Machine Learning Techniques for PAPR Reduction in Multicarrier Communication Systems

Supervised By

Dr. Michael Ibrahim
Electronics & Communication Eng. Dept.
Faculty Of Engineering
Ain Shams University

Presented By

Ahmed Fawzy
Rana Muhammad
Raneem Gaber
Amr Saeed
Metwaly Yahia
Marwan Yasser
Mustafa Ahmed
Youssef Salah

Project Overview

- OFDM technique has been widely deployed in high-speed data transmission systems, primarily for its good resistance to multipath fading.
- High PAPR is the major drawback in implementing an OFDM system.
- Various classical PAPR reduction methods have been proposed. Also, machine learning (ML), specifically deep learning (DL), has made remarkable achievements in terms of PAPR reduction.
- In this project, we aim to implement various PAPR reduction techniques proposed, starting from classical approaches and ending with ML approaches.
- We mainly used Python in simulating our models.

Outlines

1. Introduction to OFDM system.
2. PAPR in OFDM system.
3. Classical PAPR reduction techniques.
4. Introduction to Machine Learning.
5. ML PAPR reduction techniques.

Introduction to OFDM system

Single Carrier Transmission

- Single carrier transmission isn't suitable for High Data Rates, as it’s hard for SCM systems to deal with dynamic channel and inter symbol interference, as it requires complex equalizers [1].
- To overcome the frequency selectivity of the wideband channel experienced by single-carrier transmission, multiple carriers can be used for high data rate transmission.

Multi-Carrier Transmission

- In multi-carrier transmission, wide band signal of a single carrier, which is higher than the coherence bandwidth, thus suffer from frequency selective channel, can be divided into N independent flat sub-channels as in fig.1, so every flat sub channel requires a simple equalizer instead of a complex one, and the signal will not suffer from frequency selective channel as sub-band is smaller than the coherence bandwidth.

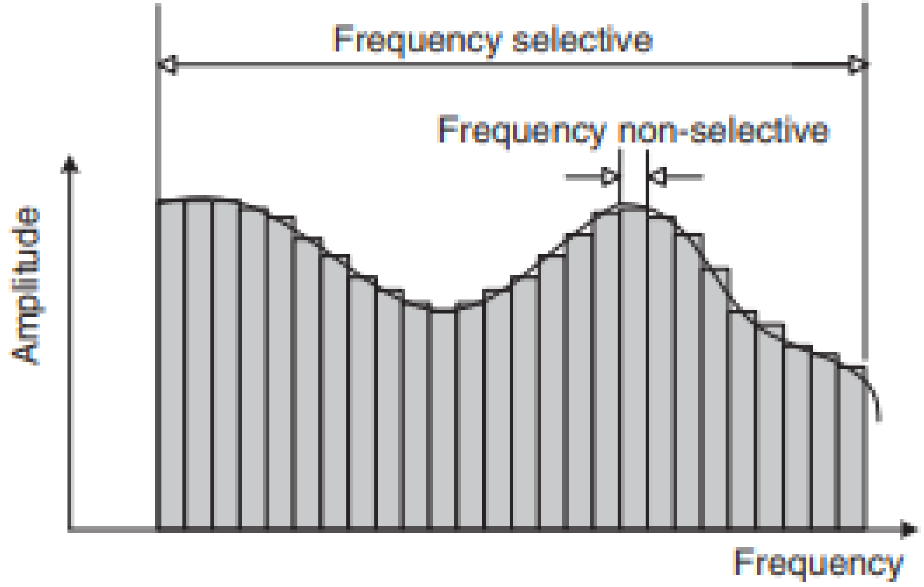


Fig.1 The frequency response of multichannel transmission system [1]

- The different symbols are transmitted over an equally separated subcarriers in parallel form as seen in fig.2, but there is a main drawback in this structure that we need multiple local oscillators to generate subcarriers.

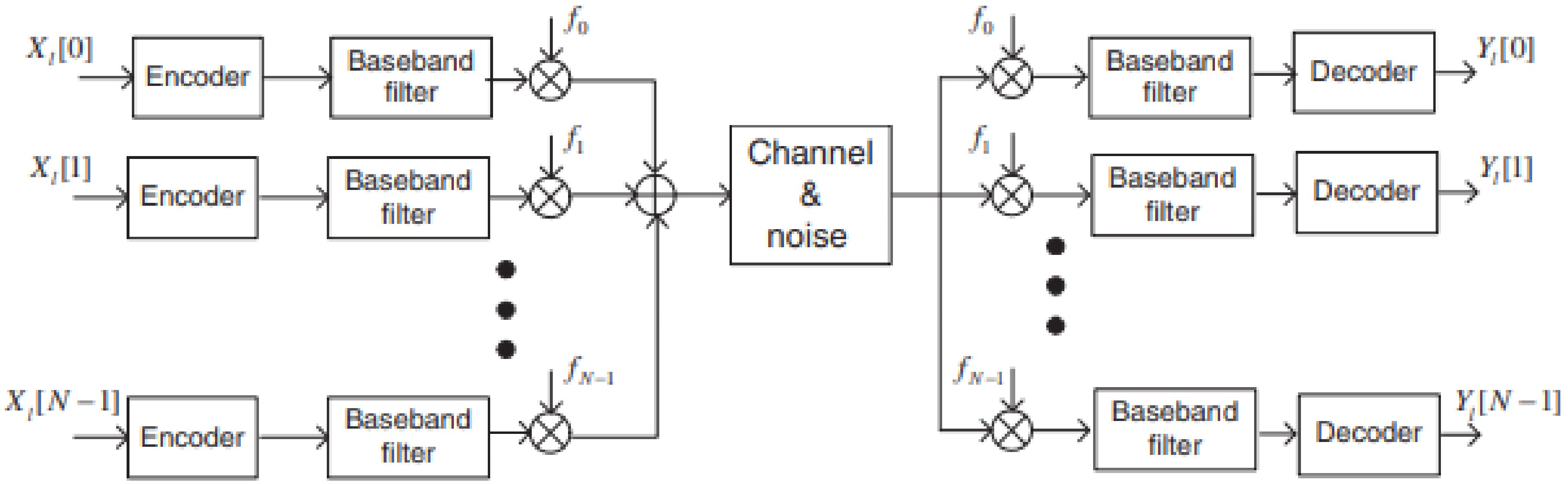


Fig.2 Basic structure of multicarrier system [1]

What is an OFDM System?

- Orthogonal frequency division multiplexing (OFDM) is a multicarrier transmission system that is suitable for high data rate transmission.

- The spectra of subcarriers are overlapped for better bandwidth efficiency, where the wideband is fully divided into N **orthogonal narrowband subchannels**.

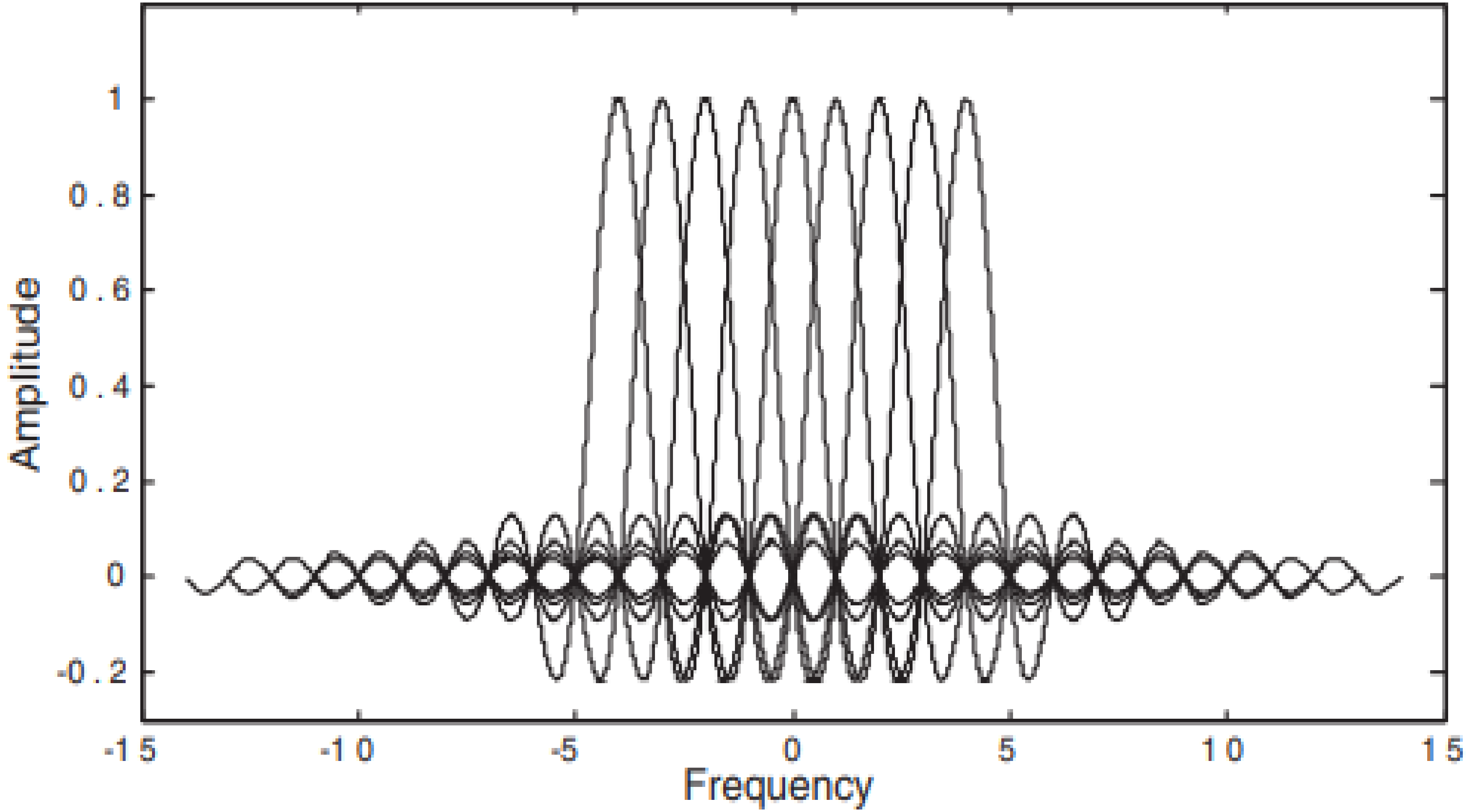


Fig.3 The spectrum of OFDM signal (linear scale) [1]

What is an OFDM System?

- The OFDM system is implemented using discrete Fourier transform (DFT) and inverse DFT (IDFT) processes. Which can be implemented efficiently by using fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT) respectively [1].
- It does not use oscillators for each subchannel, and it doesn't require filters to separate sub-bands in the receiver side, thanks to orthogonality between subcarriers.

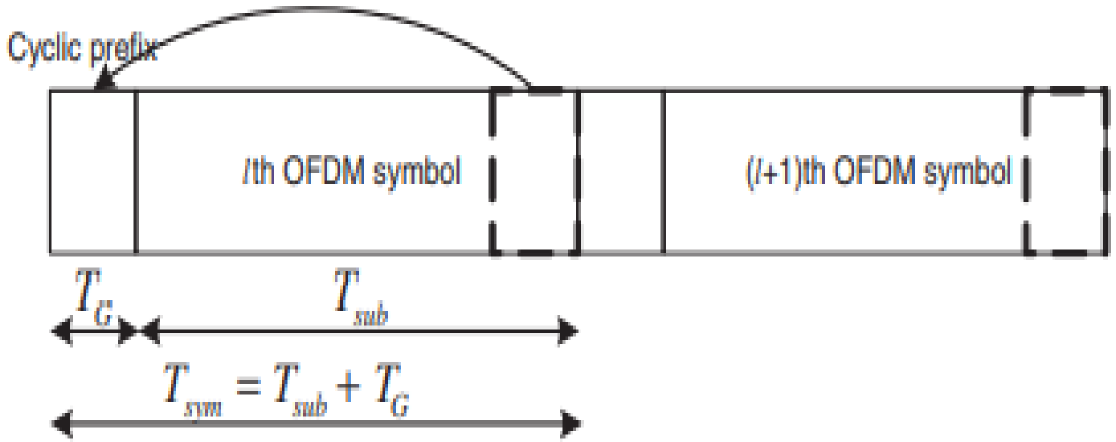


Fig.4 OFDM symbols with CP

- Due to multipath channel, the receiver has many versions of the transmitted data which causes inter symbol interference (ISI).
- OFDM can simply overcome ISI by adding a Guard interval called Cyclic Prefix, as shown in fig.4.
- CP is to extend the OFDM symbol by copying the last samples of the OFDM symbol into its front. CP must be longer than maximum delay spread of the channel.

