PUCRS

ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
MESTRADO EM ENGENHARIA ELÉTRICA

MATEUS PRAUCHNER ENRICONI

BEAMFORMING WITH PHASE TRANSMITTANCE COMPLEX VALUED RBF NEURAL NETWORK FOR STATIC AND DYNAMIC SYSTEMS

Porto Alegre

2018

PÓS-GRADUAÇÃO - *STRICTO SENSU*

Pontifícia Universidade Católica
do Rio Grande do Sul

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL

SCHOOL OF TECHNOLOGY

GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

MATEUS PRAUCHNER ENRICONI

BEAMFORMING WITH PHASE TRANSMITTANCE COMPLEX VALUED RBF NEURAL NETWORK FOR STATIC AND DYNAMIC SYSTEMS

Advisor: Prof. Fernando César Comparsi de Castro, Ph.D.

Porto Alegre – RS, Brazil

2018

MATEUS PRAUCHNER ENRICONI

BEAMFORMING WITH PHASE TRANSMITTANCE COMPLEX VALUED RBF NEURAL NETWORK FOR STATIC AND DYNAMIC SYSTEMS

Dissertation presented to the Graduate Program in Electrical Engineering of the Pontifícia Universidade Católica do Rio Grande do Sul, as requisite to obtain Master's degree in Electrical Engineering.

Concentration Area: Signals, Systems and Information Technology.

Research Area: Telecommunications.

ADVISOR: Prof. Fernando César Comparsi de Castro, Ph.D.

Porto Alegre – RS, Brazil

2018

Ficha Catalográfica

Pontifícia Universidade Católica do Rio Grande do Sul
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA - PPGEE

BEAMFORMING WITH PHASE TRANSMITTANCE COMPLEX VALUED RBF NEURAL NETWORK FOR STATIC AND DYNAMIC SYSTEMS

CANDIDATO: MATEUS PRAUCHNER ENRICONI

Esta Dissertação de Mestrado foi julgada para obtenção do título de MESTRE EM ENGENHARIA ELÉTRICA e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Pontifícia Universidade Católica do Rio Grande do Sul.
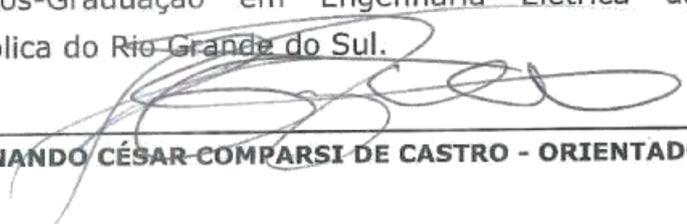
DR. FERNANDO CÉSAR COMPARSI DE CASTRO - ORIENTADOR

BANCA EXAMINADORA
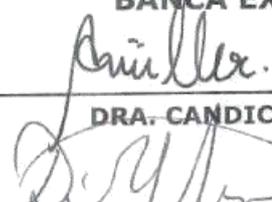
DRA. CANDICE MÜLLER - UFSM
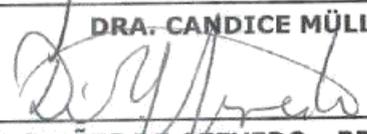
DR. DARIO F. GUIMARÃES DE AZEVEDO - PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA - PUCRS

PUCRS

This page is intentionally left blank

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Fernando César Comparsi de Castro, not only for teaching me a great deal of engineering, but also for teaching me how to attack problems and discover interesting questions. For allowing me to wander off in other directions with delightful discussion, for the subtle guidance and for making me see that we engineers are only limited by the laws of the man up there, and not by the laws of the men down here. Thanks also to my spiritual advisors Maria Cristina Felipetto de Castro and Candice Muller, who have always been willing to help me.

I am also grateful to my lab fellows for the discussions, insights, paradoxes and for building an extremely pleasant and suitable environment from which we could uncover some of the mysteries that have haunted humanity until now. I hope that this kind of spirit stays with us always.

As any great project in life, starting is the most difficult step. My special thanks to those who showed me the way.

Last but surely not least, I would like to thank my parents, who always encouraged and supported me, especially my mother Lisante Tainete Prauchner, who cared for me at every moment. This thesis is dedicated to you.

"All things by immortal power,
Near or far,
Hiddenly
To each other linked are,
That thou canst not stir a flower
Without troubling of a star..."
(Francis Thompson)

ABSTRACT

In a world of insatiable demand for data, in a limited spectrum environment, wireless communications are increasingly operating under dynamic conditions, not only regarding information traffic parameters but also regarding the time varying conditions on the propagation channel that conveys the information between the transmitter (TX) and the receiver (RX). In this context, TX and RX need dynamically to adapt its operational parameters in order to obtain maximum data transmission efficiency. Smart antennas and beamforming techniques have an essential role on this dynamic operational environment. Such antennas are arranged on arrays and are based on adaptive systems, making them capable of generating any radiation pattern when the array comprises a sufficient number of electromagnetic irradiators. This thesis proposes the implementation of a novel beamforming technique, based on a complex radial basis function artificial neural network which presents phase transmittance between the input nodes and the output node (PT-RBF). The PT-RBF is capable of adaptively adjusting the radiation pattern of a smart antenna through a learning process based on the steepest descent algorithm. The proposed beamforming technique presents significantly superior results when compared with state-of-the-art algorithms presented in literature, making it possible to operate communication links under static scenarios on self-organizing wireless networks, and in dynamic scenarios with access in motion, both with multiple interferences, thus maximizing the throughput and the spectrum efficiency.

Keywords: Beamforming, RBF, Smart Antenna, Phased-Transmittance Radial Basis Function

RESUMO

Em um mundo de crescente demanda por dados e com um espectro limitado, sistemas de comunicações sem fio operam cada vez mais em condições dinâmicas, não apenas em relação às informações de tráfego, mas também em relação às condições variáveis no canal de propagação que transmite as informações entre transmissor (TX) e o receptor (RX). Neste contexto, TX e RX precisam adaptar dinamicamente seus parâmetros operacionais em prol da máxima eficiência na transmissão de dados. Antenas inteligentes e técnicas de *beamforming* desempenham um papel fundamental neste ambiente operacional dinâmico. Distribuídas em arranjos e com operação alicerçada em sistemas adaptativos, tais antenas podem gerar qualquer diagrama de irradiação quando utilizados um número suficiente de irradiadores eletromagnéticos. Este trabalho propõe a implementação de uma nova técnica de *beamforming* baseada em uma rede neural artificial de base radial complexa com transmitância de fase (PT-RBF) entre os nós de entrada e saída da rede. A PT-RBF é capaz de ajustar de forma adaptativa o diagrama de irradiação de uma antena inteligente através de aprendizado baseado no algoritmo *steepest descent*. A nova técnica de *beamforming* proposta apresenta resultados significativamente superiores em comparação com o estado da arte, possibilitando links de comunicação em cenários estáticos em redes auto organizáveis e em cenários dinâmicos em acessos em movimento, ambos com múltiplas interferências, maximizando assim o *throughput* de dados e a eficiência do uso do espectro.

Palavras-chave: *Beamforming*, *RBF*, *Smart Antennas*, *Phase Transmittance Radial Basis Function*

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

DOA	Direction of Arrival
ANN	Artificial Neural Network
IoT	Internet of Things
IIoT	Industrial-Internet of Things
RBF	Radial Basis Function
EM	Electromagnetic
RF	Radio Frequency
UCA	Uniform Circular Array
EMF	Electro-Motive Force
DSP	Digital Signal Processing
QAM	Quadrature Amplitude Modulation
SCTE	Society of Cable Telecommunications Engineers
SDMA	Spatial Division Multiple Access
SER	Signal to Noise Ratio
SIR	Signal to Interference Ratio
AI	Artificial Intelligence
SD	Steepest Descent
LMS	Least Mean Square
C-RBF	Complex Radial Basis Function
FC-RBF	Fully-Complex Radial Basis Function
PT-RBF	Phased-Transmittance Radial Basis Function
MSEA	Mean Squared Error Average

CONTENTS

1 INTRODUCTION

Antennas and adaptive processes are as old as the physiological processes of simple life forms of our planet. Maybe the first antenna concept was extracted from the animals – some don't have eyes, ears or tongue, but have light sensors that forms images, vibration sensors that translates sounds, air or small particles, motion sensors in order to find food, among other examples. Not rarely, physiological processes for reception and transmission of information observed in nature have been mimicked into cutting-edge technological applications. Certain families of ophidians, for example, collect information through electromagnetic sensors in the infrared frequency range, called pit organs. These sensors are located at both sides of the head, towards the back and they work as two antennas, combined to the cognitive-neural ophidian system, allowing it to set up an adaptive guiding system and establish the route of the moving prey through its radiated heat - a techique that has been adopted in defense systems based on thermal target signatures.

Another equally interesting example is the gastropod orientation system. In order to locomote toward an odor source, an animal must first detect the scent concentration gradient produced by the source. Then there are two possible ways of moving up the gradient. In the first method, the animal senses the direction of the gradient by simultaneously sampling the environment at two or more places using an array of olfactory organs (tentacles, or antennas). In the second method, the neural-cognitive system of the animal uses temporal integration rather than spatial integration, making successive assessments of the chemical concentration and comparing the concentration determined at one moment with that determined at a subsequent time. If the neural-cognitive system of the animal senses that the concentration is increasing, then it will continue moving in the same direction; otherwise, if the concentrarion is decreasing, it will chance its direction (GUTIÉRREZ; MARCO, 2009).

The first method demands from the gastropods an intelligent distribution of its antennas. The signal present in the environment impinges on each antenna array element and the spatial distribution, the separation and the size of each element give the gastropods the ability of finding the most suitable boresight (main lobe direction of a antenna). Thus, the gastropod performs a sensoring method in order to obtain direction of arrival (DOA) estimates of the desired signal.

However, the second method demands from the gastropods an iterative process of sampling and comparing the samples which, in turn, accomplish by the animal's neural network. Although the gastropods do not exactly have a brain, its neural-cognitive system is incorporated as a nervous system composed by nervous nodes that cluster the neurons

responsible for this decision making. At each instant, a sample of the environment is obtain and compared with a sample obtained at a subsequent time. Such comparison generates an error which indicates to the gastropod how close the odor source is. Therefore, these animals can adapt its behaviour and store information in its neural-cognitive system.

The gastropods' orientation system is a great example of adaptive processing for determining the DOA through a neural network, artificial or not, based on the information coming from the sensor antennas.

Ahead the current trends on telecommunication area, towards an increasing demand for services and data, man try to develop its technology by understanding how nature work. The need for a system capable of focusing its radiation pattern adaptively and conveniently at some points of interest is quite high once it makes the process of transmitting and receiving more efficient and independent from other possible simultaneous links. To make full use of the RF spectrum resource and improve system performance and channel capacity, intelligent antenna technology has been widely introduced to mobile communication. Radar, sonar, seismology, radio astronomy, speech, and biomedicine, besides wireless communications are some examples of its applications. Systems like 5g and internet of things (IoT) are being developed including principles of intelligent antennas.

Two classic problems solved by intelligent antennas techniques are the DOA estimation and the beamforming. The goal of DOA is to estimate the directions of the signals from the desired users as well as the directions of interference signals. The results of DOA estimation are used to adjust the weights of an adaptive beamformer so that the radiated power is maximized towards the desired users, and radiation nulls are placed in the directions of interference signals. Hence, adaptive antenna array technology enables the antenna characteristics to be dynamically altered aiming to improve the system performance, steering the energy to an interest direction and minimizing the interference caused by other users under static or dynamic scenarios.

Beamforming is a well-known interference-suppression technique as military systems have been using such concepts for a while. Beamforming techniques result an efficient use of spectrum resources allowing multiple simultaneous transmissions using the same frequency without mutual interference, enabling an abundant spectrum reuse even when high intensity signals are needed at the receivers of network users located in a bad propagation environment. By suppressing radiation in directions into which there is no user receivers, and boosting radiation in specific directions of interest, beamforming techniques not only increases the coverage range area but also reduce the necessary transmitter power consumption. It also

allows to establish multipath links, in which the signals of each respective path are spatially combined in order to maximize the signal-interference ratio (HAGHIGHI; HOSSEIN, 2016). Finally, once the beamforming suppresses interference and noise, it also increases the throughput at the network nodes (which is limited by interference from neighboring nodes), particularly for higher order modulations.

As adaptive beamforming is basically an adaptation problem, Artificial Neural Networks (ANN) are an method to perform this procedure since they are based on non-linear adaptive filter theory and, just as the biological neural-cognitive system, have the capacity to assume and store knowledge from data analysis obtained in its operational environment, learning through the information from a desired approximation. Also, ANNs can handle non-linear behavior of physical quantities, being able to grasp the underlying random process of the data environment, presenting an extreme flexible and powerful performance over dynamic scenarios, learning fast from complex patterns and adapting itself equally quick.

Many algorithms have been proposed to solve the adaptive problem of beamforming. In (GODARA, 2009), a comprehensive overview of beamforming schemes and several adaptive algorithms are presented. A complete overview of neural methods applied to the adaptive beamforming problem is given by (DU et al., 2002). Beamforming can be implemented by the Sample Matrix Inversion algorithm (VAN TREES, 2002) method. However, in practical case, when dealing with a large array, matrix inversion is computationally expensive and mostly not implementable. (YUANJIAN; XIAOHUI, 2016) showed that the variable weights of a beamforming can be adjusted by a simple adaptive technique based on the least-mean-squares (LMS) algorithm. It showed a fast convergence rate and good tracking accuracy but the LMS algorithm uses a noisy estimate of the required gradient to adaptively estimate the weights of an optimal antenna array. Although the LMS approach performs well at linear scenarios, the estimation of weights is not accurate for non-linear scenarios. In (ZOOGHBY; CHRISTODOULOU; GEORGIOPOULOS, 1998), a Radial Basis Function (RBF) network is used in the computation of the optimum weights of a fast-tracking system, used to constantly track the users. The radiation pattern of the antenna is then adapted to place multiple narrow beams in the direction of desired users and nulls into the direction of interfering sources. In all these neural network methods, real valued networks are used. However, as will be shown in section 2.3, the beamformer must deal with complex-valued signals in order to be able to handle the phase information of an electromagnetic wave. In (HAYKIN, 2013), a Complex-valued Radial Basis Function (C-RBF) Network is used for non-linear beamforming in multiple antenna aided communication systems that employ

complex-valued quadrature phase shift keying modulation scheme. However, in the complex-valued RBF methods, the activation function is mapped on a real domain. As a result, despite the weights and centers being complex-valued, the response of each hidden neuron remains real-valued, and, thus, the hidden neurons do not effectively transmit the phase of the complex-valued signal from the input nodes to the output nodes. In (SAVITHA; SURESH; SUNDARARAJAN, 2009) a Fully Complex-valued Radial Basis Function (FC-RBF) network, using a hyperbolic secant activation function has been proposed. The FC-RBF has been applied on adaptive beamforming over a Uniform Linear Array (ULA) of 5 elements, presenting good results and a fast convergence rate for many scenarios, however, it fails on highly noisy scenarios.

In this context, this dissertation proposes a novel beamforming approach based on a Phase-Transmittance Radial Basis Function (PT-RBF) ANN. Different from most solutions available in literature, the PT-RBF ANN is able to handle data represented by complex numbers with no phase information loss on the neuron's transmittances. The PT-RBF is considered a curve fitting network for complex-values (LOSS et al., 2007). The proposed PT-RBF ANN beamforming is a non-blind algorithm that determines the channel response by a training sequence or pilot signal and adjusts the weights of the adaptive process according to certain criteria. Thus, the PT-RBF ANN process the desired and the measured values of the impinging signal and performs the adaptive control of the radiation pattern and demodulation process. Hence, the proposed solution adaptively controls the spatial distribution of the irradiated energy from the antenna array, or the spatial distribution of the received electrical field when the array is used as a receiving antenna, maximizing the electrical field magnitude at certain directions of interest and minimizing or even nulling the electrical field at other certain directions where there is no interest on receiving energy.

The main advantage of the proposed beamformer based on the PT-RBF ANN over other solutions found in literature based on ANNs, is its characteristic of minimizing the phase invariance issue. Since it does not strictly rely on the Euclidean distance between the input vector and the center vectors, it enables the transference of module and phase information from input to output. This is an important feature of the PT-RBF ANN, since all wireless mobile operation relies on carrier phase recovery, adaptive beamforming and further adaptive equalization of the complex valued modulation IQ symbols (LOSS et al., 2007) .

The proposed beamformer has been evaluated with an antenna array of six half-wavelength ($\lambda/2$) dipoles distributed along a circle at the plane xy. The Uniform Circular Array (UCA) distribution provides a good control of the boresight lobe according to the spherical

coordinate ϕ, making possible to aim the radiation pattern boresight of the array to any direction of the azimuth ϕ.

The proposed beamformer has been developed in Matlab (MATHWORKS, 2016) and the results are compared to the state of the art beamforming algorithms. The performance has been evaluated over different static and dynamic scenarios. The metrics used to evaluate the proposed approach are: (1) the interference suppression capability measured by the signal to interference ratio (SIR), (2) the SER (Symbol Error Rate) resultant from the demodulation process and (3) the MSEA (Mean Square Error Average) of the ANN. Simulations show that the proposed approach presents significantly superior results in terms of energy radiation steering and suppression and demodulation accuracy over low SIR (Signal to Interference Ratio) scenarios.

The rest of this dissertation is organized as follows. Section 0 brings forward concepts and analytical definitions used to implement the beamforming algorithm proposed in this thesis. The theoretical framework starts with Section 2.1, where the concept of antenna is developed and explored, so we can convert variations in propagating electromagnetic waves to and from conducted electrical signals. Section 2.2 allows us to travel both ways from the analog world of electrical signals to the digital world of bits. Section 2.3 is responsible for the main concept of this thesis, which is to focus all the energy transmitted and received on a link communication in order to maximize its performance and Section 2.4 transforms all the previous concepts into an intelligent system. Finally, Section 3 is where the author of this thesis presents the proposal for the new beamforming algorithm and Section 4 is where the results are shown and explained. Section 5 is then the closure, and some final remarks and an overview is presented. This Section is also directed to those who would like to continue this study and go even further than the author.

2 THEORETICAL FRAMEWORK

The purpose of Section 0 is to provide a description of the concepts needed in this thesis, providing an overview of the main contributions in the area as well. Sections 2.1, 2.2, and 2.3 consist of a brief of the major elements necessary to build this work, making it possible to follow the development of the proposed technology. As this work aims on the algorithm implementation, Section 2.4 develops the math behind the ANNs responsible for solving the beamforming problem. It should also be mentioned that parts of the theoretical framework presented in this section are based on the available literature, that include submitted journal manuscripts, books, technical reports, and conference papers, which are properly identified in the REFERENCES section. The material presented in Section 0 establish the basis in order to provide a better understanding of the state of the art, so the contributions of this thesis can be better interpreted in Sections 3, 4 and 5.

2.1 ANTENNAS

Basically, an antenna can be viewed as a matching network that takes the power from a transmission line with specific characteristic impedance and matches it to the free space impedance (377Ω). As previously mentioned, work on antennas started many years ago based on natural behavior of animals and based on the interest of demonstrating the existence of electromagnetic radiation. In 1889, a German called Heinrich Rudolf Hertz (1857 - 1894) built a system to produce and detect radio waves. Between 1893 and 1894, the Brazilian Roberto Landell de Moura, in São Paulo, transmitted human voice over 8km using electromagnetic waves. However, it is traditionally assigned to the Italian Guglielmo Marconi the invention of radio, which, in 1894, managed to ring a bell at the other side of a room by pressing a telegraph button.

The IEEE Standard Definitions of Terms for Antennas (IEEE, 2013) defines the antenna as "a means for radiating or receiving radio waves." In other words, an antenna is actually a transformer that transforms electrical signals, i.e. voltage and currents from a transmission line, into electromagnetic waves, i.e. electric and magnetic fields, or vice versa. A satellite dish antenna, for example, receives the radio wave from a satellite and transforms it into electrical signals, which are further processed. Our eyes may be viewed as another example of antennas. In this case, the wave is not a radio wave but an optical wave, another form of electromagnetic wave, which has much higher frequencies.

In this work, a receiving antenna array is employed in the beamforming system proposed. However, whatever is said about reception is also true for transmission. The subject of antennas is about how to design a suitable device which will be well matched with its feed line and radiate/receive the radio waves in an efficient and desired manner. The next sections will cover some design steps in order to achieve the desired array later used along with the ANN to perform the beamforming algorithm. First, the geometry of the dipole is defined, and the radiation intensity is calculated. Then a simple circuit model of the dipole is introduced, the coupling between elements are discussed and some parameters of this model are calculated. Finally, the dipole is used as a receiving antenna and expressions for the induced voltage from an incident electromagnetic (EM) wave are obtained.

2.1.1 The Dipole Element

The dipole antenna is one of the most important, cheap, simple and commonly used types of RF antenna. It is widely used on its own, but it is also incorporated into many other RF antenna designs where it forms the radiating or driven element for the overall antenna. Its shape is of a form of a cylinder, built in electrically conductive material in a way that a time variant electric current spatial spread along the cylinder is associated to an EM field in the space that encircles the dipole. If an external voltage generator forces a current in the dipole so that the surrounding EM field is consequence of this current, then it is said that the dipole is a transmitter. If there is an EM field at the dipole's neighborhood so that a current is induced in the dipole because of the EM field around it, then it is said that the dipole is a receiver.

An EM wave field defined in any region of space $\mathbb{R}^3$ where the electrical field $\underline{E}$ and magnetic field $\underline{H}$ both vary with time and space, thus establishing a region into with electrical field $\underline{E}$ waves and magnetic field $\underline{E}$ waves are able to propagate. Any wavefront in an EM wave field carries a true power that is proportional to the product $\underline{E} \times \underline{H}^*$ in the wavefront coordinates. This true power causes the EM wave to be radiated by the transmitter antenna, conveying information between the transmitter antenna and the receiver antenna. Figure 1 shows a transmitting dipole antenna and the EM vectors associated. Most practical dipoles have a physical size of half wavelength $(\lambda/2)$, in order to optimize the compromise between radiation efficiency and dipole size.

The time variant electrical current spatially spread along the dipole's conductive cylinder is associated to the EM field in the space that encircles the dipole through Maxwell's equations, shown at Table 1.

Table 1 – Maxwell's equations

Maxwell's Equations	Integral Form	Differential Form	
Gauss	$\iint\limits_{S} \underline{E}\cdot d\underline{S} = \dfrac{1}{\varepsilon}\iiint\limits_{V}\rho dV$	$\underline{\nabla}\cdot\underline{E} = \dfrac{\rho}{\varepsilon}$	(2.1)
Faraday	$\oint\limits_{C} \underline{E}\cdot d\underline{l} = -\mu\iint\limits_{S}\dfrac{\partial \underline{H}}{\partial t}\cdot d\underline{S}$	$\underline{\nabla}\times\underline{E} = -\mu\dfrac{\partial \underline{H}}{\partial t}$	(2.2)
Ampère	$\oint\limits_{C} \underline{H}\cdot d\underline{l} = \iint\limits_{S}(\underline{J}+\underline{J}_D)\cdot d\underline{S}$	$\underline{\nabla}\times\underline{H} = \underline{J}+\underline{J}_D = \sigma\underline{E}+\varepsilon\dfrac{\partial \underline{E}}{\partial t}$	(2.3)

Source: (AUTHOR, 2018)

where ε is the electrical permittivity of the medium $[F/m]$, μ is the magnetic permeability of the medium $[H/m]$, ρ stands for electric charge density $[C/m^3]$, σ is the conductivity of the medium [℧], and $\underline{J}$ and $\underline{J}_D$ are the conduction current density and the displacement current density respectively.

By interpreting equations of Table 1, it is possible to see that the EM radiation process involves a periodic chain reaction where originally a time variant conduction current $\underline{J}$ generates a time variant magnetic field $\underline{H}$ on its proximity, which generates a time variant electric field $\underline{E}$ on its proximity, which in turn generates a time variant magnetic field $\underline{H}$ on its proximity and so on so forth in a way that, step by step, an EM wave keeps propagating in the space that encircles the time variant current $\underline{J}$ which originated the chain reaction.

From these equations that Maxwell wisely interpreted and organized, it is possible to infer more precisely the radiation process of an EM wave. By energizing a radiating conductor, as the one in Figure 1, with an excitation voltage $V(t)$, a current $I(t)$ manifests as consequence and flows through the conductor with a conduction current density $\underline{J}$. If the applied excitation voltage is a steady state sinusoid in the time domain, $\underline{J}, \underline{J}_D$, $\underline{E}$ and $\underline{H}$ will also exhibit a steady state sinusoidal time variation. Also, if the conductor diameter d is much smaller than its length ℓ, $\underline{J}$ will exhibit a sinusoidal variation along the conductor length. In this work we assume steady state sinusoidal excitation. In this context, the EM radiation process can be summarized as follows:

I. The flow of the sinusoidal current inside the conductor imposes an acceleration to its free electrical charges, in a way that a non-zero electric charge density ρ varies along the conductor.

II. The sum of all electrical charges inside the radiator volume V generate an electrical field $\underline{E_\rho}$ in each point of the imaginary surface S which wraps the radiator external surface (Eq.(2.1)).

III. Simultaneously, a magnetic field $\underline{H_J}$ is generated as consequence of the conduction current density $\underline{J}$, in other words, as a consequence of the sum of all conduction/displacement current lines that pass through S (Eq.(2.3)).

IV. If $\underline{J}$ varies over the time, necessarily a magnetic field $\underline{H_J}$ generated by $\underline{J}$ also varies over time. The sum of the temporal variation of the lines of $\underline{H_J}$ that pass through S induces a perturbation that gives origin to an electric field $\underline{E_H}$, also variant over time (Eq.(2.2)).

V. This new $\underline{E_H}$ sums with $\underline{E_\rho}$ giving rise to an electrical field $\underline{E}$ which, in turn, gives rise to a displacement current density $\underline{J_D}$.

VI. Once again, just as in III, a magnetic field $\underline{H_D}$ is generated as consequence of $\underline{J_D}$ Eq.(2.3).

VII. Finally, just as in IV, an electric field $\underline{E_H}$ is generated as a consequence of the time variation of $\underline{H_D}$.

From Eq.(2.2), it is possible to see that the partial time derivative of $\underline{H}$ gives origin to a spatial variation for each Cartesian component of $\underline{E}$ as a result of the vector product between $\underline{E}$ and the Nabla $\underline{\nabla}$ operator. The vector product between $\underline{E}$ and the Nabla $\underline{\nabla}$ operator in Eq. (2.2) also establishes that the fields $\underline{E}$ e $\underline{H}$ will always be perpendicular in all points of the space $\mathbb{R}^3$. Similarly, from Eq.(2.3), the partial time derivative of $\underline{E}$ gives origin to a spatial variation for each Cartesian component of $\underline{H}$ as a result of the vector product between $\underline{H}$ and the Nabla $\underline{\nabla}$ operator. The vector product between $\underline{E}$ and the Nabla $\underline{\nabla}$ operator in Eq. (2.3) also establishes that the fields $\underline{H}$ e $\underline{E}$ will always be perpendicular in all points of the space $\mathbb{R}^3$.

This chain of processes in which a time varying magnetic field $\underline{H}$ generates a time varying spatial variation of electric field $\underline{E}$, followed by the time varying electric field $\underline{E}$ generating a time varying spatial variation of magnetic field $\underline{H}$, followed by the time varying magnetic field $\underline{H}$ generating a time varying spatial variation of electric field $\underline{E}$, and so on ad infinitum, is the overall radiation process of an electromagnetic wave (SADIKU, 2018).

Thus, the analysis of a transmitter antenna basically consists of determining the values of $\underline{E}$ and $\underline{H}$ that encircles the antenna as a consequence of the time variant current spatial distribution that flows through the antenna conductive structure. Similarly, the analysis of a receiver antenna basically consists in determining the time variant current spatial distribution

that flows in the antenna conductive structure as a consequence of the $\underline{E}$ and $\underline{H}$ fields at the antenna neighborhood.

Figure 1 – For the analysis of the dipole and the irradiated fields, the spherical coordinate system is more convenient, where (r, θ, ϕ) are the orthogonal unit vectors that defines $\mathbb{R}^3$
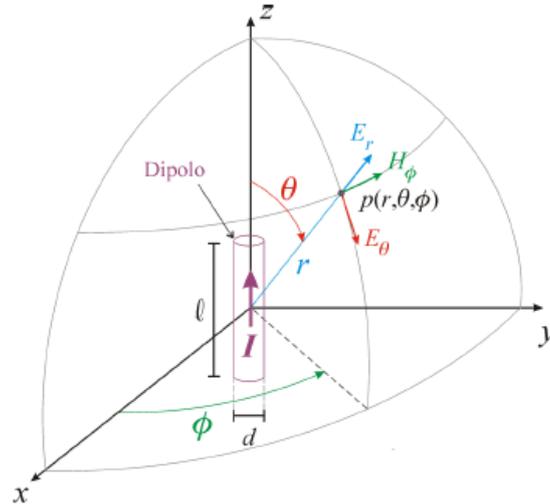
Source: (DE CASTRO; FRANCO, 2005).

