x_i)-y_i)^2}{2\sigma^2}\right) \tag{13}$$

In the field of statistical learning, the loss function due to the correntropy criterion is as (14).

$$\psi_\sigma(e) = \sigma^2\left(1 - \mathbb{G}_\sigma\,\left(\frac{e^2}{2\sigma^2}\right)\right) = \sigma^2\left(1 - exp\left(-\frac{e^2}{2\sigma^2}\right)\right) \tag{14}$$

where $\sigma > 0$ scale parameter. The loss function can be viewed as a variant of the Welsch function, and the estimator on the hypothesis space $\mathcal{H}$ can be expressed as (15):

$$\min_{f\in\mathcal{H}} \frac{1}{N}\sum_{i=1}^N \psi_\sigma(f(x_i) - y_i) \tag{15}$$

3. PROPOSED MODEL

In this paper, we use the proposed loss function (15) in a detector model based on deep neural networks. Our proposed model is based on an OAMP-Net and DetNet network as well as MMNet. With powerful learning capabilities, the data-driven DL detector can create a robust and accurate model for achieving comparable performance to traditional detectors and be used to detect MIMO. In this section, the performance of a data-driven DL detector is analyzed. In general, the main idea of machine learning based signal detectors is based on learning algorithms in which the output of the model $\hat{x}_{ML}$ is an estimate transmitted signal vector x with high accuracy and can be formulated as (16).

$$\hat{x}_{ML} = \mathbb{T}(f(x \,،\, \mathcal{P})) \tag{16}$$

This includes x information received and CSI, if any, and $\mathcal{P}$ a set of learnable parameters. The function $\mathbb{T}$ was mapped to, $f(\cdot)$, which is nonlinear. In contrast to classical models, which are quite complex, unfolding detectors have been proposed utilizing machine learning techniques. Deep neural networks are used in these detectors because they train using nonlinear activation functions. In DetNet and OAMP-Net models, as well as MMNet, MSE-based loss function is used in each layer, and as mentioned, in non-Gaussian noise, instead of an MSE-based loss function, we use loss theory-based information theory criteria. In this paper, in the above detection models, we use the loss function based on MCC criterion, so we examine the loss function in the following way (17) and find the following results in the presence of non-Gaussian noise:

$$loss\left(x \,،\, \hat{x}^{(l)}\right) = \frac{1}{N}\sum_{i=1}^{N} \psi_\sigma\left(x_i - \hat{x}_i{}^{(l)}\right) = \frac{1}{N}\sum_{i=1}^{N} \sigma^2 \,(1 - \exp\left(-\frac{\left(x_i - \hat{x}_i{}^{(l)}\right)^2}{2\sigma^2}\right)) \tag{17}$$

where $\hat{x}^{(l)}$ is the output of the network layer ($l^{th}\ layer$) and x is the signal sent and N is the number of training samples. In this paper, the number of samples is per BATCH. ψ_σ kernel Gaussian and in this article, based on different values of bandwidth, the results are reviewed, and the best bandwidth is selected. The relationships in a DetNet network based on the proposed loss function are summarized as (18):

$$\begin{aligned} q^{(l)} &= \hat{x}^{(l-1)} - \delta_1{}^{(l)} H^T y + \delta_2{}^{(l)} H^T H \hat{x}^{(l-1)} \\ z^{(l)} &= Relu(W_1{}^{(l)} q^{(l)} + b_1{}^{(l)}) \\ \hat{x}^{(l)} &= W_2{}^{(l)} z^{(l)} + b_2{}^{(l)} \end{aligned} \tag{18}$$

where, δ_1 and δ_2 are step sizes and $\hat{x}^{[l]}$ is lst layer output. And the network output is as (19):

$$\hat{x}_{ML}{}^{(L)} = sign(\hat{x}^{(L)}) \tag{19}$$

and the proposed loss function is expressed as (20):

$$loss = \frac{1}{N}\sum_{i=1}^{N} \sigma^2 \,(1 - \exp\left(-\frac{\left(x_i - \hat{x}_i{}^{(l)}\right)^2}{2\sigma^2}\right))\log\,(i) \tag{20}$$

where in x is transmitted signal vector and $\hat{x}^{(l)}$ lst layer output. We now consider a single-layer network and assume the channel matrix is fixed. We know the input-output relationship of the wireless channel is as (21).

$$y = Hx + n \tag{21}$$

By placing in the DetNet network relations and assuming that the initial value of the network output is zero:

$$\begin{aligned} q^{(1)} &= \hat{x}^{(0)} - \delta_1{}^{(1)} H^T y + \delta_2{}^{(1)} H^T H \hat{x}^{(0)} = 0 - \delta_1{}^{(1)} H^T y + 0 \\ &= -\delta_1{}^{(1)} H^T H x - \delta_1{}^{(1)} H^T n \\ z^{(1)} &= Relu\left(W_1{}^{(1)}\left(-\delta_1{}^{(1)} H^T H x - \delta_1{}^{(1)} H^T n\right) + b_1{}^{(1)}\right) \end{aligned} \tag{22}$$

Assuming $W_1{}^{(1)}\left(-\delta_1{}^{(1)} H^T H x - \delta_1{}^{(1)} H^T n\right) + b_1{}^{(1)}$ is greater than zero, we have (23).