Implementation of OFDM system

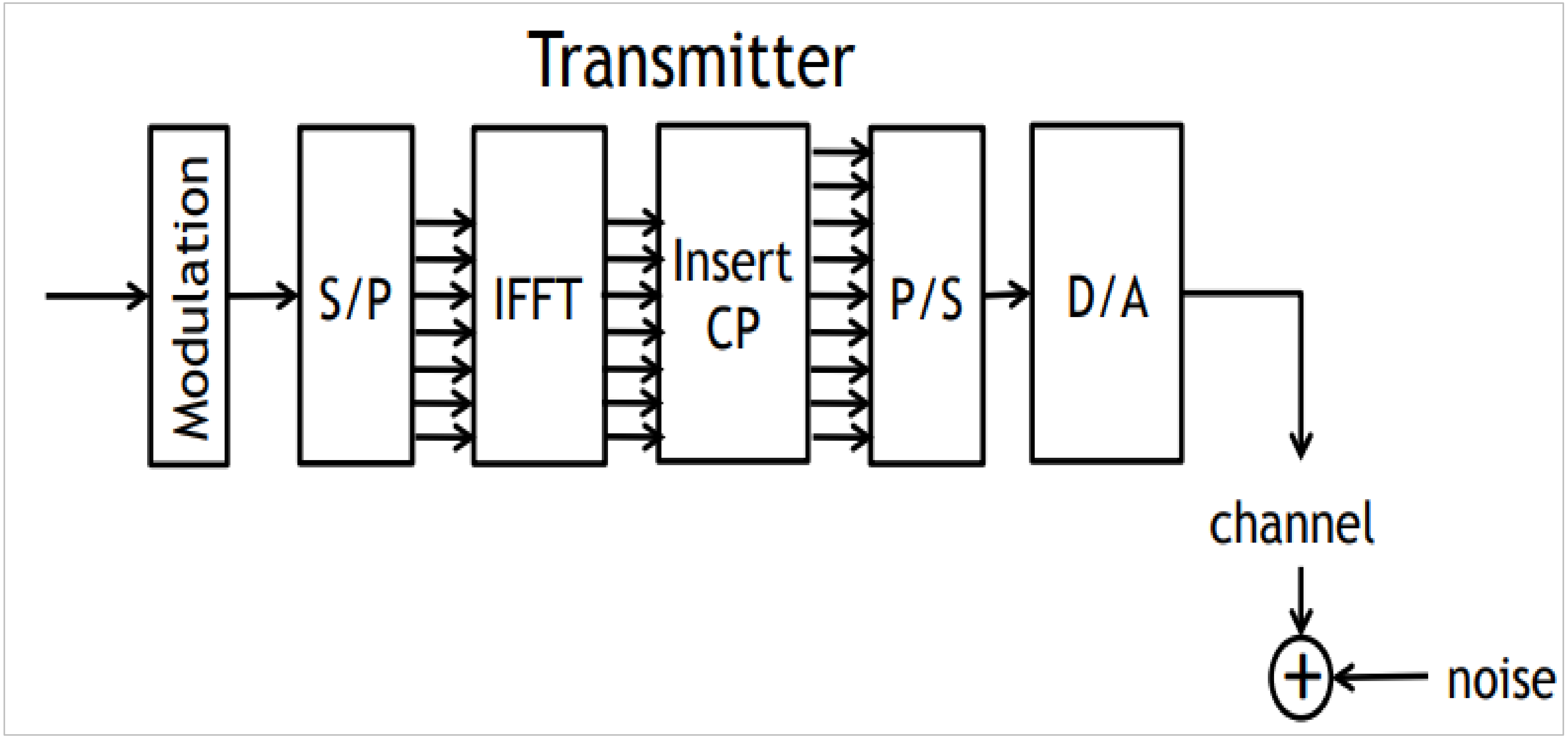


Fig.5 Block diagram of transmitter in an OFDM system [11]

Two channel models

1. Rayleigh distribution.

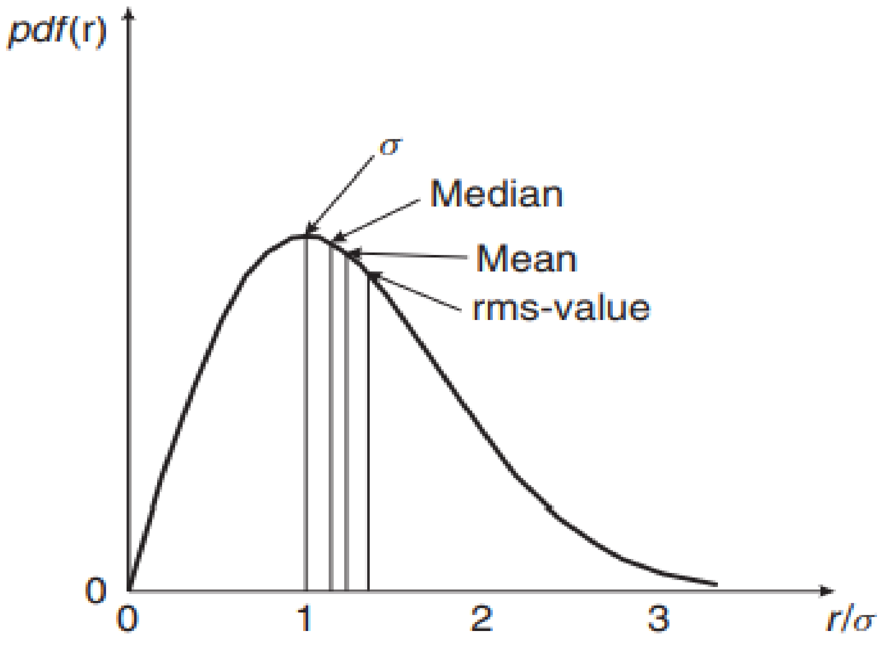


Fig.6 Pdf of a Rayleigh distribution [14]

2. Rice distribution.

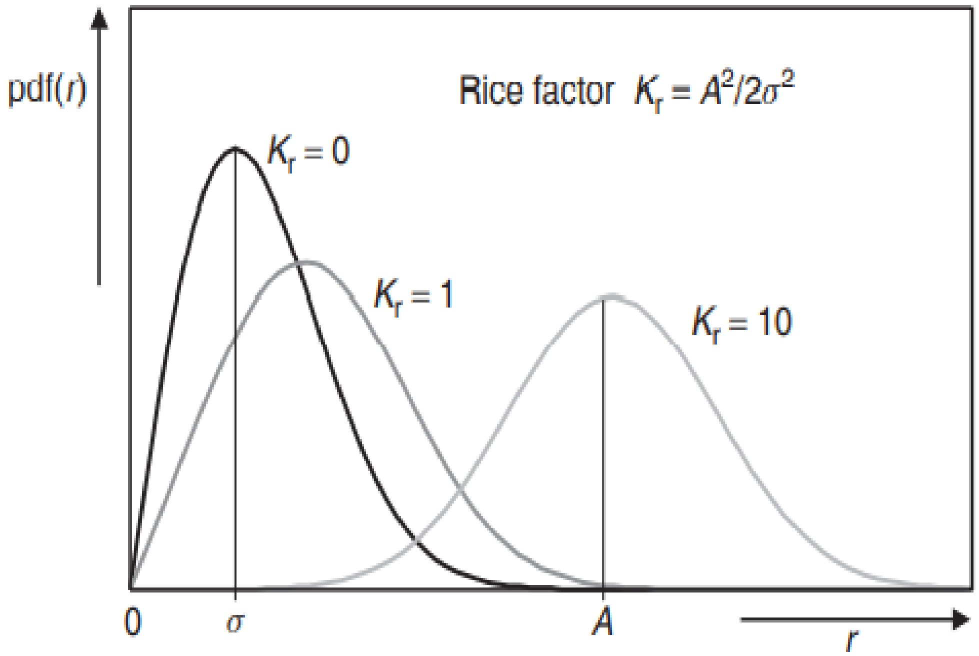


Fig.7 Rice distribution for three different values of Kr [14]

Implementation of OFDM system

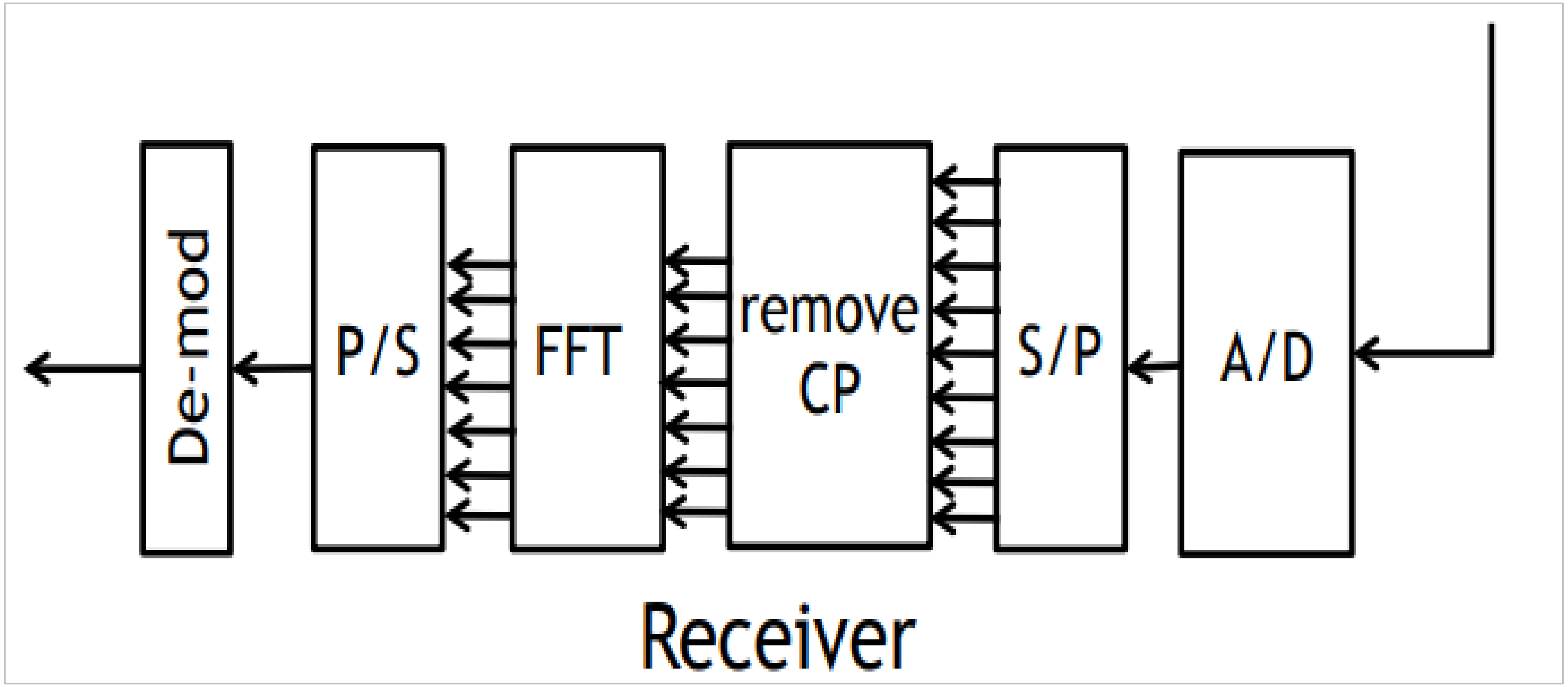


Fig.8 Block diagram of receiver in an OFDM system [11]

PAPR in OFDM system

- Peak to average power ratio is the ratio between the maximum power and the average power of the complex passband signal [1].

$$PAPR\{x[n]\} = \frac{max_{0 \le n \le N-1} \ |x[n]|^2}{E[|x[n]|^{\ 2}]} \quad (2)$$

Where $x[n]$ is the complex passband signal and N is the number of subcarriers. From equation, it’s clear that increasing N will result in higher PAPR.

- As mentioned before, one of the major problems of OFDM system is high PAPR.

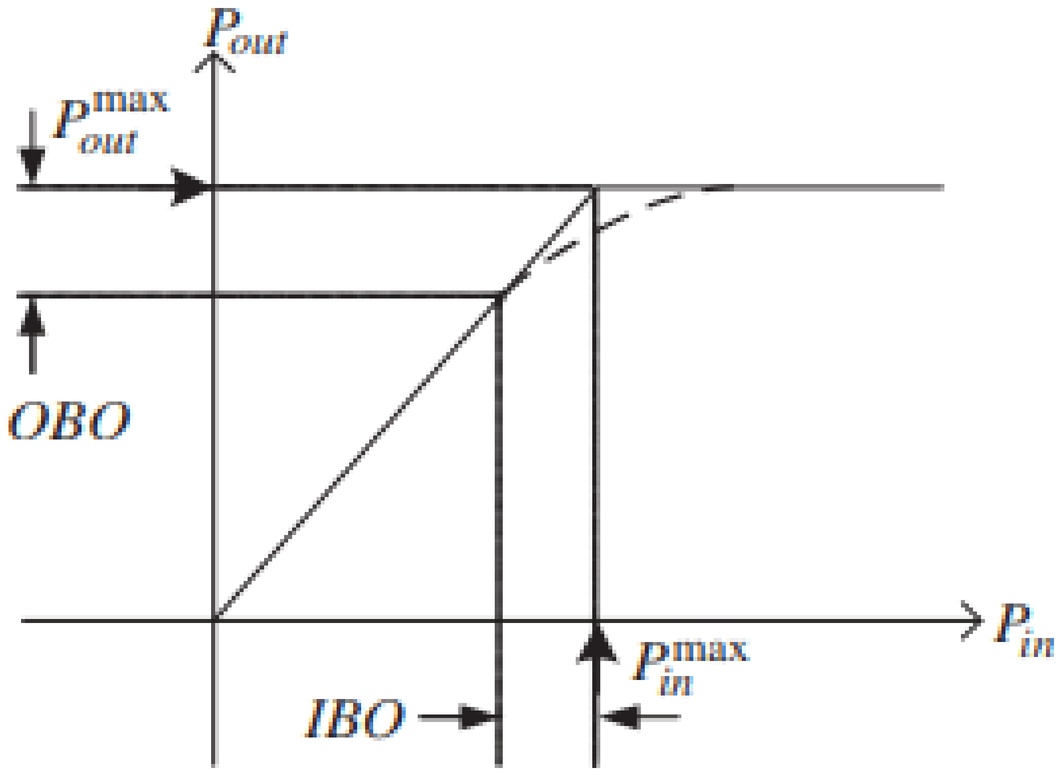
- Mainly this issue originates from the fact that an OFDM signal is the superposition of N sinusoidal signals on different subcarriers, since all subcarriers are added together in time domain via IFFT operation, this may result in adding all subcarriers constructively.
- High PAPR decreases SQNR (Signal-to-Quantization Noise Ratio) of ADC (Analog-to-Digital converter) and DAC (Digital-to-Analog Converter). And the main drawback of PAPR is **degrading the efficiency of the high-power** amplifier in the transmitter.

High Power Amplifier

- Most radio systems employ the HPA (High Power Amplifier) in the transmitter to obtain sufficient transmission power.
- Practical power amplifiers are linear only over a finite range of input amplitudes, so if input becomes much larger than its nominal value, output will be distorted.

High Power Amplifier

- Note that the nonlinear characteristic of HPA, causes the out-of-band radiation that affects signals in adjacent bands, and in-band distortions that result in rotation, attenuation, and offset on the received signal [1].