From Maxwell's equations shown at Table 1, for the specific case of a dipole with a length ℓ, operating under sinusoidal steady state with maximum instantaneous current I_0 flowing in the dipole cylinder, the relation between $\underline{E}$ and $\underline{H}$ at a point $p(r, \theta, \phi)$ in the space that encircles the dipole and the spatial distribution of the time variant current are given by Eq. (2.4) and (2.5). These equations are derived by using vector potentials and can be found in many antenna textbooks (BALANIS, 2016).

$$E_\theta = \frac{I_0 60}{r} e^{j\left(\omega t - \beta r + \frac{\pi}{2}\right)} \left\{ \frac{\cos\left(\frac{L}{\lambda} \pi \cos\theta\right) - \cos\left(\frac{\pi L}{\lambda}\right)}{\sin\theta} \right\} \tag{2.4}$$

$$H_\phi = \frac{E_\theta}{Z_0} \tag{2.5}$$

where E_θ is the component of $\underline{E}$ along the unit vector of direction θ of the spherical coordinate system and H_ϕ is the component of $\underline{H}$ along the unit vector of direction ϕ of the spherical coordinate system. In Eq. (2.4), $I_0 [A]$ is the current phasor representing the phase and the maximum instantaneous magnitude of the current that flows through the cylinder of the dipole, $r [m]$ is the distance between the dipole and the point $p(r, \theta, \phi)$ where E_θ and H_ϕ are being determined, $L [m]$ is the dipole length, $\beta [rad/m]$ is the wave propagation constant, $\lambda [m]$ is the wavelength and $Z_0 [\Omega]$ is the free space impedance. Equations (2.4) and (2.5) are valid for

$r > 10\lambda$, i.e., point $p(r, \theta, \phi)$ must lie in the dipole farfield region (DE CASTRO; FRANCO, 2005).

2.1.1.1 <u>Radiation Pattern</u>

One of the most important characteristics of an antenna is the directional properties of the radiated energy, i.e. the radiation intensity pattern which is obtained from the radiated field of the antenna (SVANTESSON, 1999). In order to calculate the radiated field, the current distribution must be known. Given a voltage $V(t) = V_0 \cos(2\pi f t)$ applied to the dipole, where V_0 is the instantaneous voltage and f is the voltage frequency, the resulting current is therefore, $I(t) = V_0 \cos(2\pi f t + \phi)/R$. The current distribution along the axis z' of the dipole can be approximated by

$$\underline{I}\left(z^{'}\right) = \underline{z} I_0 \sin\left[\beta\left(\frac{L}{2} - \left| z^{'} \right| \right) \right] \tag{2.6}$$

where $-L/2 \leq z' \leq L/2$, L is the dipole length and β is the propagation constant.

The radiation pattern $F(\theta, \phi)$ is the value of E_θ from Eq. (2.4) normalized by its maximum value obtained in all possible directions (θ, ϕ) of space $\mathbb{R}^3$. The normalized $E_\theta(\theta, \phi)$ 3D surface is the dipole's radiation diagram $F(\theta, \phi)$ plot. It is used to analyze the spatial magnitude distribution of field E_θ. An important spatial parameter of any antenna is the boresight direction (θ_b, ϕ_b), which is the direction in the $F(\theta, \phi)$ plot that occurs the maximum E_θ magnitude.

Eq. (2.4) is independent of angle ϕ, i.e., it is valid for $0° < \phi \leq 360°$. Thus, for the specific case of a cylindrical dipole in the z axis, the 3D surface $F(\theta, \phi)$ reduces to the surface $F(\theta)$ defined by the ratio between E_θ and the maximum value $E_{\theta\,máx}$, with angle θ varying in the range $0° < \theta \leq 180°$:

$$F\left(\theta\right) = \frac{E_\theta}{E_{\theta\max}} \tag{2.7}$$

Since Eq. (2.7) is valid for $0° < \phi \leq 360°$, the magnitude of the radiation pattern $F(\theta)$ is omni-directional, i.e., it is non-directional in the azimuth plane, although being directional in the elevation plane. Figure 2 shows half of the radiation pattern 3D surface for a thin dipole of length $\lambda/2$, aligned with the z-axis and with center of the dipole located in the origin of the Cartesian coordinate axes.

Figure 2 – Radiation pattern of a half wavelength dipole
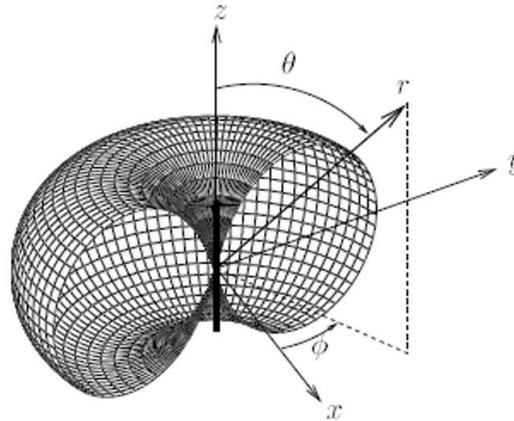
Source: (SVANTESSON, 1999).

Most dipole antennas are cylindrical, with length $\lambda/2$ and with cylinder diameter being much smaller than its length. Thus, for an excitation voltage $V(t) = V_0\cos(2\pi ft + \phi)$ applied at the dipole center, the spatial distribution of the current along the dipole cylinder will be sinusoidal with maximum magnitude at the center of the dipole and zero magnitude at the dipole ends. It is important to note that the radiated field depends directly on the current and not directly on the applied voltage, so, in order to perform a transmitting beamformer, a control over the current on each dipole of the array is necessary.

2.1.2 Dipole Array

Until now, a single dipole element has been discussed and fundamental properties have been introduced. In this section, several dipoles are assembled together in an array of dipoles, also known as a phased array. Figure 3 shows the effects of using a different number of elements for a linear array of half have dipoles parallel to the z axis and distributed along the y axis. Each dipole is spaced from each other by a distance of $\lambda/4$.

Figure 3 – Radiation pattern of a linear dipole array with (a) 1 dipole; (b) 2 dipoles; (c) 3 dipoles; (d) 4 dipoles

(a)   (b)

(c)   (d)
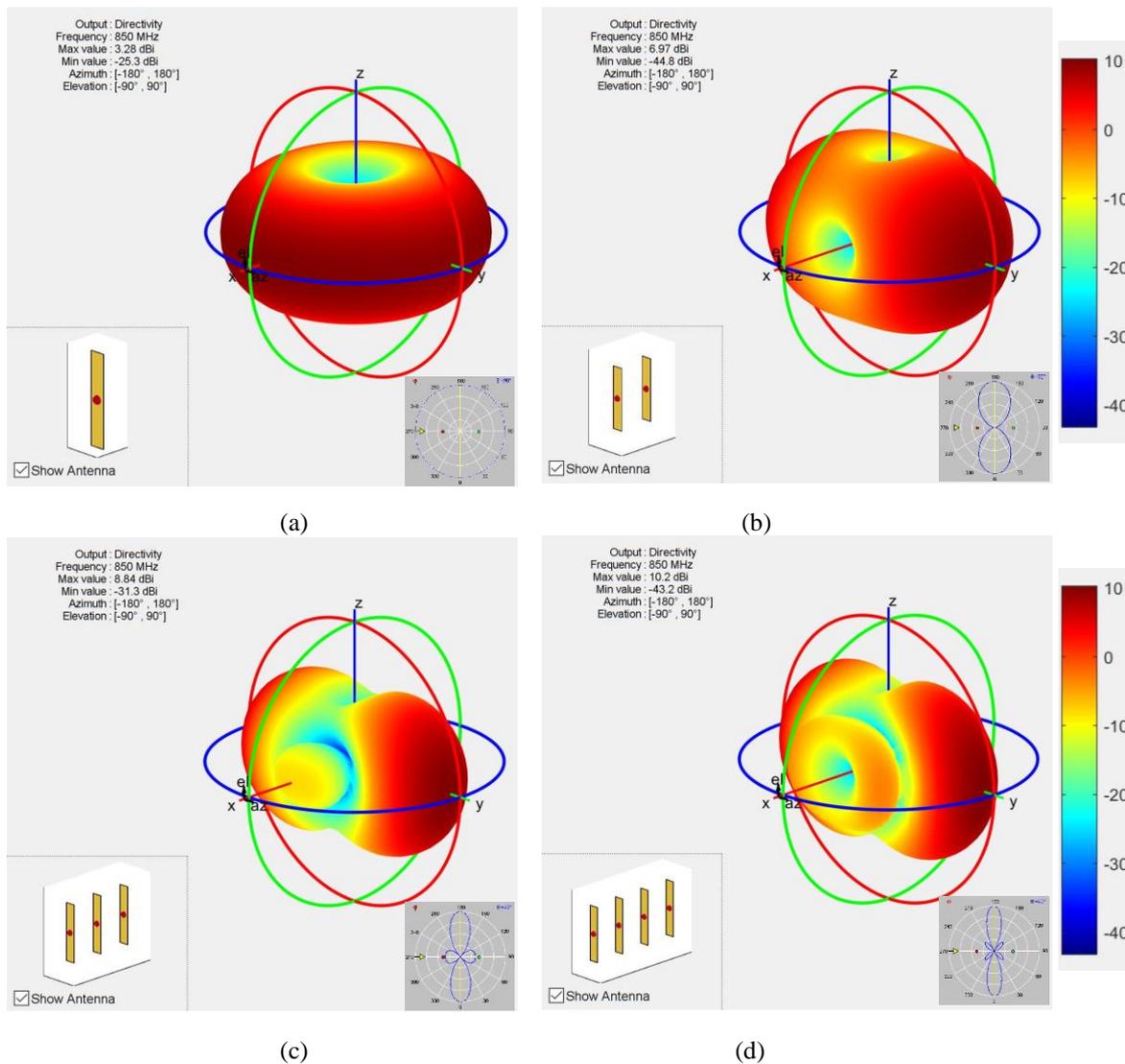
Source: (AUTHOR, 2018).

Just as a Fourier series, which makes possible to form any wave shape in a period by only summing a sufficient number of magnitudes and phases of sines and cosines functions, an antenna array can form any radiation pattern by summing the individual electric field of each individual element (dipole). The principle of the phased array is to synthesize a specified electric field across an aperture. In this work, the individual antennas are spaced by $s = \lambda/4$, or a quarter wavelength apart. Adding a phase shift to the signal received or transmitted by each antenna in an array allows the superimposed signals of all individual antennas to act as the signal of a single antenna with performance vastly different from the individual antennas in the array.

By arraying K antennas, it is possible to achieve a series of results:

I. The signal resulting from a receiving array is the properly combined signals from all individual antennas, and so, the resulting signal presents a greater magnitude than any individual signal.

II. Due to the same superposition principle, the radiation pattern of a transmitting array can be much narrower than the radiation pattern of any individual antenna of the array.

III. The direction of the peak sensitivity of a receiving array can be altered without mechanically re-positioning the individual antenna elements. For an array with a combiner able to deal with the signal magnitude and phase (electronically variable phase shifters), it is possible to alter the beam direction according to the combined gains and phase shifts. Same is valid for a transmitting array, for which the magnitude and phase of each individual current is controlled.

IV. For a single antenna, if the mechanical positioning system fails, it is not possible to point to anything except to the boresight of the antenna. For an antenna array, if one antenna element fails, all the rest continue to work, and the collective pattern is modified slightly.

V. For airborne applications the weight of a phased array is less than that of a comparable rapidly-steerable, mechanically gimbaled, single antenna.

VI. A very large mechanically steered antenna may be replaced with a collection of less expensive smaller antennas without losing resolution.

VII. By using the wide range of control provided by the phase shifters, it is possible to synthesize multiple beam responses if desired.

VIII. Individual antenna performs the same even when surrounded by other antennas when considering the mutual coupling between the elements. In this case, the array design must consider such influences of mutual coupling.

In other words, individually, each kth dipole of the array exhibits a specific radiation pattern. However, an array with a sufficient number of K dipoles, where the kth array dipole, $k = 1, \dots, K$ is located at coordinates (x_k, y_k, z_k) and fed by the current phasor I_k, generates an interference pattern given by the sum of all K fields in each point $p(r, \theta, \phi)$ of space $\mathbb{R}^3$, allowing to set up any desired radiation diagram just by tuning the modules and phases of the currents I_k.

The parameters of a dipole array that affect the generated wave interference pattern are the center coordinates (x_k, y_k, z_k) of each kth dipole and the magnitude and phase of the current I_k that feed the respective k^{th} dipole. In this context, from (2.4), the field E_θ at a point $p(r, \theta, \phi)$ of the farfield zone become a superposition of individual electric fields E_θ of the individual dipoles of the array. Such superposition of individual electric fields E_θ, results in equation (2.8).

$$E_\theta = \sum_{k=1}^{K} \frac{I_k\, 60}{r} e^{j\left(\omega t - \beta r + \frac{\pi}{2}\right)} \left\{ \frac{\cos\left(\dfrac{L}{\lambda}\pi\cos\theta - \cos\dfrac{\pi L}{\pi}\right)}{\sin\theta} \right\} e^{\frac{j2\pi}{\lambda}\left(x_k \sin(\theta)\cos(\phi) + y_k \sin(\theta)\sin(\phi) + z_k \cos(\theta)\right)} \quad (2.8)$$

where x_k, y_k e z_k are the cartesian coordinates of the kth array element and θ, ϕ e r are the spherical coordinates of point p.

Therefore, given the magnitudes and phases of the currents I_k , $k = 1, \dots, K$, in a array of K dipoles parallel to the z axis, each kth dipole center located at coordinates (x_k, y_k, z_k) , Eq. (2.8) determines the radiation pattern $F(\theta, \phi)$ for the E_θ electrical field component.

The spatial distribution of the elements of an antenna array can be spherical, linear, rectangular, hexagonal, planar, concentric, cylindrical or can be of arbitrary geometry. In this work we adopt a Uniform Circular Array (UCA) with six half-wave dipoles parallel to the z axis, which is a classic and popular array for azimuthal steering (IOANNIDES; BALANIS, 2005). It is a phased array whose main lobe can be electrically steerable by controlling magnitude and phase of the dipoles current. The six dipoles are equally spaced around a circle of radius s, as shown in Figure 4.

Figure 4 – Architecture of a six elements UCA
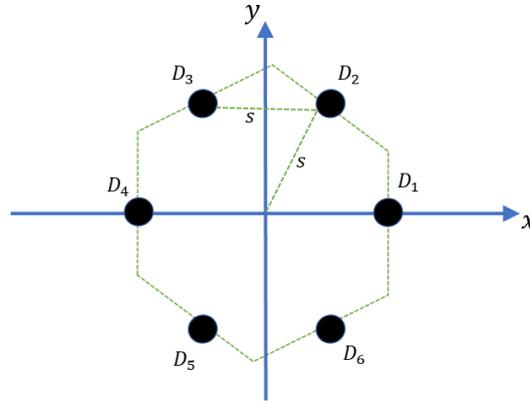
Source: (AUTHOR, 2018).

In Figure 4, the dipoles center lies in the xy-plane of the xyz coordinate system, whose origin lies at the center of the circle of radius s. The azimuth angle is defined as the angle, in the xy-plane, from the x-axis toward the y-axis. The elevation angle is defined as the angle from the xy-plane toward the z-axis. Elevation angles for all array elements are zero, (IOANNIDES; BALANIS, 2005). The center coordinates (x_k, y_k, z_k) of each kth dipole of the UCA, $k = 1, \dots, K$, with $K = 6$, can be summarized in the matrix form by Eq. (2.9).

$$XYZ = \begin{bmatrix} s & 0 & 0 \\ s\cos(\alpha) & s\sin(\alpha) & 0 \\ -s\cos(\alpha) & s\sin(\alpha) & 0 \\ -s & 0 & 0 \\ -s\cos(\alpha) & -s\sin(\alpha) & 0 \\ s\cos(\alpha) & -s\sin(\alpha) & 0 \end{bmatrix} \qquad (2.9)$$

where, in this case of a six elements UCA, $\alpha = 60°$ is the angle between the dipoles.

2.1.2.1 <u>Mutual Coupling and Circuit Model</u>

In the majority of transmitting or receiving dipole arrays, each dipole lies in the near field region of all other dipoles. Thus, in the near field region, each dipole is magnetically coupled to all other dipoles, resulting in a mutual impedance between each pair of dipoles. Therefore, the radiation pattern for both transmitting or receiving dipole arrays is affected by the mutual coupling. Lets analyze the mutual coupling for a receiving dipole array. The analysis is also valid for transmitting arrays, with minor adaptations, as we shall see.

If an EM wave impinges upon a dipole array from the (θ, ϕ) direction, the wave will reach each dipole at different instants of time. By measuring the induced voltages on each $k = 1, \dots, K$ dipole of the array, the DOA can be estimated using these time instants, assuming a signal of known frequency. However, the total induced voltage on each dipole will generate a current on the dipole, which in turn radiates an EM field, which affects the surrounding dipoles, inducing additional voltages. The induced voltage at dipole k, V_{ind_k}, consists of a direct induced voltage from the wave V_{wave_k} and a voltage induced from the neighboring dipoles due to the mutual impedance coupling (SVANTESSON, 2001), as further shown in Eq. (2.15).

Figure 5 represent the circuit of dipole k. The applied voltage is denoted V_g, which is zero, since it is a receiving array. The induced voltage V_{ind} is the total voltage (considering the mutual coupling to the other dipoles) over the dipole k. Z_T is the network terminating impedance, usually 50Ω (given the commercial standard). Z_A is the dipole self-impedance, determined without any other dipoles in the neighborhood, so that there is no nearfield coupling as a consequence of the mutual impedance. There will be as many circuits identical to the circuit of Figure 5 as there are dipoles in the receiving/transmitting array. The dipole self-impedance Z_A is mostly calculated on the literature (KUN-CHOU LEE; TAH-HSIUNG CHU, 1995) using the concept of induced Electro-Motive Force (EMF). The derivation for the self-impedance can be found in most antenna handbooks like (POZAR, 2012).

Figure 5 – Schematic diagram of a terminating or generating network and antenna and equivalent circuit
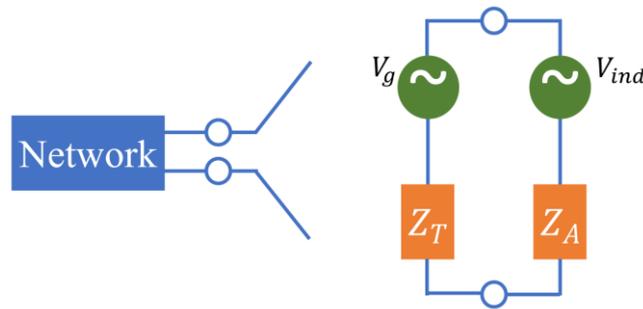
Source: (AUTHOR, 2018).

Another classical method for the determination of the self-impedance Z_A of each dipole in an array is the Pocklington method (POCKLINGTON, 1897), which establishes that the real part of Z_A can be found by integrating the real part of the Poynting vector $\underline{S}$ over a spherical surface on the farfield region. Note that this is just the law of conservation of energy being followed - "In a closed system, i.e., a system isolated from its surroundings, the total energy of the system is conserved." The imaginary part of Z_A is found by integrating the imaginary part of $\underline{S}$ over a surface that closely wraps the volume of the dipole, keeping an infinitesimal distance from the closed surface that delimits the volume.

There are several other methods for the determination of the self-impedance of dipoles that can be found in classic antennas textbooks (STUTZMAN; GARY, 2013).

As for the mutual coupling, each dipole couples its magnetic field to all other dipoles in the nearfield region, which results in a mutual impedance between each pair of dipoles in the array. The current in dipole j induces a voltage in the other dipole k which is proportional to the mutual impedance Z_{kj} between dipole j and dipole k. Thus, the resulting current in each dipole is dependent upon the voltages induced by the currents in all other dipoles, therefore, affecting the radiation pattern for both transmitting or receiving dipole arrays. Thus, it is important to account for the coupling correctly. A known coupling can even increase the performance of DOA estimation if compensated correctly.

For a transmitting array, each dipole k in the array is connected to a voltage source V_g. This voltage gives rise to a current which in turn results in an EM field that propagates, as described in Section 2.1.1 according to Maxwell equations. But the magnetic field from other dipoles will induce a respective voltage in dipole k as a result of the respective mutual impedance, similarly to the receiving array analysis. The dipoles are thus mutually coupled, and the mutual impedances must be taken into account in order to correctly determine the radiation pattern.

Considering Z_{kk} as the self-impedance of dipole D_k of the array and considering Z_{kj} as the mutual impedance between dipoles j and k, with $k = 1, ..., K$ and $j = 1, ..., K$, Eq. (2.10) defines mutual impedance Z_{kj} as the ratio between the induced voltage V_{kj} at dipole k due to a current in dipole j, as follows:

$$Z_{kj} = \frac{V_{kj}}{I_j} \qquad (2.10)$$

Since all dipoles in the UCA have the same length and size, it is possible to write:

$$Z_{kj} = Z_{jk} = Z_m \qquad (2.11)$$

where V_{kj} is the voltage appearing at the terminals of dipole D_k due to a current I_j in dipole D_j.

One precise analytical method for determining the mutual impedance between two symmetric dipoles was developed by (SCHELKUNOFF, 1955) and is known as the method of the biconical antenna perturbation. By expanding the unknown electromagnetic fields of a coupled system in terms of the known modes of an uncoupled system, Schelkunoff obtained a set of generalized telegrapher equations directly from Maxwell equations. The coupling coefficients are determined unambiguously once the modes of the uncoupled system are defined. The coupled mode theory remains an approximate, yet insightful and often accurate mathematical description of electromagnetic oscillation and wave propagation in a coupled system.

The mutual impedance Z_{kj} between two generic antennas has an analytical solution that usually presents a high complexity, if not intractable. For the simple case of two finite thin dipoles placed side by side, an approximate result for the mutual impedance Z_m can be found by the method of induced EMF force as follows:

$$Z_m = R + jX \qquad (2.12)$$

where

$$R = \frac{\eta}{4\pi \sin^2\left(\frac{\beta L}{2}\right)} \left[-2\left(\int_{\beta s|k-j|}^{\infty} \frac{\cos y}{y} dy \right) + \int_{\beta\left(\sqrt{s^2|k-j|^2+L^2}+L\right)}^{\infty} \frac{\cos y}{y} dy + \int_{\beta\left(\sqrt{s^2|k-j|^2+L^2}-L\right)}^{\infty} \frac{\cos y}{y} dy \right] \quad (2.13)$$

$$X = -\frac{\eta}{4\pi \sin^2\left(\frac{\beta L}{2}\right)} \left[2\left(\int_{0}^{\beta s|k-j|} \frac{\cos y}{y} dy \right) - \int_{0}^{\beta\left(\sqrt{s^2|k-j|^2+L^2}+L\right)} \frac{\cos y}{y} dy - \int_{0}^{\beta\left(\sqrt{s^2|k-j|^2+L^2}-L\right)} \frac{\cos y}{y} dy \right] \quad (2.14)$$

A more detailed approach of this mutual impedance technique can be found in (BALANIS, 2016).

Thus, as a consequence of the self and mutual impedances, the induced voltage in dipole k is given by (see

Figure 5):

$$V_{ind_k} = V_{wave_k} - \sum_{j \neq k} I_j Z_{kj} \qquad (2.15)$$

where $Z_{kj} = Z_m$ is the impedance between dipoles j and k. This equation shows that the voltage that drives the current of the dipole k is the direct voltage that results from the impingent wave minus the voltages that the respective current in other dipoles induce in dipole k due to the nearfield mutual coupling.

The direct voltage induced from the wave V_{wave_k} can be derived in a similar manner as to the mutual impedance determination, but using the incident E field, and can be determines by Eq. (2.16) (SVANTESSON, 2001).

$$V_{wave_k} = \frac{\rho E_{0_k} \lambda}{\pi \sin\left(\frac{\beta L}{2}\right)} \left[\frac{\cos\left(\frac{\beta L}{2}\cos\theta\right) - \cos\left(\frac{\beta L}{2}\right)}{\sin\theta} \right] \qquad (2.16)$$

where E_{0_k} is the field strength at dipole k and ρ is the polarization mismatch factor.

Using the circuit model in

Figure 5, Eq. (2.17) for V_{wave_k} is obtained

$$V_{wave_k} = I_k (Z_A + Z_T) + \sum_{j \neq k} I_j Z_{kj} \qquad (2.17)$$

Eq. (2.16) and (2.17) along with the circuit model from

Figure 5 form a system of equations, from which the received or measured voltage over Z_T can be obtained according to Eq. (2.18).

$$\underline{V}_T = \underline{C} H(\theta, \phi) \begin{bmatrix} 1 \\ e^{j\beta d(m,k)} \\ \vdots \\ e^{j\beta d(m,K)} \end{bmatrix} \underline{E}_0 \qquad (2.18)$$

where $\underline{E}_0$ is the magnitude and phase of the instantaneous electrical field phasor, $\underline{C}$ is the mutual impedance coupling matrix given by Eq. (2.19), and $H(\theta, \phi)$ corresponds to the spatial response of the dipole k given by Eq. (2.20).

$\underline{C}$ and $H(\theta, \phi)$ can be found and deduced in many textbooks like (BALANIS, 2016). In Eq. (2.18), the time delay between each dipole becomes a simple phase shift and it is represented by the term in square brackets, also known as the geometrical array factor.

$$\underline{C} = \frac{(Z_T + Z_A)}{(\underline{Z} + Z_T \underline{I})} \qquad (2.19)$$

$$H(\theta,\phi) = \frac{\lambda}{\pi \sin\left(\frac{\beta L}{2}\right)} \left[\frac{\cos\left(\frac{\beta L}{2}\cos\theta\right) - \cos\left(\frac{\beta L}{2}\right)}{\sin\theta}\right] \tag{2.20}$$

where $\underline{I}$ stands for the identity matrix and $\underline{Z}$ stands for the mutual impedance matrix.

$\underline{Z}$ is a matrix with all the self-impedances in the diagonal elements and the mutual impedances in the remaining elements, as shown in Eq. (2.21).

$$\underline{Z} = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1K} \\ Z_{21} & Z_{22} & \cdots & Z_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{K1} & Z_{K2} & \cdots & Z_{KK} \end{bmatrix} \tag{2.21}$$

When using an UCA of K antenna elements, K different voltages will be measured. These voltages can be shown in Eq. (2.22) on a vector form.

$$V_{ind}(t) = \begin{bmatrix} V_{ind_1}(t) \\ V_{ind_2}(t) \\ \vdots \\ V_{ind_K}(t) \end{bmatrix} = \begin{bmatrix} H_1(\theta,\phi)e^{j\beta d(m,1)} \\ H_2(\theta,\phi)e^{j\beta d(m,2)} \\ \vdots \\ H_K(\theta,\phi)e^{j\beta d(m,K)} \end{bmatrix} E_0(t) = \underline{\Phi}(\theta,\phi)E_0(t) \tag{2.22}$$

where d is the distance from the phase reference plane to the center of dipole k, and finally, vector $\underline{\Phi}$ models the spatial response of the array due to an incident EM wave from (θ,ϕ) direction. The phase reference plane contains the point $p(0,0,0)$ located at the origin of the Cartesian axes and it is perpendicular to the line between the center of dipole k and point $p(0,0,0)$.

Vector $\underline{\Phi}$ is fundamentally important for this thesis. It is the geometrical array factor combined with the array response and is known as steering vector. Eq. (2.22) can finally be expressed by Eq. (2.23).

$$\underline{V_T} = \underline{C}\Phi(\theta,\phi)E_0(t) \tag{2.23}$$

The concept of steering vector is further discussed in Section 2.3.

2.2 MODULATION

Modulation is the process of transforming a signal into a more suitable form so it can be transmitted over a particular channel (PROAKIS; SALEHI, 2001).

Usually, modulation involves two waveforms, a modulating signal that represents the message to be transmitted and a sinusoidal carrier signal with a specific frequency f_c. The purpose of the modulation process is to shift the information bearing message signal whose

spectrum is centered around the frequency zero, called the baseband signal, to a signal whose spectrum centered around the carrier frequency f_c. A modulator is an important stage of any transmitter for wireless communication systems. The output signal of the modulator, referred to as the modulated signal, is fed into the antenna for the transmission of the modulated EM wave across the communication channel. The resulting modulated EM wave thereby carries the message information. It is generally required that modulation be a reversible operation, so the message can be retrieved by the complementary process of demodulation in the receiver.

Shifting the message spectrum to the carrier frequency f_c is necessary because many channels, both wired and wireless, can convey only a narrow range of frequencies centered in some frequency. Also, the size of the antenna is proportional to the wavelength $\lambda = c/f_c$ of the sinusoidal carrier wave of frequency f_c. Therefore, a carrier frequency f_c much higher than the center frequency of the message signal spectrum, or the baseband signal, is also required to keep the size of the antenna at an acceptable limit. Finally, for digital message cases, the message signal is a block of discrete samples over time. Thus, it is necessary to represent these blocks under a continuous time signal by means of a modulation process, so that the transmission can be performed over the communication channel.