$$z^{(1)} = W_1{}^{(1)}\left(-\delta_1{}^{(1)} H^T H x - \delta_1{}^{(1)} H^T n\right) + b_1{}^{(1)} \tag{23}$$

In this case, the network output is as (24).

$$\hat{x}^{(1)} = W_2{}^{(1)} z^{(1)} + b_2{}^{(1)} = W_2{}^{(1)}(\, W_1{}^{(1)}\left(-\delta_1{}^{(1)} H^T H x - \delta_1{}^{(1)} H^T n\right) + b_1{}^{(1)}) + b_2{}^{(1)} \tag{24}$$

In this case, the network error is as (25):

$$\begin{aligned} e &= x - \hat{x}^{(1)} = x - {W_2}^{(1)} z^{(1)} - {b_2}^{(1)} \\ &= x - {W_2}^{(1)}\left({W_1}^{(1)}\left(-{\delta_1}^{(1)} H^T H x - {\delta_1}^{(1)} H^T n\right) + {b_1}^{(1)} \right) - {b_2}^{(1)} \\ &= \left(I + {W_2}^{(1)} {W_1}^{(1)} {\delta_1}^{(1)} H^T H \right) x + {W_2}^{(1)} {W_1}^{(1)} {\delta_1}^{(1)} H^T n - {W_2}^{(1)} {W_1}^{(1)} {b_1}^{(1)} - {b_2}^{(1)} \\ &= {\theta_1}^{(1)} x + {\theta_2}^{(1)} n - {\theta_3}^{(1)} \end{aligned} \tag{25}$$

where in (26):

$$\begin{aligned} {\theta_1}^{(1)} &= I + {W_2}^{(1)} {W_1}^{(1)} {\delta_1}^{(1)} H^T H \\ {\theta_2}^{(1)} &= {W_2}^{(1)} {W_1}^{(1)} {\delta_1}^{(1)} H^T \\ {\theta_3}^{(1)} &= {W_2}^{(1)} {W_1}^{(1)} {b_1}^{(1)} + {b_2}^{(1)} \end{aligned} \tag{26}$$

we have a proposed loss function as (27).

$$E = \mathbb{E}[\, \sigma^2 (1 - exp\,(-\frac{{\|e\|_2}^2}{2\sigma^2}))] \tag{27}$$

According to the gradient descent (GD) criterion, to calculate the model parameters, we must calculate the derivative of the loss function based on each of the parameters, so we have the direction of the parameter ${\theta_2}^{(1)}$.

$$\frac{\partial E}{\partial {\theta_2}^{(1)}} = \frac{\partial E}{\partial e}\frac{\partial e}{\partial {\theta_2}^{(1)}} = \mathbb{E}[\frac{\partial}{\partial e}\left(\sigma^2 \left(1 - exp\left(-\frac{{\|e\|_2}^2}{2\sigma^2}\right)\right)\right] n = -\mathbb{E}[exp\left(-\frac{{\|e\|_2}^2}{2\sigma^2}\right) e] n \tag{28}$$

Now if we do the derivation according to the MSE criterion:

$$E = \mathbb{E}[\, {\|e\|_2}^2] \tag{29}$$

we have (30).

$$\frac{\partial E}{\partial {\theta_2}^{(1)}} = \frac{\partial E}{\partial e}\frac{\partial e}{\partial {\theta_2}^{(1)}} = \mathbb{E}[2e] n \tag{30}$$

Based on (28) and (30), we see that in (30), the presence of the exponential coefficient of error reduces the effect of non-Gaussian noise impact changes, and therefore the effect of impact noise is reduced based on the proposed loss function in the DetNet network model. In the same way, the summary of OAMP-Net network simplification relations can be expressed as (31).

$$\begin{aligned} z^{(l)} &= \hat{x}^{(l-1)} + \gamma^{(l)} w^{(l)} (y - H z^{(l-1)}) \\ \hat{x}^{(l)} &= \eta^{(l)}(z^{(l)}; \theta) \end{aligned} \tag{31}$$

where it γ_t normalizes coefficient and $w^{(l)}$ trainable matrix and denoising function is as follows:
η softmax function and the loss function is expressed as (32):

$$loss = \frac{1}{N}\sum_{i=1}^{N} \sigma^2 \,(1 - exp\left(-\frac{\left(x_i - \hat{x}^{(l)}\right)^2}{2\sigma^2}\right)) \tag{32}$$

and for MMNet network we have (33):

$$z^{(l)} = \hat{x}^{(l-1)} + {\theta_1}^{(l)} H^T (y - H\hat{x}^{(l-1)}) \tag{33}$$

$$\hat{x}^{(l)} = \eta(z^{(l)}; \sigma^2)$$

and denoising function is as (34):

$$loss = \frac{1}{N}\sum_{i=1}^{N} \sigma^2 \,(1 - exp\,(-\frac{\left(x_i - x^{(l)}\right)^2}{2\sigma^2})) \tag{34}$$

In the following, based on the non-Gaussian noise models and the Gaussian channel model, independent of the transmitted signal vector and assuming that the wireless channel is linear, we examine the detector models.