- Due to the saturation characteristic of the amplifier, the maximum possible output is limited by P_{out}^{max} ,when the corresponding input power is given by P_{in}^{max}. Therefore, the nonlinear region can be described by IBO (Input Back-Off) or OBO (Output Back-Off) [1].

$$IBO = 10\, log_{10} \frac{P_{in}^{max}}{P_{in}}\,,\ OBO = 10\, log_{10} \frac{P_{out}^{max}}{P_{out}} \quad (3)$$

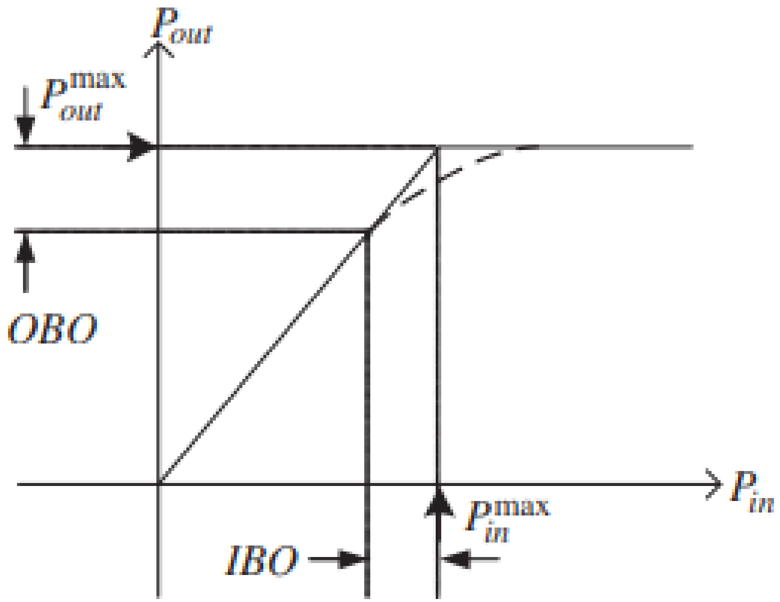


Fig.10 input-output characteristic of an HPA

Note: To prevent saturation and clipping of the OFDM signal peaks, the amplifiers must be operated with sufficient back-off. However, increasing back-off will reduce the efficiency of the power amplifier [4].

How to deal with high PAPR?

1. Use a power amplifier with a very high linear rang that can amplify linearly the possible peak value of the transmit signal. But this is not practical as it requires expensive and power-consuming amplifiers.

2. Use a non-linear amplifier and accept the fact that amplifier characteristics will lead to distortions in the output signal. Those nonlinear distortions destroy orthogonality between subcarriers, and lead to increased out-of-band emissions.

3. Apply PAPR reduction techniques.

Classical PAPR reduction techniques

1. **Tone reservation PAPR reduction technique**
2. **Selective mapping PAPR reduction technique**

Tone reservation PAPR reduction technique

- In this proposed technique, some OFDM subcarriers are reserved. These reserved subcarriers don’t carry any data information, they are only used for reducing PAPR. This method is called Tone Reservation [10].
- There are different methods for tone reservation technique, such as Kernel TR method and Scaling Signal To Clipping Noise Ratio (TR-S-SCR) method.
- Each method will be explained in detail with the performance analysis.

- **Kernel TR** method creates a reference kernel vector p, which is an impulse function with FFT size, in the kernel vector the reserved tones positions will be set to one, and all other values are set to zero.

$$X = \begin{bmatrix} modulated \\ modulated \\ reserved \\ modulated \\ reserved \\ modulated \\ . \\ . \\ . \end{bmatrix}, \qquad P = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ . \\ . \\ . \end{bmatrix}$$

…Then we do IFFT for X & P [6].

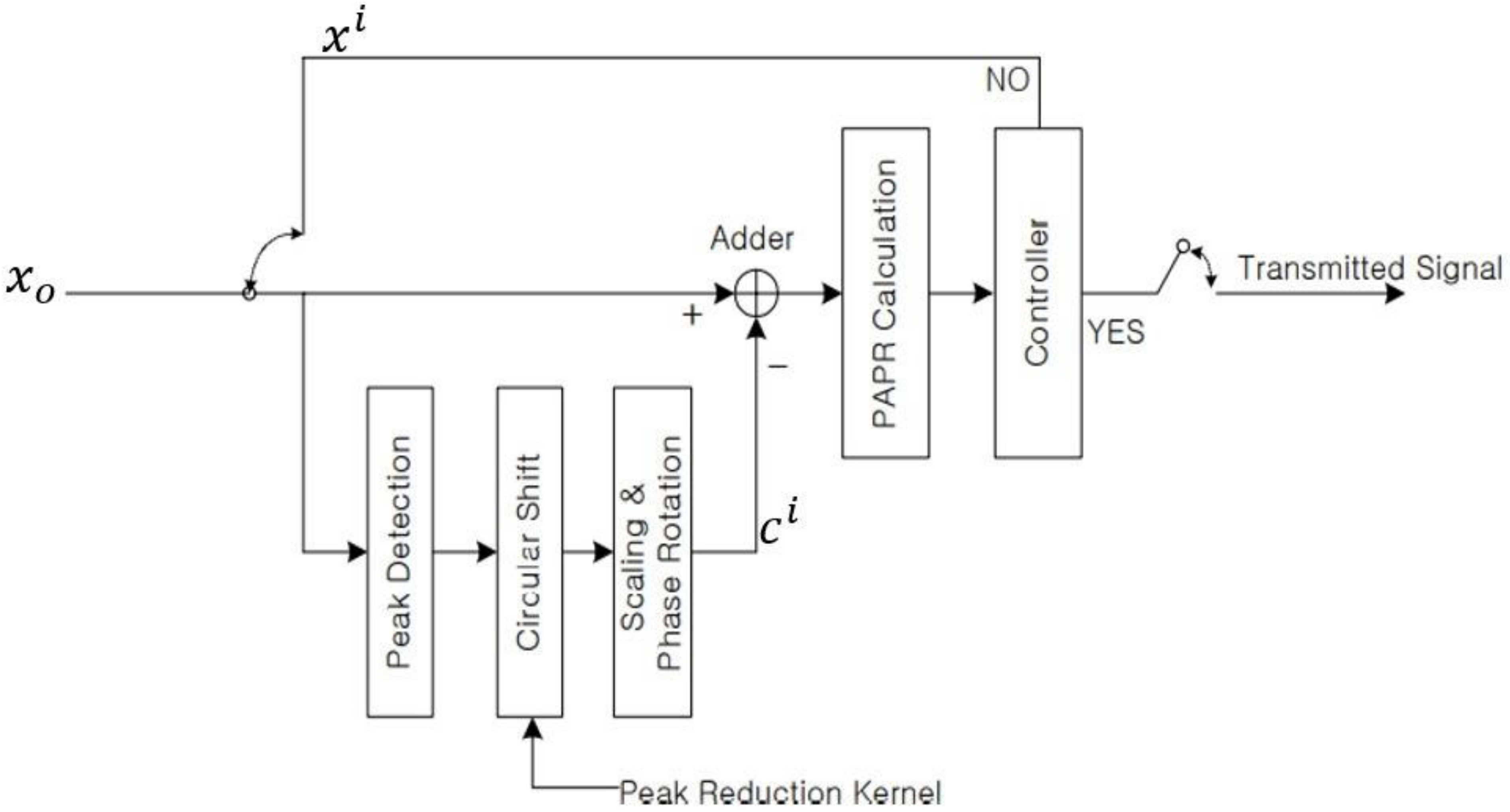


Fig. 11 Procedure of Gradient Algorithm [7]

- In every iteration i the maximum amplitude A^i of OFDM signal x^i and its position m^i must be found. The position of the maximal amplitude of the kernel vector is circularly shifting to the position m^i, scaled and phase rotated. The subtract of kernel vector from x^i reduces the peak to a previously determined wanted clipping level [7].

- The following equations explain the method:

$$x^{i+1} = x^i - c^i$$

$$c^i = \alpha^i p(m^i) \qquad (4)$$

$$\alpha^i = \frac{x(m^i)}{A^i}\left(A^i - A_{max}\right)$$

$$A_{max} = CR * \sqrt{\frac{1}{N}\sum_{n=0}^{N-1}|x^i(n)|^2}\,, CR = 1.5dB \qquad (5)$$

where A_{max} is the clipping amplitude, $p(m^i)$ is the circularly shifted time domain kernel and CR is the clipping ratio, then the PAPR is calculated [5,6]. If the number of iteration reaches predetermined maximum iteration number or PAPR reaches the threshold, control escapes the process and resulting signal is transmitted. If not, clipping operation is executed iteratively [7].

- To justify the performance of the TR-K method we compare our simulation results with research results in [8].
- Simulation parameters →QPSK modulation
  - →1024 sub carrier
  - →10k symbols
  - →32 Peak reduction tones
  - →20 iterations
  - →PAPR threshold = 5 dB
  - →Random set optimization for PRTs

- Research results are smoother as the use of 1M symbols in research simulation, and the use of only 10K symbols in our simulation.
- From results, we can see a reduction in PAPR value by almost 3dB.

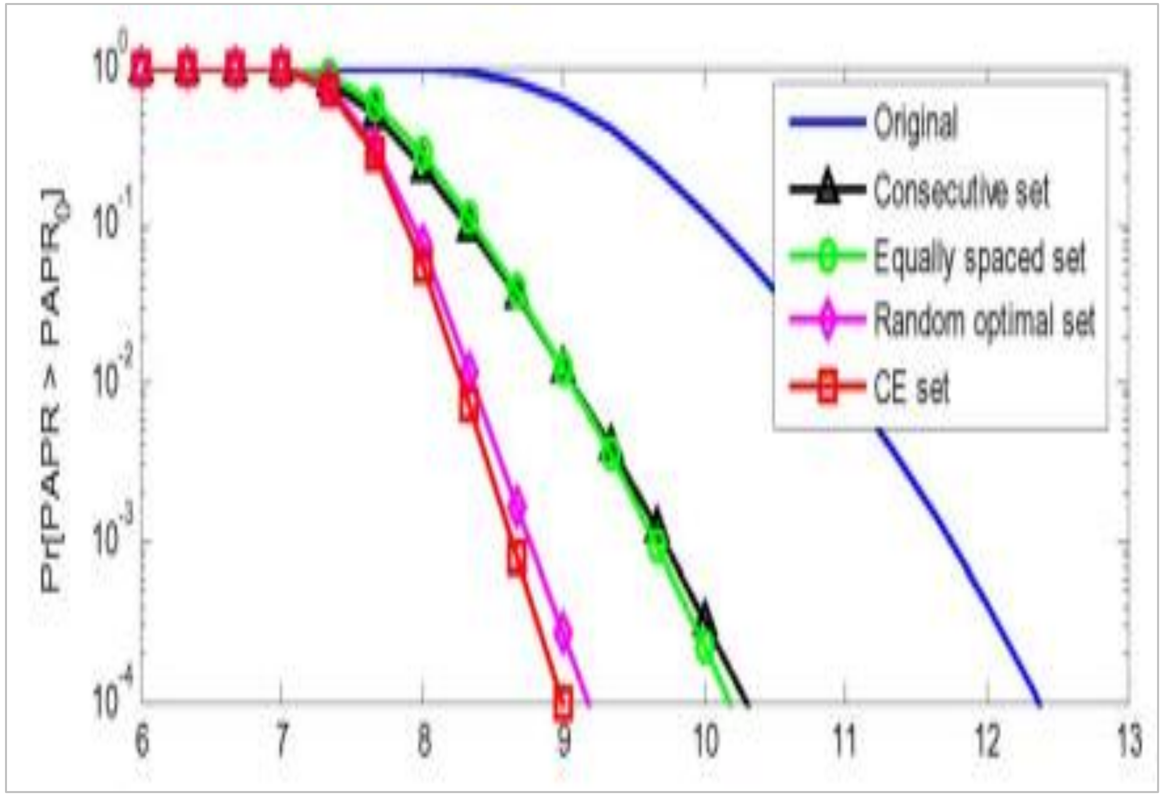


Fig. (a) Research simulation results [8]

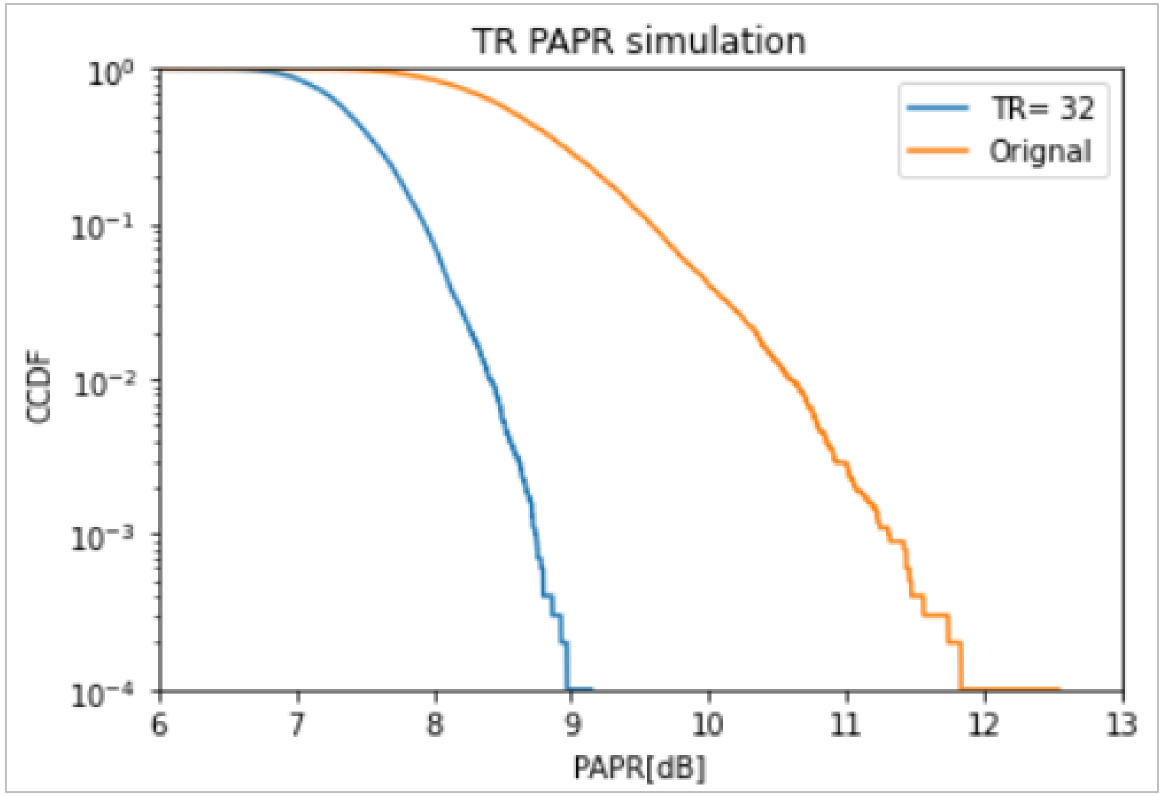


Fig. (b) Team simulation results

- In the **SCR** scheme, the peak reduced signal $x(n)$ is iteratively updated by using a simple gradient algorithm given by:

$$x^{m+1}(n) = x^{m}(n) - \mu.\, c_{max}^{m}(n) \qquad (6)$$

$$c_{max}^{m}(n) = \left(x^{m}(n_{max}^{m}) - A.\, e^{j\theta_{max}}\right)p(n - n_{max}^{m}) \qquad (7)$$

→ $\mu.\, c_{max}^{m}(n)$ is the peak reduction signal

→ μ is a constant scaling factor

→ $A = CR * \sqrt{\frac{1}{LN}\sum_{n=0}^{LN-1}|x^{m}(n)|^{2}}$ is the required clipping threshold

→ CR is the clipping ratio

→ $p(n - n_{max}^{m})$ denotes the circularly shifted sequence of time domain kernel to max n_{max}^{m} which is the position of the peak amplitude of $x^{m}(n)$ in the m-th iteration [9].