As an example, incidentally, humans act as a modulator whenever they speak. The transmission of voice through air is accomplished by generating carrier tones in the vocal cords and modulating these tones with muscular actions of the oral cavity. Thus, what the ear hears as speech is a modulated acoustic wave similar to an Amplitude Modulation (AM) signal (FARUQUE, 2017).

The information to be transmitted can be either analog or digital, where the carrier is a high frequency sinusoidal waveform. For digital signals, as the one used in this thesis, there are several modulation techniques available. The three main digital modulation techniques are the Amplitude shift keying (ASK), the Frequency shift keying (FSK) and the Phase shift keying (PSK). With the exception of audio broadcasting that still use analog Amplitude Modulation (AM) and Frequency Modulation (FM), most modern communications systems use digital modulation. The advantages of digital modulation include:

I. Energy efficiency: better audio or video quality with the same or lower power consumption;

II. Bandwidth efficiency: more users can be supported in the same bandwidth;

III. Flexibility: digital modulation can transmit both data and digitized waveforms of any type;

IV. Security: data can be effectively encrypted.

Modern communication devices are implemented using digital signal processing (DSP) techniques both at the transmitter and at the receiver. The DSP algorithms at the receiver present a higher computational complexity than the algorithms at the transmitter. Any digital receiver has an analog radiofrequency (RF) front end with a low noise amplifier (LNA) and a mixer, which convert the analog received RF signal to an intermediate frequency (IF). Then an A/D converter samples and digitizes the IF signal, followed by a digital downconverter which brings the digital IF signal spectrum to the baseband. The baseband signal is then processed by a DSP processor to recover the original transmitted data. A reverse process is performed at the digital transmitter.

When any device at the analog RF front end is not linear, it generates multiple intermodulation products spread not only in the modulated signal spectrum band (i.e., in-band) but also in the adjacent spectrum bands. The intermodulation products of third and fifth order are particularly insidious since they are responsible for the in-band interference, which interfere with the desired signal causing distortion and increasing the receiver bit error rate (BER).

A largely adopted parameter to characterize the non-linear relationship between input and output of an analog RF device is the intercept point (IPn) (LI et al., 2012). Determining the IPn of third (IP3) and fifth (IP5) order are enough to characterize the operation dynamic range of such devices since the third order intermodulation products and the fifth order intermodulation products cause significant in-band interference, i.e., distorting the modulated signal. The others high order intermodulation products vanish rapidly in the off-band spectrum and can be neglected.

This thesis adopts the Quadrature Amplitude Modulation (QAM), since it is quite usual in wideband wireless systems. Specifically, in order to simulate a real information transmission, we adopt a 16-QAM digital modulation, which is described in the next section.

2.2.1 Quadrature Amplitude Modulation 16QAM

16-QAM controls the amplitude of two sinusoidal carriers of frequency f_c with phase shift of 90° between them, so that the sum of the amplitude modulated in-phase carrier with the amplitude modulated quadrature carrier results in a sinusoidal modulated output signal of frequency f_c with amplitudes and phases that are a function of the input message. The modulating signal (input message) is a four bits binary word, with two of the bits controlling the amplitude of the in-phase carrier and with the other two bits controlling the amplitude of the quadrature carrier (HANZO; WEBB; KELLER, 2000).

It is interesting to note that the emerging fifth generation (5G) wireless access network, aiming at providing ubiquitous and high data rate connectivity, is envisaged to deploy large number of base stations with higher density and smaller sizes, where inter-cell interference (ICI) becomes a critical problem. Derivations of the QAM modulation have been shown to reduce the ICI at the cell edge therefore achieve a higher transmission rate for cell edge users (WU et al., 2016).

The in-phase carrier is denoted I signal, and the quadrature carrier is denoted Q signal. Mathematically, the in-phase carrier can be represented by a cosine waveform, and the quadrature carrier can be represented by a sine waveform.

Thus, QAM is expressed by an alphabet set of I and Q symbols, called IQ symbols. The IQ symbols represent the different discrete levels of amplitude and phase shifts. Figure 6 shows the mapping between each binary 4-bit word and the in-phase (I) and quadrature (Q) components of each one of the 16 IQ symbols. This graphic representation of the mapping shown in Figure 6 is called the constellation of the digital modulation. As $M = 16$ symbols are used in 16-QAM, each symbol can express a binary word of $\log_2 M = 4$ bits. In order to minimize the error probability of the demodulation process, the most usual way is to use the Gray code, mapping the neighborhood symbols of the QAM constellation with binary words different only by $n = 1$ bit.

Figure 6 – 16QAM constellation, which is the IQ symbols' graph. This constellation shows 12 possible phases and 3 possible amplitudes.
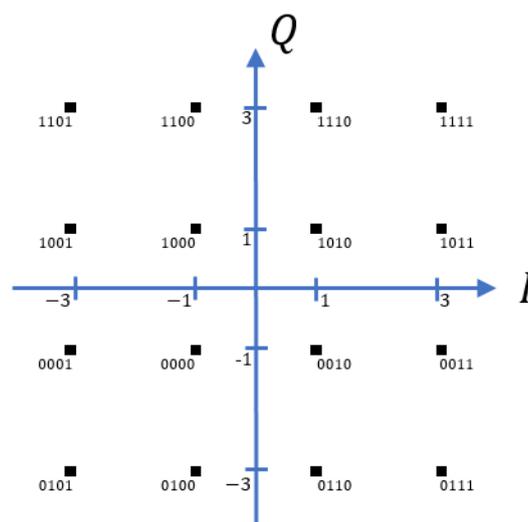
A QAM modulator has mainly 4 stages:

I.Mapping: A block of n bits is transformed on a complex representation with magnitude and phase corresponding to a symbol, as presented in Figure 6;

II. Filtering: The complex baseband I and Q signals are filtered by a low-pass filter which performs the spectral containment of both signals. The low-pass filter, called shaping filter (WU et al., 2016), smooths the edge of the I and Q pulses in the time domain, minimizing the spectral spurious components intrinsically generated by the abrupt transition of the pulse's edges;

III. Multiplying (up-conversion): Both, real I and imaginary Q signals, are multiplied by a carrier with angular frequency $\omega_c = 2\pi f_c$;

IV. Adding: Both signals are then added, resulting on the modulated signal $x(t)$.

$$x(t) = \left[\sum_{k=-\infty}^{\infty} Re\{a_m(kT)\}\sqrt{2}g(t - kT)\right] cos(\omega_c t)$$
$$- \left[\sum_{k=-\infty}^{\infty} Im\{a_m(kT)\}\sqrt{2}g(t - kT)\right] sin(\omega_c t)$$

(2.24)

where $Re\{\cdot\}$ and $Im\{\cdot\}$ return the real and imaginary part of the argument $\{\cdot\}$ respectively, $a_m = Re\{a_m\} + Im\{a_m\} = I_m + jQ_m$ is the complex value of the mth IQ symbol, k is the time index for the discrete sequence of IQ symbols, T is the IQ symbol duration and $g(\cdot)$ is the shaping filter impulse response.

The terms in square brackets represent the convolution between the shaping filter impulse response $g(t)$ with the sequence of pulses a_m. Note, from Eq. (2.24), that the in-phase carrier represented by the cosine waveform with amplitude proportional to $Re\{a_m\}$, and the quadrature carrier represented by the sine waveform with amplitude proportional to $Im\{a_m\}$ are subtracted from each other (or added with a minus signal multiplying the quadrature carrier). It can be shown, by simple trigonometric relationships, that the resulting modulated signal $x(t)$ is a cosine waveform of frequency $\omega_c = 2\pi f_c$, with amplitude given by $|a_m| = \sqrt{(I_m)^2 + (Q_m)^2}$ and with phase given by $\sphericalangle\{a_m\} = atan\left(\frac{Q_m}{I_m}\right)$ (HANZO; WEBB; KELLER, 2000).

The modulated signal $x(t)$ is then transmitted across the transmission channel, being corrupted by additive Gaussian noise. The noise in the channel adds to the real and imaginary part of the transmitted symbol so that each received IQ symbol differs from the original transmitted one that was picked up from the reference constellation as a function of the input binary word. However, if the noise power is small compared to the signal power, there is a high probability of the received symbols to be close to the constellation reference symbols and a Gaussian cluster of received symbols are formed around the reference symbols, as shown in Figure 7

Figure 7 – 16QAM constellation with a random level of additive noise
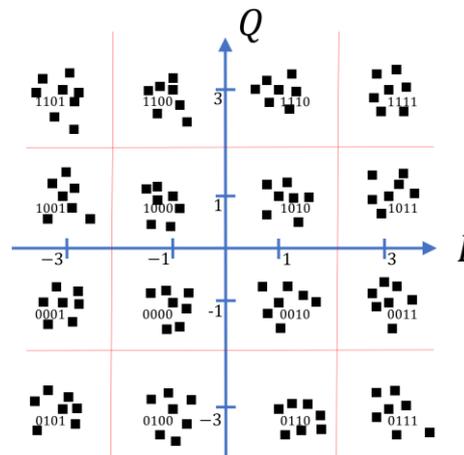
Source: (AUTHOR, 2018).

Higher constellation densities (higher number of IQ symbols) allow the increasing of the data throughput, since more bits are sent per transmitted IQ symbol. On the other hand, denser constellations are more sensitive to noise and interference, since the constellation symbols are closer to each other and a small shift in the symbol phase or in the symbol amplitude can push the symbol into a decision region which is adjacent to the region respective to the original transmitted IQ symbol, which may increase the BER to unacceptable levels. A decision region respective to a specific IQ symbol is any of the squares delimited by the red lines in

Figure 7, with the IQ symbol located at the center of the square.

64 QAM and 256 QAM are often used in digital cable television and cable modem applications. Note that a coaxial cable is a wired channel with a much lower noise level than the noise level in a wireless channel. In the UK, 16 QAM and 64 QAM are currently used for digital terrestrial television using Digital Video Broadcasting. In the US, 64 QAM and 256 QAM are mandatory modulation schemes for digital cable as standardized by the Society of Cable Telecommunications Engineers (SCTE).

2.3 BEAMFORMING

Beamforming is a technique largely adopted for the precise beam steering of smart antennas. Deployed at the base station of the existing wireless infrastructure, smart antennas can bring outstanding system capacity improvement, particularly in urban and densely populated areas, providing an efficient reuse of frequency. This unique feature has been made feasible through the impressive advances in the field of digital signal processing in the past few

years, which enable smart antennas to dynamically tune out interference while focusing on the intended user. Nevertheless, the development of AI, as described in Section 2.4, grants the beamforming to go even further, allowing self-optimizing networks (SON) to adapt to different scenarios, without any human intervention.

With the direction-finding ability of smart antennas, value-added services, such as position location services for an emergency call, fraud detection, intelligent transportation systems, law enforcement, accident reporting, etc., have been developed and improved (BLAUNSTEIN; CHRISTODOULOU, 2014; IOANNIDES; BALANIS, 2005). Smart antennas are also deployed in ad hoc networks, mobile networks or wireless local-area networks (WLANs), for example, with mobile terminals (notebooks, PDAs, etc.) in a wireless network. The direction-finding ability supports the design of the packet-routing protocol, which decides the way packets are relayed. The beamforming or interference-suppression ability makes possible to increase the throughput at the network nodes, which is limited by interference from neighboring nodes.

Mobile users, for example, continue to demonstrate an insatiable demand for data. The next decade will experience the full impact of next generation 5G cellular networks with gigabit throughput, low-latency, and connectivity to billions of devices (IoT). The explosive demand for wireless access will surpass the data transfer capacity of existing broadband links. The transition from analog to digital communications marks the boundary between 1G and 2G. 3G introduced CDMA techniques for significant spectral efficiency gains. 4G LTE is nearly at the theoretical limits of wireless resource utilization, leaving little room for significant network improvement where frequency and time are the only free dimensions. Beamforming techniques will open space for these improvements for 5G. Figure 8 express such scenario (HAGHIGHI; HOSSEIN, 2016).

Figure 8 – Mobile technology evolution.

Source: (BLACK, 2016).

The wireless world is working feverishly to open a new degree of freedom (space) so the network capacity and the performance for all the services mentioned before can grow explosively. Actively leveraging the spatial dimension is frequently called Space Division Multiple Access (SDMA).

As the very name suggests, beamforming is an algorithm that aims to establish high intensity radiation lobes in the radiation pattern of an antenna array at directions (θ, ϕ) which it is desirable "illuminate" by means of the EM radiation of the array. A beamforming system also minimize the EM radiation intensity at directions (θ, ϕ) in which is desirable "to leave at dark" (POZAR, 2012). This SDMA process performed by the beamformer maximizes the signal intensity at the desired reception points, spatially separating the signal of system users so that one does not interfere in each other, allowing multiple transmissions sharing the same frequency. The outcome is the rise of the signal to noise ratio (SNR), allowing more bits per Hertz to be transmitted until reaching the Shannon's limit. This approach can be applied to wireless links of all kinds.

Different algorithms have been proposed for the estimation of the directions of arrival (DOAs) of signals arriving at the array (SVANTESSON, 1999), and several adaptive techniques (BLAUNSTEIN; CHRISTODOULOU, 2014) have been examined for shaping the radiation pattern under different constraints imposed by the wireless environment. A key component of smart antenna technology is the adaptive beamforming, which simultaneously places the maximum beam lobe on the antenna radiation pattern towards the intended user or signal of interest, and ideally places nulls toward directions of interfering signals or signals that are not of interest. In this work, it is proposed an adaptive beamforming for an uniform circular array. The performance of the proposed algorithm built upon a uniform circular array is compared with three others beamforming algorithms which are references in the literature, already mentioned in the previous sections.

Current cellular and some internet systems utilize antennas that form static 60 to 90-degree sector beams to spread energy. By contrast, beamforming permits a more focused communications protocol between base station and user, as illustrates Figure 9. Highly directive beamforming enhances the SIR (signal-to-interference ratio) of the communication channel. Signal strength is increased by focusing power at the intended receiver. Furthermore, beamforming hardens the channel by suppressing multipath components. Interference and noise are reduced by minimizing the angular field of view. Interfering signals from other channel systems outside the main beam are attenuated by beamforming. Wireless operators can benefit from the SIR improvement in three ways: increase of coverage range, increase of

throughput by using higher order modulation schema (16 QAM to 64 QAM), and/or reduced transmit power. Beamforming preserves the spectral orthogonality of the sector and allows multiple concurrent transmissions using the same frequency without interference as shown in Figure 9, thus allowing for abundant spectrum reuse with higher intensity signals delivered to both stationary and mobile users.

Figure 9 – Basestation with no beamforming (low gain, low capacity, low spectral efficiency) and basestation with beamforming (maximized throughput, increased spectral efficiency)

Source: (BLACK, 2016).

In a dynamic scenario, wireless operators can continuously reuse the same band of spectrum, at the same time, within a given spatial region or refocus the beam of one underutilized cell site reaching out to provide extra capacity to another cell site that is saturated or impaired (BLACK, 2016). In 4G networks, for example, optimal coverage patterns for a cell are rarely sector-shaped and never static due to urban canyons, commuter traffic and special events. Dynamic beamforming allows a mobile operator to manipulate the shape of coverage in azimuth and elevation and to reposition it electronically. This is crucial for small cells, which are often mounted on municipal furniture that can twist and sway in the wind.

The beamforming algorithm employs a set with M sensors of field E_θ, each sensor being a receiving dipole with polarization lined up to the E_θ direction. These sensors are spatially distributed and located at fixed coordinates at which is desired specific values of E_θ, representing a way of informing the beamforming algorithm about the values of the E_θ field radiated by the array into direction (θ_m, ϕ_m) of each mth sensor. The E_θ field value measured by each mth sensor at direction (θ_m, ϕ_m) is compared with the desired E_θ at direction (θ_m, ϕ_m) so that the resulting error signal can be processed by the adaptive beamformer algorithm. Note that the M sensors are in the farfield region of the array, i.e, they are spatially distributed in points over a spherical surface of radius $r > 10\lambda$ in whose center is found the array. For simulation purposes, the values of the E_θ field radiated by the array into direction (θ_m, ϕ_m) of each mth sensor are given by the radiation pattern $F(\theta_m, \phi_m)$ calculated at

direction (θ_m, ϕ_m) for each mth sensor. The radiation pattern $F(\theta_m, \phi_m)$ is obtained by normalizing Eq. (2.8) by the term $\frac{60}{r} e^{j(\omega t - \beta r + \frac{\pi}{2})}$, resulting on Eq. (2.25).

$$F_m = \sum_{k=1}^{K} I_k \left\{ \frac{\cos\left(\frac{L}{\lambda} \pi \cos\theta_m\right) - \cos\left(\frac{\pi L}{\lambda}\right)}{\sin\theta_m} \right\} e^{j\frac{2\pi}{\lambda}\left(x_k \sin(\theta_m)\cos(\phi_m) + y_k \sin(\theta_m)\sin(\phi_m) + z_k \cos(\theta_m)\right)} \quad (2.25)$$

The distance d in the direction (θ_m, ϕ_m) of the mth sensor that separate the kth dipole located at coordinates (x_k, y_k, z_k) from the plane ψ of zero phase is given by Eq. (2.26).

$$d(m,k) = \left(x_k \sin(\theta_m)\cos(\phi_m) + y_k \sin(\theta_m)\sin(\phi_m) + z_k \cos(\theta_m) \right) \quad (2.26)$$

The plane ψ of zero phase is the plane that contains the origin of the Cartesian system being perpendicular to the unit vector that defines the direction of the radius r that starts at the origin of the Cartesian system and points at the direction (θ_m, ϕ_m) of the mth sensor. The plane ψ defines the geometric phase reference for the beamforming algorithm.

From Eq. (2.25), let (2.27) be the relative intensity of the radiated E_θ field by a single dipole of the array and measured at the coordinates of the mth sensor.

$$\chi_m = \left\{ \frac{\cos\left(\frac{L}{\lambda} \pi \cos\theta_m\right) - \cos\left(\frac{\pi L}{\lambda}\right)}{\sin\theta_m} \right\} \quad (2.27)$$

From these definitions, the radiation diagram F_m value, at the corresponding direction (θ_m, ϕ_m) of the mth sensor is given by

$$F_m = \sum_{k=1}^{K} I_k \chi_m e^{j\frac{2\pi}{\lambda}d(m,k)} = \sum_{k=1}^{K} I_k \Phi_{m,k} = \underline{I} \cdot \underline{\Phi}_m \quad (2.28)$$

where

$$\Phi_{m,k} = \chi_m e^{j\frac{2\pi}{\lambda}d(m,k)} \quad (2.29)$$

is the phasor that represents the magnitude and phase of the EM wave radiated by the kth dipole and that propagates along direction (θ_m, ϕ_m) over the vector radius r at the direction of the mth sensor. $\underline{I} = [I_1, I_2, \cdots, I_K]^T$ is the vector which defines the corresponding currents of the K dipoles. $\underline{\Phi}_m = [\chi_m e^{j\frac{2\pi}{\lambda}d(m,1)}, \chi_m e^{j\frac{2\pi}{\lambda}d(m,2)}, \cdots, \chi_m e^{j\frac{2\pi}{\lambda}d(m,K)}]^T$ is named the steering vector, a vector whose each component expresses the respective magnitude and phase of the EM wave radiated by the kth dipole along the direction (θ_m, ϕ_m) of the propagation path, at the end of

which the wave impinges the mth sensor, where $m = 1, ..., M$ and , $k = 1, ..., K$. (DE CASTRO; FRANCO, 2005).

The beamforming algorithm iteratively adjusts the current phasors $\underline{I} = [I_1, I_2, \cdots, I_K]^T$ in Eq. (2.28). Given a set of M steering vectors $\underline{\Phi}_m$, each one referring to the direction (θ_m, ϕ_m) of the mth sensor, Eq. (2.28) yields a set of M values of normalized E_θ field, i.e., it yields a set of M values for F_m , $m = 1, ..., M$. When operating in real time, the set of M normalized values of the E_θ field are respective measured by the M dipoles sensors. Then, for each iteration n, are calculated M errors, each one obtained by the difference between the desired F_m at the mth sensor position and the F_m value obtained from Eq. (2.28) (or measured) at discrete time instant n. This error is then used to iteratively, iteration by iteration, to tune up the current phasors of vector $\underline{I}$. After a sufficient number of iterations, the error is minimized, and the process converges when, finally, the components of the vector $\underline{I} = [I_1, I_2, \cdots, I_K]^T$ are settled.

Cutting edge technology have just started exploring beamforming technology. On the other hand, much have been developed and published about it. Furthermore, with the advent of ANNs being employed for adaptive beamforming, an issue that turns out to be important is the selection of the appropriate ANN and the appropriate activation function for it. In this work we adopt a complex valued RBF ANN with phase transmittance between input nodes and output.

The Least Mean Square (LMS), a well-known algorithm based on the adaptive filtering proposed by (WIDROW; HOFF, 1960) showed that the variable weights of a linear combiner can be automatically adjusted by a simple adaptive technique, and (YUANJIAN; XIAOHUI, 2016) tested this algorithm and some variations for the beamforming problem. The Complex-Valued Radial Basis Function (C-RBF) was proposed by (HAYKIN, 2013) and used by (CHEN et al., 2008) for non-linear beamforming in multiple antenna aided communication systems that employ complex-valued quadrature phase shift keying modulation scheme. The Fully Complex-Valued Radial Basis Function was proposed and used by (SAVITHA; SURESH; SUNDARARAJAN, 2009) to solve an adaptive beamforming problem with symmetric steering and nulling. Finally, (LOSS et al., 2007) presented the Phase Transmittance Radial Basis Function (PT-RBF) with phase transmittance between the input and output nodes for the context of channel equalization on quadrature digital modulation. The next section analytically describes each one of the referred state-of-the-art algorithms, so that they can be implemented in order to assess their performances.

An analytical description of each ANN type is given in sections 2.4.2, 2.4.4, 2.4.5 and 2.4.6 in order to fairly evaluate and compare different beamforming algorithms, since their

authors publishes the results under different scenarios. Different scenarios mean different antennas array architecture, different number of impinging signals to suppress or steer, different complexity problems the ANNs must solve, which leads to different training sets, different learning rates and different number of neurons. On the other hand, some important assumptions can be withdrawn from the referred papers.

It will be seen in section 2.4.2, that a simple adaptive filtering architecture can solve the beamforming problem. In (YUANJIAN; XIAOHUI, 2016), an ULA of $k = 9$ elements spaced by $s = \lambda$ performs the beamforming for a desired signal located at $\phi = 15°$ and two interference signals located at $\phi = 45°$ and $75°$ with SIR $= 10dB$ and $3dB$ respectively. However, the LMS algorithm uses a noisy estimate of the required gradient to adaptively estimate the weights of an optimal antenna array. Hence the estimation of weights is not accurate. It also fails to solve more complex problems, since its architecture cannot handle non-linearity and it is real-valued, while, as will be shown in section 2.4, the beamformer must deal with complex-valued signals

Also, it will be seen in section 2.4.4 that a partially complex network can also solve the beamforming problem. In (CHEN et al., 2008), an ULA of $k = 3$ elements spaced by $s = 1/\lambda$ was employed to support four QPSK users located at $\phi = 15°, 340°, 45°$ and $290°$ with equal power and a SIR varying from $0dB$ to $20dB$. For each SNR scenarios, a training data set consisting of $N = 600$ samples was adopted. However, in these methods of partially complex-valued, the activation function maps the synapses between the neurons on a real domain. As a result, despite the weights and centers being complex-valued, the response of the hidden neuron remains real-valued, and therefore, does not transmit the complex-valued signal between the input and the output nodes effectively, losing information.

In section 2.4.5 it will be seen that a fully complex-valued network using an $sech$ activation function has been proposed and tested to solve the beamforming problem. In (SAVITHA; SURESH; SUNDARARAJAN, 2009), an ULA of $k = 5$ elements spaced by $s = \lambda/2$ was trained to look at desired signals coming from $\phi = 30°$ and $330°$, and to suppress interferences from $\phi = 15°$ and $345°$ with SNR $= 50dB$. A training data set consisting of $N = 250$ samples was adopted. A fully complex-valued network preserves the magnitude and phase information of complex-valued signals, and so, it performs approximations and classifications of complex-valued signals better than real and imaginary split complex networks. However, although presenting a fast convergence rate and a good performance result, this network fails over high noisy scenarios.

2.4 NEURAL NETWORK

Artificial neural networks (ANN) are one of the main tools used in machine learning. As their name suggests, they are brain inspired systems which are intended to replicate the way that humans process information. Thus, an ANN is capable of learning, inferring and decision making, inspired in the human brain. In the same way that people learn from experience in their lives, neural networks require data to learn. In most cases, the more data can be given to a neural network, the more accurate it will become. Just as humans, it can be understood as any task someone do over and over. Over time, this someone gradually get more efficient and make fewer mistakes.

The exact behavior of the human brain is still a mystery. Yet, some aspects of this amazing processor are known. In particular, the most basic element of the human brain is a specific type of cell that, unlike the rest of the body, does not appear to regenerate. Because this type of cell is the only part of the body that is not slowly replaced, it is assumed that these cells are what provides humans with abilities to remember, think, and apply previous experiences to every action. These cells, around 100 billion of them, are known as neurons. Each of these neurons can connect with up to 200,000 other neurons, although 1,000 to 10,000 is typical. By understanding how elements of our brain works is easy to understand the mathematical model hereafter presented.

ANNs are then composed by neurons. As any other biological cell, the neuron is covered by a thin cell membrane responsible for the electrical operation of the nerve cell. The neuron has a cell body, called soma, which is the center of the metabolic processes of the nerve cell. Various structures extend from the cell body. These include many short, branching structures, known as dendrites, and a separate structure that is typically longer than the dendrites, known as the axon, as shown in Figure 10.

Figure 10 – Neuron biological model
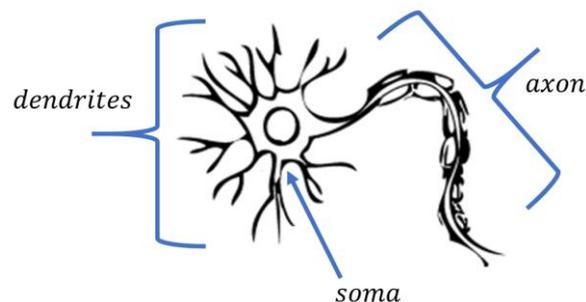
Source: (AUTHOR, 2018).

Basically, neurons have three functions. The first two neuronal functions (receiving and processing incoming information) generally take place in the dendrites and cell body. Incoming signals can be excitatory, which means they tend to make the neuron to fire up, generating an electrical impulse. Or can be inhibitory, which means that they tend to keep the neuron in an idle state (SOLOMON, 2015).

A single neuron may have more than one set of dendrites and may receive many thousands of input signals. Whether or not a neuron is excited into firing up an impulse, it depends on the sum of all of the excitatory and inhibitory signals it receives. If the neuron does fire up, the nerve impulse, or action potential, is conducted down the axon. The axon terminals are finally responsible for making connections to target cells.

Neuron-to-neuron connections are made onto the dendrites and cell bodies of other neurons. These connections, known as synapses, are the structures through which the information is carried from the first neuron to the target neuron. Figure 11shows where the synapses take place.

Figure 11 Neuron`s connection biological model
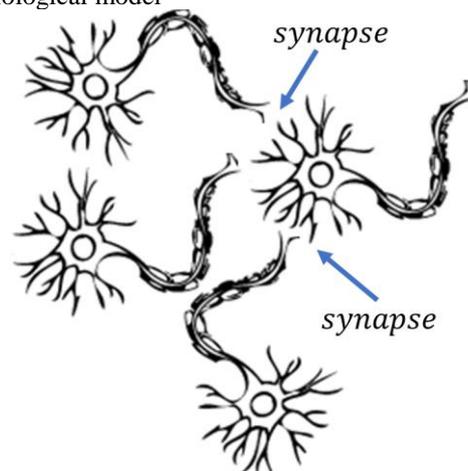
Source: (AUTHOR, 2018).

The information is transmitted in the form of neurotransmitters. When an action potential travels down an axon and reaches the axon terminal, it triggers the release of neurotransmitter from the first neuron. Neurotransmitter molecules cross the synapse and bind to membrane receptors on the target neuron, conveying an excitatory or inhibitory signal. Thus, the third basic neuronal function, which is communicating information to target cells, is carried out by the axon.

Basically, a biological neuron receives inputs from other sources, combines them in some way, performs a generally non-linear operation, and then outputs the result. The artificial analogy of the biological neuron must do the same. Figure 12 shows the diagram of an artificial

neuron. It is possible to tell apart: The dendrites, which receives the data provided to the network in the form of a vector $\underline{u}$; the soma, where the process information takes place, known as the activation function φ on the artificial models; the synapses, in a form of synaptic weights vector $\underline{w}$; and the axon, where the finished computations of the network are placed for use in a form of an output signal y.