4. EXPERIMENTS

In this section, a computer simulation for testing model-based DL detectors in linear MIMO systems is presented. In addition, in the simulation, we assume CSI is known for the model-based DL detector to achieve comparable performance. In addition, the simulation results show that the iterative detection algorithm is the determining factor that affects the performance of the model-based DL detector.

4.1. Simulation setting

We know that SNR is expressed in terms of the ratio of signal power to noise power, we assume that the signal strength is given by (35).

$$\mathbb{E}[|H|^2]\mathbb{E}[|x|^2] \tag{35}$$

In which H channel matrix and x Sent signal are independent of each other. We also assume that the noise model based on (36) is in the form of a Gaussian mixture and is presented as (36).

$$n = \sum_{i=1}^{N} \lambda_i \, \mathcal{N}(\mu_i، \sigma_i{}^2) \tag{36}$$

Based on this, and using the calculations related to the variance of the Gaussian mixture of noise power, it can be shown as (37):

$$\mathbb{E}[n^2] = \sum_{i=1}^{N} \lambda_i \sigma_i{}^2 + \sum_{i=1}^{N} \lambda_i \, (\mu_i{}^2) - \mu^2 \tag{37}$$

where in (38).

$$\mu = \sum_{i=1}^{N} \lambda_i \, \mu_i \tag{38}$$

Therefore, the SNR can be displayed as (39).

$$SNR = \frac{\mathbb{E}[|H|^2]\mathbb{E}[|x|^2]}{\mathbb{E}[n^2]} \tag{39}$$

To calculate different values of SNR, the values λ_i are assumed to be constant and we also assume that the following condition is satisfied:

$$\lambda_1 > \lambda_2 > \cdots > \lambda_N \tag{40}$$

Assuming constant Gaussian distribution variance and modifying the transmitted signal strength, or assuming constant Gaussian distribution variance and varying the transmitted signal power, we compute various SNR values. In this article, we look at how the second approach is applied. Using KDE, we take into account an impulse noise model and produce the noise model based on a Gaussian mixture model KD).

The performance measures commonly used when working with a MIMO detection problem are BER and SER at different SNR values. Both criteria are divisible by the number of errors in the estimated message $\hat{x}$ compared to the original message sent x and the number of values sent. BER can be defined as:

$$BER = \frac{nnumber\ of\ bits\ in\ error}{total\ number\ of\ transmitted\ bits}$$

and SER can be expressed as:

$$SER = \frac{number\ of\ symbols\ in\ error}{total\ number\ of\ transmitted\ symbols}$$

Thus, BER works at the bit level, while SER works in constellation symbols. In the article, only the SER metric will be used.

4.2. Simulation results

We use the channel model with Gaussian distribution and independent of the transmitted signal and M-QAM modulation and check the results in both $M = 4$ and $M = 16$. We use traditional algorithms based on the ZF and AMP models to test against data-driven DL detectors. In Figure 3, we evaluate the SER performance of DetNet, OAMP-Net, MMENT detectors and traditional algorithms in the 8×8 Gaussian channel and in the presence of non-Gaussian noise with the following model.

$$levy\_stable(\alpha = 1.2 ، \beta = -0.7) \tag{41}$$

We assume that a complete CSI is available in the receiver. Figure 3 shows that the SER performance of the DetNet detector is significantly better than the other detectors under consideration. Interestingly, Figure 4 shows that the MMNet detector has the same response at different SNRs because it is implemented based on MMSE and in the presence of noise. Non-Gaussian does not converge to an appropriate response, even when using the proposed denoise function. Figure 5 shows a comparison between the two loss functions, MSE and correntropy in DetNet and OAMP-Net. As shown in Figure 5, the results show that at low SNRs, or in other words, at high noise powers, the correntropy loss function has given better results. Figure 5(a) shows the results of the comparison of the signal detection of MSE and correntropy loss functions in the DetNet model and Figure 5(b) shows the result of the comparison in the OAMP-Net model.

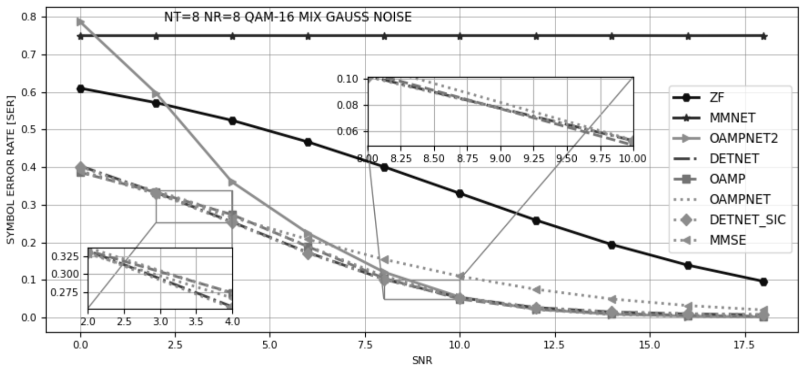


Figure 3. The SER performance of the model-driven DL detector versus SNR compared to other MIMO detectors in QAM_16 modulation