- To improve the convergence rate, we present an **S-SCR** scheme which employs the gradient algorithm with additional peak regeneration restraint to obtain the optimized scaling factor.
- By using the LSA algorithm, the optimal solution of μ can be calculated as

$$\mu = \frac{\sum_{n\in s}|x^m(n)||f^m(n)|}{\sum_{n\in s}|x^m(n)|^2} \qquad (8)$$

- Given the clipping threshold A, the directly clipped signal can be expressed as

$$\dot{x}^m(n) = \begin{cases} x^m(n) & |x^m(n)| < A \\ A.e^{j\theta_n} & |x^m(n)| > A \end{cases} \qquad (9)$$

then the clipping noise is defined: $f^m(n) = x^m(n) - \dot{x}^m(n)$ (10)

- Then, we express the position of the peaks which satisfy $|f^m(n)| > 0$ with U entries as $S = \{S_0 S_1, \dots, S_{U-1}\}$
- Furthermore, for the purpose of avoiding an undesirable peak regeneration, we design a 1×LN scaling vector Ψ that restricts the scaling adjusting only on the peak value in one iteration. Thus, we get

$$\Psi(i) = \begin{cases} \mu & i = n_{max} \\ 1 & i \neq n_{max} \end{cases} \qquad (11)$$

- Accordingly, we can rewrite the peak reduced signal $x(n)$ with optimal solution of the scaling factor as

$$x^{m+1}(n) = x^m(n) - \Psi\left(x^m(n_{max}^m) - A.e^{j\theta_{max}}\right)p(n - n_{max}^m) \qquad (12)[9].$$

- To justify the performance of the S-SCR method we compare our simulation results with research results in [9].
- Simulation parameters → 16-QAM modulation
  → 256 sub carrier
  → T/N=1/8
  → 10k symbols.
  → Random set optimization for PRTs
  → Iterations = 18
  → CR= 2dB
  → Oversampling factor (W)=4

- From results, we can see a reduction in PAPR value by almost 3.5dB.

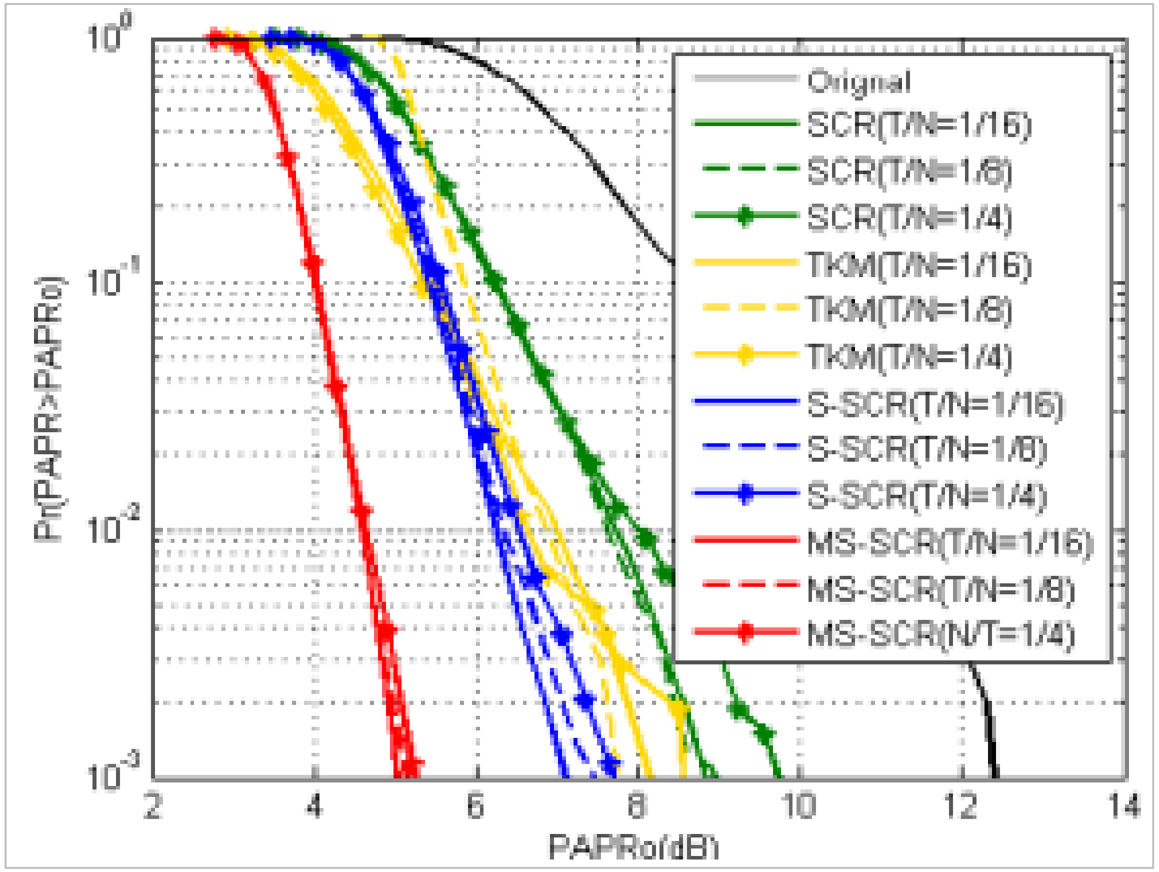


Fig. (a) Research simulation results [9]

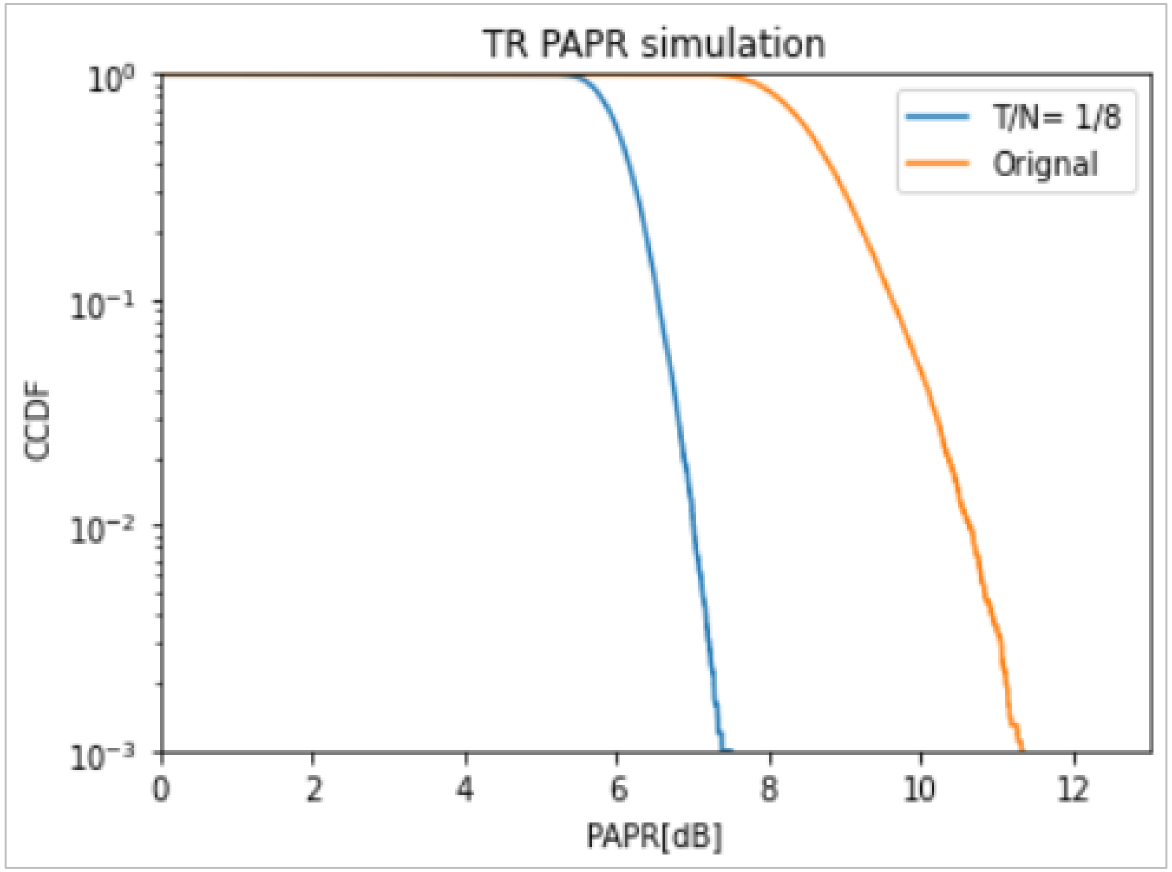
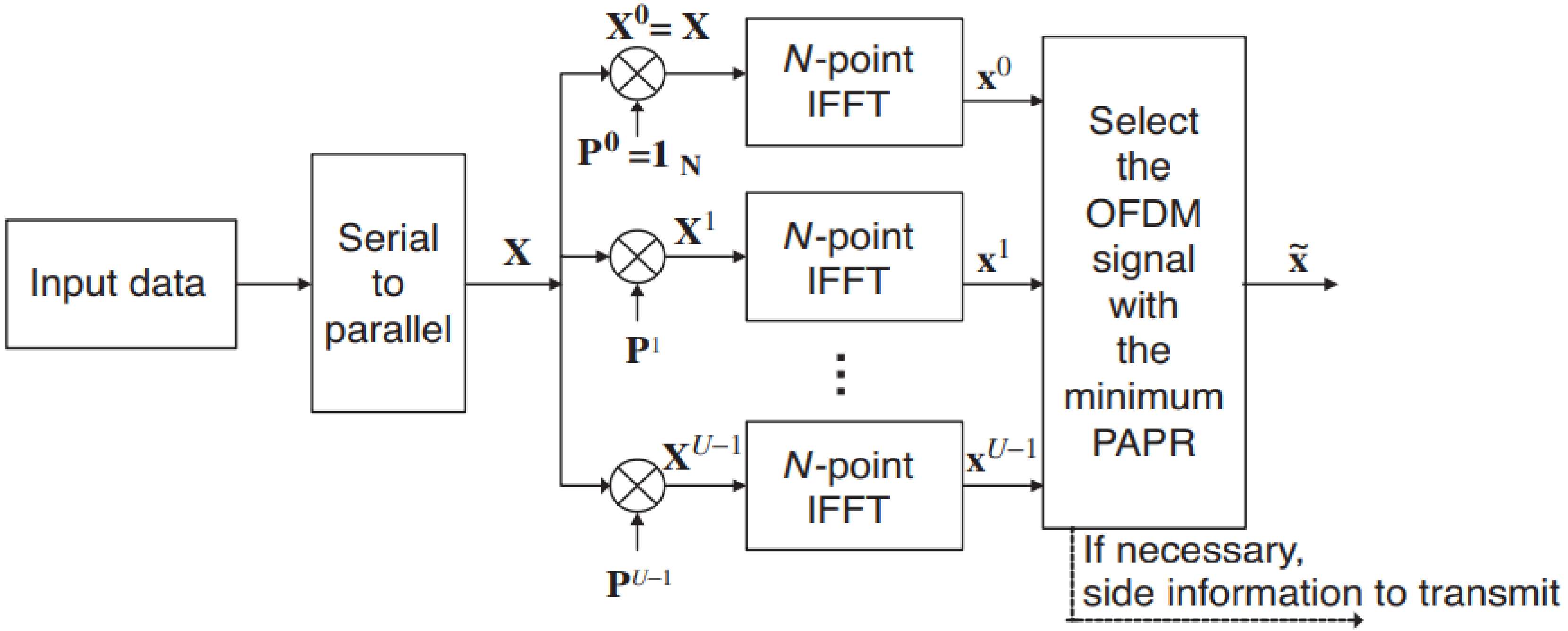


Fig.(b) Team simulation results

Selective mapping PAPR reduction technique



- In selective mapping technique, the input data block $X = \left[x[0], x[1] \ldots, x[N-1]\right]$ is multiplied with U different phase sequences $p^u = [p_0^u, p_1^u, \ldots, p_{N-1}^u]^{\,T}$ where $p_v^u = e^{\vartheta_v^u}, v = 0,1, \ldots, N-1$, and $u = 0,1, \ldots \ldots, U-1$

- This produce a modified data blocks $X^u = \left[x^u[0], x^u[1], \ldots\ldots, x^u[N-1]\right]^T$ which represent the same data, These are then forwarded into IFFT operation simultaneously. And then the PAPR is calculated for each vector separately.
- The sequence with the smallest PAPR is selected for final transmission. For the receiver to be able to recover the original data block, the information about the selected phase sequence P^u should be transmitted as a side information [1].