Figure 12 - Neuron artificial model. Because our knowledge of neurons is incomplete and our computing power is limited, our models are necessarily gross idealizations of real networks of neurons
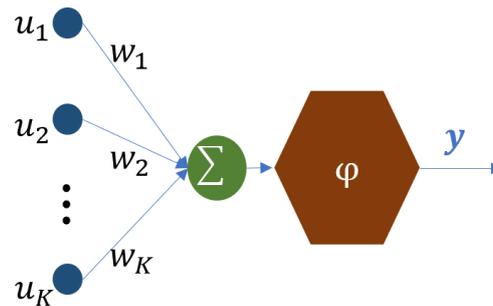
Source: (AUTHOR, 2018)

Although an ANN processes a number much smaller of neurons than the biological neural network, the digital gates of an ANN works up to six orders of magnitudes faster than the biological neuron. In mathematical terms, a neuron can be described as (2.30).

$$y = \varphi \sum_{k=1}^{K} w_k u_k \qquad (2.30)$$

By using a set of artificial neurons like the one shown in Figure 12, an ANN can be built. In the ANN model, layers are connected by transmittances called synaptic weights. As the information flows through the model, the synaptic weights and the parameters that define the activation function φ are used to store knowledge. These are the only free parameters that an ANN can modify in order to adapt and to reach a desired goal. The procedure responsible to the process of reaching the goal is known as the Learning Algorithm.

The design of an ANN includes the way the neurons are connected and the way the free parameters are adjusted. The ANN performance is deeply related to the learning algorithm used to train the network. It is also important how the interneuron connections are arranged, and the nature of the connections which determine the network structure. But, basically, the way that the connections are adjusted or trained to achieve a desired overall behavior of the network is governed by its learning algorithm. There are multiple types of ANN, each one with their own specific use, cases and levels of complexity. The most basic type of ANN is known as feedforward, in which information travels in only one direction, from input to output.

An interesting type of ANN is the recurrent neural network, in which data can flow in multiple directions. These neural networks possess greater learning abilities and are widely employed for more complex tasks such as learning handwriting or language recognition (YANG; ZHANG; FU, 2016).

There are also convolutional neural networks, Boltzmann machine networks, Hopfield networks, and a variety of others (HAYKIN, 2009). Picking the right network for a task depends on the data that will be used to train it, and the specific problem that is desirable to solve. In some cases, it may be desirable to use multiple approaches, such as would be the case with a challenging task like voice recognition.

Broadly speaking, all these networks are designed for spotting patterns in data. Specific tasks could include classification (classifying data sets into predefined classes), clustering (classifying data into different undefined categories), and prediction (using past events to guess future ones, like the prediction of time series for stock market and for financial market). The old school digital signal processing (DSP) do routine things well, like keeping ledgers or performing complex math. But DSP have trouble recognizing even simple patterns, let alone generalizing those patterns of the past into actions of the future.

While ANNs have been around since the 1940s, it is only in the last several decades when they have become a major part of Artificial Intelligence (AI). This is due to the introduction of a technique called backpropagation, presented on Section 2.4.1, which allows networks to adjust their hidden layers of neurons in a learning process that fits the outcome to the desired one (CASTRO; CASTRO, 2005; HAYKIN, 2009). Finally, an ANN can learn by three distinct means: with a tutor, without a tutor and with a judge. The algorithm proposed in this thesis makes use of the tutor, thus its learning process is called supervised learning (or non-blind).

2.4.1 Supervised Learning Process and Backpropagation

A supervised learning-based ANN is an ANN that has a knowledge of the environment, with that knowledge being represented by a set of input-output examples. The environment is, however, unknown to the neural network of interest. It has two modes of operation: the training mode and the execution mode.

In the training mode, the neurons are iteratively trained to fire up (or not) for particular input value patterns defined in the ANN training set. Each input value patterns and the

corresponding desired values at the ANN output nodes is stored in the training set (TS) as a vector. The set of vectors stored in the TS must be representative of the numerical mapping problem to be solved, mapping which is usually analytically unknown. At iteration n the nth vector of the TS containing a specific input values pattern and the corresponding desired values at the ANN output nodes is processed by the learning algorithm, so that an instant error is obtained from the difference between the desired values (stored in the TS) at the output nodes of the ANN and the resulting values at the output nodes of the ANN at iteration n. The learning algorithm uses the obtained instant error at iteration n to adjust the ANN free parameters aiming to reduce the instant error at the next iteration $n+1$. The training process goes on, iteration by iteration, until the instant error for all training vectors stored in the TS are sufficiently small, situation in which the ANN is said to be trained. Once trained, the ANN can now operate in the execution mode.

In the execution mode, when a taught input pattern is detected at the input nodes, the ANN yields the corresponding learnt output value. If a particular input pattern does not belong to the taught list of input patterns specified in the TS, the ANN tries to infer a proper output based on the extrapolation of the underlying stochastic process learnt from the TS. This is the reason the ANN training set must be thoroughly representative of the numerical mapping problem to be solved, mapping which is usually analytically unknown.

One of the simplest supervised learning processes is the adaptive process of a single linear neuron operating as an adaptive filter (HAYKIN, 2013). The classical problem to be solved with adaptive filtering is to determine, with a single linear neuron, the model that rules the behavior of an unknown dynamic system Γ. The neuron operates under the influence of an algorithm A which controls the necessary adjustments to the synaptic weights w so, as the adjustments succeed, the output of the neuron proceed towards the output of the system Γ. Filtering consists in two processes, the filtering itself and adapting.

Adaptive filtering involves two signals. An output y, which is a consequence of an input vector $\underline{u}$, and an error e, which is a consequence of the comparation between the output y and a desired output Υ (a target that the network must mimic).

Adapting involves the adjustments of the synaptic weights w by means of an algorithm A, having the error e as base.

$$e = \Upsilon - y \tag{2.31}$$

These two processes combined describes the feedback loop of a neuron. In the training mode, once all vectors $\underline{u}$ from the TS have been applied to the neuron input nodes, it is said that an epoch have passed. The two processes repeat over and over until the error e gets

sufficiently small. The most famous algorithm for the feedback loop is known as backpropagation. Its popularity is because its relative simplicity and the fact that it is a powerful device to store information on the synaptic weights $\underline{w}$ (HAYKIN, 2013).

In order to measure how efficient the process of filtering is and in order to determine the iterative adjustment of the synaptic weights w it is useful to define a function J, known as cost function, which states how unfit is the neuron when approximating its output y to d.

In a similar manner to the single neuron adaptive filtering, the main goal of an ANN is to minimize the cost function J. In this thesis, the algorithm responsible for that is the Steepest Descent (SD), which presents low computational complexity (HAYKIN, 2009).

In the SD algorithm, the successive adjustments on the synaptic weights $\underline{w}$ are in the direction of the steepest descent of the surface S generated by the values of the cost function J. Since J depends directly on w, i.e., $J = J(\underline{w})$, surface S will have as many dimensions as synaptic weights $\underline{w}$ on the input layer of the network. Thus, the problem of minimizing J has turned to a simple "find the minimum of a surface S", as shows Figure 13.

Figure 13 - Cost function $J(x, y)$ for a M-dimensional domain $\mathbb{R}^M$, $M = 2$
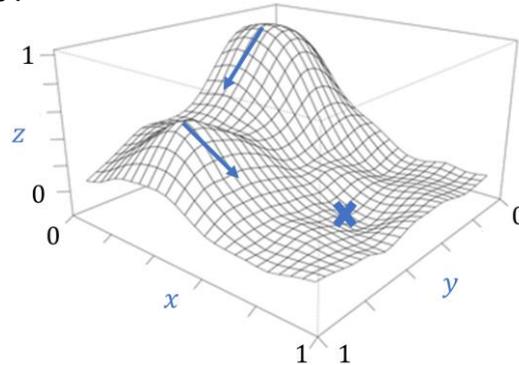
Source: (AUTHOR, 2018).

Mathematically, the minimum of the surface can be found by applying the gradient vector to the surface, i.e., $\underline{\nabla}J(\underline{w})$, and taking a step to the opposite direction. Thus, the algorithm SD can be expressed as

$$\underline{w}(n+1) = \underline{w}(n) - \mu\underline{\nabla}J\left(\underline{w}(n)\right) \tag{2.32}$$

where μ is the learning rate and n is any discrete time instant of the adjustment process.

When μ is small, the transient response of the algorithm is overdamped, in that the trajectory traced by $w(n)$ follows a smooth path in the w-plane, as illustrated in Figure 13. When μ is large, the transient response of the algorithm is underdamped, in that the trajectory

of $w(n)$ follows a zigzagging (oscillatory) path. When μ exceeds a certain critical value, the algorithm becomes unstable (i.e., it diverges).

$\underline{\nabla}J(\underline{w})$ is the gradient vector of the cost function J. Note that the gradient $\underline{\nabla}$ always points to the direction of largest increase of the function (in this case, function $J(\underline{w}(n))$), and has a magnitude equal to the maximum rate of increase at the point in which gradient $\underline{\nabla}$ is determined.

$$\nabla J(w) = \left(\frac{\partial J}{w_1}, \frac{\partial J}{w_2}, \cdots, \frac{\partial J}{w_M} \right)^T \tag{2.33}$$

To minimize $J(\underline{w})$, the process is started with an initial guess denoted by $w(0)$, generating a sequence of weight vectors $w(1)$, $w(2)$, … , such that the cost function J of Eq. (2.31) is reduced at each iteration of the algorithm. Figure 13 shows two vectors (blue arrows) determined at two respective distinct points in the surface of the cost function $J(\underline{w}(n)) = J(x, y)$. Note that the arrows point to the opposite direction of the steepest variation of the surface defined by function $J(\underline{w}(n)) = J(x, y)$, or, equivalently, point to the opposite direction of the gradient vector $\underline{\nabla}J(\underline{w})$. The SD algorithm aims at the minimization of $J(\underline{w}(n))$ in order to reach the global minimum at the blue "x" in Figure 13. Since the gradient $\underline{\nabla}J(\underline{w})$ represents the direction of the steepest slope at the point it is determined, the SD algorithm iteratively takes a step backwards on the opposite direction as shown by the blue arrows in order to fulfill the goal to reach the global minimum at the blue "x".

In conclusion, it is necessary to essentially ask the ANN a large amount of questions and provide it with answers. This is a field called supervised learning. With enough examples of question answer pairs, the calculations and values stored at each neuron and synapse are slowly adjusted. When it is wrong, an error is calculated and the values at each neuron and synapse are propagated backwards through the ANN for the next epoch using backpropagation. This process takes a plenty of examples. Once an entire epoch is completed, then backpropagation is applied. Finally, once trained, the network shall be able to, once received any input that had never been part of the TS, to return an output close enough to what would be expected of the unknown process' scope.

2.4.2 Least Mean Square

The Least Mean Square (LMS) algorithm is the simplest (and for most problems, the fastest) way to minimize the error of an ANN. Because of that, it is considered the leading algorithm in ANN and in this thesis, it is used as a reference to the novel algorithm proposed.

LMS are a class of adaptive filter used to mimic a desired filter process by finding the filter coefficients, or weights w, that relate to producing the least mean square of the error signal (difference between the desired and the actual signal). It is a stochastic gradient descent method in that the filter is only adapted based on the error at the current time. The adaptive filtering model can be described by a signal-flow graph of Figure 14.

Figure 14 – Signal-flow graph of LMS adaptive model for a system
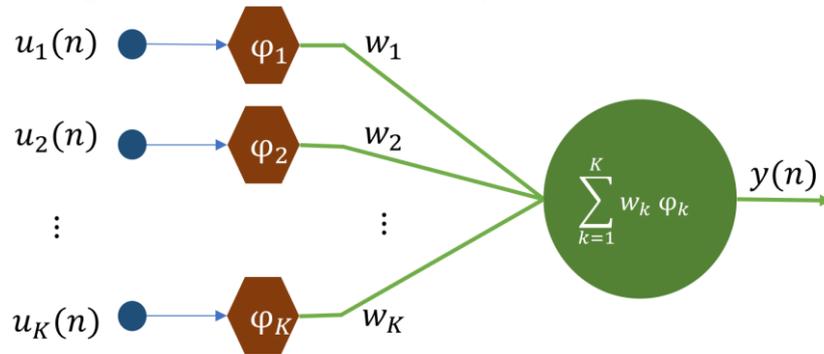
Source: (AUTHOR, 2018).

The activation function φ on LMS is linear, so it can be considered equal to 1. Any other number would represent just a gain. Based on the neurons model of Eq. (2.30), the output signal of the LMS is given by Eq. (2.34). It is possible to see in this equation the linearity of the activation function.

$$y(n) = \sum_{k=1}^{K} w_k(n)u_k(n) = \underline{u}^T(n)\underline{w}(n) \tag{2.34}$$

where the superscript T denotes the transpose operation.

The LMS's main characteristic is to minimize the cost function J of Eq. (2.35).

$$J\left(\underline{w}(n)\right) = \frac{1}{2}e^2(n) \tag{2.35}$$

In order to find the direction of greatest increase of the function J, the gradient vector $\underline{\nabla}$ is applied on Eq. (2.35) by varying $J(\underline{w}(n))$ in response to an infinitesimal variation in coordinate $\underline{w}(n)$, and the following Eq. (2.36) is obtained.

$$\underline{\nabla}J\left(\underline{w}(n)\right) = \frac{\partial\left\{\frac{1}{2}e^2(n)\right\}}{\partial\underline{w}(n)} = e(n)\frac{\partial e(n)}{\partial\underline{w}(n)} \tag{2.36}$$

Bringing back Eq. (2.31) into Eq. (2.34) and since Υ does not depend on $\underline{w}(n)$, it is possible to reach Eq. (2.37).

$$\frac{\partial e(n)}{\partial \underline{w}(n)} = -\underline{u}(n) \tag{2.37}$$

Applying Eq. (2.37) on Eq. (2.36), it is possible to reach Eq. (2.38).

$$\underline{\nabla} J\left(\underline{w}(n)\right) = -e(n)\underline{u}(n) \tag{2.38}$$

And finally, replacing Eq. (2.38) into Eq. (2.32) for $\underline{w}(n + 1)$, comes Eq. (2.39).

$$\underline{w}(n+1) = \underline{w}(n) + \mu e(n)\underline{u}(n) \tag{2.39}$$

This final equation defines the process of adjusting the synaptic weights $\underline{w}$ of a linear neuron in order to minimize J. The greatest advantage of the LMS algorithm is that it does not need to know the correlation matrix R and the cross-correlation vector $\underline{p}$, in contrast to the SD algorithm (based on the Wiener filter deduction, but not mentioned in this thesis). This feature of the LMS algorithm is important from a practical perspective.

In consequence, on LMS, the vector $\underline{w}$ follows a random path in the process of minimizing J (and for this reason, the LMS algorithm is sometimes referred to as a gradient descent algorithm), different from the SD algorithm of Figure 13.

In the SD algorithm, the weight vector $\underline{w}(n)$ follows a well-defined trajectory in the $J(x, y, z)$ M-dimensional space $\mathbb{R}^M$ for a prescribed μ. It is also noteworthy that the inverse of the learning rate parameter μ acts as a measure of the memory of the LMS algorithm: The smaller μ, the longer the memory span over which the LMS algorithm remembers past data. Consequently, when μ is small, the LMS's accuracy is higher, but the convergence rate is slower (HAYKIN, 2013) . As the number of iterations in the LMS algorithm approaches infinity, it performs a random walk (Brownian motion) towards the solution.

The LMS, or gradient descent algorithm, plays a major role in the ANN algorithms further presented.

2.4.3 Radial Basis Function

The Radial Basis Function (RBF) ANN uses a non-linear model for its neuron (the activation function φ is non-linear, as later explored in this section). RBFs are supervised ANNs, and are considered universal approximators. Such combination of learning algorithms plus universality results in an attractive mix with many interesting theorems as ingredients.

The universal approximation theorem states that a feed-forward network with a single hidden layer containing a finite number of neurons, can approximate continuous functions on compact subsets of $\mathbb{R}^M$, under mild assumptions on the activation function. The higher the

number of hidden neurons, more improved is the approximation. The theorem thus states that neural networks with a single hidden layer can be used to approximate any continuous function to any desired precision. Figure 15 shows a simple approximation example. For every possible input x, there is always a neural network to output the value $y = f(x)$ (or some close approximation), no matter the x to approximate.

Figure 15 – Orange line is the function $y = x^3 + x^2 - x - 1$. Blue line is an example of function approximation
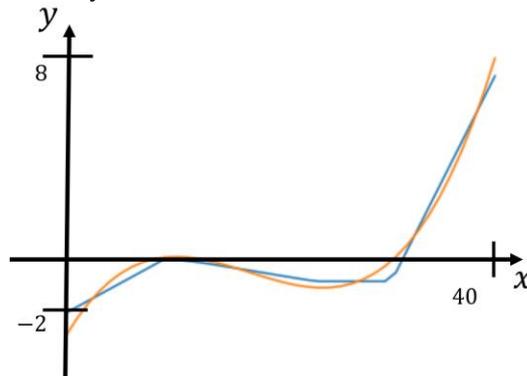
Source: Adapted from (NIELSEN, 2015).

Therefore, using an universal approximator and introducing non linearity via an activation function allows to approximate almost any function (FORTUNER, 2017; WINKLER, 2017).

After understanding what an approximator does, it is now possible to define the ANN used in this thesis. The RBF is an ANN that transforms the problem into a n-dimensional space $\mathbb{R}^n$ in a non-linear manner (usually n is quite high). The high dimensional space is a multidimensional plot of the output as a function of the input. Bringing up another theorem: "A complex pattern classification problem cast in a high dimensional space nonlinearly is more likely to be linearly separable than in a low-dimension space" (COVER, 1965). In other words, once there are linearly separable patterns, the classification problem is relatively easy to solve. Figure 16 illustrates the pattern separation by different complexity methods.

Figure 16 - Three examples of φ-separable dichotomies of different sets of five points in two dimensions: linearly separable dichotomy; spherically separable dichotomy; quadrically separable dichotomy
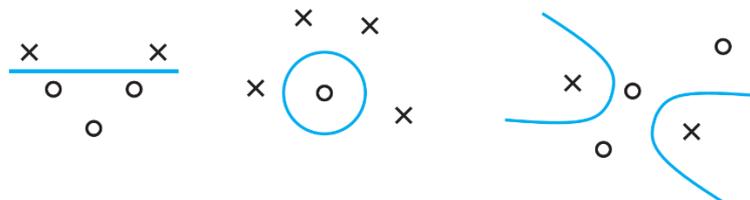
Source: (HAYKIN, 2013).

The RBF is a multilayer ANN with a single hidden layer, where each neuron in the hidden layer operates as a computational kernel, providing a set of functions that constitute an arbitrary basis for the input vectors when they are expanded into the hidden space. These kernels are radial basis functions, with a compact domain, and are responsible to apply a non-linear transformation from the input space to the hidden space. Basically, each kernel computes the distance between the nth input data vector $\underline{u}(n)$ and its own center vector $\underline{t}_k$. The input layer is made up of source nodes that connect the network to its environment. These source nodes can be connected, for example, to the outputs of an array of transducers or any other source of information representative of the numerical mapping problem to be solved. The output layer combines the kernels output φ_k through a linear combiner, as shown in Figure 17.

Figure 17 – RBF ANN architecture
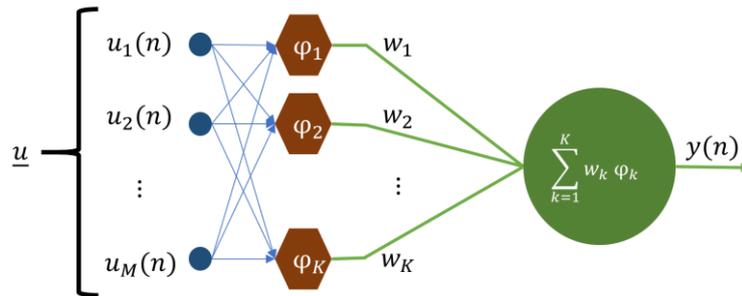
Source: (AUTHOR, 2018).

The RBF learning process is divided in two phases, the training and the generalization.

The training phase constitutes the optimization of a fitting procedure from the surface J, based on known data in the form of vectors $\underline{u} = [u_1, u_2, \cdots, u_M]^T$, presented to the network as input output examples. As M is the dimension of the vector $\underline{u}$, and as every vector can be seen as a coordinate on a M-dimensional space $\mathbb{R}^M$, the input data $\underline{u}$ can also be seen as a point in this M-dimensional space $\mathbb{R}^M$. The TS presented to an ANN can be organized as a matrix composed by a collection of vectors $\underline{u}$, in other words, a matrix full of points in the M-dimensional space $\mathbb{R}^M$. Each one of these points partially describes an unknown process. It is desired that the ANN represents each one of these unknown processes as better as possible.

$$
TrainingSet = \begin{bmatrix} u_1(1) & u_1(2) & \cdots & u_1(N) \\ u_2(1) & u_2(2) & \cdots & u_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ u_M(1) & u_M(2) & \cdots & u_M(N) \end{bmatrix} \tag{2.40}
$$

The generalization phase is where interpolation between the data points are performed along the constrained surface generated by the fitting procedure as the optimum approximation to the true surface. The interpolation is done between points in order to estimates the value of a function for intermediate values between the set of known data. Figure 18 illustrates the interpolation process of a 2-dimensional function from a random set of samples.

Figure 18 – Example of data interpolation
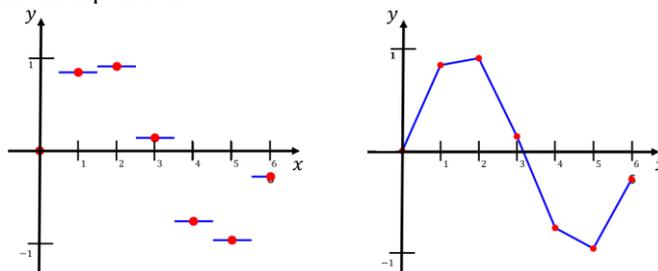
Source: (AUTHOR, 2018)

Another important issue in this context is the multivariable interpolation. According to Davis theorem of 1963, "Given a set of N different points and a corresponding set of N real numbers, it is possible to find a function that satisfies the interpolation condition where the strict interpolating surface is constrained to pass through all the training data points". Finally, Micchelli covers in his theorem of 1986 a large class of radial-basis functions broadly used in RBF ANNs to perform the interpolation, like multiquadric, inverse multiquadric and gaussian functions. These are all continuous functions bell shaped over a compact domain with limited radius ζ in order to make feasible the interpolation and the extrapolation of numerical values through linear combination. Each shape has a better performance on different scenarios. The main difference on the three RBF ANNs models lately presented in the next sections is the radial basis function adopted in the ANN hidden layer.

This thesis adopts a Gaussian function for the real and imaginary parts of complex valued radial basis functions (kernels whose output is φ_k in Figure 17) in the hidden layer of the network.

The Gaussian function is defined by Eq. (2.41). Note that it is a localized function, i.e. $\varphi(\zeta) \to 0$ as $\zeta \to \infty$. For each vector $\underline{u}$ with size K presented to the network, K signals are sent towards every single neuron at the hidden layer. Each neuron (kernel) with transmittance $\varphi(\zeta)$ return values between 0 and 1. The higher the neurone output the higher the neuron activation. Thus, each neuron transmittance $\varphi(\zeta)$ is also the neuron activation function.

$$\varphi\left(\zeta\right) = e^{\left(-\frac{\zeta^2}{2\sigma^2}\right)}$$

(2.41)

where σ is the standard deviation parameter which represents how large is the basis of the bell shaped Gaussian kernel.

The input data points $\underline{u}$ in the M-dimensional space $\mathbb{R}^M$ form clusters in the space $\mathbb{R}^M$ in a greater or lesser extent. The center vector $\underline{t}_k$ of each bell-shaped activation function φ_k is also located somewhere in the M-dimensional space $\mathbb{R}^M$. Each center vector $\underline{t}_k$ is a free parameter of the RBF ANN to be adjusted during the training phase. Note that the closer the coordinates of each center vector $\underline{t}_k$ are of the coordinates of a cluster of input data $\underline{u}$, the better the approximation performance of the RBF ANN. Note also that the distance between the coordinates of each center vector $\underline{t}_k$ and the coordinates of an input data vector $\underline{u}$ is given by

$$\zeta = \left\| \underline{u} - \underline{t}_k \right\| \tag{2.42}$$

where $\|.\|$ is the Euclidian operator, defined by

$$\|\underline{u}\| = \sqrt{\sum_{m=1}^{M} \left(u_m \right)^2} \tag{2.43}$$

The standard deviation σ of the kth Gaussian kernel φ_k controls the reach radius (the width of the domain) of the corresponding bell-shaped activation function. In the context of an RBF ANN, σ defines the average Euclidian distance (average radius) that measures the data spreading represented by the kernel $\varphi_k(\zeta)$ around its center $\underline{t}_k$. The standard deviation σ is also a free parameter of the RBF ANN to be adjusted during the training phase.

The data which an RBF ANN represent are, therefore, expanded with reference to a finite set of activation functions φ_k, $k = 1, 2 \cdots K$, where K is chosen based on the universally approximation theory (the higher the number of neurons on the hidden layer, the higher the precision on the approximation). Each one of these K functions are centered in a particular coordinate $\underline{t}_k$ at the multidimensional space $\mathbb{R}^M$ which contains the set of the input vectors $\underline{u}$. Thus, the activation function φ_k of the kth neuron, with center at coordinates $\underline{t}_k$, and with standard deviation σ_k, is specifically given by

$$\varphi_k \left(\underline{t}_k, \sigma_k^2, \underline{u} \right) = e^{\left(-\frac{\left\| \underline{u} - \underline{t}_k \right\|^2}{2\sigma_k^2} \right)} \tag{2.44}$$

Note that Eq. (2.44) can be interpreted as a Probability Density Function (PDF) in the context of the data clustering technique called Gaussian Mixture Model (GMM) (BISHOP, 2006; DIAZ-ROZO; BIELZA; LARRANAGA, 2018; LI et al., 2018).

Thus, the set of input vectors $\underline{u}$ is represented by the set of K kernels $\varphi\left(t_k, \sigma_k^2, \underline{u} \right)$, $k = 1, 2 \cdots K$, with the kth kernel output being weighted by the respective ANN synaptic weight w_k. The synaptic weight is the memory of the RBF ANN since it stores information during the state

transitions over the learning process. The learning process is an iterative process, which means that it will repeat over and over again along time. Each time it repeats it is called an epoch. Each epoch consists of the randomly ordered presentation of all data vectors $\underline{u}(n)$ stored in the TS to the RBF input nodes. In other words, each vector $\underline{u}$ belonging to the TS is randomly and periodically presented to the RBF input.

The output y is the outcome of all K neurons activation functions (kernels) and is obtained through the sum of all activation functions φ_k weighted by the corresponding synaptic weights w_k. For each iteration n, in which the nth vector $\underline{u}(n)$ of the TS is presented to the RBF input nodes, the RBF yields the output $y(n)$

$$y(n) = \sum_{k=1}^{K} w_k(n)\varphi_k(n) \tag{2.45}$$

Simultaneously, for each iteration n, whereby the nth vector $\underline{u}$ of the TS is presented to the RBF input nodes, there is a desired value $\Upsilon(n)$ for the output $y(n)$. In other words, the training is performed with a set of "question and answer" pairs. At each iteration n, the RBF output $y(n)$ is compared to the desired output $\Upsilon(n)$ and an error $e(n)$ is obtained by

$$e(n) = \Upsilon(n) - y(n) \tag{2.46}$$

The SD algorithm uses the obtained instant error $e(n)$ at iteration n to update the RBF ANN free parameters aiming to reduce the instant error at the next iteration $n+1$. The training process goes on, iteration by iteration, comprising an integer number of epochs, until the instant error for all training vectors stored in the TS are sufficiently small, situation in which the ANN is said to be trained, or, equivalently, the process of minimization of the cost function J has achieved convergence. The SD updating equations for the RBF ANN free parameters are presented in the next sections.

The effectiveness of the process of minimization of function J is assessed by the Mean Square Error Average (MSEA), given by Eq. (2.47).

$$MSEA = \frac{1}{N} \sum_{n=1}^{N} (\Upsilon(n) - y(n))^2 \tag{2.47}$$

where N is the number of the last training vectors $\underline{u}(n)$ presented to the ANN input nodes, being N such that it comprises an integer number of epochs

When the MSEA reaches an acceptable small value, the ANN learning process is said to have converged and the ANN is fit to be used as an approximation of the process represented by the TS.

The last stage on defining an RBF ANN consists of determining the SD updating equations for the RBF ANN. As each one of the ANNs presented in the next sections are defined by different update equations, these will be explored in their own sections.

As already mentioned before in the context of curve approximation algorithms, the dimension of the RBF ANN hidden layer is directly related to the ANN capacity to approximate a smooth input-output mapping. The higher the number K of kernels in the RBF hidden layer, the higher the approximation accuracy (MHASKAR, 1996; NIYOGI; GIROSI, 1996). The number K of neurons in the hidden layer of any ANN is arbitrary and it usually determined by performance tests of the ANN network (HAYKIN, 2013). The higher the K, the better the output approximation. Nevertheless, K can be also determined by the proposed number of clusters which the set of input vectors $\underline{u}$ forms.