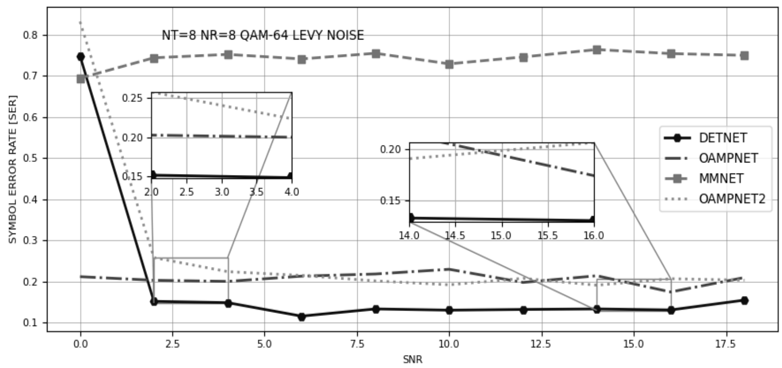
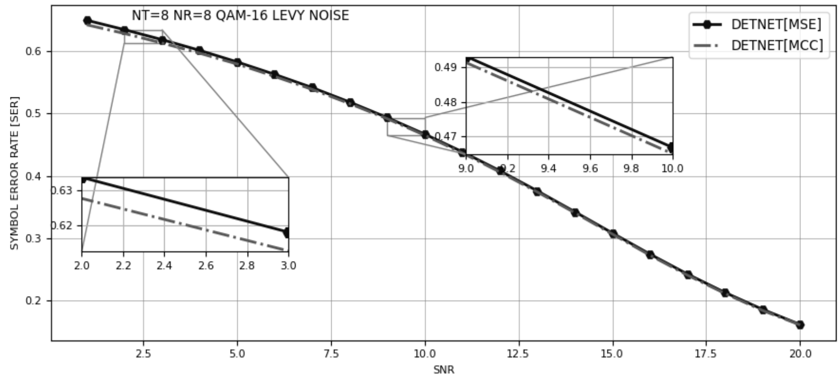
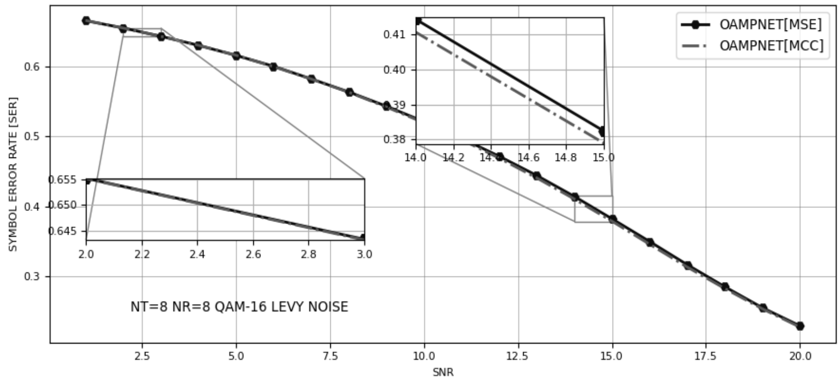


Figure 4. The network convergence of the model-driven DL detector versus SNR compared to model

We repeat the experiments based on QAM-16 modulation in OAMP-Net, the simulation results of which are shown in Figure 6. Next, we perform the simulation based on QAM-64 modulation, which is shown in the results in Figure 7. We also compare the results for different modulations in the DetNet network and using the MCC loss function Figure 8. Based on Figures 3 and 7, in impulse noise and based on MCC loss function, DetNet network offers the best performance, so we conclude that in impulse noise, it is better to use DetNet model and MCC loss function.



(a)



(b)

Figure 5. Compare MCC and MSE loss function for SER performance in (a) DetNet and (b) OAMP-Net

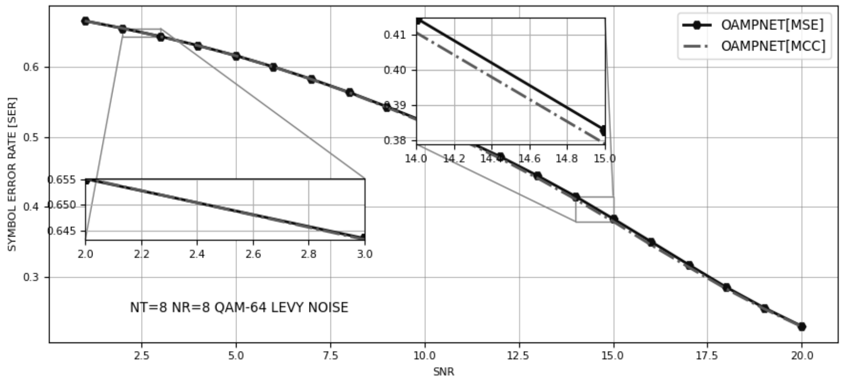


Figure 6. The SER performance of the OAMP-Net detector versus SNR in QAM_64 modulation

We continue the results based on the Middleton Class A noise model, which is a Gaussian mixture, as follows: equation (8) and examine the performance of the various detector models. We know that the Gaussian mixed model is as (42).