- To justify the performance of the SLM method we compare our simulation results with research results in [10].
- Simulation parameters → 4-QAM modulation
  → 64 sub carrier
  → 10k symbols
  → U=2, 4, 8, 16, 32, 64

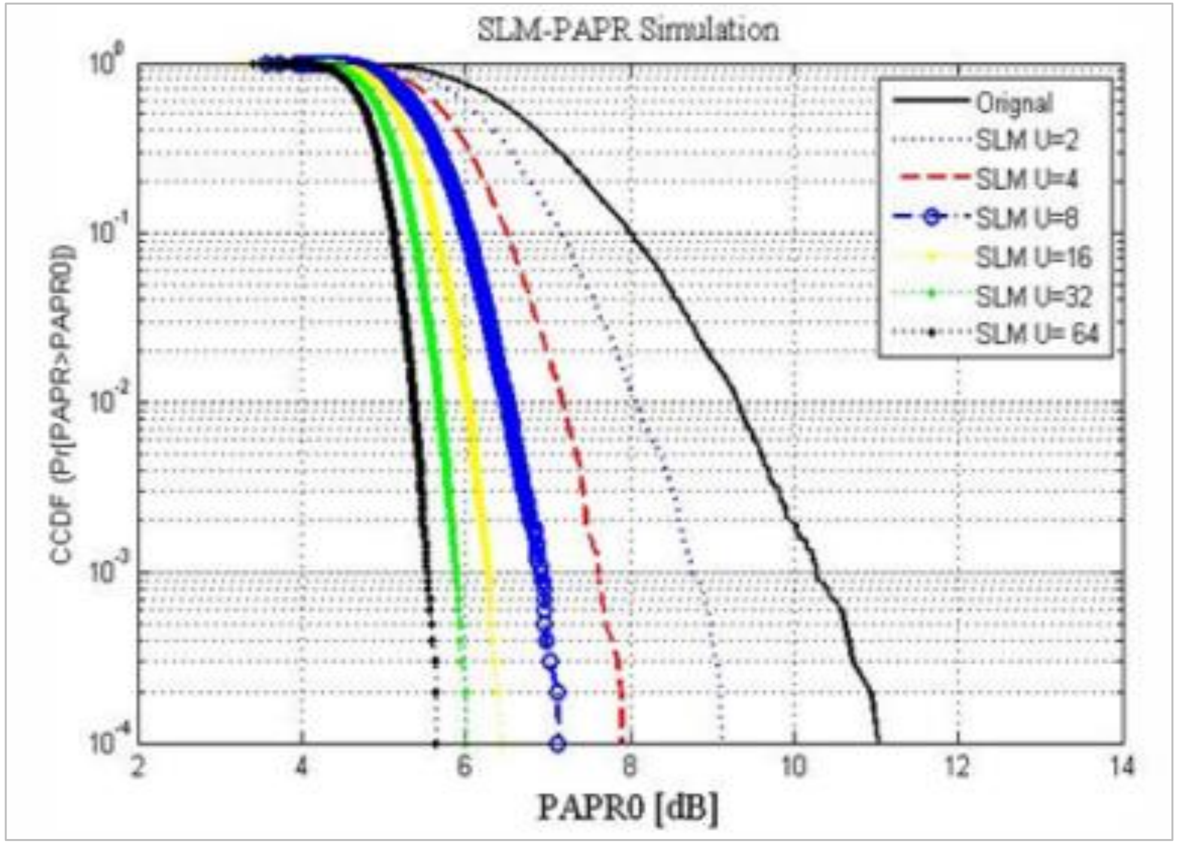


Fig. (a) Research simulation results [10]

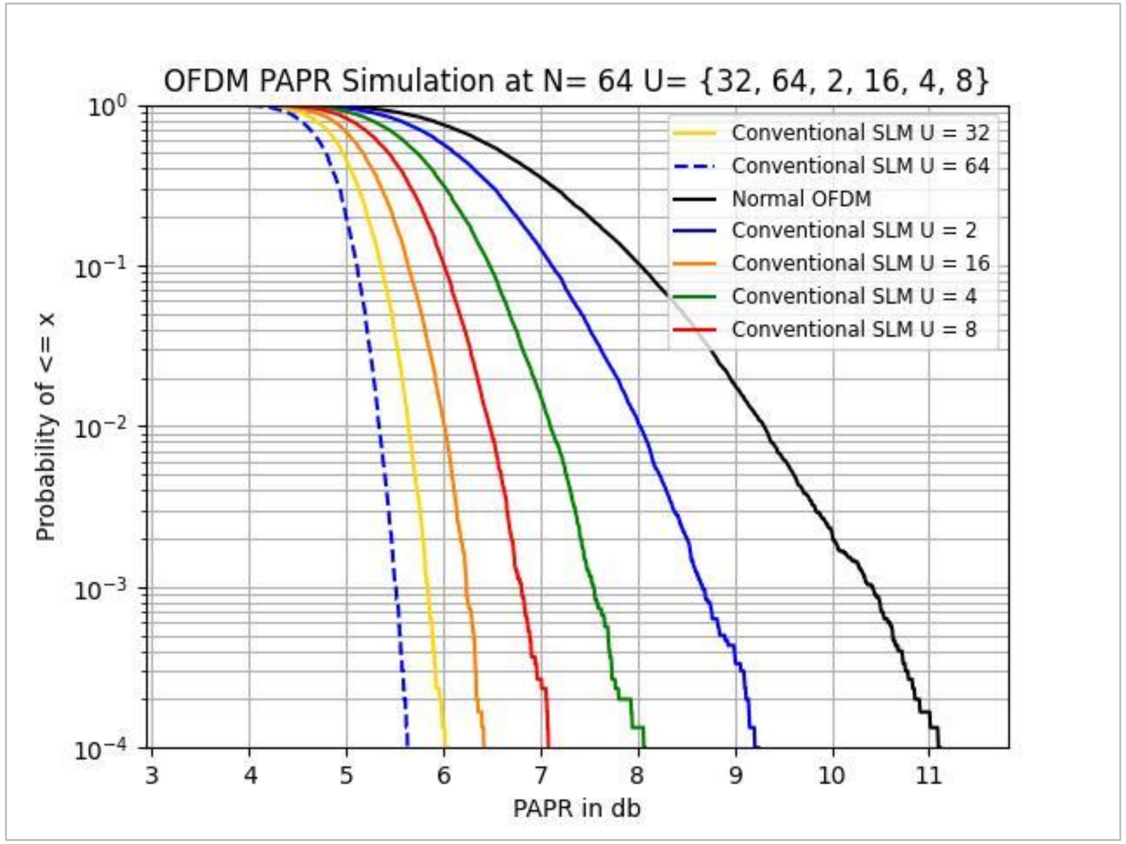


Fig.(b) Team simulation results at different values of U

- Note: We notice that as the number of phase sequence increases, reduction of PAPR also increases. However, this will lead to more IFFT operation, and increasing the side information that should be sent to the receiver to recover OFDM symbols.

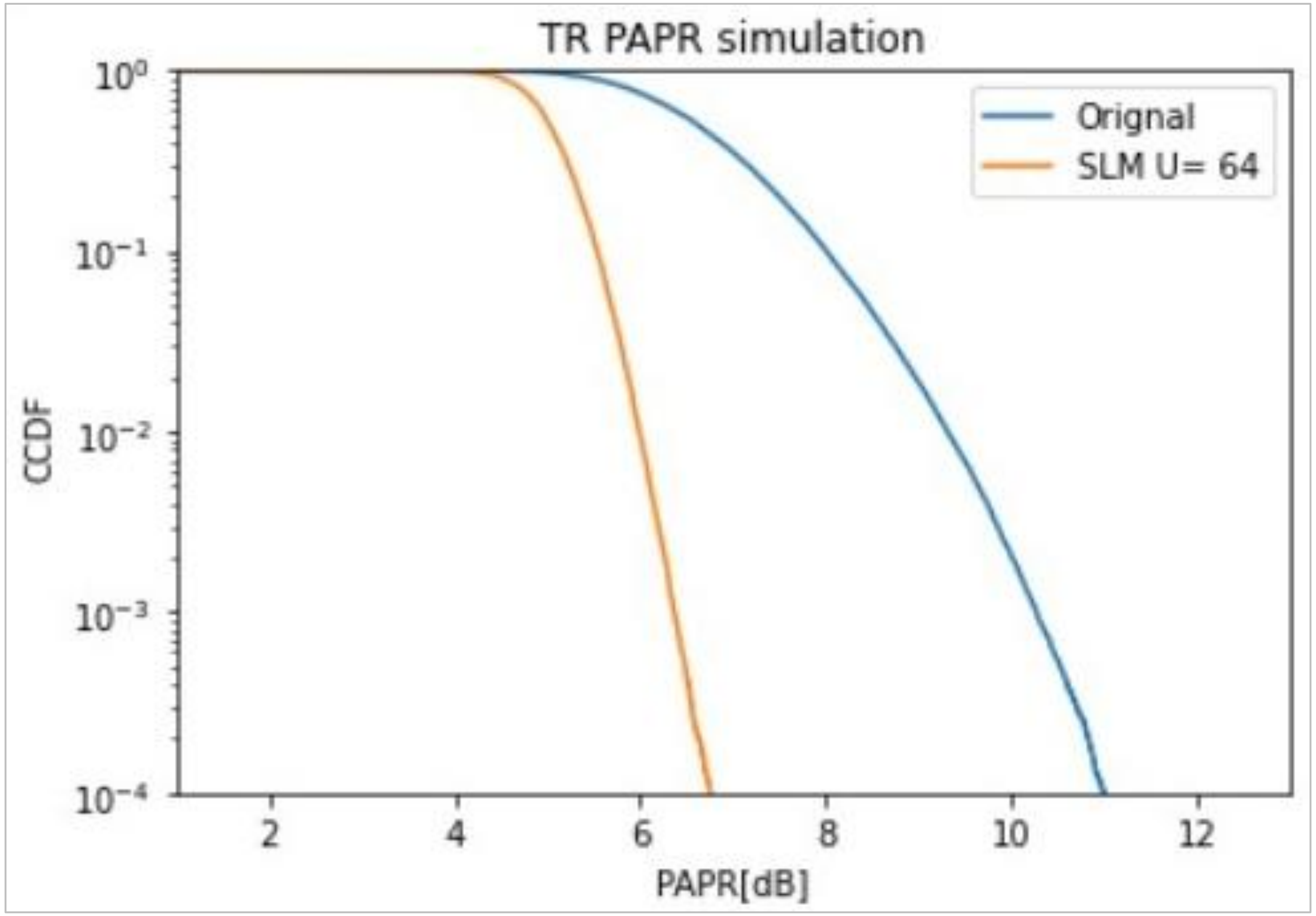


Fig.(c) Team simulation results with real Hadamard Matrix Sequence

Introduction to Machine Learning

What is Machine Learning?

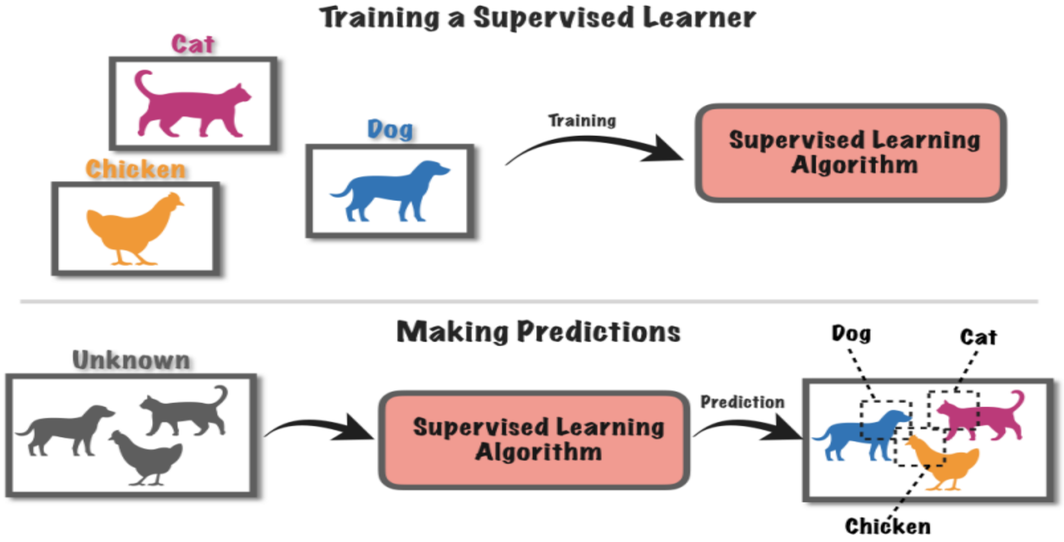
- Machine learning is an application of AI that provides systems the ability to learn on their own and improve from experiences without being programmed explicitly.
- It is about extracting knowledge from data and thus making predictions upon the collected information.
- Machine learning has 2 main types:

➢ **Supervised learning** in which training data is labelled.

➢ **Unsupervised learning** in which training data has no labels.

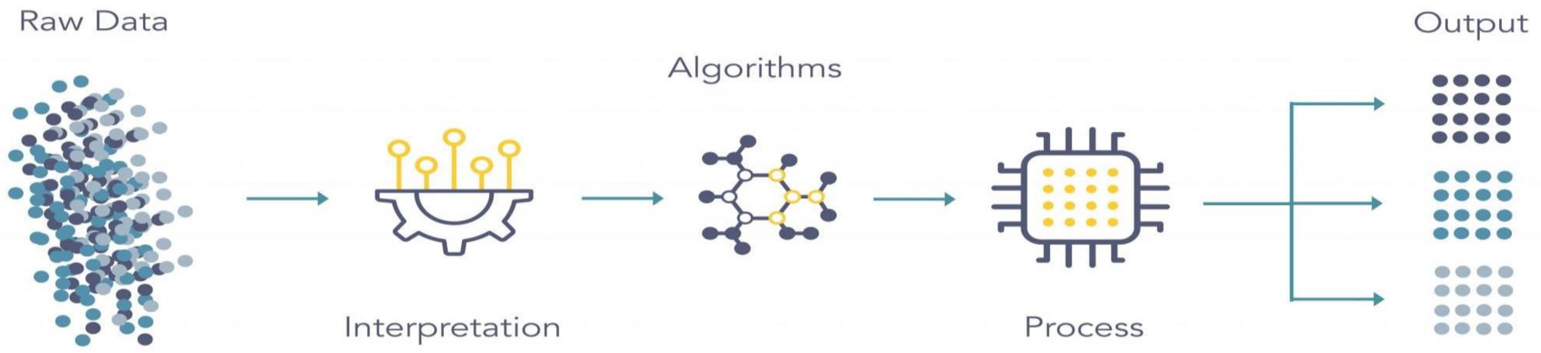
Supervised Learning

- There are two main types of supervised learning.
- **Classification**, the goal here is to predict a class label from a predefined list of probabilities. It is also separated into two main categories **Binary Classification** and **Multiclass classification**.
- **Regression**, the goal is to predict a continuous number.
- An easy way to distinguish between classification and regression to ask whether there is continuity in the output.



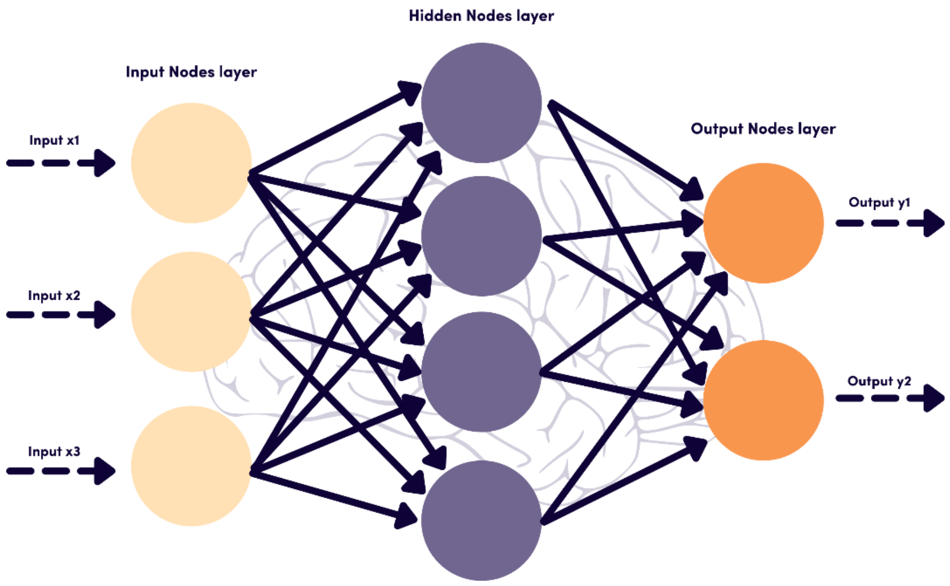
Unsupervised Learning

- There are three main architectures for the Unsupervised Learning. **Clustering**, **Recurrent NN** and **Autoencoders.**
- **Autoencoders** are regenerative models that tries to generate the input data after it has been corrupted and this type of autoencoders is called **Denoising Autoencoders**.
- We used a Denoising Autoencoder to reconstruct the signal after it has been distorted due to the channel.