The only way to infer this knowledge is to use a clustering algorithm. Several clustering algorithms are available in literature, like the popular K-means algorithm (KANUNGO et al., 2002). Besides inferring an approximation for the number K of kernels with respective centers a coordinate $\underline{t}_k$, the clustering algorithm allows to infer another important parameter: a good approximation for the standard deviation σ initialization, given by

$$\sigma = \frac{d_{t\max}}{\sqrt{2K}}$$

(2.48)

where $d_{t\,max}$ is the maximum distance between the radial basis centers $\underline{t}_k$ of the RBF ANN.

Most other ANN has multiple hidden layers where each layer is responsible for a characteristic of the information being processed, just as the biological neural network. The RBF has a different approach, in fact, it is classified as an ANN only because it behaves as a curve fitting (approximation) problem in a high-dimensional space. In summary, for an RBF ANN, learning is equivalent to finding a surface in a multidimensional space that provides a best fit to the training data set, with criterion for best fit being measured in some statistical sense (in this work, the MSEA analysis). Correspondingly, generalization is equivalent to the use of this multidimensional surface to interpolate and extrapolate the test data set (HAYKIN, 2009).

As this work handles with module and phase representative of phasors of EM waves, the data are all complex physical quantities. Bearing it in mind, the RBF network must be able to handle complex numbers on the best possible way. Therefore, the RBFs chosen to implement the algorithm proposed in this thesis are all complex networks and their respective update rules for the free parameters t, σ and w, based on the LMS algorithm (gradient descent), will be described in the next sections.

2.4.4 C-RBF

One of the first RBF network to consider problems with complex values was published by (HAYKIN, 2013), named Complex-valued Radial Basis Function (C-RBF).

Complex-valued RBF networks are the extensions of real valued RBF networks for operating on complex-valued signals. The C-RBF network has M input nodes, K hidden neurons, and O output neurons. The input nodes convey the input signals to all K neurons in the hidden layer. The neurons in the hidden layer employ the Gaussian activation function of Eq. (2.44). Each hidden neuron estimates the localized response for a given input. Finally, the O neurons in the output layer use the weighted sum of the hidden layer responses to generate the complex-value output y of Eq. (2.45). Figure 17 shows the architecture for a complex-valued RBF ANN with one single output y. The selection of a complex-valued RBF is in fact the same as that for a real-valued RBF, with some minor modifications. As complex quantities will be handled by the network, the centers of the activation functions φ and the output weights w are complex-valued, however, the response of the hidden neurons remains real-valued.

This section presents the gradient descent-based update rules for the C-RBF (CASTRO; CASTRO, 2005; HAYKIN, 2013).Gradient descent based update rule for $w_k(n+1)$

For convenience, the update Eq. (2.32) through the gradient descent is repeated above.

$$w(n+1) = w(n) - \mu_w \underline{\nabla} J\big(w(n)\big) \tag{2.49}$$

The cost function J is given by Eq. (2.35) and the error equation is given by Eq. (2.31). From Eq. (2.49), in order to determine $w(n+1)$, it is necessary to find $\underline{\nabla} J$, so we write

$$\nabla_k J = \frac{\partial J}{\partial w_k} = \frac{1}{2}\frac{\partial}{\partial w_k}(\Upsilon - y)^2 \tag{2.50}$$

where

$$\frac{\partial}{\partial w_k}(\Upsilon - y)^2 = 2(\Upsilon - y)\frac{\partial}{\partial w_k}(\Upsilon - y) \tag{2.51}$$

Substituting Eq. (2.51) into Eq. (2.30), it is possible to write

$$\frac{\partial}{\partial w_k}(\Upsilon - y)^2 = 2(\Upsilon - y)\frac{\partial}{\partial w_k}\left(\Upsilon - \sum_{k=1}^{K} w_k \varphi_k\right) \tag{2.52}$$

Which can be expanded into

$$\frac{\partial}{\partial w_k}(\Upsilon - y)^2 = 2(\Upsilon - y)\frac{\partial}{\partial w_k}(\Upsilon - w_1 \varphi_1 - w_2 \varphi_2 - \cdots - w_K \varphi_K) \tag{2.53}$$

As the desired output Υ does not depend on the synaptic weights w, and the derivative of $\partial/\partial w_k$ will only exist for the current neuron k, then it is possible to write

$$\frac{\partial}{\partial w_k}\left(\Upsilon - y\right)^2 = 2\left(\Upsilon - y\right) - \varphi_k \tag{2.54}$$

Substituting Eq. (2.54) into Eq. (2.50), it is possible to find

$$\nabla_k = \frac{\partial J}{\partial w_k} = \frac{1}{2}\frac{\partial}{\partial w_k}\left(\Upsilon - y\right)^2 = \frac{1}{2}\left[2\left(\Upsilon - y\right) - \varphi_k\right] = \left(\Upsilon - y\right)\left(-\varphi_k\right) \tag{2.55}$$

Finally, replacing Eq. (2.55) into Eq. (2.49), it is possible to obtain the update equation for the kth synaptic weight, given by

$$w_k\left(n+1\right) = w_k\left(n\right) - \mu_w\left(\Upsilon - y\right)\left(-\varphi\right) = w_k\left(n\right) + \mu_w e\left(n\right)\varphi_k\left(n\right) \tag{2.56}$$

The gradient descent-based update rule for the kth kernel center $\underline{t}_k(n+1)$ is obtained as follows. Similarly to the Eq. (2.32) , the gradient descent update rule for each $\underline{t}_k$ center is

$$\underline{t}\left(n+1\right) = \underline{t}(n) - \mu_t \underline{\nabla} J\left(\underline{t}\left(n\right)\right) \tag{2.57}$$

Just as for the synaptic weights, the cost function J is given by Eq. (2.35) and the error equation is given by Eq. (2.31). From Eq. (2.57), in order to determine $\underline{t}(n + 1)$, it is necessary to find $\underline{\nabla} J$, so we write

$$\underline{\nabla}_k J = \frac{\partial J}{\partial \underline{t}_k} = \frac{1}{2}\frac{\partial}{\partial \underline{t}_k}\left(\Upsilon - y\right)^2 \tag{2.58}$$

where

$$\frac{\partial}{\partial \underline{t}_k}\left(\Upsilon - y\right)^2 = 2\left(\Upsilon - y\right)\frac{\partial}{\partial \underline{t}_k}\left(\Upsilon - \sum_{k=1}^{K} w_k \varphi_k\right) \tag{2.59}$$

Eq. (2.44), gives the expression for the activation function φ. Substituting Eq. (2.44) into Eq. (2.59), it is possible to write

$$\frac{\partial}{\partial \underline{t}_k}\left(\Upsilon - y\right)^2 = 2\left(\Upsilon - y\right)\frac{\partial}{\partial \underline{t}_k}\left[\Upsilon - \sum_{k=1}^{K} w_k e^{\left(-\frac{\|u - t_k\|^2}{2\sigma_k^2}\right)}\right] \tag{2.60}$$

By expanding the summation, it is possible to write

$$\frac{\partial}{\partial \underline{t}_k}\left(\Upsilon - y\right)^2 = 2\left(\Upsilon - y\right)\frac{\partial}{\partial \underline{t}_k}\left[\Upsilon - w_1 e^{\left(-\frac{\|u - t_1\|^2}{2\sigma_1^2}\right)} - w_2 e^{\left(-\frac{\|u - t_2\|^2}{2\sigma_2^2}\right)} - \cdots - w_k e^{\left(-\frac{\|u - t_k\|^2}{2\sigma_k^2}\right)}\right] \tag{2.61}$$

As the desired output Υ does not depend on the radial basis centers $\underline{t}_k$, and the derivative of $\partial/\partial \underline{t}_k$ will only exist for the current neuron k, then it is possible to write

$$\frac{\partial}{\partial \underline{t}_k}(\Upsilon - y)^2 = 2(\Upsilon - y)\frac{\partial}{\partial \underline{t}_k}\left[-w_k e^{\left(-\frac{\|u-t_k\|^2}{2\sigma_k^2}\right)}\right] = 2(\Upsilon - y)(-2)w_k\left(\frac{u-t_k}{2\sigma_k^2}\right)e^{\left(-\frac{\|u-t_1\|^2}{2\sigma_1^2}\right)} \quad (2.62)$$

Replacing Eq. (2.62) into Eq. (2.58), comes

$$\underline{\nabla}_k J = \frac{\partial}{\partial \underline{t}_k} = \frac{1}{2}\left\{2(\Upsilon - y)(-2)w_k\left(\frac{u-t_k}{2\sigma_k^2}\right)e^{\left(-\frac{\|u-t_k\|^2}{2\sigma_k^2}\right)}\right\}$$

$$= -2w_k(\Upsilon - y)\left(\frac{u-t_k}{2\sigma_k^2}\right)e^{\left(-\frac{\|u-t_k\|^2}{2\sigma_k^2}\right)} \quad (2.63)$$

Finally, replacing Eq. (2.63) into Eq. (2.57), it is possible to obtain the update equation for the radial basis centers, given by

$$\underline{t}_k(n+1) = \underline{t}_k(n) + 2\mu_t w_k(n)e(n)\varphi_k(n)\frac{u-t_k}{2\sigma_k^2} \quad (2.64)$$

The gradient descent-based update rule for the kth Gaussian kernel variance $\sigma_k^2(n+1)$ is obtained as follows. Similarly to the Eq. (2.32), the gradient descent update rule for each variance σ_k^2 is given by

$$\alpha(n+1) = \alpha(n) - \mu_\sigma \underline{\nabla}J(\alpha(n)) \quad (2.65)$$

where $\sigma^2 = \alpha$, for convenience.

Just as for the synaptic weights, the cost function J is given by Eq. (2.35) and the error equation is given by Eq. (2.31). From Eq. (2.65), in order to determine $\alpha(n+1)$, it is necessary to find $\underline{\nabla}J$, so we write

$$\underline{\nabla}_k J = \frac{\partial J}{\partial \alpha_k} = \frac{1}{2}\frac{\partial}{\partial \alpha_k}(\Upsilon - y)^2 \quad (2.66)$$

where

$$\frac{\partial}{\partial \alpha_k}(\Upsilon - y)^2 = 2(\Upsilon - y)\frac{\partial}{\partial \alpha_k}\left(\Upsilon - \sum_{k=1}^{K}w_k\varphi_k\right) \quad (2.67)$$

Eq. (2.44) gives the expression for the activation function φ. Substituting Eq. (2.44) into Eq. (2.67), it is possible to write

$$\frac{\partial}{\partial \alpha_k}(\Upsilon - y)^2 = 2(\Upsilon - y)\frac{\partial}{\partial \alpha_k}\left[\Upsilon - \sum_{k=1}^{K}w_k e^{\left(-\frac{\|u-t_k\|^2}{2\alpha_k}\right)}\right] \quad (2.68)$$

By expanding the summation, it is possible to write

$$\frac{\partial}{\partial \alpha_k}(\Upsilon - y)^2 = 2(\Upsilon - y)\frac{\partial}{\partial \alpha_k}\left[\Upsilon - w_1 e^{\left(-\frac{\|u-t_1\|^2}{2\alpha_1}\right)} - w_2 e^{\left(-\frac{\|u-t_2\|^2}{2\alpha_2}\right)} - \cdots - w_K e^{\left(-\frac{\|u-t_K\|^2}{2\alpha_K}\right)}\right] \tag{2.69}$$

As the desired output Υ does not depend on the variances α_k, and the derivative of $\partial/\partial \alpha_k$ will only exist for the current neuron k, then it is possible to write

$$\frac{\partial}{\partial \alpha_k}(\Upsilon - y)^2 = 2(\Upsilon - y)\frac{\partial}{\partial \alpha_k}\left[-w_k e^{\left(-\frac{\|u-t_k\|^2}{2\alpha_k}\right)}\right] \tag{2.70}$$

As $\frac{\partial}{\partial x}e^{\left(-\frac{c}{x}\right)} = \frac{c}{x^2}e^{\left(-\frac{c}{x}\right)}$, Eq. (2.70) can be re-written as

$$\frac{\partial}{\partial \alpha_k}(\Upsilon - y)^2 = 2(\Upsilon - y)(-2)w_k\left(\frac{\|u-t_k\|^2}{(2\alpha_k)^2}\right)e^{\left(-\frac{\|u-t_k\|^2}{2\alpha_k}\right)} \tag{2.71}$$

Replacing Eq. (2.71) into Eq. (2.66), comes

$$\nabla_k J = \frac{\partial J}{\partial \alpha_k} = \frac{1}{2}\left\{2(\Upsilon - y)(-w_k)\left(\frac{\|u-t_k\|^2}{(2\alpha_k)^2}\right)e^{\left(-\frac{\|u-t_k\|^2}{2\alpha_k}\right)}\right\}$$
$$= -w_k(\Upsilon - y)\left(\frac{\|u-t_k\|^2}{(2\alpha_k)^2}\right)e^{\left(-\frac{\|u-t_k\|^2}{2\alpha_k}\right)} \tag{2.72}$$

Finally, replacing Eq. (2.72) into Eq. (2.65), it is possible to obtain

$$\alpha_k(n+1) = \alpha_k - \mu_\sigma(\Upsilon - y)(-w_k)\left(\frac{\|u-t_k\|^2}{(2\alpha_k)^2}\right)e^{\left(-\frac{\|u-t_k\|^2}{2\alpha_k}\right)} \tag{2.73}$$

And considering that $e(n) = \Upsilon(n) - y(n)$ and $\alpha_k = \sigma_k^2$, the update equation for the variance is given by

$$\sigma_k^2(n) = \sigma_k^2(n) + \mu_\sigma e(n)w_k(n)\varphi_k(n)\frac{\|u(n)-t_k(n)\|^2}{(\sigma_k^2(n))^2} \tag{2.74}$$

Moreover, during the RBF ANN training phase, the real part of the error is used to update the real part of the network parameters and the imaginary part of the error is used to update the imaginary part of the network parameters. Nevertheless, the gradients used are not a true representation of the gradient of the target function and hence, will not approximate phase accurately, as explained in (SAVITHA; SURESH; SUNDARARAJAN, 2009).

From Eq. (2.56), (2.64) and (2.74) one can note that though the centers and weights of the CRBF network are all complex-valued, the response of the hidden neuron is real-valued,

and hence, does not transmit the phase of the complex-valued input signal to the output neurons. This affects the phase approximation ability of the network. Besides, it may also be observed from the same equations, the gradients use the Real and Imaginary components of the complex-valued error and the output weights to update the free parameters of the network. This does not capture the correlation between the real-imaginary components of the complex-valued gradient. Hence, the gradient thus derived, is not a true representation of the true complex-valued gradient. This is also due to the fact that the responses at the hidden neurons are real-valued.

2.4.5 FC-RBF

The Fully Complex-Radial Basis Function (FC-RBF) is able to process complex values keeping phase information.

Like the C-RBF network, the FC-RBF network has one input layer, one hidden layer and one output layer. The architecture of FC-RBF network is also given by Figure 17. As can be observed from the figure, the network has M input nodes, K hidden neurons and O output neurons. The input nodes convey the input signals to all K neurons in the hidden layer. The main difference here is that, at the hidden layer, the localized responses of the inputs are computed with a non-linear fully complex-valued activation function φ. In the article (SURESH; SUNDARARAJAN; SAVITHA, 2013) the activation function φ_k for the kth hidden neuron is a hyperbolic secant "sech($\cdot$)" function, as follows

$$\varphi_k\left(\underline{t}_k, \underline{\sigma}_k, \underline{u}\right) = \operatorname{sech}\left(\underline{\sigma}_k^T\left(\underline{u} - \underline{t}_k\right)\right) \tag{2.75}$$

where the superscript T is the transpose operator and $\underline{\sigma}$ is a vector with M elements, each element corresponding to the complex scaling factor applied to each dimension of the Mth dimensional complex valued input data space $\underline{u}$. Note that each one of the M complex-valued scaling factor in vector $\underline{\sigma}$ plays a role similar to the deviation σ^2 in the C-RBF Gaussian function.

The complex-valued $sech$ satisfies the desirable properties of activation functions φ for an RBF ANN. It is analytic and bounded almost everywhere, pretty much similar to the Gaussian activation function of Eq. (2.44). The small difference in its shape (that defines how the network approximates its radial basis centers t to the input data $\underline{u}$) is the fundamental difference on the network.

Figure 19 – Gaussian and sech activation functions
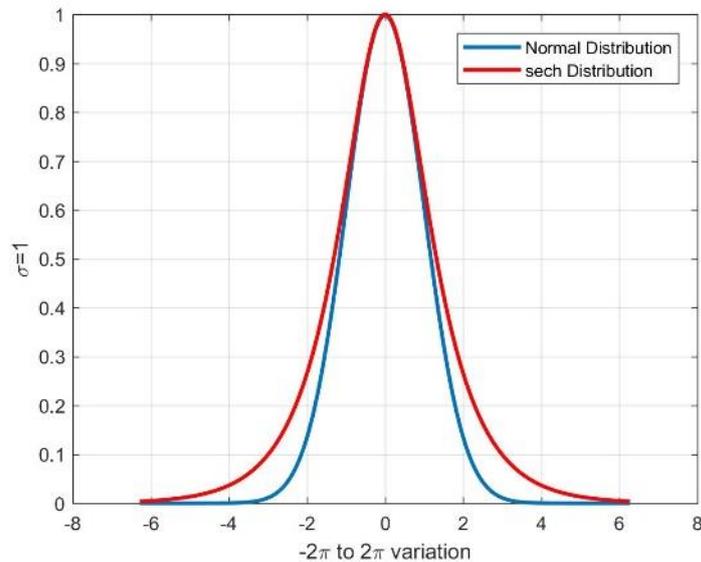
As this ANN is fully complex, the first important difference is on how the output $y(n)$ of Eq. (2.30) is written. Eq. (2.76) express this difference

$$
\begin{aligned}
y(n) &= \sum_{k=1}^{K} w_k(n)\varphi_k(n) = \sum_{k=1}^{K}\left(\operatorname{Re}\{w_k\} + j\operatorname{Im}\{w_k\}\right)\left(\operatorname{Re}\{\varphi_k\} + j\operatorname{Im}\{\varphi_k\}\right) \\
&= \sum_{k=1}^{K}\left(\operatorname{Re}\{w_k\}\operatorname{Re}\{\varphi_k\} - j\operatorname{Im}\{w_k\}\operatorname{Im}\{\varphi_k\}\right) + j\left(\operatorname{Re}\{w_k\}\operatorname{Im}\{\varphi_k\} + j\operatorname{Im}\{w_k\}\operatorname{Re}\{\varphi_k\}\right)
\end{aligned}
\tag{2.76}
$$

This section presents the gradient descent based update rules for the FC-RBF (SURESH; SUNDARARAJAN; SAVITHA, 2013). As this ANN is fully complex, the second important difference is on how the cost function J of Eq. (2.35) is written. Eq. (2.77) express that difference.

$$
J(n) = \frac{1}{2}e^2(n) = \frac{1}{2}e(n)e^*(n)
\tag{2.77}
$$

The error $e(n)$ is still given by Eq. (2.31), except this time both y and Y are complex values.

Similarly to the Eq. (2.32), the gradient descent update rule for the kth synapses vector $\underline{w}_k(n+1)$ is given by

$$
w(n+1) = w(n) - \mu_w \underline{\nabla} J(w(n))
\tag{2.78}
$$

The cost function J is given by (2.77) and the error equation is given by Eq. (2.31). From Eq. (2.78), in order to determine $w(n+1)$, it is necessary to find $\underline{\nabla}J$, so it is possible to write

$$\nabla_k J = \frac{\partial J}{\partial w_k} = \frac{\partial J}{\partial \text{Re}\{w_k\}} + j\frac{\partial J}{\partial \text{Im}\{w_k\}} \tag{2.79}$$

Using the chain rule, the derivative of the cost function with respect to the real part of w_k is given by

$$\frac{\partial J}{\partial \text{Re}\{w_k\}} = \frac{\partial J}{\partial \text{Re}\{y\}}\frac{\partial \text{Re}\{y\}}{\partial \text{Re}\{w_k\}} + \frac{\partial J}{\partial \text{Im}\{y\}}\frac{\partial \text{Im}\{y\}}{\partial \text{Re}\{w_k\}} \tag{2.80}$$

And the derivative of the cost function with respect to the imaginary part of w_k is given by

$$\frac{\partial J}{\partial \text{Im}\{w_k\}} = \frac{\partial J}{\partial \text{Re}\{y\}}\frac{\partial \text{Re}\{y\}}{\partial \text{Im}\{w_k\}} + \frac{\partial J}{\partial \text{Im}\{y\}}\frac{\partial \text{Im}\{y\}}{\partial \text{Im}\{w_k\}} \tag{2.81}$$

Defining $\text{Re}\{e(n)\} = -\partial J/\partial \text{Re}\{y\}$ and $\text{Im}\{e(n)\} = -\partial J/\partial \text{Im}\{y\}$ and using the following partial derivatives obtained from Eq. (2.76), it is possible to write

$$\frac{\partial \text{Re}\{y\}}{\partial \text{Re}\{w_k\}} = \text{Re}\{\varphi_k\}$$

$$\frac{\partial \text{Im}\{y\}}{\partial \text{Re}\{w_k\}} = \text{Im}\{\varphi_k\}$$

$$\frac{\partial \text{Re}\{y\}}{\partial \text{Im}\{w_k\}} = -\text{Im}\{\varphi_k\}$$

$$\frac{\partial \text{Im}\{y\}}{\partial \text{Im}\{w_k\}} = \text{Re}\{\varphi_k\} \tag{2.82}$$

$$\frac{\partial J}{\partial \text{Re}\{w_k\}} = -\text{Re}\{e\}\text{Re}\{\varphi_k\} - \text{Im}\{e\}\text{Im}\{\varphi_k\}$$

$$\frac{\partial J}{\partial \text{Im}\{w_k\}} = -\text{Re}\{e\}\left(-\text{Im}\{\varphi_k\}\right) - \text{Im}\{e\}\text{Re}\{\varphi_k\}$$

From all Eqs. (2.82) the gradient of the error can be written as

$$\nabla_k J = \left(-\text{Re}\{e\}\text{Re}\{\varphi_k\} - \text{Im}\{e\}\text{Im}\{\varphi_k\}\right) + j\left(\text{Re}\{e\}\text{Im}\{\varphi_k\} - \text{Im}\{e\}\text{Re}\{\varphi_k\}\right)$$
$$= -\left(\text{Re}\{e\} + j\text{Im}\{e\}\right)\varphi_k^* \tag{2.83}$$

where the superscript * is the complex-conjugate operator.

Finally, replacing Eq. (2.83) into Eq, (2.78), it is possible to obtain the update equation to the synaptic weight, given by

$$w_k(n+1) = w_k(n) + \mu_w \varphi_k^* e(n) \tag{2.84}$$

Similarly to the Eq. (2.32), the gradient descent update rule for the kth kernel scaling factor vector $\underline{\sigma}_k(n+1)$ is given by

$$\underline{\sigma}(n+1)=\underline{\sigma}(n)-\mu_t\underline{\nabla}J\left(\underline{\sigma}(n)\right) \tag{2.85}$$

Similarly, the update for the scaling factor σ requires the gradient of the real cost function J with respect to the real and imaginary components of $\underline{\sigma}$ is given by Eq. (2.86). Let $p_k + jq_k = \underline{\sigma}_k(\underline{u} - \underline{t}_k)$

$$\nabla_{\underline{k}}J = \frac{\partial J}{\partial \underline{\sigma}_k} = \frac{\partial J}{\partial \mathrm{Re}\{\underline{\sigma}_k\}} + j\frac{\partial J}{\partial \mathrm{Im}\{\underline{\sigma}_k\}} \tag{2.86}$$

where the derivative of the cost function with respect to the real part of $\underline{\sigma}$ is given by

$$
\begin{aligned}
\frac{\partial J}{\partial \mathrm{Re}\{\underline{\sigma}_k\}} = &\frac{\partial J}{\partial \mathrm{Re}\{y\}}\left[\frac{\partial \mathrm{Re}\{y\}}{\partial \mathrm{Re}\{\varphi_k\}}\frac{\partial \mathrm{Re}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \mathrm{Re}\{\underline{\sigma}_k\}} + \frac{\partial \mathrm{Re}\{y\}}{\partial \mathrm{Re}\{\varphi_k\}}\frac{\partial \mathrm{Re}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \mathrm{Re}\{\underline{\sigma}_k\}} \right.\\
&\left. + \frac{\partial \mathrm{Re}\{y\}}{\partial \mathrm{Im}\{\varphi_k\}}\frac{\partial \mathrm{Im}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \mathrm{Re}\{\underline{\sigma}_k\}} + \frac{\partial \mathrm{Re}\{y\}}{\partial \mathrm{Im}\{\varphi_k\}}\frac{\partial \mathrm{Im}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \mathrm{Re}\{\underline{\sigma}_k\}} \right]\\
&+ \frac{\partial J}{\partial \mathrm{Im}\{y\}}\left[\frac{\partial \mathrm{Im}\{y\}}{\partial \mathrm{Re}\{\varphi_k\}}\frac{\partial \mathrm{Re}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \mathrm{Re}\{\underline{\sigma}_k\}} + \frac{\partial \mathrm{Im}\{y\}}{\partial \mathrm{Re}\{\varphi_k\}}\frac{\partial \mathrm{Re}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \mathrm{Re}\{\underline{\sigma}_k\}} \right.\\
&\left. + \frac{\partial \mathrm{Im}\{y\}}{\partial \mathrm{Im}\{\varphi_k\}}\frac{\partial \mathrm{Im}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \mathrm{Re}\{\underline{\sigma}_k\}} + \frac{\partial \mathrm{Im}\{y\}}{\partial \mathrm{Im}\{\varphi_k\}}\frac{\partial \mathrm{Im}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \mathrm{Re}\{\varphi_k\}} \right]
\end{aligned}
\tag{2.87}
$$

And the derivative of the cost function with respect to the imaginary part of $\underline{\sigma}$ is given by

$$
\begin{aligned}
\frac{\partial J}{\partial \mathrm{Im}\{\underline{\sigma}_k\}} = &\frac{\partial J}{\partial \mathrm{Re}\{y\}}\left[\frac{\partial \mathrm{Re}\{y\}}{\partial \mathrm{Re}\{\varphi_k\}}\frac{\partial \mathrm{Re}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \mathrm{Im}\{\underline{\sigma}_k\}} + \frac{\partial \mathrm{Re}\{y\}}{\partial \mathrm{Re}\{\varphi_k\}}\frac{\partial \mathrm{Re}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \mathrm{Im}\{\underline{\sigma}_k\}} \right.\\
&\left. + \frac{\partial \mathrm{Re}\{y\}}{\partial \mathrm{Im}\{\varphi_k\}}\frac{\partial \mathrm{Im}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \mathrm{Im}\{\underline{\sigma}_k\}} + \frac{\partial \mathrm{Re}\{y\}}{\partial \mathrm{Im}\{\varphi_k\}}\frac{\partial \mathrm{Im}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \mathrm{Im}\{\underline{\sigma}_k\}} \right]\\
&+ \frac{\partial J}{\partial \mathrm{Im}\{y\}}\left[\frac{\partial \mathrm{Im}\{y\}}{\partial \mathrm{Re}\{\varphi_k\}}\frac{\partial \mathrm{Re}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \mathrm{Im}\{\underline{\sigma}_k\}} + \frac{\partial \mathrm{Im}\{y\}}{\partial \mathrm{Re}\{\varphi_k\}}\frac{\partial \mathrm{Re}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \mathrm{Im}\{\underline{\sigma}_k\}} \right.\\
&\left. + \frac{\partial \mathrm{Im}\{y\}}{\partial \mathrm{Im}\{\varphi_k\}}\frac{\partial \mathrm{Im}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \mathrm{Im}\{\underline{\sigma}_k\}} + \frac{\partial \mathrm{Im}\{y\}}{\partial \mathrm{Im}\{\varphi_k\}}\frac{\partial \mathrm{Im}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \mathrm{Im}\{\varphi_k\}} \right]
\end{aligned}
\tag{2.88}
$$

Identifying the partial derivatives of Eq. (2.87) and (2.88), and from Eq. (2.82), (2.76) and from the Cauchy-Riemann equation (ALPAY, 2016) that states the following relation $f_a'(z) = \mathrm{Re}\{f_a\} = -j\mathrm{Im}\{f_a\}$, leads to

$$\frac{\partial J}{\partial \underline{\sigma}_k} = e(n)w_k^*f_a'^*\left(\underline{\sigma}_k^T\left(\underline{u}-\underline{t}_k\right)\right)\left(\underline{u}-\underline{t}_k\right)^* \tag{2.89}$$

where $f_a'^*$ is the complex-conjugate of the derivative of the function f_a.