$$f(n) = (1-\epsilon)\eta(n) + \epsilon h(n)) \tag{42}$$

where ϵ is small positive constant, η is a standard Gaussian function and h is another density function with heavier tails. It is clearly visible that, f expressed by (42) is a valid density function if ϵ in the interval $[0,1]$. The behavior of f near the origin is dominated by η for suitably small values of ϵ and under the assumption that h is a bounded function, but for large values of $|n|$ h dominates the behavior of f because its tail decays more slowly than the tail of η. The noise model below assumes (43) that:

$$n = (1-\epsilon)\mathcal{N}(0,1) + \epsilon\mathcal{N}(1,10)$$
$$\epsilon = 0.02 \quad (43)$$

Based on the choice of SNR, we examine the results in non-Gaussian noise distributions and several signal detection models primarily based totally at the proposed loss function (14) and compare with the loss function based on MSE. Based on Figures 9 to 11, we see that the proposed loss function performs better in the OAMP-Net2 network, and in total, the proposed loss function has better results in all three detection models. The results of comparing the performance of the two loss functions in the presence of Gaussian mixed noise and different detector models are summarized in Table 1. Based on the results of Table 1, we see that it performed better in almost all SNR values of the proposed loss function.

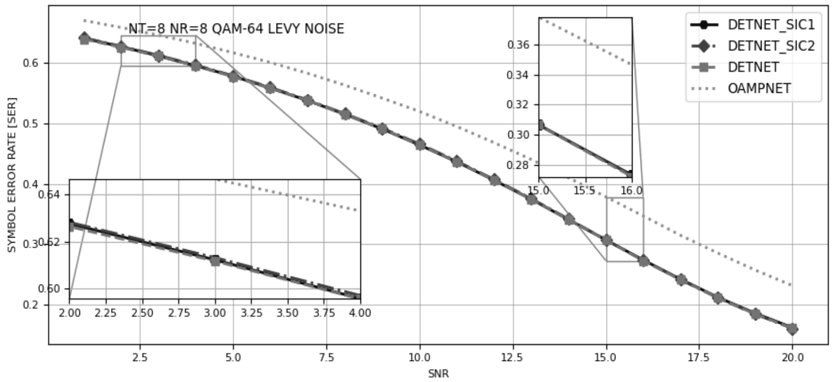


Figure 7. The SER performance of the DL detector versus SNR in QAM_64 modulation

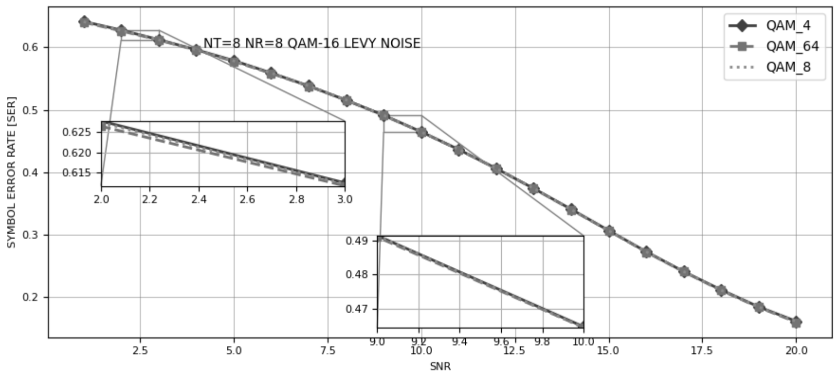


Figure 8. The SER performance of the DetNet detector versus SNR in QAM_64-4-8 modulation

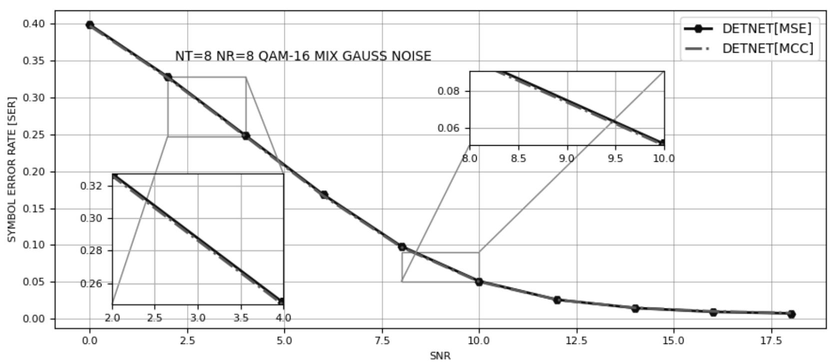


Figure 9. Compare MCC and MSE loss function for SER performance in DetNet and Gaussian mixed noise

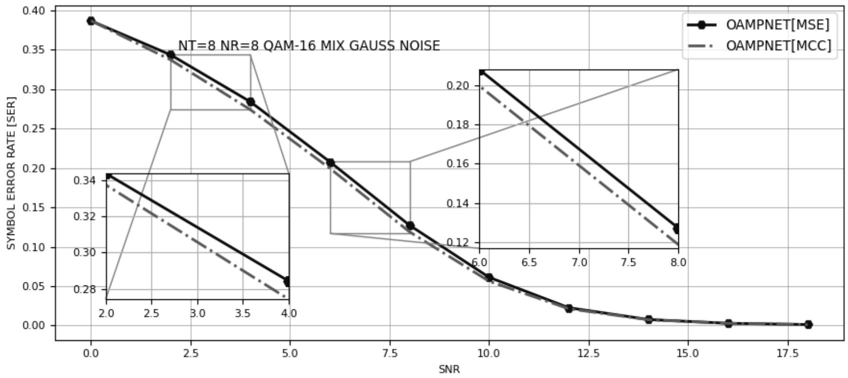


Figure 10. Compare MCC and MSE loss function for SER performance in OAMP-Net and mixed noise