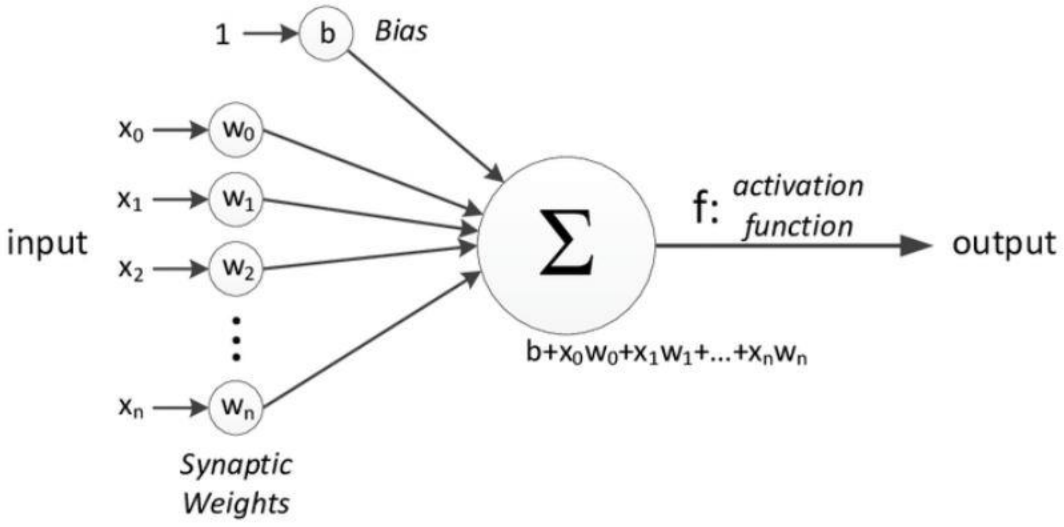
What are Neural Networks?

- We can describe a neural network as a mathematical model for information processing.
- Neural network consists of **layers** ,each layer contains number of neurons.
- Information processing occurs over elements called **Neurons**.
- Neurons are connected together, and they exchange information through information link.
- They look like neurons in the brain, that's why they are called Neural Networks



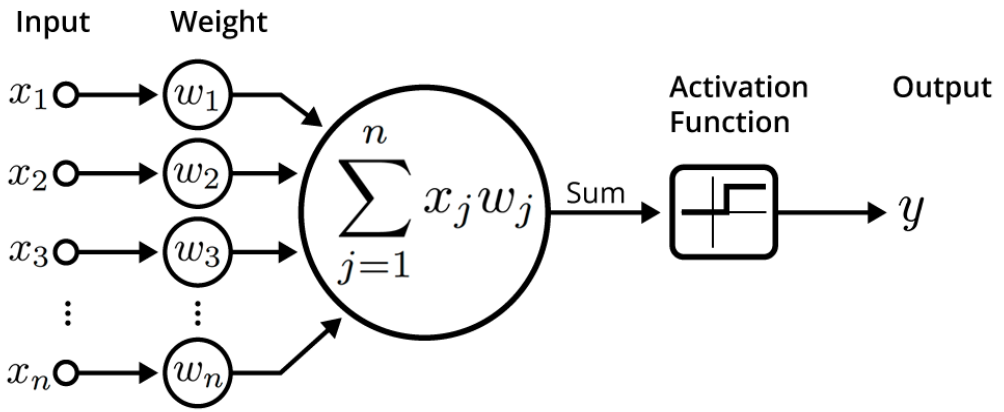
What are Neural Networks?

- Their idea are mainly based on the **Logistic Regression.**
- Logistic regression is simply **Linear Regression** but with activation function.
- Linear regression is simply matrix multiplication between weights and inputs and then a bias is added.



Training Neural Networks

1. Randomly initialize the weights.
2. Put your inputs in the input layer.
3. **Forward-Propagation** from left to right, Propagate through network until getting the predicted value.
4. Compare the predicted result to the actual result and measure the generated error (Cost function).



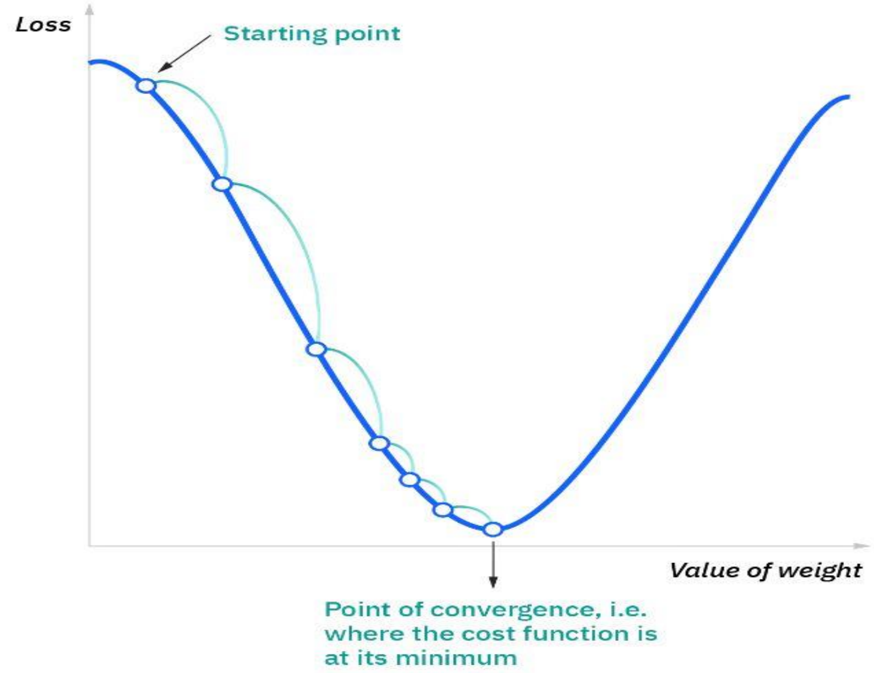
An illustration of an artificial neuron. Source: Becoming Human.

Training Neural Networks

5. **Back-Propagation** from right to left, the error is backpropagated. Update the weights according to how much they are responsible for the error.

   "The learning rate decides how much we update weights"

6. Repeat until the cost function is minimized as much as possible.

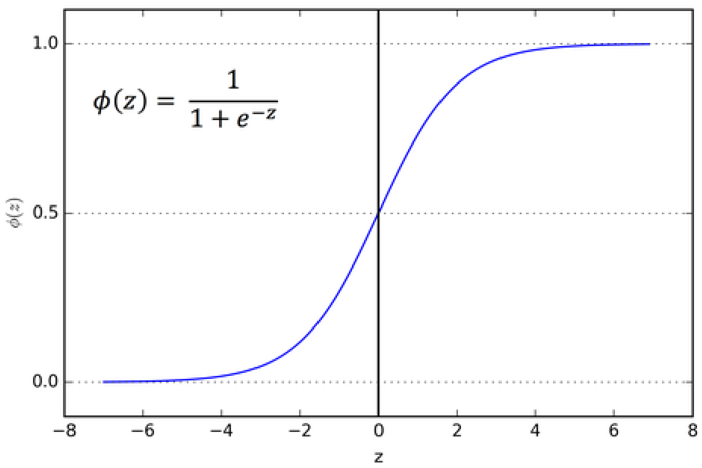


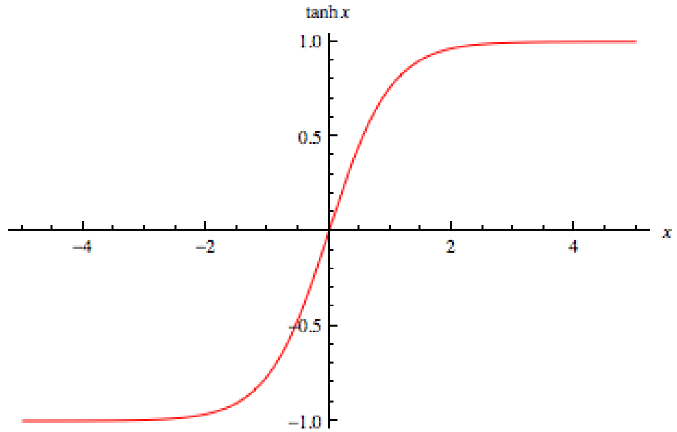
Activation Functions

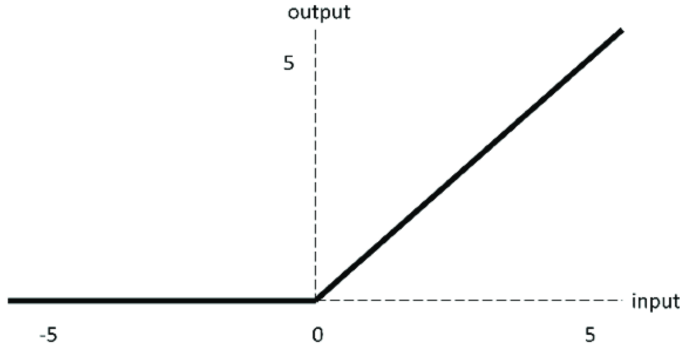
- Activation function decides, whether a neuron should be activated or not.
- The purpose of the activation function is to introduce **non-linearity** into the output of a neuron.
- A neural network without an activation function is essentially just a linear regression model. The activation function does the **non-linear transformation**. to the input making it capable to learn and perform more **complex tasks**.

Activation Functions Types

- **Sigmoid** is a logistic function, and the output is ranging between 0 and 1.
- **tanh** Usually used in hidden and output layers of a neural network, as it's values lies between [-1,1].
- **ReLU** is less computationally expensive than tanh and sigmoid because it involves simpler mathematical operations, it's used mainly in the hidden layers, to speed up training.





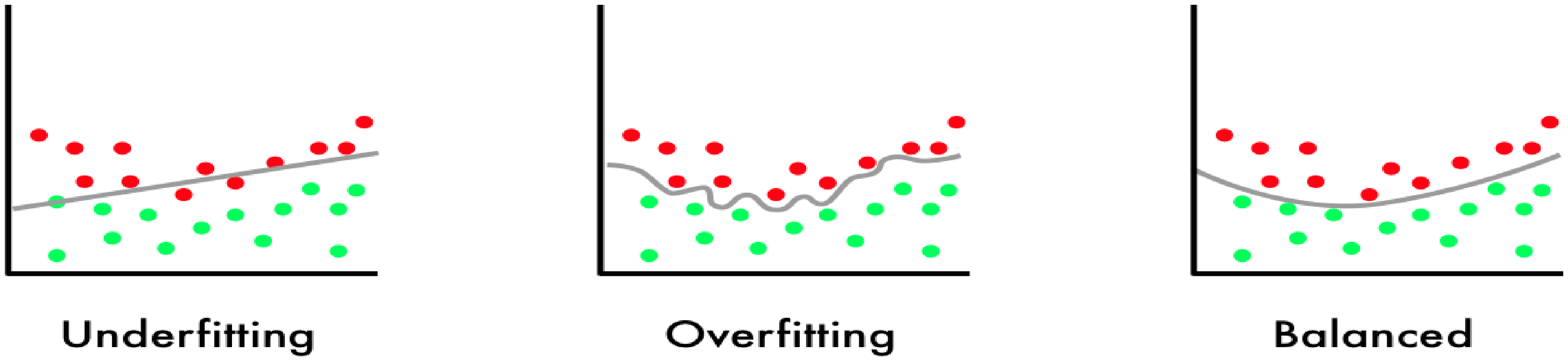


Loss Function

- The loss function in a neural network is the difference between the expected outcome y_i and the output produced by the model $\hat{y}_i$.
- The main goal of a Deep Neural Network is to minimize loss function.
- There are many types, but we are only concerned about the Log Loss, as we used it in our PRNet model.
- Log Loss formula is $L = -\frac{1}{k}\sum_{i=1}^{k} y_i \log(\hat{y}_i) - (1 - y_i)\log(1 - \hat{y}_i)$.
- We also used the Loss function concept and made it to be the PAPR that we want the Network to minimize it, $L = PAPR$, and it’s called customized Loss Function.

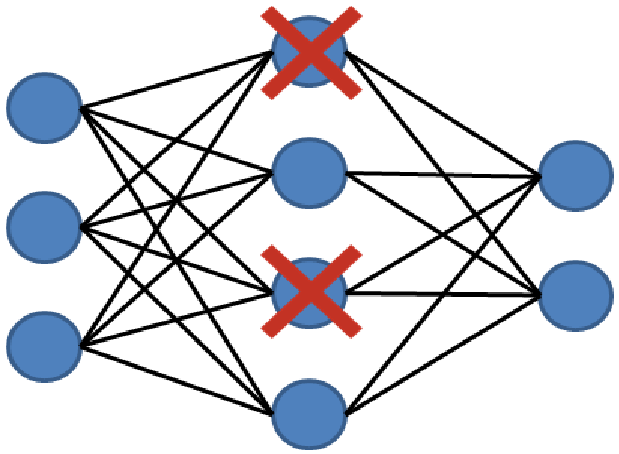
Overfitting

- **Overfitting** occurs when the model gives the predicted results in a good manner while training, but it behaves badly on new unseen data.
- One way to solve this is by using **Regularization** as it reduces the complexity of the model.
- Another way is by using **Dropout**.



Overfitting

- **Dropout** is a technique where randomly selected neurons are ignored during training. They are “dropped-out”.
- Generally, use a small dropout value of 20% to 50%, value too high results in under-learning by the network.
- We widely used the concept of Regularization and Dropout in our models to avoid overfitting and to get better Generalization.



What is TensorFlow ?