And hence

$$\underline{\sigma}_k\left(n+1\right)=\underline{\sigma}_k\left(n\right)+\mu_\sigma e\left(n\right)w_k^* f_a^{*'}\left(\underline{\sigma}_k^T\left(\underline{u}-\underline{t}_k\right)\right)\left(\underline{u}-\underline{t}_k\right)^* \tag{2.90}$$

Similarly to the Eq. (2.32), the gradient descent update rule for the kth kernel center $\underline{t}_k(n+1)$ is given by

$$\underline{t}\left(n+1\right)=\underline{t}\left(n\right)-\mu_w\underline{\nabla}J\left(\underline{t}\left(n\right)\right) \tag{2.91}$$

Similar derivation for the update of the radial basis centers $\underline{t}$ follows to

$$\nabla_k J = \frac{\partial J}{\partial \underline{t}_k} = \frac{\partial J}{\partial \operatorname{Re}\{\underline{t}_k\}} + j\frac{\partial J}{\partial \operatorname{Im}\{\underline{t}_k\}} \tag{2.92}$$

Where the derivative of the cost function with respect to the real part of $\underline{t}_k$ is given by Eq. (2.93).

$$
\begin{aligned}
\frac{\partial J}{\partial \operatorname{Re}\{\underline{t}_k\}} = \frac{\partial J}{\partial \operatorname{Re}\{y\}}&\left[\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{\varphi_k\}}\frac{\partial \operatorname{Re}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \operatorname{Re}\{\underline{t}_k\}}+\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{\varphi_k\}}\frac{\partial \operatorname{Re}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \operatorname{Re}\{\underline{t}_k\}}\right.\\
&\left.+\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{\varphi_k\}}\frac{\partial \operatorname{Im}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \operatorname{Re}\{\underline{t}_k\}}+\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{\varphi_k\}}\frac{\partial \operatorname{Im}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \operatorname{Re}\{\underline{t}_k\}}\right]\\
+\frac{\partial J}{\partial \operatorname{Im}\{y\}}&\left[\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{\varphi_k\}}\frac{\partial \operatorname{Re}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \operatorname{Re}\{\underline{t}_k\}}+\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{\varphi_k\}}\frac{\partial \operatorname{Re}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \operatorname{Re}\{\underline{t}_k\}}\right.\\
&\left.+\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{\varphi_k\}}\frac{\partial \operatorname{Im}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \operatorname{Re}\{\underline{t}_k\}}+\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{\varphi_k\}}\frac{\partial \operatorname{Im}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \operatorname{Re}\{\underline{t}_k\}}\right]
\end{aligned}
\tag{2.93}
$$

And the derivative of the cost function with respect to the imaginary part of $\underline{t}_k$ is given by Eq. (2.94).

$$
\begin{aligned}
\frac{\partial J}{\partial \operatorname{Im}\{\underline{t}_k\}} = \frac{\partial J}{\partial \operatorname{Re}\{y\}}&\left[\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{\varphi_k\}}\frac{\partial \operatorname{Re}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \operatorname{Im}\{\underline{t}_k\}}+\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{\varphi_k\}}\frac{\partial \operatorname{Re}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \operatorname{Im}\{\underline{t}_k\}}\right.\\
&\left.+\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{\varphi_k\}}\frac{\partial \operatorname{Im}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \operatorname{Im}\{\underline{t}_k\}}+\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{\varphi_k\}}\frac{\partial \operatorname{Im}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \operatorname{Im}\{\underline{t}_k\}}\right]\\
+\frac{\partial J}{\partial \operatorname{Im}\{y\}}&\left[\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{\varphi_k\}}\frac{\partial \operatorname{Re}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \operatorname{Im}\{\underline{t}_k\}}+\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{\varphi_k\}}\frac{\partial \operatorname{Re}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \operatorname{Im}\{\underline{t}_k\}}\right.\\
&\left.+\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{\varphi_k\}}\frac{\partial \operatorname{Im}\{\varphi_k\}}{\partial p_k}\frac{\partial p_k}{\partial \operatorname{Im}\{\underline{t}_k\}}+\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{\varphi_k\}}\frac{\partial \operatorname{Im}\{\varphi_k\}}{\partial q_k}\frac{\partial q_k}{\partial \operatorname{Im}\{\underline{t}_k\}}\right]
\end{aligned}
\tag{2.94}
$$

Identifying the partial derivatives of Eq. (2.93) and (2.94), and from Eq. (2.82) and from the Cauchy-Riemann equation that states the following relation $f_a'(z) = \operatorname{Re}\{f_a\} = -j\operatorname{Im}\{f_a\}$, leads to Eq. (2.95).

$$\underline{t}_k\left(n+1\right)=\underline{t}_k\left(n\right)-\mu_t e\left(n\right)w_k^* f_a^{*'}\left(\underline{\sigma}_k^T\left(\underline{u}-\underline{t}_k\right)\right)\underline{\sigma}_k^* \tag{2.95}$$

Thus, the complex-valued gradient update rule for the three free parameters $(\underline{t}, w, \underline{\sigma})$ of the ANN FC-RBF have been summarized as Eq. (2.84), (2.90) and (2.95).

2.4.6 PT-RBF

Developing efficient neural network algorithms for function approximation, classification and array signal processing requires a fully complex-valued networks, capable of handling complex-valued inputs with complex-valued weights and thresholds, where not only the free parameters but the output $y(n)$ and the instantaneous error $e(n)$, are complex numbers. The Phase Transmittance Radial Basis Function (PT-RBF) is another ANN fully complex RBF of this kind. It was proposed by (LOSS et al., 2007) in order to perform the process of channel equalization on quadrature digital modulation systems over wireless communications. This thesis proposes the use of an ANN PT-RBF suitable to process the complex valued information necessary for adaptive beamforming algorithms. When compared to the ordinary C-RBF, The PT-RBF allows operations with fully complex numbers on the activation functions and on the update equations. When compared to the FC-RBF, the PT-RBF operates with a different kernel (a Gaussian one). The activation function φ used in the kth hidden neuron of the PT-RBF, described by Eq. (2.96), is also fully complex, and although it has the real and imaginary part separated, both parts consider each other.

$$\varphi_k(n) = e^{\left(\frac{-1}{\text{Re}\{\sigma_k^2(n)\}}\left\|\text{Re}\{\underline{u}(n)\} - \text{Re}\{\underline{t}_k(n)\}\right\|^2\right)} + je^{\left(\frac{-1}{\text{Im}\{\sigma_k^2(n)\}}\left\|\text{Im}\{\underline{u}(n)\} - \text{Im}\{\underline{t}_k(n)\}\right\|^2\right)} \tag{2.96}$$

And once again, just like for the FC-RBF in Eq. (2.76), the output y of the ANN can be written as

$$\begin{aligned} y = \sum_{k=1}^{K} w_k \varphi_k &= \sum_{k=1}^{K} (Re\{w_k\} + jIm\{w_k\})\,(Re\{\varphi_k\} + jIm\{\varphi_k\}) \\ &= \sum_{k=1}^{K} (Re\{w_k\}Re\{\varphi_k\} - jIm\{w_k\}Im\{\varphi_k\}) + j\,(Re\{w_k\}Im\{\varphi_k\} + jIm\{w_k\}Re\{\varphi_k\}) \end{aligned} \tag{2.97}$$

In order to use this algorithm in a beamforming context, it is necessary to go over each part. This section presents the gradient descent-based update rules for the PT-RBF (LOSS et al., 2007).

Remembering from Eq. (2.35), the cost function J can be written as

$$J(n) = \frac{1}{2}e^2(n) \tag{2.35}$$

Once again, the error $e(n)$ is still given by Eq. (2.31) and both y and Y are complex values.

Similarly to the Eq. (2.32), the gradient descent update rule for the kth synapses vector $\underline{w}_k(n+1)$ is given by

$$w(n+1) = w(n) - \mu_w \underline{\nabla} J(w(n))$$

(2.98)

The cost function J is given by (2.35) and the error equation is given by Eq. (2.31). From Eq. (2.98), in order to determine $w(n+1)$, it is necessary to find $\underline{\nabla} J$, so it is possible to write

$$\nabla_k J = \frac{\partial J}{\partial w_k} = \frac{\partial J}{\partial \operatorname{Re}\{w_k\}} + j \frac{\partial J}{\partial \operatorname{Im}\{w_k\}}$$

(2.99)

Using the chain rule, the derivative of the cost function with respect to the real part of w_k is given by

$$\frac{\partial J}{\partial \operatorname{Re}\{w_k\}} = \frac{\partial J}{\partial \operatorname{Re}\{y\}} \frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{w_k\}} + \frac{\partial J}{\partial \operatorname{Im}\{y\}} \frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{w_k\}}$$

(2.100)

Substituting Eq. (2.97) and (2.31) into Eq. (2.100), it is possible to write

$$\frac{\partial}{\partial \operatorname{Re}\{w_k\}}(\Upsilon - y)^2 = -2(\operatorname{Re}\{\Upsilon\} - \operatorname{Re}\{y\})\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{w_k\}} - 2(\operatorname{Im}\{\Upsilon\} - \operatorname{Im}\{y\})\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{w_k\}}$$

(2.101)

As the desired output y does not depend on the synaptic weights w, and the derivative of $\partial/\partial w_k$ will only exist for the current neuron k, then it is possible to expand Eq. (2.101) into

$$\frac{\partial}{\partial \operatorname{Re}\{w_k\}}(\Upsilon - y)^2 = -2(\operatorname{Re}\{\Upsilon\} - \operatorname{Re}\{y\})\operatorname{Re}\{\varphi_k\} - 2(\operatorname{Im}\{\Upsilon\} - \operatorname{Im}\{y\})\operatorname{Im}\{\varphi_k\}$$

(2.102)

Now, the derivative of the cost function with respect to the imaginary part of w_k using the chain rule is given by

$$\frac{\partial J}{\partial \operatorname{Im}\{w_k\}} = \frac{\partial J}{\partial \operatorname{Re}\{y\}} \frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{w_k\}} + \frac{\partial J}{\partial \operatorname{Im}\{y\}} \frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{w_k\}}$$

(2.103)

Substituting Eq. (2.97) and (2.31) into Eq. (2.100), it is possible to write

$$\frac{\partial}{\partial \operatorname{Im}\{w_k\}}(\Upsilon - y)^2 = -2(\operatorname{Re}\{\Upsilon\} - \operatorname{Re}\{y\})\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{w_k\}} - 2(\operatorname{Im}\{\Upsilon\} - \operatorname{Im}\{y\})\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{w_k\}}$$

(2.104)

As the desired output Υ does not depend on the synaptic weights w, and the derivative of $\partial/\partial w_k$ will only exist for the current neuron k, then it is possible to expand Eq. (2.104) to

$$\frac{\partial}{\partial \operatorname{Im}\{w_k\}}(\Upsilon - y)^2 = -2(\operatorname{Re}\{\Upsilon\} - \operatorname{Re}\{y\})\operatorname{Im}\{\varphi_k\} - 2(\operatorname{Im}\{\Upsilon\} - \operatorname{Im}\{y\})\operatorname{Re}\{\varphi_k\}$$

(2.105)

Finally, replacing both Eq. (2.102) and (2.105) which are the derivatives of the real and imaginary error e of the cost function J of Eq. (2.35) into Eq. (2.99), it is possible to find

$$\nabla_k J = -\left[\operatorname{Re}\{e(n)\}\operatorname{Re}\{\varphi_k(n)\} + \operatorname{Im}\{e(n)\}\operatorname{Im}\{\varphi_k(n)\}\right] \\ - j\left[\operatorname{Im}\{e(n)\}\operatorname{Re}\{\varphi_k(n)\} - \operatorname{Re}\{e(n)\}\operatorname{Im}\{\varphi_k(n)\}\right]$$

(2.106)

but

$$e(n)\varphi_k^*(n) = \left(\operatorname{Re}\{e(n)\}\operatorname{Re}\{\varphi_k(n)\} + \operatorname{Im}\{e(n)\}\operatorname{Im}\{\varphi_k(n)\}\right) \\ + j\left(\operatorname{Im}\{e(n)\}\operatorname{Re}\{\varphi_k(n)\} - \operatorname{Re}\{e(n)\}\operatorname{Im}\{\varphi_k(n)\}\right)$$

(2.107)

It is possible then to rewrite Eq. (2.106) as

$$\nabla_k J = e(n)\varphi_k^*(n)$$

(2.108)

Finally, replacing Eq. (2.108) into Eq. (2.98), it is possible to obtain the update equation to the synaptic weight, given by

$$w_k(n+1) = w_k(n) - \mu_w e(n)\varphi_k^*(n)$$

(2.109)

Similarly to the Eq. (2.32), the gradient descent update rule for the kth kernel center $\underline{t}_k(n+1)$ is given by

$$\underline{t}(n+1) = \underline{t}(n) - \mu_t \nabla J(\underline{t}(n))$$

(2.110)

Similar derivation for the update of the radial basis centers $\underline{t}$ follows to

$$\nabla_k J = \frac{\partial J}{\partial \underline{t}_k} = \frac{\partial J}{\partial \operatorname{Re}\{\underline{t}_k\}} + \frac{\partial J}{\partial \operatorname{Im}\{\underline{t}_k\}}$$

(2.111)

In order to solve Eq. (2.111), it is necessary to find its two derivatives of the cost function J with respect to the real and imaginary part of the radial basis center $\underline{t}_k$. As the Cost function J is written as a function of the output y, it is then necessary to find its two derivatives of the output y with respect to the real and imaginary part of the radial basis center $\underline{t}_k$.

Using the chain rule on Eq. (2.111), the derivative of the cost function J with respect to the real part of $\underline{t}_k$ is given by:

$$\frac{\partial J}{\partial \operatorname{Re}\{\underline{t}_k\}} = \frac{\partial J}{\partial \operatorname{Re}\{y\}}\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{\underline{t}_k\}} + \frac{\partial J}{\partial \operatorname{Im}\{y\}}\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{\underline{t}_k\}}$$

(2.112)

Where is evident the need for the derivative of the output function y with respect to the real part of the radial basis center $\underline{t}_k$.

Using the chain rule again on Eq. (2.111), the derivative of the cost function J with respect to the imaginary part of $\underline{t}_k$ is given by:

$$\frac{\partial J}{\partial \operatorname{Im}\{\underline{t}_k\}} = \frac{\partial J}{\partial \operatorname{Re}\{y\}}\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{\underline{t}_k\}} + \frac{\partial J}{\partial \operatorname{Im}\{y\}}\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{\underline{t}_k\}} \qquad (2.113)$$

Where it is again evident the need for the derivative of the output function y with respect to the imaginary part of the radial basis center $\underline{t}_k$.

Now, considering (2.97) and (2.31), the derivatives of the cost function J with respect to the real and imaginary part of $\underline{t}_k$ can be written as following

$$\frac{\partial}{\partial \operatorname{Re}\{\underline{t}_k\}}\left(\Upsilon - y\right)^2 = -2\left(\operatorname{Re}\{\Upsilon\} - \operatorname{Re}\{y\}\right)\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{\underline{t}_k\}} - 2\left(\operatorname{Im}\{\Upsilon\} - \operatorname{Im}\{y\}\right)\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{\underline{t}_k\}} \qquad (2.114)$$

$$\frac{\partial}{\partial \operatorname{Im}\{\underline{t}_k\}}\left(\Upsilon - y\right)^2 = -2\left(\operatorname{Re}\{\Upsilon\} - \operatorname{Re}\{y\}\right)\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{\underline{t}_k\}} - 2\left(\operatorname{Im}\{\Upsilon\} - \operatorname{Im}\{y\}\right)\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{\underline{t}_k\}} \qquad (2.115)$$

Moving forward, ìn order to solve Eq. (2.111), the four derivatives of the output y with respect to the radial basis center $\underline{t}_k$ must derived.

Derivatives of the output y with respect to the real part of the radial basis center $\underline{t}_k$ on

$$\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{\underline{t}_k\}} = \frac{\partial}{\partial \operatorname{Re}\{\underline{t}_k\}}\left(\sum_{k=1}^{K} \operatorname{Re}\{w_k\} e^{-\frac{\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right)^2}{\operatorname{Re}\{\sigma_k^2\}}}\right) = \frac{2\operatorname{Re}\{w_k\}\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right)}{\operatorname{Re}\{\sigma_k^2\}}\operatorname{Re}\{\varphi_k\} \qquad (2.116)$$

$$\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{\underline{t}_k\}} = \frac{\partial}{\partial \operatorname{Re}\{\underline{t}_k\}}\left(\sum_{k=1}^{K} \operatorname{Im}\{w_k\} e^{-\frac{\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right)^2}{\operatorname{Re}\{\sigma_k^2\}}}\right) = \frac{2\operatorname{Im}\{w_k\}\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right)}{\operatorname{Re}\{\sigma_k^2\}}\operatorname{Re}\{\varphi_k\} \qquad (2.117)$$

Derivatives of the output y with respect to the imanigary part of the radial basis center $\underline{t}_k$ on

$$\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{\underline{t}_k\}} = \frac{-\partial}{\partial \operatorname{Im}\{\underline{t}_k\}}\left(\sum_{k=1}^{K} \operatorname{Im}\{w_k\} e^{-\frac{\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)^2}{\operatorname{Im}\{\sigma_k^2\}}}\right) = \frac{-2\operatorname{Im}\{w_k\}\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)}{\operatorname{Im}\{\sigma_k^2\}}\operatorname{Im}\{\varphi_k\}$$

$$(2.118)$$

$$\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{\underline{t}_k\}} = \frac{\partial}{\partial \operatorname{Im}\{\underline{t}_k\}}\left(\sum_{k=1}^{K} \operatorname{Re}\{w_k\} e^{-\frac{\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)^2}{\operatorname{Im}\{\sigma_k^2\}}}\right) = \frac{2\operatorname{Re}\{w_k\}\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)}{\operatorname{Im}\{\sigma_k^2\}}\operatorname{Im}\{\varphi_k\} \qquad (2.119)$$

Formulating Eq. (2.114) with Eq. (2.116) and (2.117)

$$\frac{\partial}{\partial \operatorname{Re}\{\underline{t}_k\}} (\Upsilon - y)^2$$

$$= -4 \operatorname{Re}\{\varphi_k\} \frac{\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right)}{\operatorname{Re}\{\sigma_k^2\}} \left[\operatorname{Re}\{w_k\}\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right) + \operatorname{Im}\{w_k\}\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)\right]$$

$$(2.120)$$

And similarly, formulating Eq. (2.115) with Eq. (2.118) and (2.119)

$$\frac{\partial}{\partial \operatorname{Im}\{\underline{t}_k\}} (\Upsilon - y)^2$$

$$= -4 \operatorname{Im}\{\varphi_k\} \frac{\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)}{\operatorname{Im}\{\sigma_k^2\}} \left[\operatorname{Im}\{w_k\}\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right) - \operatorname{Re}\{w_k\}\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)\right]$$

$$(2.121)$$

Continuing on solving Eq. (2.111), considering Eq. (2.120) and (2.121), it is then possible to write

$$\nabla_k J = 2 \operatorname{Re}\{\varphi_k\} \frac{\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right)}{\operatorname{Re}\{\sigma_k^2\}} \left[\operatorname{Re}\{w_k\}\operatorname{Re}\{e\} + \operatorname{Im}\{w_k\}\operatorname{Im}\{e\}\right]$$

$$+ j2 \operatorname{Im}\{\varphi_k\} \frac{\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)}{\operatorname{Im}\{\sigma_k^2\}} \left[\operatorname{Re}\{w_k\}\operatorname{Im}\{e\} - \operatorname{Im}\{w_k\}\operatorname{Re}\{e\}\right]$$

$$(2.122)$$

Finally, writing Eq. (2.122) into Eq. (2.110), it is possible to obtain the update equation to the radial basis center $\underline{t}_k$, given by

$$\underline{t}_k(n+1) = \underline{t}_k(n) + 2\mu_t \left\{ \operatorname{Re}\{\varphi_k(n)\} \frac{\operatorname{Re}\{\underline{u}(n)\} - \operatorname{Re}\{\underline{t}_k(n)\}}{\operatorname{Re}\{\sigma_k^2(n)\}} \left(\begin{array}{l} \operatorname{Re}\{w_k(n)\}\operatorname{Re}\{e(n)\} \\ + \operatorname{Im}\{w_k(n)\}\operatorname{Im}\{e(n)\} \end{array} \right) \right.$$

$$\left. + j \operatorname{Im}\{\varphi_k(n)\} \frac{\operatorname{Im}\{\underline{u}(n)\} - \operatorname{Im}\{\underline{t}_k(n)\}}{\operatorname{Im}\{\sigma_k^2(n)\}} \left(\begin{array}{l} \operatorname{Re}\{w_k(n)\}\operatorname{Im}\{e(n)\} \\ - \operatorname{Im}\{w_k(n)\}\operatorname{Re}\{e(n)\} \end{array} \right) \right\}$$

$$(2.123)$$

Similarly to the Eq. (2.32), the gradient descent update rule for the kth Gaussian kernel variance $\sigma_k^2(n+1)$ is given by

$$\alpha(n+1) = \alpha(n) - \mu_\sigma \underline{\nabla} J(\alpha(n))$$

$$(2.124)$$

where $\sigma^2 = \alpha$, for convenience.

Similar derivation for the update of the radial basis centers $\underline{t}$ follows to

$$\nabla_k J = \frac{\partial J}{\partial \alpha_k} = \frac{\partial J}{\partial \operatorname{Re}\{\alpha_k\}} + j \frac{\partial J}{\partial \operatorname{Im}\{\alpha_k\}}$$

$$(2.125)$$

In order to solve Eq. (2.125), it is necessary to find its two derivatives of the cost function J with respect to the real and imaginary part of α. As the Cost function J is written as

a function of the output y, it is then necessary to find its two derivatives of the output y with respect to the real and imaginary part of α.

Using the chain rule on Eq. (2.125), the derivative of the cost function J with respect to the real part of α_k is given by

$$\frac{\partial J}{\partial \operatorname{Re}\{\alpha_k\}} = \frac{\partial J}{\partial \operatorname{Re}\{y\}} \frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{\alpha_k\}} + \frac{\partial J}{\partial \operatorname{Im}\{y\}} \frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{\alpha_k\}} \tag{2.126}$$

Where is evident the need for the derivative of the output function y with respect to the real part of α.

Using the chain rule again on Eq. (2.125), the derivative of the cost function J with respect to the imaginary part of α_k is given by

$$\frac{\partial J}{\partial \operatorname{Im}\{\alpha_k\}} = \frac{\partial J}{\partial \operatorname{Re}\{y\}} \frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{\alpha_k\}} + \frac{\partial J}{\partial \operatorname{Im}\{y\}} \frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{\alpha_k\}} \tag{2.127}$$

Where it is again evident the need for the derivative of the output function y with respect to the imaginary part of α.

Now, considering Eq. (2.97) and (2.31), the derivatives of the cost function J with respect to the real and imaginary part of α_k can be written as following

$$\frac{\partial}{\partial \operatorname{Re}\{\alpha_k\}} (\Upsilon - y)^2 = -2\big(\operatorname{Re}\{\Upsilon\} - \operatorname{Re}\{y\}\big) \frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{\alpha_k\}} - 2\big(\operatorname{Im}\{\Upsilon\} - \operatorname{Im}\{y\}\big) \frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{\alpha_k\}} \tag{2.128}$$

$$\frac{\partial}{\partial \operatorname{Im}\{\alpha_k\}} (\Upsilon - y)^2 = -2\big(\operatorname{Re}\{\Upsilon\} - \operatorname{Re}\{y\}\big) \frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{\alpha_k\}} - 2\big(\operatorname{Im}\{\Upsilon\} - \operatorname{Im}\{y\}\big) \frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{\alpha_k\}} \tag{2.129}$$

Moving forward, just as the derivations for the radial basis center $\underline{t}$, in order to solve Eq. (2.125), the four derivatives of the output y with respect to α_k must derived.

Derivatives of the output y with respect to the real part of α_k on

$$\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Re}\{\alpha_k\}} = \frac{\partial}{\partial \operatorname{Re}\{\alpha_k\}} \left(\sum_{k=1}^{K} \operatorname{Re}\{w_k\} e^{-\frac{(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\})^2}{\operatorname{Re}\{\alpha_k\}}} \right)$$

$$= \frac{\operatorname{Re}\{w_k\}\big(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\big)^2}{\operatorname{Re}\{\alpha_k^2\}} e^{-\frac{(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\})^2}{\operatorname{Re}\{\alpha_k\}}} \tag{2.130}$$

$$\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Re}\{\alpha_k\}} = \frac{\partial}{\partial \operatorname{Re}\{\alpha_k\}} \left(\sum_{k=1}^{K} \operatorname{Im}\{w_k\} e^{-\frac{(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\})^2}{\operatorname{Re}\{\alpha_k\}}} \right)$$

$$= \frac{\operatorname{Im}\{w_k\}\big(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\big)^2}{\operatorname{Re}\{\alpha_k^2\}} e^{-\frac{(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\})^2}{\operatorname{Re}\{\alpha_k\}}} \tag{2.131}$$

Derivatives of the output y with respect to the imanigary part of the radial basis center $\underline{t}_k$ on

$$\frac{\partial \operatorname{Re}\{y\}}{\partial \operatorname{Im}\{\alpha_k\}} = \frac{-\partial}{\partial \operatorname{Im}\{\alpha_k\}}\left(\sum_{k=1}^{K} \operatorname{Im}\{w_k\} e^{-\frac{(\operatorname{Im}\{\underline{u}\}-\operatorname{Im}\{\underline{t}_k\})^2}{\operatorname{Im}\{\alpha_k^2\}}}\right)$$

$$= -\frac{\operatorname{Im}\{w_k\}\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)^2}{\operatorname{Im}\{\alpha_k^2\}} e^{-\frac{(\operatorname{Im}\{\underline{u}\}-\operatorname{Im}\{\underline{t}_k\})^2}{\operatorname{Im}\{\alpha_k\}}}$$

(2.132)

$$\frac{\partial \operatorname{Im}\{y\}}{\partial \operatorname{Im}\{\alpha_k\}} = \frac{\partial}{\partial \operatorname{Im}\{\alpha_k\}}\left(\sum_{k=1}^{K} \operatorname{Re}\{w_k\} e^{-\frac{(\operatorname{Im}\{\underline{u}\}-\operatorname{Im}\{\underline{t}_k\})^2}{\operatorname{Im}\{\alpha_k^2\}}}\right)$$

$$= \frac{\operatorname{Re}\{w_k\}\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)^2}{\operatorname{Im}\{\alpha_k^2\}} e^{-\frac{(\operatorname{Im}\{\underline{u}\}-\operatorname{Im}\{\underline{t}_k\})^2}{\operatorname{Im}\{\alpha_k\}}}$$

(2.133)

Formulating Eq. (2.128) with Eq. (2.130) and (2.131)

$$\frac{\partial}{\partial \operatorname{Re}\{\alpha_k\}}(\Upsilon - y)^2 = -2\operatorname{Re}\{\varphi_k\}\frac{\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right)}{\operatorname{Re}\{\alpha_k^2\}}$$

$$\left[\operatorname{Re}\{w_k\}\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right) + \operatorname{Im}\{w_k\}\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)\right]$$

(2.134)

And similarly, formulating Eq. (2.129) with Eq. (2.132) and (2.133)

$$\frac{\partial}{\partial \operatorname{Im}\{\alpha_k\}}(\Upsilon - y)^2$$

$$= 2\operatorname{Im}\{\varphi_k\}\frac{\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)}{\operatorname{Im}\{\alpha_k^2\}}\left[\operatorname{Im}\{w_k\}\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right) - \operatorname{Re}\{w_k\}\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)\right]$$

(2.135)

Continuing on solving (2.125), considering Eq. (2.134) and (2.135), it is then possible to write

$$\nabla_k J = -\operatorname{Re}\{\varphi_k\}\frac{\left(\operatorname{Re}\{\underline{u}\} - \operatorname{Re}\{\underline{t}_k\}\right)}{\operatorname{Re}\{\alpha_k^2\}}\left[\operatorname{Re}\{w_k\}\operatorname{Re}\{e\} + \operatorname{Im}\{w_k\}\operatorname{Re}\{e\}\right]$$

$$+ j\operatorname{Im}\{\varphi_k\}\frac{\left(\operatorname{Im}\{\underline{u}\} - \operatorname{Im}\{\underline{t}_k\}\right)}{\operatorname{Im}\{\alpha_k^2\}}\left[\operatorname{Im}\{w_k\}\operatorname{Re}\{e\} - \operatorname{Re}\{w_k\}\operatorname{Re}\{e\}\right]$$

(2.136)

Finally, writing Eq. (2.136) into Eq. (2.124), and considering that $\alpha_k = \sigma_k^2$, it is possible to obtain the update equation to the radial basis center σ_k^2, given by

$$\sigma_k^2(n+1) = \sigma_k^2(n) + \mu_\sigma \left\{ \left| \text{Re}\{\underline{u}(n)\} - \text{Re}\{\underline{t}_k(n)\} \right|^2 \frac{\text{Re}\{\varphi_k(n)\}}{\text{Re}\{\sigma_k^2(n)\}^2} \left(\begin{array}{c} \text{Re}\{w_k(n)\}\text{Re}\{e(n)\} \\ +\text{Im}\{w_k(n)\}\text{Im}\{e(n)\} \end{array} \right) \right. \tag{2.137}$$

$$\left. + j \left| \text{Im}\{\underline{u}(n)\} - \text{Im}\{\underline{t}_k(n)\} \right|^2 \frac{\text{Im}\{\varphi_k(n)\}}{\text{Im}\{\sigma_k^2(n)\}^2} \left(\begin{array}{c} \text{Re}\{w_k(n)\}\text{Im}\{e(n)\} \\ -\text{Im}\{w_k(n)\}\text{Re}\{e(n)\} \end{array} \right) \right\}$$

Thus, the complex-valued gradient update rule for the three free parameters $(\underline{t}, w, \underline{\sigma})$ of the ANN FC-RBF have been summarized as Eq. (2.109), (2.123) and (2.137).