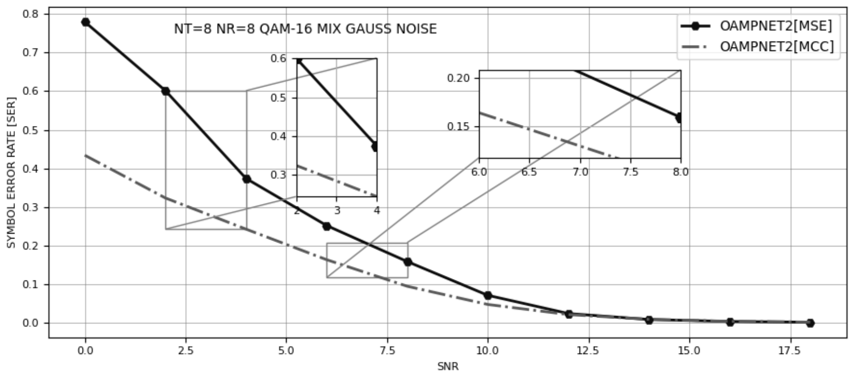


Figure 11. Compare MCC and MSE loss function for SER performance in OAMP-Net2 and mixed noise

Table 1. SER Performance for different models of detector in mixed noise

Detector model	LOSS function	SNR=0	SNR=2	SNR=4	SNR=6	SNR=8	SNR=10	SNR=12	SNR=14	SNR=16
DetNet	MSE	0.39783	0.32740	0.2480	0.16811	0.0983	0.0512	0.0261	0.0149	0.0098
DetNet	MCC	0.3962	0.32604	0.2467	0.16635	0.0971	0.0504	0.0260	0.0152	0.0104
OAMP-Net	MSE	0.3867	0.3436	0.2841	0.2078	0.1273	0.0612	0.0226	0.0076	0.00288
OAMP-Net	MCC	0.3869	0.3373	0.2741	0.1995	0.1189	0.0562	0.0213	0.0074	0.00284
OAMP-Net2	MSE	0.779	0.6014	0.3740	0.2520	0.1587	0.0711	0.0239	0.0086	0.0034
OAMP-Net2	MCC	0.434	0.3233	0.2426	0.1635	0.0946	0.0465	0.0213	0.0089	0.0038

5. CONCLUSION

In this article, different detectors based on deep learning when non-Gaussian additive noise is present are reviewed, and the results show that the use of cost functions based on information theory criteria has provided better performance. We showed that the use of the MCC criterion in detector models based on deep learning in non-Gaussian environments and at various signal-to-noise ratios provided better results. In the future, the proposed method can be developed for other channel models with other criteria for Gaussian and non-Gaussian noise.

REFERENCES

[1] J. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 5, pp. 871–878, Jun. 1987, doi: 10.1109/JSAC.1987.1146600.

[2] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Aug. 2002, doi: 10.1002/bltj.2015.

[3] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov. 1999, doi: 10.1002/ett.4460100604.

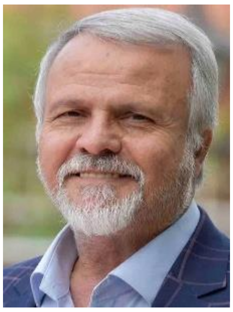
[4] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010, doi: 10.1109/TWC.2010.092810.091092.