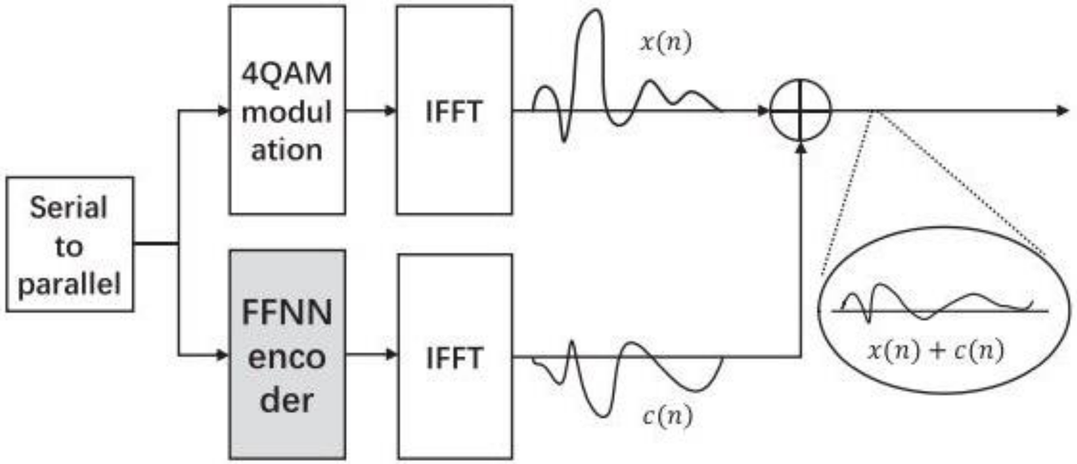
- Open Source software library for numerical computation using data flow graphs.
- Based on Neural Networks.
- Built in C++ with Python interface.
- Supports CPU computations as well as GPU.
- Developed by **Google** and is being used by many companies.
- We mainly depended on TensorFlow in developing our models.

ML PAPR reduction techniques

Tone Reservation Scheme based on Deep Learning

Main Idea

- We introduce a Tone Reservation Model based on Deep Neural Network to enhance the classical Tone Reservation technique.
- We utilize the Feed Forward Neural Network to adaptively generate the peak cancelling signal according to the input signal characteristics.
- The baseband signal is copied into two copies. One copy is used for generating the time domain signal.
- The other copy is fed into the network to generate the peak cancelling signal.
- Finally, we add them together in the time domain and they are ready to be transmitted [12].



TRNet Loss Function

- Every Neural Network is trained to minimize a certain Loss Function.
- We took advantage of that and made our Loss function the PAPR of the data to be transmitted including the reserved tones.
- $Loss = PAPR(x + c) = PAPR\big(IFFT(X + C)\big)$, which is a customized Loss function.
- Neural Networks in general are unable to process complex data.
- Since the modulated symbols are complex, we split the complex numbers into real part and imaginary part in a process called **Data pre-processing**.

TRNet Structure

- We have: Input layer, 3 identical sub-blocks, and One Output layer.
- Input layer contains number of nodes equal to the input data subcarriers but without the reserved tones.
- 3 identical sub-blocks, each contains a FC layer, BN layer, Dropout layer and tanh activation function.

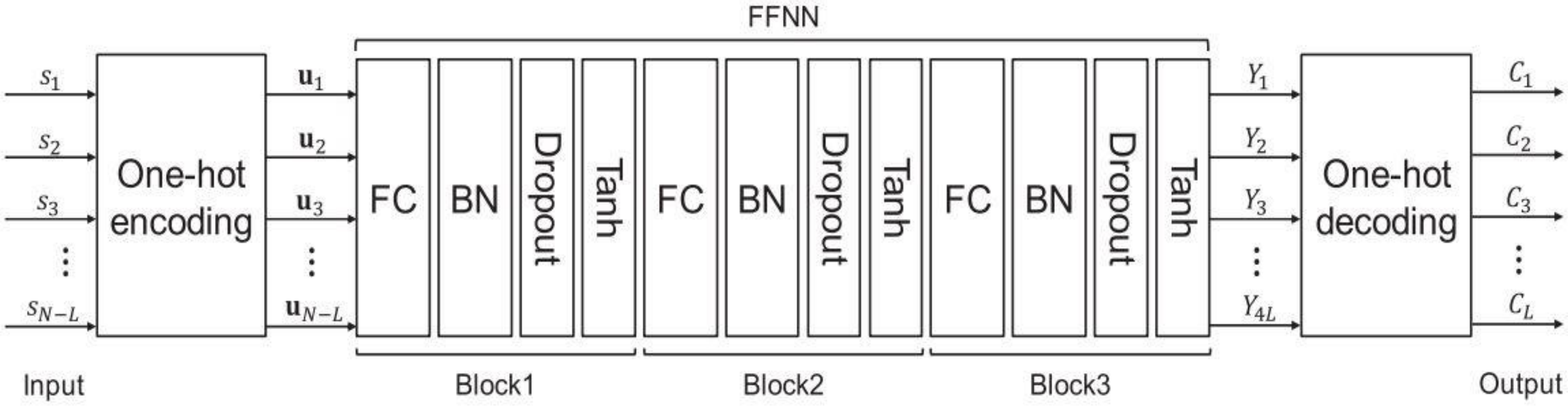


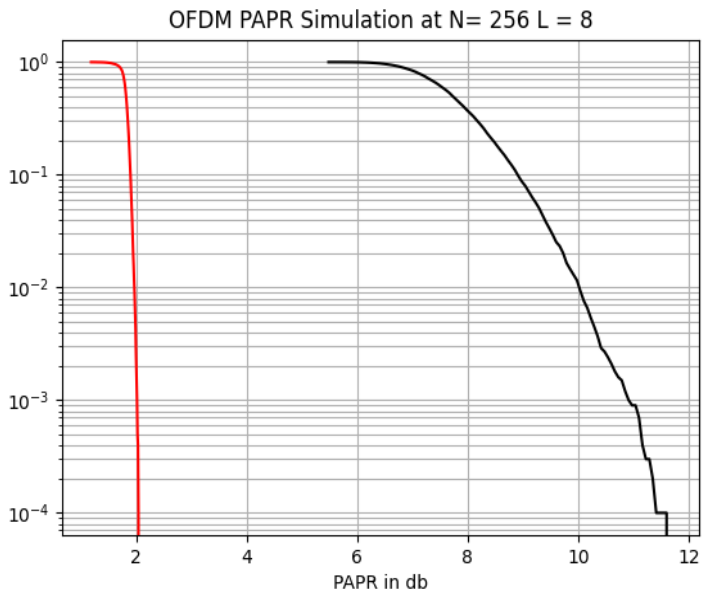
Fig. 12 TRNet structure [12]

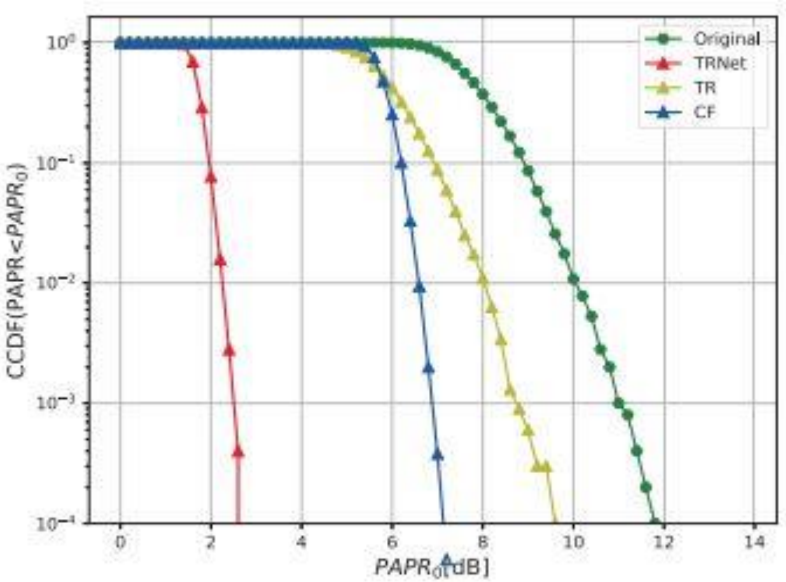
TRNet Structure

- Batch normalization layer normalizes the output of the FC layer to have zero mean and standard deviation of one.
- This normalization helps in solving the vanishing gradient problem and speeds up the training process.
- The data after the BN layer will be limited in the interval of [-1,1].
- The tanh activation function is very close to $y = x$ in this interval and can directly perform the matrix multiplication, which improves training efficiency.
- Dropout layer is to prevent overfitting and make the model more generalized.

Simulation Results

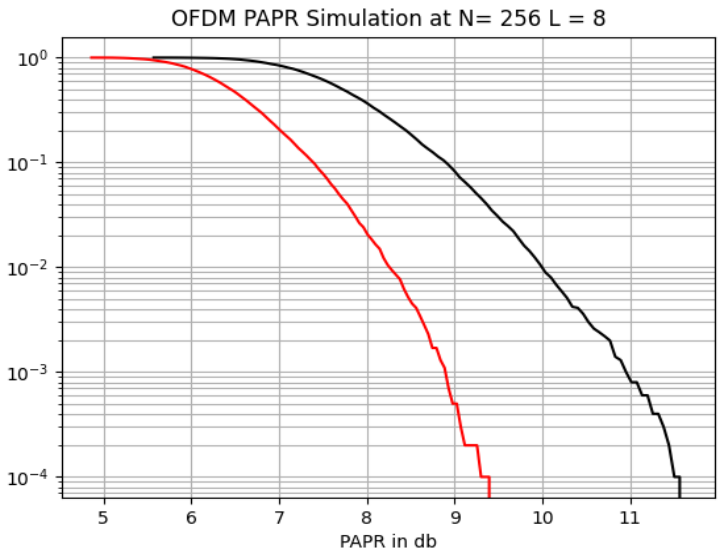
- We considered 256 subcarriers with just 8 reserved tones used for PAPR reduction, 150,000 symbols and 4-QAM modulation scheme.
- We used ADAM optimizer with learning rate of 0.001 and 90 epochs.
- The execution time took around 13 seconds for 10,000 symbols.
- On the left is our model compared to the model found in [12] on the right.

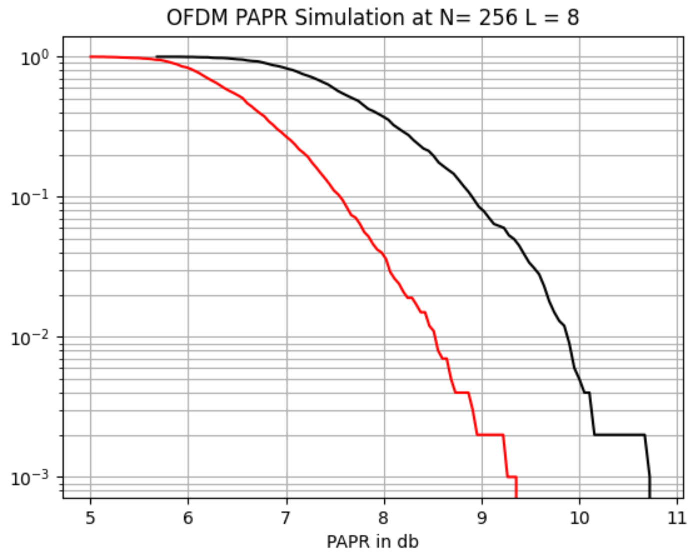




Pros and Cons

- The major drawback is that if we remove the tanh from the output layer, the output won't be limited to [-1,1] anymore and will take any value.
- This leads to a significant increase in the signal power which is not practical.
- We tried another solution by using a tanh function but stretched to be in the interval [-5,5].
- TRNet on the left compared to the classical TR method on the right.





Pros and Cons

- The model gave better performance than the classical TR method. But with little increase in the signal power.
- But the main advantage here is the small execution time and better bandwidth efficiency compared to the classical TR method.
- TRNet uses only one Neural Network at the transmitter which makes the model converges faster.
- New model can be trained in a short time when waveform parameters and modulation parameters change.

ML PAPR reduction techniques

PAPR Reduction Scheme for OFDM System based on Deep Learning

Main Idea

- Autoencoder is a tool for learning data coding efficiently in an unsupervised manner, such that the input to the network is unlabelled. Used in applications such as:
  - Dimensions Reduction
  - Feature Extraction
  - Image Denoising / Compression / Search
  - Anomaly Detection

More simply, the input is encoded by the network to focus only on the most critical feature. In our case used to map the constellation to avoid high PAPR.

System Model

- **The Two main Objectives that must be taken into account:**

  i. PRNet must generate a transmission signal that shows low PAPR

  ii. PRNet must be able to reconstruct the transmitted signal from the received signal such that BER of the system does not deteriorate.

System Model

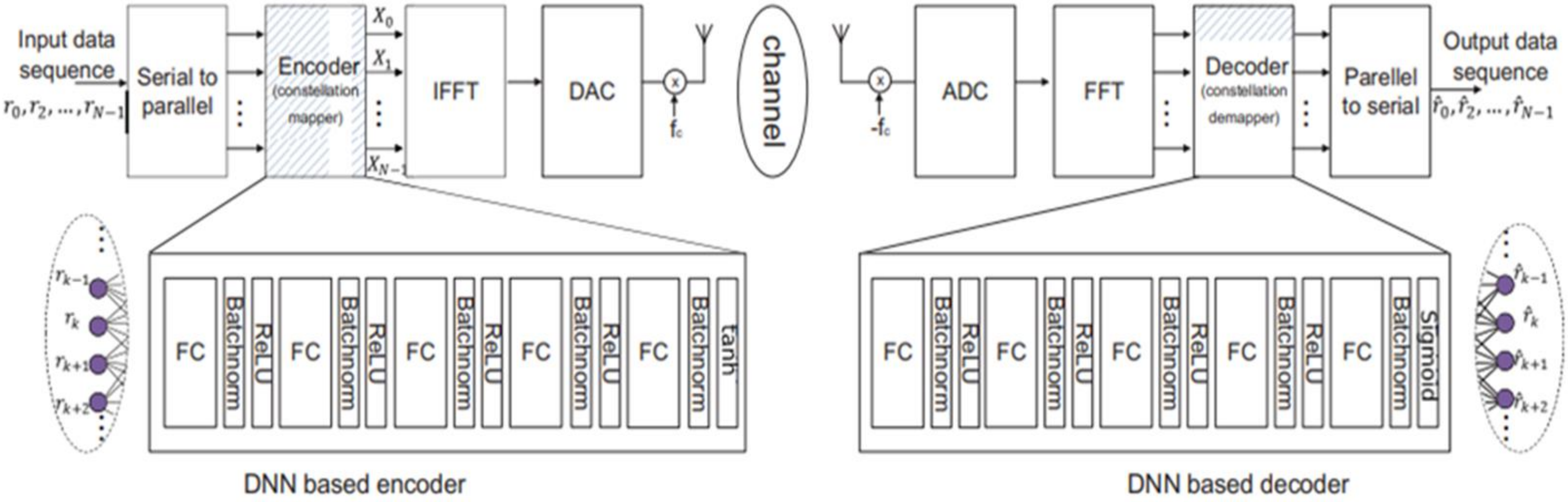


Fig. 13 PRNet system model [13]

System Model

- PAPR is reduced based on the Autoencoder Deep Learning Architecture.
- The Constellation mapping and demapping of symbols on each subcarrier is determined adaptively through a deep learning technique, such that both Bit error rate and PAPR are jointly minimized.
- The Loss function used in the Encoder is considered to minimize the PAPR.
- The Loss function used in the Decoder is considered to minimize the BER.