Once all the fundaments for this thesis have been presented, it is time to implement the proposed algorithm. The following section shows the implementation of all ANNs presented to the beamforming problem and bring foward the parallel between them through simulation.

3 PROPOSAL

This thesis proposes a new beamforming solution based on the PT-RBF ANN (LOSS et al., 2007), to handle static and dynamic wireless operational scenarios. As the PT-RBF has phase transmittance between the input and output nodes, it presents characteristics that empower its performance when compared with RBF ANNs whose activation function is based solely on the Euclidean norm, thus minimizing the phase invariance problem. The phase transmittance between the input and output nodes of the PT-RBF ANN allows the beamforming algorithm to suppress energy radiation coming from interfering sources and to demodulate the received signal with higher gain and lower SER.

We compare the performance of the PT-RBF based beamforming algorithm with state of the art beamformers under the same operational scenarios. The heuristic followed to assess and to test the proposed PT-RBF based beamformer is detailed in this section. Furthermore, it is also applicable to all other state-of-the-art ANNs. For convenience, some equations will be brought back.

The results are gathered over distinct static and dynamic operational scenarios. The test heuristic starts by setting the DOA of the desired signal and the DOAs of the interfering signals, so establishing the operational scenario. All these signals impinge the antenna array from distinct DOA angles. Then, each scenario is exposed to three levels of SIR (low = -10dB, medium = 10dB and high = 35dB), with the SIR expressing the relationship between the desired signal power and the interference signals power. For each SIR level, non-linear distortion in the analog RF front end of the RX is also introduced, and white Gaussian noise is added to the received signal (AWGN).

The assessment of the comparative results is based on the polar diagram of the array gain, showing the gain curve in dB as a function of the DOA bearing angle. The polar diagram also includes the DOA of the desired signal and the DOAs of the interfering signals. It is expected that, ideally, the beamforming algorithm be able to establishes a polar diagram of the array gain such that the gain curve presents nulls in the bearing angles of the DOA of the interfering signals and presents a high gain lobe in the bearing angle of the DOA of the desired signal.

The assessment of the comparative results is also based on MSEA plot versus discrete time index, in order to show the comparison of the ANNs convergence rate. The assessment also uses the plot of the received desired signal symbol set in order to show the conformity with respect to the reference constellation symbols.

The DOA angle ϕ of each signal that impinges the array varies in a dynamic operational scenario, since TXs and RX are moving with respect to each other. Therefore, it is possible to estimate how a RBF ANN based beamformer is able to follow varying DOAs as a result of high speed mobile operation.

This section is organized in three parts: Section 3.1 describes the array architecture and the physical construction of the antenna used to evaluate the proposed beamformer, Section 3.2 describes the beamforming architecture and the proposed algorithm construction and finally, Section 3.3 presents the criteria used to evaluate the results of this work.

3.1 ARRAY ARCHITECTURE

Figure 20 presents the UCA, a six dipoles antenna array. Each dipole has a load impedance of $Z_T = 50\Omega$, a length of $L = \lambda/2$ and are spaced between each other by a distance of $s = \lambda/4$. The UCA arrangement allows to sweep over the spherical coordinate angle ϕ, so that it is easy to point the boresight lobe to any direction on the plane xy.

As the current trends on wireless telecommunication area are the 5G mobile network and the IoT service, and as the spectrum allocation for 5G in Brazil is still being investigated, the frequency chosen to model this thesis is based on the IoT service $f = 850MHz$. For this frequency, the dipoles length is $L = 17,634cm$ and the distance between each other dipole is $s = 8,817cm$.

Figure 20 – UCA's architecture. Drawn and simulated in CST Microwave Studio
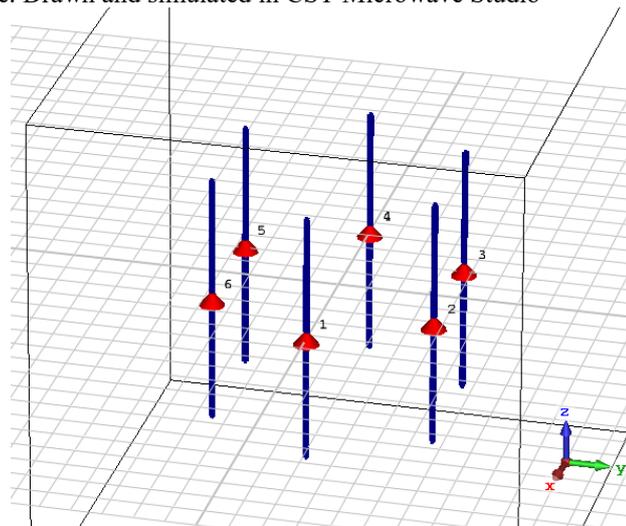
Source: (AUTHOR, 2018).

As explained in Section 2.1.2.1, the self and mutual impedances of the array must be known so the correct voltage on each element can be further calculated. Following the procedures on the mentioned section, the following impedance matrix is obtained

$$\underline{Z} = \begin{bmatrix} 78.424 + j45.545 & 46.9401 - j32.6392 & -0.791 - j41.3825 & -14.4422 - j34.4374 & -0.791 - j41.3825 & 46.9401 - 32.6392i \\ 46.9401 - j32.6392 & 78.424 + j45.545 & 46.9401 - j32.6392 & -0.791 - j41.3825 & -14.4422 - j34.4374 & -0.791 - 41.3825i \\ -0.791 - j41.3825 & 46.9401 - j32.6392 & 78.424 + j45.545 & 46.9401 - j32.6392 & -0.791 - j41.3825 & -14.4422 - 34.4374i \\ -14.4422 - j34.4374 & -0.791 - j41.3825 & 46.9401 - j32.6392 & 78.424 + j45.545 & 46.9401 - j32.6392 & -0.791 - 41.3825i \\ -0.791 - j41.3825 & -14.4422 - j34.4374 & -0.791 - j41.3825 & 46.9401 - j32.6392 & 78.424 + j45.545 & 46.9401 - 32.6392i \\ 46.9401 - j32.6392 & -0.791 - j41.3825 & -14.4422 - j34.4374 & -0.791 - j41.3825 & 46.9401 - j32.6392 & 78.424 + 45.545i \end{bmatrix}$$

where the self-impedance of each dipole is $\underline{Z}[1,1] = 78.424 + j45 \ \Omega$ and where the mutual impedance between dipoles i and j is given by $\underline{Z}[i,j]$. All matrix elements are expressed in ohms (Ω).

With the mutual impedance matrix $\underline{Z}$, the coupling matrix $\underline{C}$ can be calculated, followed by normalized response of an individual dipole to the incident wave $H(\theta, \phi)$ and finally, the received voltage $\underline{V}_T$ over Z_T.

As the UCA has 6 dipoles, then vector $\underline{V}_T$ in Eq. (2.18) has 6 components, each one representing the signal received by each respective dipole. As we shall see in the next section, these 6 voltages from $\underline{V}_T$ are associated to the vector $\underline{u}$ which represents the $M = 6$ input nodes of the RBF ANN:

$$\underline{u} = \left[u_1, u_2, \cdots, u_M \right]^T \tag{2.138}$$

3.2 BEAMFORMER IMPLEMENTATION PROCEDURES

The sequence of procedures executed by the beamforming algorithm can be summarized by the high-level flow chart of Figure 21 for the static scenario.

A desired received signal and some interference received signals impinge the UCA of 6 dipoles. The 6 resulting analog voltage signals at the terminals of each dipole are determined by Eq. (2.18), each signal being a component of vector $\underline{V}_T$.

Note that each one of the 6 received signals is firstly sampled by the respective ADC, and then it is down-converted and demodulated by the respective RX. Thus, at discrete time n, vector $\underline{V}_T$ actually stores 6 complex valued samples of the respective 6 down-converted and demodulated baseband sequences, each sample corresponding to the IQ symbol of the digital modulation that has been received at instant n by the respective dipole of the UCA.

Each baseband sequence is then distorted by a non-linear function which emulates the non-linear transmittance of the respective receiver analog RF front end. Next, these distorted

sequences are assigned to the components of vector $\underline{u}$, which represent the $M = 6$ input nodes of the RBF ANN. The RBF ANN is then trained with the TS represented by the set of vectors $\underline{u}$. For each vector $\underline{u}$ at the RBF ANN input nodes there is a reference IQ symbol as desired output. After training, the RBF ANN based beamformer is (ideally) able to generalize, yielding a radiation pattern with nulls in the bearing angle of the interference signals and a constellation of received IQ symbols from the desired signal that free of interfering signals distortion. The flow chart presented in Figure 21 summarizes the high-level procedure adopted for static scenarios.

Figure 21 – High level flow chart
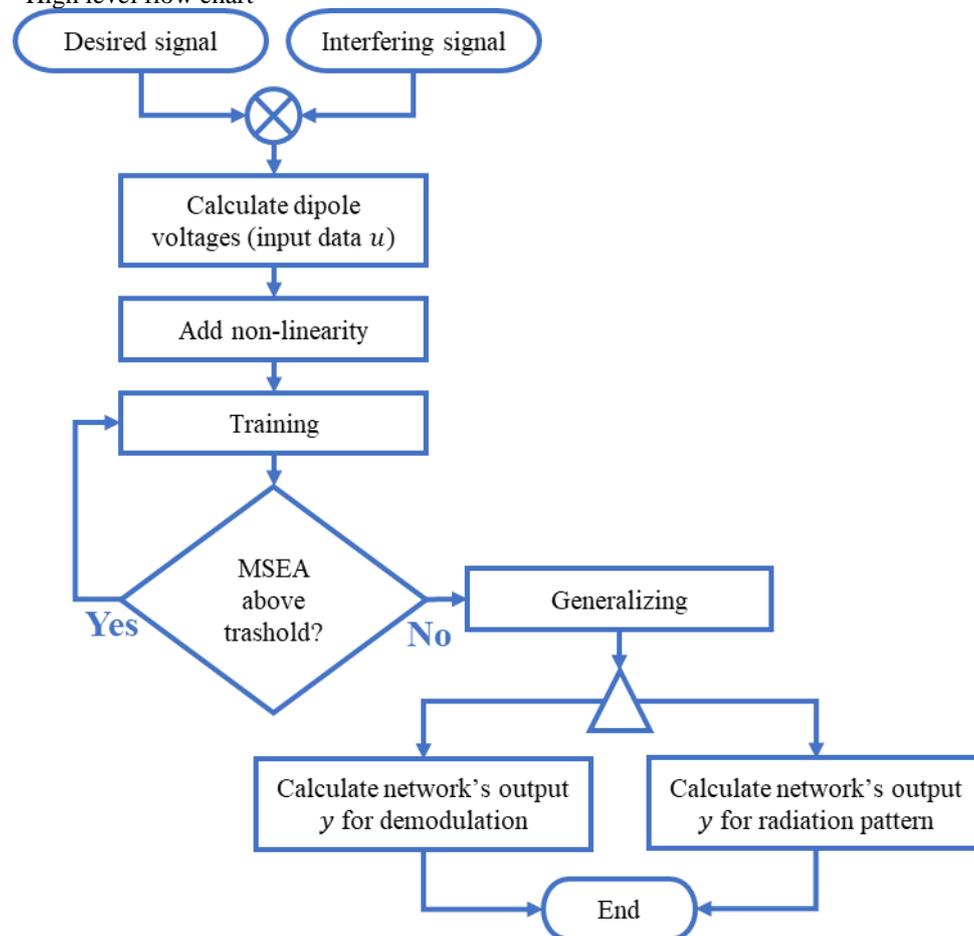
Source: (AUTHOR, 2018).

For dynamic scenarios the procedure follows the pseudocode of Table 2.

Table 2 – Dynamic PT-RBF Beamforming Pseudocode

Step	INICIALIZATION
1	Define array distribution

	$$XYZ = \begin{bmatrix} s & 0 & 0 \\ s\cos(\alpha) & s\sin(\alpha) & 0 \\ -s\cos(\alpha) & s\sin(\alpha) & 0 \\ -s & 0 & 0 \\ -s\cos(\alpha) & -s\sin(\alpha) & 0 \\ s\cos(\alpha) & -s\sin(\alpha) & 0 \end{bmatrix}$$
2	Calculate self and mutual impedances of the array elements based upon The Perturbed Biconical Method by Schelkunoff $$\underline{Z} = Z_{schelkunoff}(radius, \lambda, length)$$
3	Calculate coupling matrix $$\underline{C} = \frac{Z_T + Z_A}{\underline{Z} + Z_T\underline{I}}$$
4	Calculate Steering vectors between the K array elements and $m = 360$ points of the ϕ coordinate $$\underline{\Phi} = \left\{ \frac{\cos\left(\frac{L}{\lambda}\pi\cos\theta_m\right) - \cos\left(\frac{\pi L}{\lambda}\right)}{\sin\theta_m} \right\} e^{j\frac{2\pi}{\lambda}(x_k\sin\theta_m\cos\phi_m + y_k\sin\theta_m\sin\phi_m + z_k\cos\theta_m)}$$
5	Define training set size $$N$$
6	Define number of neurons $$K$$
7	Define learning rates $$\mu_w, \mu_\sigma, \mu_t$$
8	Initialize synaptic weights $$w = zeros[K]$$ $$t = \frac{rand[K,K]}{\max(rand[K,K])}$$ $$\sigma^2 = \frac{max(\|t[K,K]\|^2)}{max\{max(\|t[K,K]\|^2)\}}$$
10	Define number X of desired and interfering signals that will impinge on the array $$n\,D.S. = X$$ $$n\,I.S. = X$$
11	Define sequence of desired and interfering signals 16QAM $$D.S = \frac{rand(IQ\ symbols)}{max\{rand(IQ\ symbols)\}}$$ $$I.S = \frac{rand(IQ\ symbols)}{max\{rand(IQ\ symbols)\}}$$
12	Define pilot signal reference to the PT-RBF $$ref = D.S[1:N]$$
13	Define which impinging signal has its DOA varying $$D.S. \leftarrow \{DOA\ varying, DOA\ constant\}$$

	I.S. ← {DOA varying, DOA $constant$}
14	Define the DOA variation $\Delta\,DOA$ $= \pm 1^{\circ}$ shift on the ϕ coordinate of the impinging signal every R $IQ\,symbols$ received
15	Define number of IQ symbols received for a DOA variation $$R = 4000$$
16	Detect impinging signal over the array

Step	TRAINING
17	Calculate the voltage upon the array given the impinging signal at instant n considering the magnitude and phase of the instantaneous electrical field phasor $E_0(n)$ normalized to 1. $$\underline{V}(n) = \underline{\Phi}E_0(n)$$
18	Add a non-linearity to the previous voltage $$\underline{V}(n) = \underline{V}(n) - 0.1\underline{V}(n)^3 - 0.05\underline{V}(n)^5$$
19	Insert the voltage as input data to the PT-RBF $$\underline{u}(n) = \underline{V}(n)$$
20	Calculate the activation function of each neuron k at instant n $$\varphi_k(n) = e^{\left(\frac{-1}{Re\{\sigma_k^2\}}\|Re\{\underline{u}(n)\}-Re\{\underline{t}_k(n)\}\|^2\right)} + j\,e^{\left(\frac{-1}{Im\{\sigma_k^2\}}\|Im\{\underline{u}(n)\}-Im\{\underline{t}_k(n)\}\|^2\right)}$$
21	Calculate PT-RBF output $$y(n) = \sum_{k=1}^{K} w_k\varphi_k$$
22	Calculate PT-RBF error $$e(n) = ref(n) - y(n)$$
23	Update free parameters based on the error $$w_k(n+1) = w_k(n) + \mu_w e(n)\varphi_k^*(n)$$ $$\underline{t}_k(n+1) = \underline{t}_k + \mu_t\left\{ Re\{\varphi_k(n)\}\frac{[Re\{\underline{u}(n)\}-Re\{\underline{t}_k(n)\}]}{Re\{\sigma^2(n)\}}\begin{pmatrix} Re\{w_k(n)\}Re\{e(n)\} \\ +Im\{w_k(n)\}Im\{e(n)\} \end{pmatrix}\right.$$ $$\left. -Im\{\varphi_k(n)\}\frac{Im\{\underline{u}(n)\}-Im\{\underline{t}_k(n)\}]}{Im\{\sigma^2(n)\}}\begin{pmatrix} Re\{w_k(n)\}Im\{e(n)\} \\ -Im\{w_k(n)\}Re\{e(n)\} \end{pmatrix}\right\}$$ $$\sigma_k^2(n+1) = \sigma_k^2(n)$$ $$+ \mu_\sigma\left\{ |Re\{\underline{u}(n)\}-Re\{\underline{t}_k(n)\}|^2\frac{Re\{\varphi_k(n)\}}{Re\{\sigma_k^2(n)\}^2}\begin{pmatrix} Re\{w_k(n)\}Re\{e(n)\} \\ +Im\{w_k(n)\}Im\{e(n)\} \end{pmatrix}\right.$$ $$\left. +j\,|Im\{\underline{u}(n)\}-Im\{\underline{t}_k(n)\}|^2\frac{Im\{\varphi_k(n)\}}{Im\{\sigma_k^2(n)\}^2}\begin{pmatrix} Re\{w_k(n)\}Im\{e(n)\} \\ -Im\{w_k(n)\}Re\{e(n)\} \end{pmatrix}\right\}$$
24	If n < N, go back to step 17, otherwise, go further to step 25

Step	GENERALIZATION
25	Calculate the voltage upon the array given the impinging signal at instant n considering the magnitude and phase of the instantaneous electrical field phasor $E_0(n)$ $$\underline{V}(n) = \underline{\Phi}E_0(n)$$
26	Add a non-linearity to the previous voltage $$\underline{V}(n) = \underline{V}(n) - 0.1\underline{V}(n)^3 - 0.05\underline{V}(n)^5$$
27	Insert the voltage as input data to the PT-RBF $$\underline{u}(n) = \underline{V}(n)$$
28	Calculate the activation function of each neuron k at instant n $$\varphi_k(n) = e^{\left(\frac{-1}{Re\{\sigma_k^2\}}\|Re\{\underline{u}(n)\}-Re\{\underline{t}_k(n)\}\|^2\right)} + j\, e^{\left(\frac{-1}{Im\{\sigma_k^2\}}\|Im\{\underline{u}(n)\}-Im\{\underline{t}_k(n)\}\|^2\right)}$$
29	Calculate signal demodulation through the PT-RBF output $$IQ\ symbols(n) = y(n) = \sum_{k=1}^{K} w_k\varphi_k$$
30	Insert the Steering Vectors as input data to the PT-RBF $$\underline{u}(n) = \underline{\Phi}(n)$$
31	Calculate the activation function of each neuron k at instant n $$\varphi_k(n) = e^{\left(\frac{-1}{Re\{\sigma_k^2\}}\|Re\{\underline{u}(n)\}-Re\{\underline{t}_k(n)\}\|^2\right)} + j\, e^{\left(\frac{-1}{Im\{\sigma_k^2\}}\|Im\{\underline{u}(n)\}-Im\{\underline{t}_k(n)\}\|^2\right)}$$
32	Calculate the radiation pattern through the PT-RBF output $$F(n) = y(n) = \sum_{k=1}^{K} w_k\varphi_k$$
33	If DOA doesn't change (n < R), go back to step 25. Otherwise, if DOA changes (n = R) go back to step 16.

Source: (AUTHOR, 2018)

3.2.1 Input data

In order to approximate a real-world operating scenario, the desired signal and the interfering signals are generated from a stream of random bits, which are transformed into IQ symbols of a 16QAM digital modulation. Additive white Gaussian noise (AWGN) has been added to the desired signal to obtain a $35dB$ of Signal to Noise Ratio (SNR). Some interferences have also been simulated as pure WGN (no information modulated, just random noise). Furthermore, all received signals are distorted by a non-linear function which emulates the non-linear transmittance intrinsic to a real-world receiver analog RF front end. The non-linear transmittance is given in terms of the third order and fifth order intermodulation spurious,

intercept points IP3 and IP5 (LI et al., 2012; POZAR, 2012), and represented by the following polynomial:

$$OutSignal = InpSignal - 0.1 InpSignal^3 - 0.05 InpSignal^5 \qquad (2.139)$$

Figure 22 characterizes the non-linear transmittance expressed by Eq. (2.139) in terms of the intercept points IP3 and IP5 for the third order and fifth order intermodulation spurious. It is important to consider these nonlinearity parameters because the intermodulation interference caused by them cannot be solved with a simple filter as these frequencies are close to the carrier frequency and this is a realistic operational condition since linear front ends do not exist on real world.

Figure 22– Nonlinear Distortion
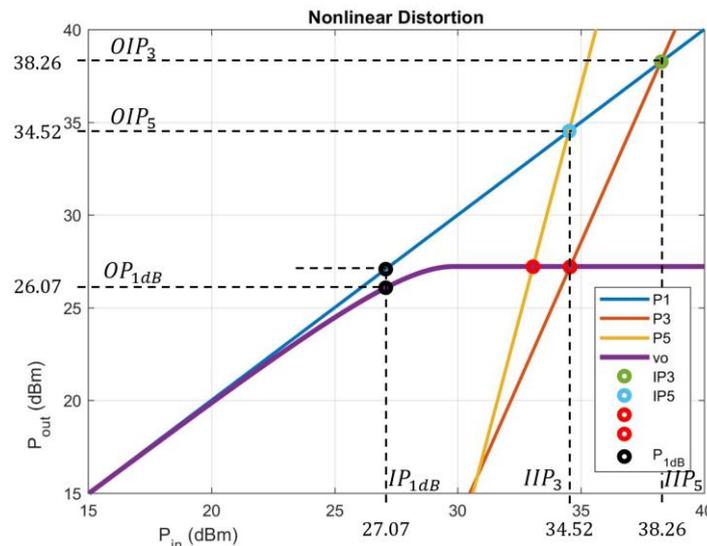
Source: (AUTHOR, 2018).

Note that OIP_3 is 12.19 dB above OP_{1dB} and that OIP_5 is 8.45 dB above OP_{1dB}, which characterizes a considerably non-linear RF front-end, possibly approximating the distortion levels of a jamming situation in the RX.

As M is the dimension of the vector $\underline{u}$, the input data $\underline{u}$ can be seen as a point in this M-dimensional space $\mathbb{C}^M$. The TS presented to the RBF ANN can be organized as a matrix composed by a collection of vectors $\underline{u}$, in other words, a matrix full of points in the M-dimensional space $\mathbb{C}^M$. Each one of these points partially describes an unknown process to be represented by the RBF ANN.

The training data set is composed by N vectors $\underline{u}$, which can vary for each scenario. The higher the complexity of the scenario (low SIR, high number of signals to null, to beam and moving signals), the higher N must be in order to the ANN to be able to reach a minimum error based on the MSEA criterion.

Once defined the number N of training vectors $\underline{u}$, it is necessary to provide the ANN with the desired output Y to the nth input vector $\underline{u}$. The desired output Y is the transmitted symbol of the 16QAM constellation normalized by the factor $(1/\sqrt{10})$, so the maximum energy per symbol never exceeds 1. Such normalization speeds up the convergence of the ANN training process.

3.2.2 Training Mode

The proposed RBF ANN based beamformer employs a reference (=desired) signal to estimate the ANN free parameters $(\underline{t}_k, w_k, \sigma_k^2)$ as shown in Figure 23. The array output (ANN input) $\underline{u}(n) = [u_1(n), u_2(n) \cdots u_6(n)]$ is processed by the non-linear mapping $\varphi_k(n)$ of the K neurons in the ANN hidden layer, whose outputs are linearly combined by the $w_k(n)$ weights yielding output $y(n)$. The output $y(n)$ is subtracted from the reference (=desired) output $Y(n)$ to generate the error signal $e(n)$, which is used to control the free parameters update. The parameters are iteratively adjusted such that the MSEA between the RBF ANN output and the reference signal is minimized.

Each discrete time instant n the input data vector $\underline{u}$ feeds the RBF ANN, all M elements are distributed to all K neurons. The number of neurons is chosen experimentally based upon a process of best fit, where multiple processes are performed with different parameters. The best result (faster convergence, lowest error), is chosen as the best fit and its neuron number is set as K.

Figure 23 - Proposed ANN architecture
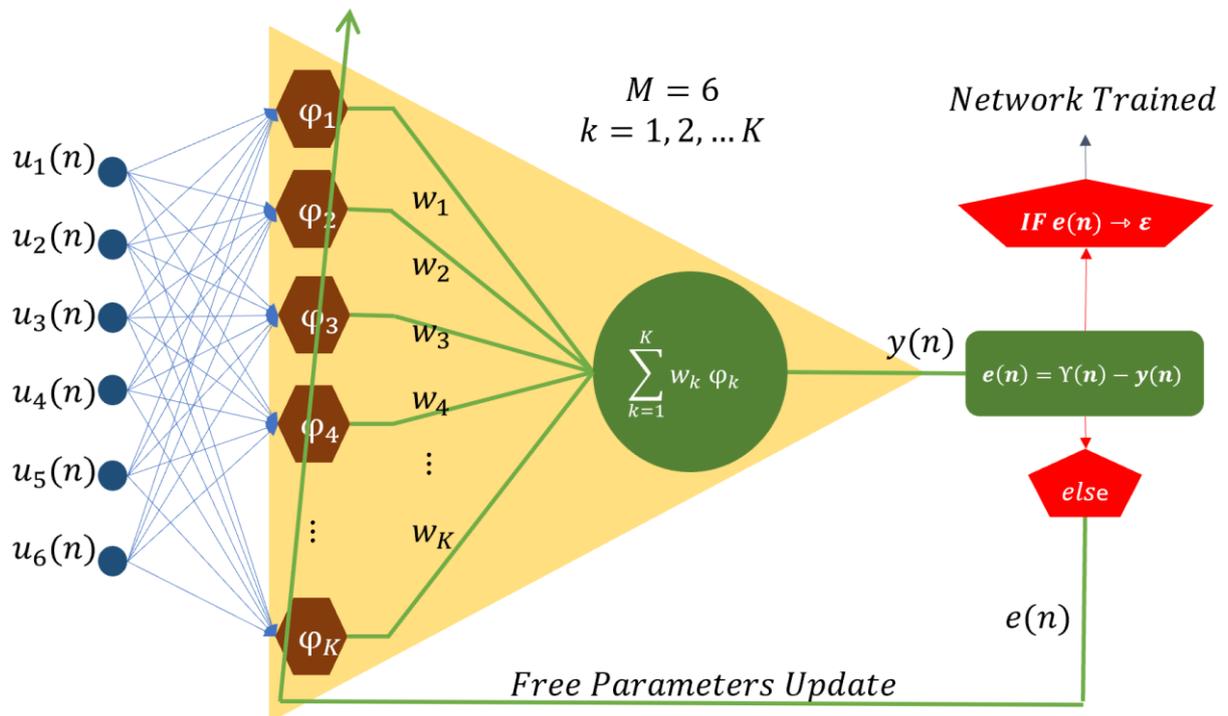
Source: (AUTHOR, 2018)

The diagram of Figure 23 expresses the training process for all three RBF ANN compared in this thesis: PT-RBF (here proposed for beamforming), C-RBF and FC-RBF.

The number K of neurons for each RBF ANN, obtained experimentally as discussed above, is:

I.   PT-RBF: $K = 16$

II.   C-RBF: $K = 11$

III.   FC-RBF: $K = 22$

IV.   The LMS algorithm, as shown in Figure 14, has a number $K = 6$ of neurons, which is equal to the size of the input vector $\underline{u}$.

The performance of any RBF ANN depends largely on the activation function adopted, the number of neurons in the hidden layer, the learning rate parameters and the initialization of the weights and radial basis centers.

For the initialization of the radial basis centers $\underline{t}_k$, it is usual to run a clustering algorithm, so it can get as close as possible to the optimum position on the M-dimensional space $\mathbb{C}^M$ (KANUNGO et al., 2002). However, as this process hasn't shown the best result, thus, in this work real and imaginary parts of each component of vector $\underline{t}_k$ have been initialized randomly, on the range $[-1,1]$. For the parameter σ_k^2, which define the radius of

influence of the kernel φ_k, a good initialization approach is to find the largest distance between all the kernel centers $\underline{t}_k$, so that the radius of influence of all kernels φ_k can encompass the maximum number of data points distributed in $\mathbb{C}^M$. The synaptic weights w_k are all initialized with $0 + j0$, so they can gradually change as the RBF ANN learns.