[5] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014, doi: 10.1109/MCOM.2014.6736761.
[6] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, “Spatial- and frequency-wideband effects in millimeter-wave massive MIMO systems,” *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3393–3406, Jul. 2018, doi: 10.1109/TSP.2018.2831628.
[7] S. Yang and L. Hanzo, “Fifty years of MIMO detection: the road to large-scale MIMOs,” *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 1941–1988, 2015, doi: 10.1109/COMST.2015.2475242.
[8] C. Jeon, R. Ghods, A. Maleki, and C. Studer, “Optimality of large MIMO detection via approximate message passing,” in *2015 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2015, pp. 1227–1231. doi: 10.1109/ISIT.2015.7282651.
[9] Renqiu Wang and G. B. Giannakis, “Approaching MIMO channel capacity with soft detection based on hard sphere decoding,” *IEEE Transactions on Communications*, vol. 54, no. 4, pp. 587–590, Apr. 2006, doi: 10.1109/TCOMM.2006.873064.
[10] X. D. Wang and H. V. Poor, “Iterative (turbo) soft interference cancellation and decoding for coded CDMA,” *IEEE Transactions on Communications*, vol. 47, no. 7, pp. 1046–1061, Jul. 1999, doi: 10.1109/26.774855.
[11] T. Diamandis, “Survey on deep learning techniques for wireless communications,” *EE 359 Final Project Stanford university*, 2017.
[12] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, “Deep learning for wireless physical layer: opportunities and challenges,” *China Communications*, vol. 14, no. 11, pp. 92–111, Nov. 2017, doi: 10.1109/CC.2017.8233654.
[13] C. Zhang, P. Patras, and H. Haddadi, “Deep learning in mobile and wireless networking: A survey,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019, doi: 10.1109/COMST.2019.2904897.
[14] D. Gunduz, P. de Kerret, N. D. Sidiropoulos, D. Gesbert, C. R. Murthy, and M. van der Schaar, “Machine learning in the air,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2184–2199, Oct. 2019, doi: 10.1109/JSAC.2019.2933969.
[15] H. Huang *et al.*, “Deep learning for physical-layer 5G wireless techniques: Opportunities, challenges and solutions,” *IEEE Wireless Communications*, vol. 27, no. 1, pp. 214–222, Feb. 2020, doi: 10.1109/MWC.2019.1900027.
[16] Y. Hong, Y. Zhou, Q. Li, W. Xu, and X. Zheng, “A deep learning method for short-term residential load forecasting in smart grid,” *IEEE Access*, vol. 8, pp. 55785–55797, 2020, doi: 10.1109/ACCESS.2020.2981817.
[17] L. Dai, R. Jiao, F. Adachi, H. V. Poor, and L. Hanzo, “Deep learning for wireless communications: an emerging interdisciplinary paradigm,” *IEEE Wireless Communications*, vol. 27, no. 4, pp. 133–139, Aug. 2020, doi: 10.1109/MWC.001.1900491.
[18] Y. Zhang *et al.*, “DeepWiPHY: deep learning-based receiver design and dataset for IEEE 802.11ax systems,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 1596–1611, Mar. 2021, doi: 10.1109/TWC.2020.3034610.
[19] J. Liao, J. Zhao, F. Gao, and G. Y. Li, “A model-driven deep learning method for massive MIMO detection,” *IEEE Communications Letters*, vol. 24, no. 8, pp. 1724–1728, Aug. 2020, doi: 10.1109/LCOMM.2020.2989672.
[20] J. Jiao, X. Sun, L. Fang, and J. Lyu, “An overview of wireless communication technology using deep learning,” *China Communications*, vol. 18, no. 12, pp. 1–36, Dec. 2021, doi: 10.23919/JCC.2021.12.001.
[21] N. Farsad and A. Goldsmith, “Neural network detection of data sequences in communication systems,” *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5663–5678, Nov. 2018, doi: 10.1109/TSP.2018.2868322.
[22] N. Shlezinger, R. Fu, and Y. C. Eldar, “DeepSIC: deep soft interference cancellation for multiuser MIMO detection,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 1349–1362, Feb. 2021, doi: 10.1109/TWC.2020.3032663.
[23] J. Sun, Y. Zhang, J. Xue, and Z. Xu, “Learning to search for MIMO detection,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, pp. 7571–7584, Nov. 2020, doi: 10.1109/TWC.2020.3012785.
[24] H. He, S. Jin, C.-K. Wen, F. Gao, G. Y. Li, and Z. Xu, “Model-driven deep learning for physical layer communications,” *IEEE Wireless Communications*, vol. 26, no. 5, pp. 77–83, Oct. 2019, doi: 10.1109/MWC.2019.1800447.
[25] N. Samuel, T. Diskin, and A. Wiesel, “Learning to detect,” *IEEE Transactions on Signal Processing*, vol. 67, no. 10, pp. 2554–2564, May 2019, doi: 10.1109/TSP.2019.2899805.
[26] Z. Zhang, Y. Li, X. Yan, and Z. Ouyang, “A low-complexity AMP detection algorithm with deep neural network for massive MIMO systems,” *Digital Communications and Networks*, Nov. 2022, doi: 10.1016/j.dcan.2022.11.011.
[27] S. Zhang, M. Zhang, S. Jin, and C.-K. Wen, “Low-complexity detection for MIMO C-FBMC using orthogonal approximate message passing,” *IEEE Signal Processing Letters*, vol. 26, no. 1, pp. 34–38, Jan. 2019, doi: 10.1109/LSP.2018.2879545.
[28] J. Zhang, C.-K. Wen, and S. Jin, “Adaptive MIMO detector based on hypernetwork: design, simulation, and experimental test,” *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 1, pp. 65–81, Jan. 2022, doi: 10.1109/JSAC.2021.3126064.
[29] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Mathematics of Control, Signals, and Systems*, vol. 2, no. 4, pp. 303–314, Dec. 1989, doi: 10.1007/BF02551274.
[30] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, no. 5, pp. 359–366, Jan. 1989, doi: 10.1016/0893-6080(89)90020-8.
[31] G. Montúfar, R. Pascanu, K. Cho, and Y. Bengio, “On the number of linear regions of deep neural networks,” in *NIPS’14: Proceedings of the 27th International Conference on Neural Information Processing Systems*, 2014, vol. 2, pp. 2924–2932.
[32] R. Arora, A. Basu, P. Mianjy, and A. Mukherjee, “Understanding deep neural networks with rectified linear units,” *Prepr. arXiv.1611.01491*, Nov. 2016.
[33] Q. Hu, F. Gao, H. Zhang, S. Jin, and G. Y. Li, “Deep learning for channel estimation: interpretation, performance, and comparison,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2398–2412, Apr. 2021, doi: 10.1109/TWC.2020.3042074.
[34] X. Tan, W. Xu, Y. Be’ery, Z. Zhang, X. You, and C. Zhang, “Improving massive MIMO belief propagation detector with deep neural network,” *Prepr. arXiv1804.01002*, Apr. 2018.
[35] M. Khani, M. Alizadeh, J. Hoydis, and P. Fleming, “Adaptive neural signal detection for massive MIMO,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5635–5648, Aug. 2020, doi: 10.1109/TWC.2020.2996144.
[36] N. Samuel, T. Diskin, and A. Wiesel, “Deep MIMO detection,” in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2017, pp. 1–5. doi: 10.1109/SPAWC.2017.8227772.
[37] G. Gao, C. Dong, and K. Niu, “Sparsely connected neural network for massive MIMO detection,” in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, Dec. 2018, pp. 397–402. doi: 10.1109/CompComm.2018.8780959.
[38] J. Choi, “MIMO detection with successive interference cancellation,” in *Optimal Combining and Detection*, Cambridge University Press, 2010, pp. 217–245. doi: 10.1017/CBO9781139193535.009.
[39] R. He, B-G. Hu, W-S. Zheng, and X-W. Kong, “Robust principal component analysis based on maximum correntropy criterion,” *IEEE Transactions on Image Processing*, vol. 20, no. 6, pp. 1485–1494, Jun. 2011, doi: 10.1109/TIP.2010.2103949.
[40] A. Singh, R. Pokharel, and J. Principe, “The C-loss function for pattern classification,” *Pattern Recognition*, vol. 47, no. 1, pp. 441–453, Jan. 2014, doi: 10.1016/j.patcog.2013.07.017.
[41] E. Hasanbelliu, L. S. Giraldo, and J. C. Principe, “Information theoretic shape matching,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 12, pp. 2436–2451, Dec. 2014, doi: 10.1109/TPAMI.2014.2324585.
[42] W. Ma, H. Qu, G. Gui, L. Xu, J. Zhao, and B. Chen, “Maximum correntropy criterion based sparse adaptive filtering algorithms for