Joint Loss

- The Loss function to minimize the PAPR is:

$$L_1(r) = PAPR\left(IFFT\big(f(r)\big)\right)$$

- Loss function used to recover the transmitted signal from the received one is the binary cross-entropy (log-loss) function:

$$L_2(r,\hat{r}\,) = -\frac{1}{N}\sum_{i=1}^{N} y_i\,.\log\big(p(\hat{r})\big) + (1-\hat{r}).\log\big(1\ -\ p(\hat{r})\big)$$

$$\hat{r} = \ \left\| r - g(FFT(\mathbb{H} \star IFFT\big(f(r)\big) + \in) \right\|$$

Where r is input signal ,∈ is the AWGN and $\mathbb{H}$ is the Channel response.

- The Joint Loss function can be written as:

$$L_{joint}(r,\hat{r}) = L_2(r,\hat{r}) + \lambda\, L_1(r)$$

- λ is the weight parameter that determines which Loss is dominant.

Channel Estimation

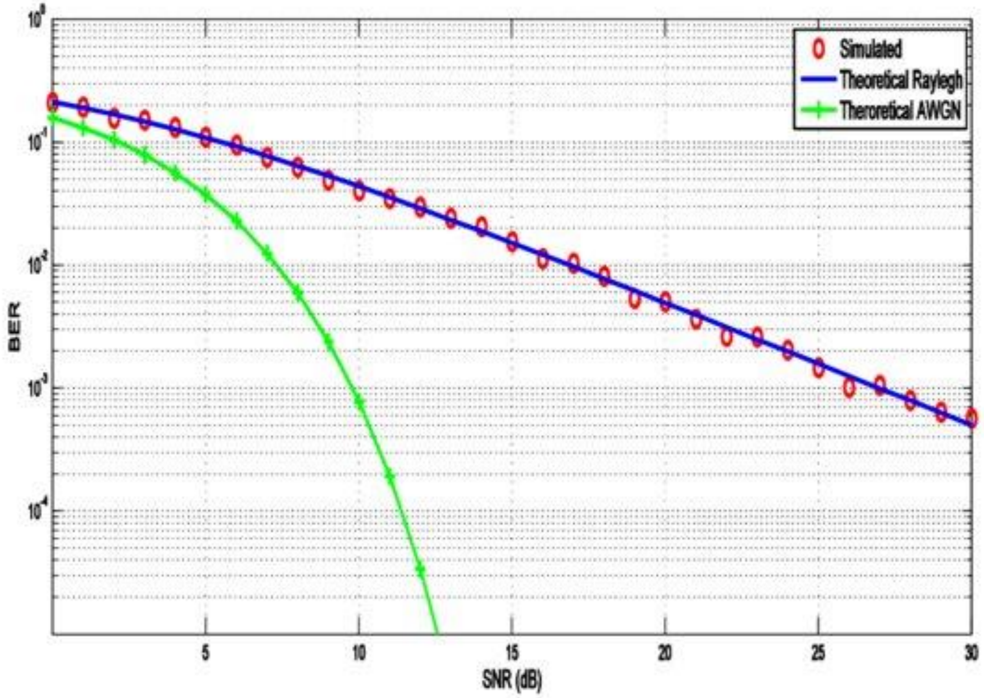
- First, the signal is first affected by channel ($\mathbb{H}$) : $\boldsymbol{R} = \mathbb{H} * \boldsymbol{Signal}$
- Then, AWGN noise ($\in$) is added by the receiver: $\boldsymbol{R}' = \mathbb{H} * \boldsymbol{Signal} + \in$
- Then, normally the equalizer step is applied here:

  We assume the equalizer is ideal and managed to capture the exact fading characteristics

- **Note:** In our approach we assumed the channel remains static throughout the entire symbol duration

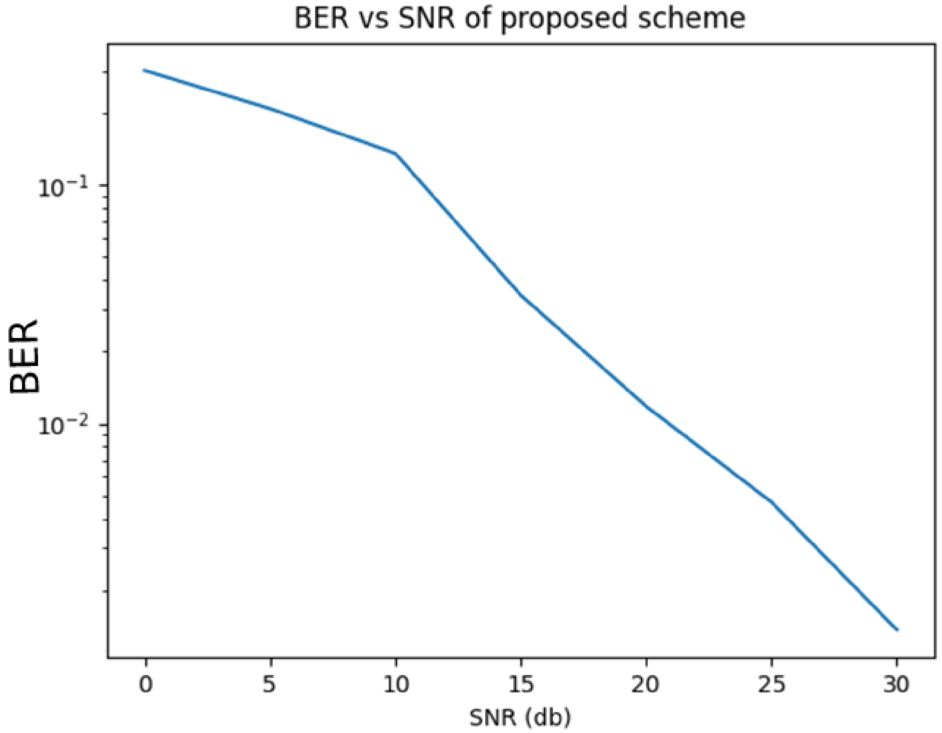
$$\boldsymbol{R}_{\boldsymbol{Equalized}} = \frac{\boldsymbol{R}'}{\mathbb{H}} = \boldsymbol{Signal} + \frac{\in}{\mathbb{H}}$$

Channel Comparison



Rayleigh Channel effect on QPSK

Source: Ubiquitous HealthCare in Wireless Body Area Networks



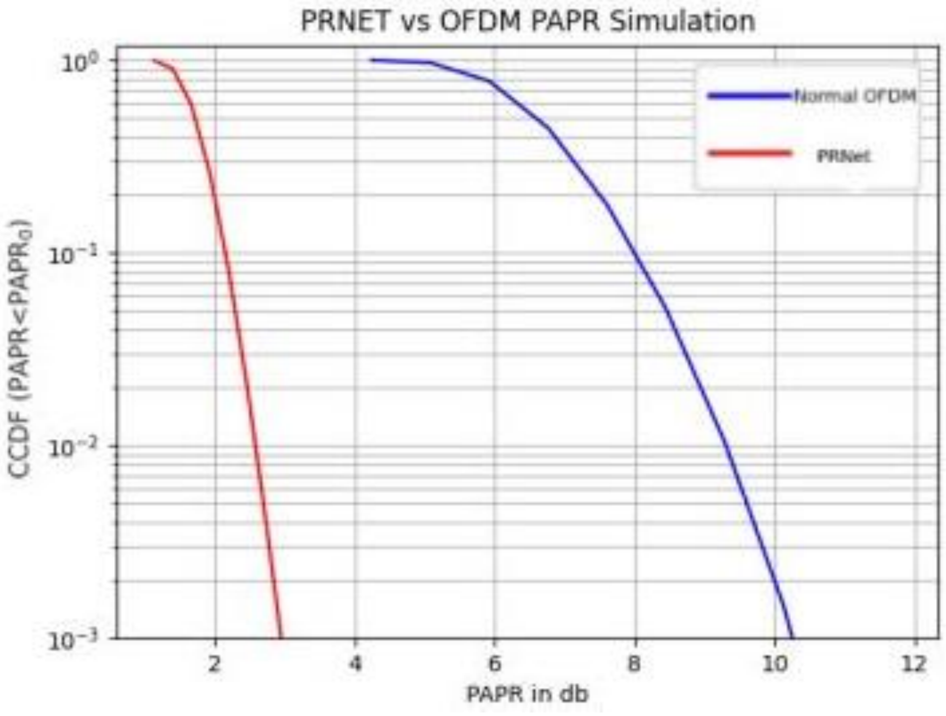
Rayleigh Channel

Our Custom Lambda Layer

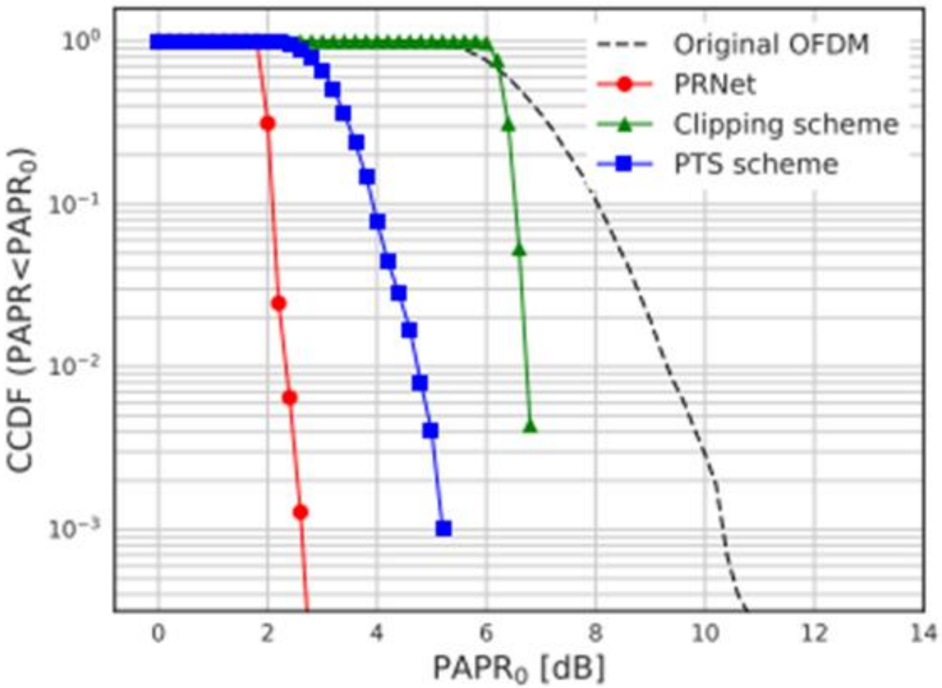
As we can see it exhibits more or less the same effect

(difference is in decoder non-linearity and training parameters effects)

Simulation Results



Our trained model simulations
Lambda = 0.002, Decoder accuracy = 0.9966



The Proposed IEEE paper results

Model Complexity:	**N = 64 (100K Symbols)**	**N= 64 (Each Symbol)**
Simulation Time	1.5291564 Seconds	15.292 µSeconds

Conclusion

- We developed two models based on Deep Learning.
- TRNet based on Tone Reservation method.
- PRNet based on Autoencoder architecture.
- The Deep Learning models gave better performance on many different aspects including greatly improved PAPR reduction and faster execution time over the classical PAPR reduction techniques.

Future Work

- We consider solving the high signal power that heavily affected the performance of TRNet.
- It would be interesting to develop less complex model with the same performance in case of the PRNet.
- We also consider applying these techniques on 5G.

References

[1] Cho, Y. S., Kim, J., Yang, W. Y., & Kang, C. G. (2010). *MIMO-OFDM wireless communications with MATLAB*. John Wiley & Sons.

[2] https://drive.google.com/drive/folders/1JECwZtMUq19K3MlQJEHvm1_j_b7toNLK?fbclid=IwAR2eFElJ-QSMAwZGsR17OZ3RQ5cx_SyA_rFWEW0F4-qI5ZlHHXiQ4bcDS94

[3] Molisch, Andreas F. *Wireless communications*. John Wiley & Sons, 2012.

[4] Li, Ye Geoffrey, and Gordon L. Stuber, eds. *Orthogonal frequency division multiplexing for wireless communications*. Springer Science & Business Media, 2006.

[5] B. Horvath and B. Botlik, "Optimization of Tone Reservation-Based PAPR Reduction for OFDM Systems," RADIOENGINEERING, vol. 26, no. 3, pp. 791-797, 2017.

[6] M. Mrou´e, et al., "Performance and Implementation Evaluation of TR PAPR Reduction Methods for DVB-T2," International Journal of Digital Multimedia Broadcasting, vol. 2010, pp. 1-10, 2010.

[7] Tone Reservation method for PAPR Reduction scheme, IEEE 802.16e-03/60r1

[8] Tone Reservation Using Near-Optimal Peak Reduction Tone Set Selection Algorithm for PAPR Reduction in OFDM Systems Jung-Chieh Chen, Member, IEEE, and Chih-Peng Li, Member, IEEE.

[9] SCR-based Tone Reservation Schemes With Fast Convergence for PAPR Reduction in OFDM System, DOI 10.1109/LWC.2018.2890596, IEEE Wireless communication letters.

[10] An Algorithm for PAPR Reduction by SLM Technique in OFDM with Hadamard Matrix Row Factor, 2nd Int'l Conf. on Electrical Engineering and Information & Communication Technology (ICEEICT) 2015.

[11] https://www.csie.ntu.edu.tw/~hsinmu/courses/_media/wn_11fall/ofdm_new.pdf

[12] A Novel Tone Reservation Scheme Based on Deep Learning for PAPR Reduction in OFDM Systems

[13] A Novel PAPR Reduction Scheme for OFDM System based on Deep Learning

[14] A. F. Molisch., Wireless communications, California, USA: A John Wiley and Sons, Ltd., Publication, 2011.

THANK YOU !