robust channel estimation under non-Gaussian environments," *Journal of the Franklin Institute*, vol. 352, no. 7, pp. 2708–2727, Jul. 2015, doi: 10.1016/j.jfranklin.2015.03.039.
[43] Y. Feng, X. Huang, L. Shi, Y. Yang, and J. A. K. Suykens, "Learning with the maximum correntropy criterion induced losses for regression," *Journal of Machine Learning Research*, vol. 16, no. 30, p. 993−1034, 2015.
[44] C. Liu, G. Wang, X. Guan, and C. Huang, "Robust M-estimation-based maximum correntropy Kalman filter," *ISA Transactions*, Oct. 2022, doi: 10.1016/j.isatra.2022.10.025.
[45] B. Chen, X. Wang, N. Lu, S. Wang, J. Cao, and J. Qin, "Mixture correntropy for robust learning," *Pattern Recognition*, vol. 79, pp. 318–327, Jul. 2018, doi: 10.1016/j.patcog.2018.02.010.
[46] H. Zhao, D. Liu, and S. Lv, "Robust maximum correntropy criterion subband adaptive filter algorithm for impulsive noise and noisy input," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 2, pp. 604–608, Feb. 2022, doi: 10.1109/TCSII.2021.3095182.
[47] J. Zhang, G. Huang, and L. Zhang, "Generalized correntropy criterion-based performance assessment for non-gaussian stochastic systems," *Entropy*, vol. 23, no. 6, Jun. 2021, doi: 10.3390/e23060764.

BIOGRAPHIES OF AUTHORS

Mohammad Reza Pourmir received the B.Sc. degree in electronics engineering from Sistan and Baluchistan University, Zahedan, Iran, in 1994, and the M.Sc. degree in computer engineering from Iran University of Science and Technology, in 2012, and he is now a Ph.D. student in computer engineering at Mashhad Ferdowsi University. His research interests are in information theory, communication theory, wireless communications and signal processing. He can be contacted at email: mo.pourmir@mail.um.ac.ir



Reza Monsefi received the B.Sc. degree in electronics engineering from Manchester University, Manchester, U.K, in 1979 and received the M.Sc. and Ph.D. degree from Salford University, U.K, in 1981 and 1983; and he is now a professor in computer engineering. His research interests are in information theory and signal processing. He can be contacted by email: monsefi@um.ac.ir.

Ghosheh Abed Hodtani received the B.Sc. degree in electronics engineering and the M.Sc. degree in communications engineering, both from Isfahan University of Technology, Isfahan, Iran, in 1985, 1987, respectively. He joined Electrical Engineering Dept., at Ferdowsi University of Mashhad, Mashhad, Iran, in 1987. He decided to pursue his studies in 2005 and received the Ph.D. degree (with excellent grade) from Sharif University of Technology, Tehran, Iran, in 2008; and he is now a professor in electrical engineering. His research interests are in multi-user information theory, communication theory, wireless communications and signal processing. Prof. Hodtani is the author of a textbook on electrical circuits and is the winner of the best paper award at IEEE ICT -2010. He can be contacted at email: hodtani@um.ac.ir.