The next step on the initialization of the RBF ANN is to define the learning rate of each free parameter $(\mu_t, \mu_w, \mu_\sigma)$. Again, the experimental best fit process is employed and the best result for a set of trials is chosen. Figure 24 (b) to (d) shows such process for the PT-RBF, where each set of test of the learning rates has 20 elements and goes from 8×10^{-1} to 1×10^{-5} and the test set for the number of neurons goes from $K = 2$ to $K = 20$. These tests were performed by calculating the residual error (MSEA) for different PT-RBF ANN configurations. It is possible to see, from Figure 24, that the beginning of the training is characterized by a high MSEA. As the PT-RBF ANN "learns" (as the training vectors $\underline{u}(n) = [u_1(n), u_2(n) \cdots u_6(n)]$ are presented to the ANN input nodes as discrete time index n increases), the residual MSEA decreases. The lower the MSEA, more precisely the PT-RBF ANN will approximate y from Y. From Figure 24 a, 11 neurons were chosen since they performed an acceptable result (a low residual error and a fast convergence rate) with fewer neurons. It is also possible to see that, with 20 neurons, the ANN presents a lower residual error, but with a cost of a much higher computational performance. The same method of choice was applied to find the best learning rate parameters on Figure 24 b, c and d.

Figure 24 - Best fit test for: (a) number of neurons; (b) variance learning step $\boldsymbol{\mu_\sigma}$; (c) radial basis center learning rate $\boldsymbol{\mu_t}$; (d) synaptic weight learning step $\boldsymbol{\mu_w}$.

(a)

(b)

(c)

(d)
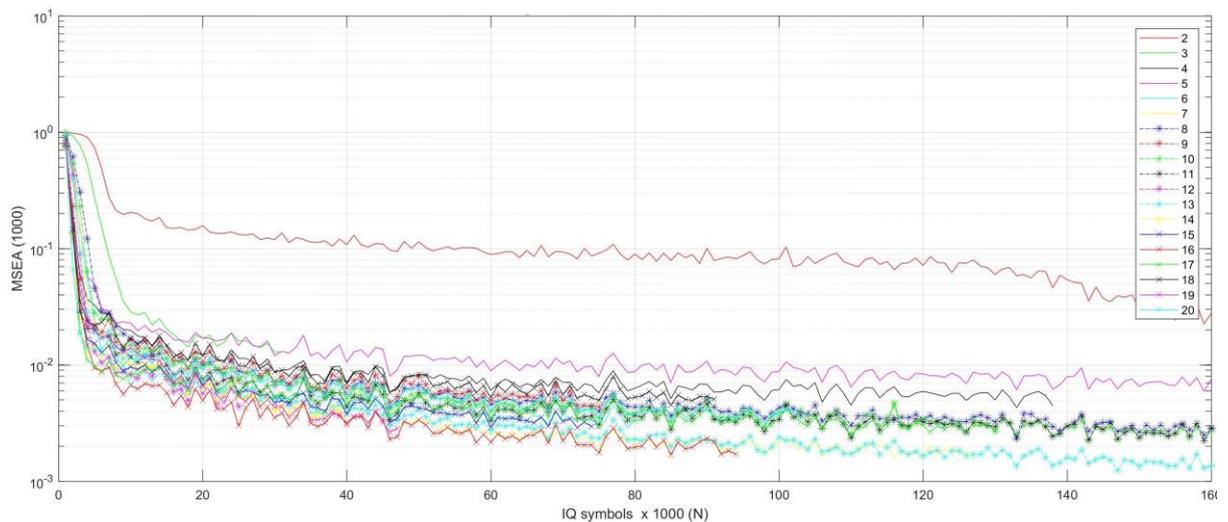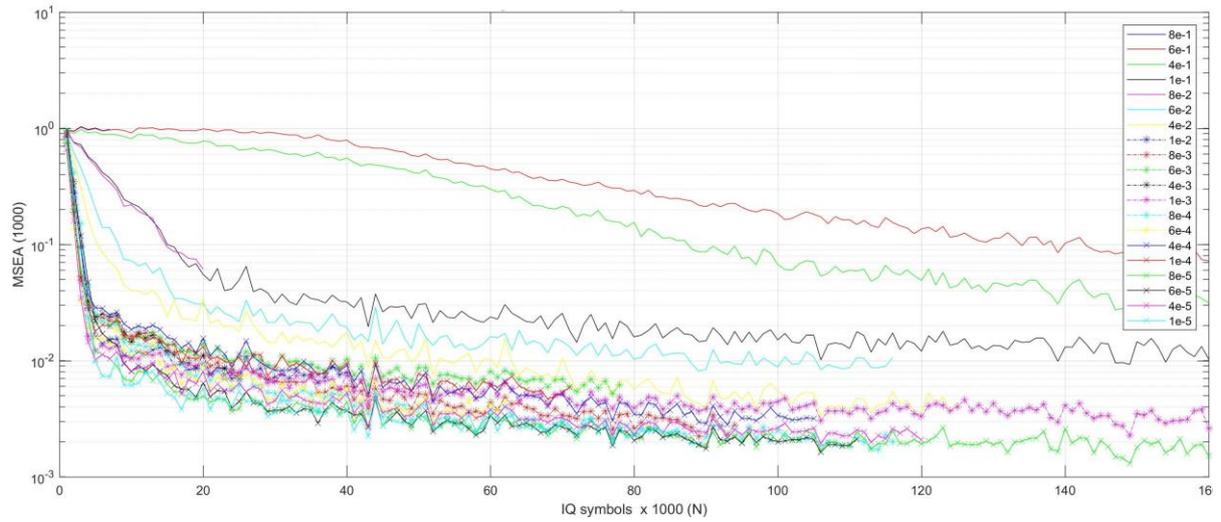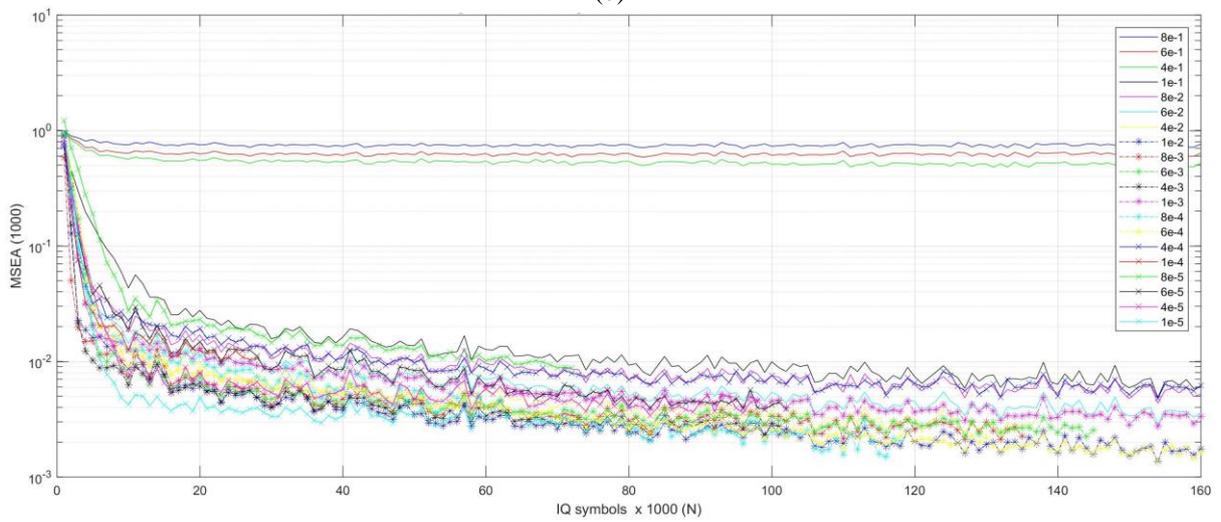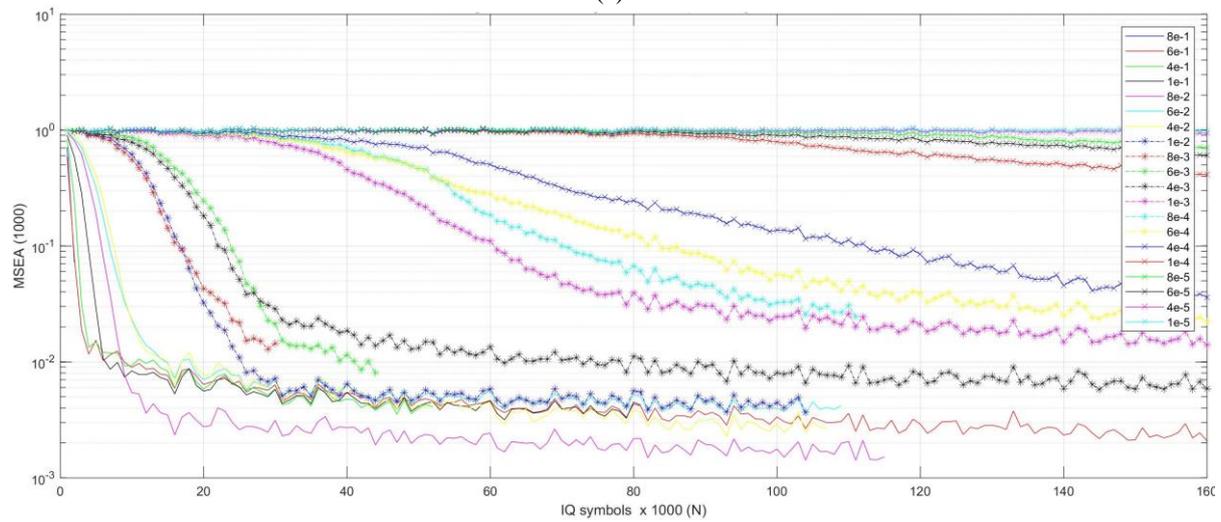
Source: (AUTHOR, 2018).

From Figure 24, Table 3 shows the best learning rates experimentally found for the PT-RBF operating in the context of the beamforming algorithm and operating under specific

scenarios. These experimental best fit processes, in which the best result for a set of trials is chosen, were also conducted for the beamformers based upon the algorithms C-RBF, FC-RBF and LMS for the same operational scenarios.

Table 3 – Learning rate constants

	C-RBF	FC-RBF	PT-RBF	LMS
μ_w	6×10^{-2}	1×10^{-1}	6×10^{-1}	1×10^{-1}
μ_t	1×10^{-3}	1×10^{-4}	4×10^{-3}	
μ_σ	1×10^{-3}	6×10^{-1}	8×10^{-5}	

Source: (AUTHOR, 2018)

After the initialization procedures, the training process goes on so that the free parameters are iteratively adjusted. Consequently, the MSEA between the RBF ANN output sequence $y(n)$ and the reference training sequence $Y(n)$ is minimized, as described before.

For the PT-RBF, at each iteration n, the activation function $\varphi_k(n)$ determines the Euclidian distance between the input data vectors $\underline{u}(n) = [u_1(n), u_2(n) \cdots u_M(n)]$ in the M-dimensional space $\mathbb{C}^M$ and each center vector $\underline{t}_k(n)$, distance that is mapped through two Gaussian functions, one real and another imaginary, as follows.

$$\varphi_k(n) = e^{\left(\frac{-1}{\text{Re}\{\sigma_k^2(n)\}}\left\|\text{Re}\{\underline{u}(n)\}-\text{Re}\{\underline{t}_k(n)\}\right\|^2\right)} + je^{\left(\frac{-1}{\text{Im}\{\sigma_k^2(n)\}}\left\|\text{Im}\{\underline{u}(n)\}-\text{Im}\{\underline{t}_k(n)\}\right\|^2\right)} \tag{2.96}$$

It is possible to see the influence of the free parameters over the activation function. Once the neurons have been activated, their individual values are weighted by the synaptic weights w_k. As in the biological model, every neuron must be weighted in order to consider a degree of importance of that neuron over the overall network answer. Following the diagram of Figure 23 , the ANN then sums all the neurons output and bring forth the network output $y(n)$. This output is tested with the reference output $Y(n)$ already provided to the ANN (the first N 16QAM symbols of the desired signal) by a simple subtraction. The difference, or the error $e(n)$, is finally responsible to feed the update equation of the free parameters, once again, repeated below for convenience.

$$w_k(n+1) = w_k(n) - \mu_w e(n) \varphi_k^*(n) \tag{2.109}$$

$$t_k(n+1) = t_k(n) + 2\mu_t \left\{ \mathrm{Re}\{\varphi_k(n)\} \frac{\mathrm{Re}\{\underline{u}(n)\} - \mathrm{Re}\{\underline{t}_k(n)\}}{\mathrm{Re}\{\sigma_k^2(n)\}} \left(\begin{array}{c} \mathrm{Re}\{w_k(n)\}\mathrm{Re}\{e(n)\} \\ +\mathrm{Im}\{w_k(n)\}\mathrm{Im}\{e(n)\} \end{array} \right) \right.$$
$$\left. + j\,\mathrm{Im}\{\varphi_k(n)\} \frac{\mathrm{Im}\{\underline{u}(n)\} - \mathrm{Im}\{\underline{t}_k(n)\}}{\mathrm{Im}\{\sigma_k^2(n)\}} \left(\begin{array}{c} \mathrm{Re}\{w_k(n)\}\mathrm{Im}\{e(n)\} \\ -\mathrm{Im}\{w_k(n)\}\mathrm{Re}\{e(n)\} \end{array} \right) \right\}$$

(2.123)

$$\sigma_k^2(n+1) = \sigma_k^2(n) + \mu_\sigma \left\{ \left| \mathrm{Re}\{\underline{u}(n)\} - \mathrm{Re}\{\underline{t}_k(n)\} \right|^2 \frac{\mathrm{Re}\{\varphi_k(n)\}}{\mathrm{Re}\{\sigma_k^2(n)\}^2} \left(\begin{array}{c} \mathrm{Re}\{w_k(n)\}\mathrm{Re}\{e(n)\} \\ +\mathrm{Im}\{w_k(n)\}\mathrm{Im}\{e(n)\} \end{array} \right) \right.$$
$$\left. + j\left| \mathrm{Im}\{\underline{u}(n)\} - \mathrm{Im}\{\underline{t}_k(n)\} \right|^2 \frac{\mathrm{Im}\{\varphi_k(n)\}}{\mathrm{Im}\{\sigma_k^2(n)\}^2} \left(\begin{array}{c} \mathrm{Re}\{w_k(n)\}\mathrm{Im}\{e(n)\} \\ -\mathrm{Im}\{w_k(n)\}\mathrm{Re}\{e(n)\} \end{array} \right) \right\}$$

(2.137)

The free parameters, responsible for storing the knowledge, are adjusted over the N training data vectors.

The PT-RBF training process can be compiled and comprehended as follows.

I. The learning process is iterative, in other words, it runs by interactions. A training will have N iteractions, where N is the number of questions/answers pairs provided to the ANN. The higher the N, more specialist the network will be;

II. At iteration n, a set of input data $\underline{u}(n)$ that partially represent an unknown process is presented to the ANN. The desired output $\Upsilon(n)$ is also presented to the ANN. It is expected the ANN to return an output $y(n)$ as close as possible from $\Upsilon(n)$;

III. The dimension M of the input data set $\underline{u}$ defines the dimension of the radial basis centers t. Both $\underline{u}$ and $\underline{t}$ are points in a M-dimensional space $\mathbb{C}^M$ and the ANN goal is to approximate its center from the input data, in order to best represent them;

IV. The real and imaginary parts of each component of vector $\underline{t}_k$ are initialized randomly, on the range $[-1,1]$. The variances σ_k^2 are initialized with the largest distance between all the kernel centers $\underline{t}_k$. The synaptic weights w_k are all initialized with $0 + j0$.

V. The ANN will have K activation functions (K neurons), where K is a number set by the network architect upon the experimental best-fit performance test. Usually, the higher K is, more faithful the ANN will approximate its output $y(n)$ to the desired output $Y(n)$ at a higher computational cost;

VI. A given input data vector $\underline{u}(n)$ feeds commonly all activation functions φ_k;

VII. The activation functions are fired. Each activation function φ_k is defined by a PDF of a Gaussian function with radius equal to the deviation pattern given by the variance σ^2 and

module equal to the Euclidian distance between the input data $\underline{u}(n)$ and the radial basis center $\underline{t}_k(n)$, of its respective neuron;

VIII. The result of each activation function φ_k is then weighted by a value called synaptic weight w_k. Such value defines how important (on the approximation process) is the result of φ_k;

IX. Then, the PT-RBF sums all Gaussians weighted from all neurons, creating a new PDF composed by multiples Gaussians between the input data $\underline{u}(n)$ and its respective radial basis center $\underline{t}_k(n)$ by using the Euclidian distance;

X. Finally, the result of this last sum is the network's output $y(n)$, and shall be as close as possible to the reference output $Y(n)$, which has already been presented to the network.

XI. A convergence test is performed, subtracting the reference output $Y(n)$ from the ANN output $y(n)$. The resultant difference is the error $e(n)$;

XII. A threshold $\mathcal{E}$ for the MSEA (an acceptable error) is defined;

XIII. If the error $\text{MSEA}(n)$ is lower than the threshold $\mathcal{E}$, then the ANN is said to be trained. Otherwise, the error $e(n)$ feeds back the free parameters (t_k, w_k, σ_k^2) update equations;

XIV. The current iteration (n) ends and a new iteration $(n+1)$ starts with the same steps adopted for the last one, until the ANN convergence;

XV. The convergence process can take many iteration processes, depending on the complexity of the function to approximate.

3.2.3 Execution Mode

The execution mode can be summarized in only one single step:

I. On execution, the free parameters (t_k, w_k, σ_k^2) are already adjusted, and the process consists of presenting the new input data vector $\underline{u}$ to the ANN, vector that has not been presented to the ANN before, and computing the ANN ouput y.

On proper initialization, configuration and training, the ANN RBF is able to generalize the learning process, so that the beamformer will be able to demodulate the received desired signal and to set up an optimum radiation pattern for the given operational scenario of the UCA array.

3.3 ASSESMENT CRITERIA

The performance comparison between the beamforming solutions is based on the resulting radiation diagram, on the resulting constellation of the received IQ symbols and on the plot of the MSEA versus discrete time.

Ideally, the resulting radiation diagram should present nulls on the bearing angles of the interfering signals and should present a high gain lobe on the bearing angle of the desired signal. Likewise, ideally, the constellation of the received IQ symbols should be minimally dispersed around the 16-QAM reference symbols in order to minimize the SER. Also, the MSEA curve should ideally drop in the shortest possible time interval, attaining the lowest possible value after convergence of the adaptive process.

The convergence occurs when the average error reaches a value small enough $e(n) \rightarrow \mathcal{E}$. Within the context of function regression, MSEA indicates the proximity that a regression line is from a set of points. Specifically, the MSEA computes the distances (errors) from the points to the regression line raised to the 2nd power. However, as the error at the PT-RBF ANN output is a complex value, the magnitude operator must be previously applied, as shown once again

$$MSEA = \frac{1}{N} \sum_{n=1}^{N} \left| d(n) - y(n) \right|^2 \tag{2.47}$$

For all LMS and RBF networks simulated in Matlab, a threshold of 10^{-4} has been set for the end of the training process. Different ANNs and different operating scenarios demand a different number N of training data IQ symbols in order reach a specific threshold. Figure 25 shows how each simulated ANN performs in order to reach a threshold of 10^{-4} on a scenario with one desired signal at $\phi = 30°$ and one interference of pure WGN with $5dBW$ (SIR $= 35dB$) of power at $\phi = 210°$. The non-linearity at the RF front end was not included, so that this represents an ideal operational situation since a practical RF front end will always present some degree of non-linearity.

Figure 25 – MSEA's performances

(a)

(b)

(c)

(d)
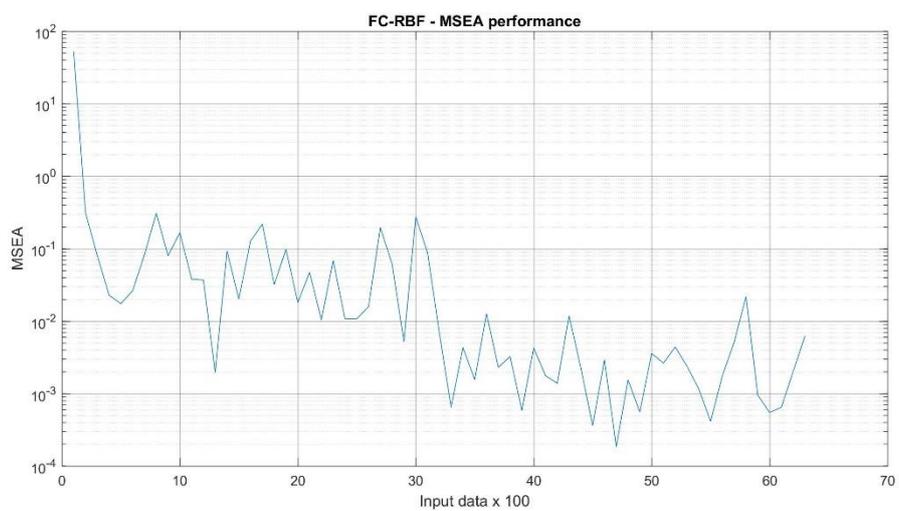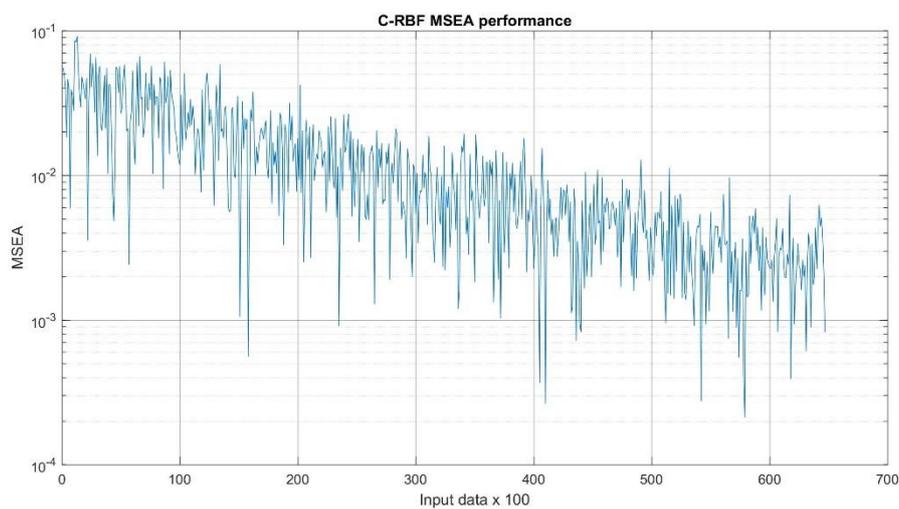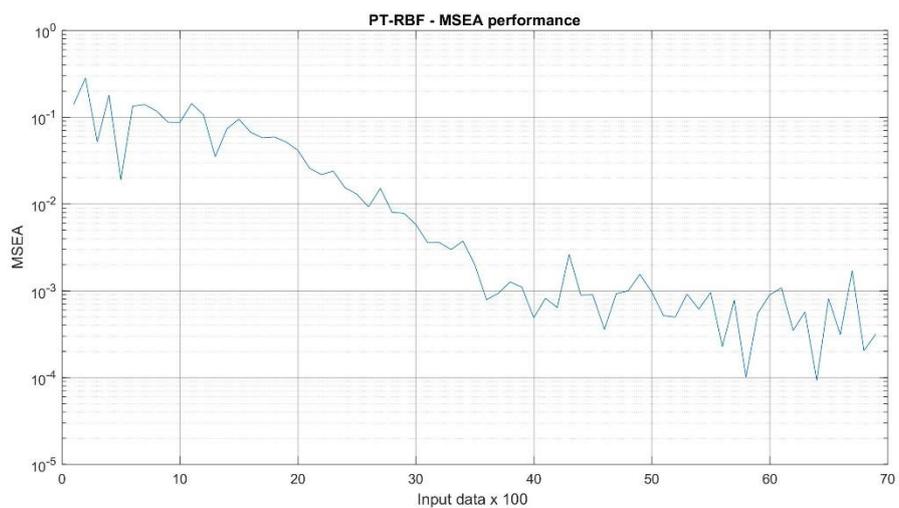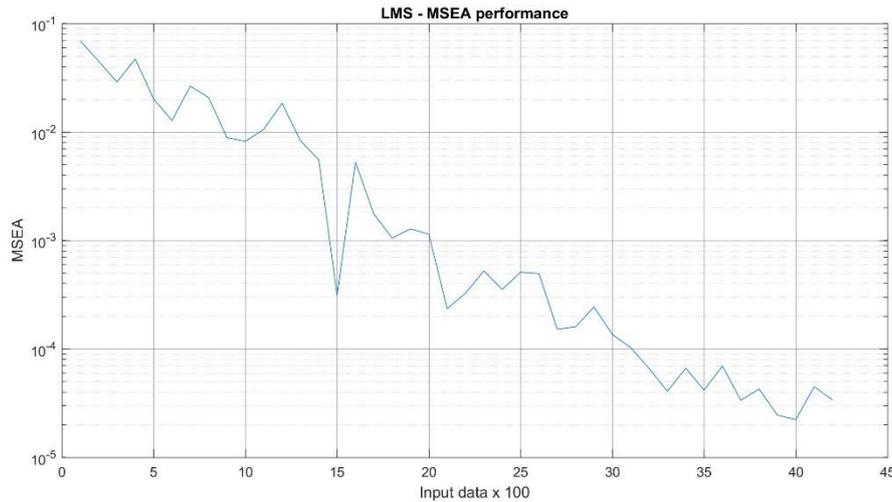
In the ideal operational situation in which the RF front end is free of non-linearity, from (b), it is already possible to infer that the C-RBF requires much more training data in order to reach the same MSEA threshold. This behavior is seen in all simulations and makes this network the slowest one. From (d), it is possible to infer that LMS needed less training data to reach the lower error level. From (c), it is possible to infer that the FC-RBF has a fast convergence at the beginning of the training, which is great for a dynamic system. Finally, from (a), it is possible to infer that the PT-RBF has performed the SD algorithm more smoothly, finding fewer local minima.

For a visual interpretation of the MSEA's behavior, Figure 26 shows the performance of the proposed algorithm for the same scenario described for Figure 25. One can see that Figure 26 (a), (b) and (c) reached the threshold of 10^{-1} with around 67 IQ symbols, Figure 26 (d), (e) and (f) reached the threshold of 20^{-2} with around 2400 IQ symbols and Figure 26 (g), (h) and (i) reached the threshold of 10^{-3} with around 7500 IQ symbols. The IQ symbols are presented to the ANN as input data $\underline{u}$.

Figure 26 – MSEA, beamformer and demodulation performance for the following thresholds: (a), (b) and (c) 70 input training data, (d), (e) and (f) 2500 input training data and (g) (h) and (i) 8000 input training data

(a)

(b)

(c)

PT-RBF- MSEA performance

(d)

PT-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 128.6815dBμV/m
Main Signal = 150°
Interference = 240°

(e)

PT-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = 1.0146dB SNR main signal = 0dB IQ symbols = 2300
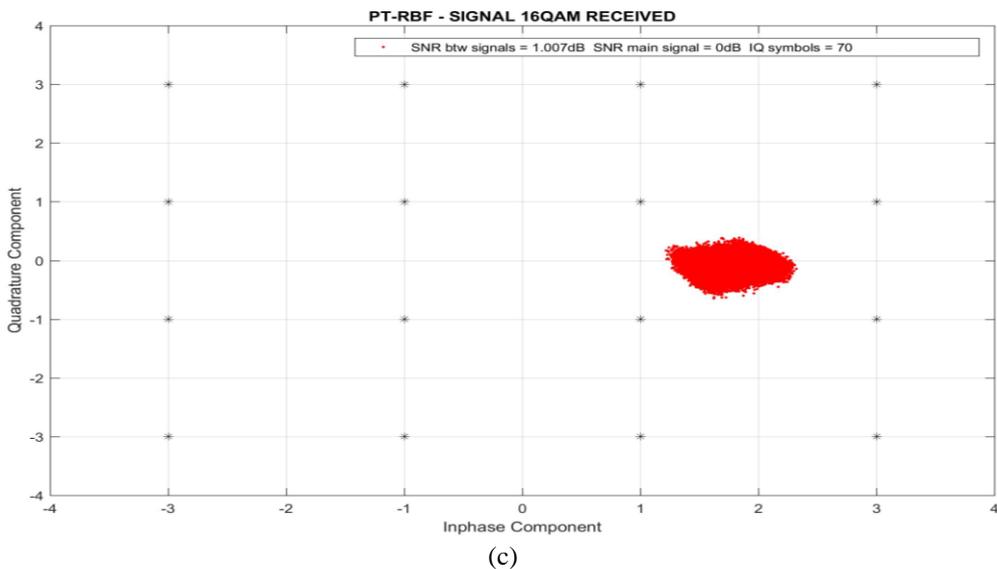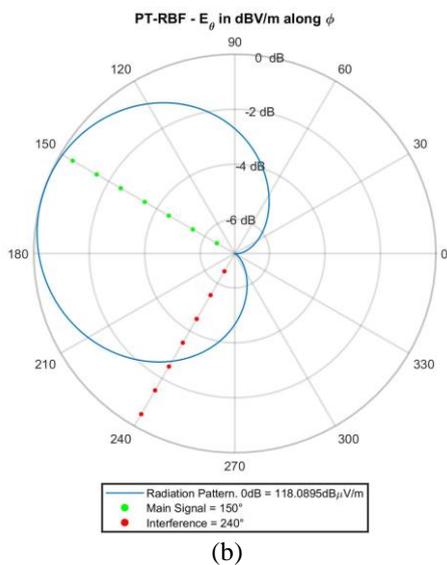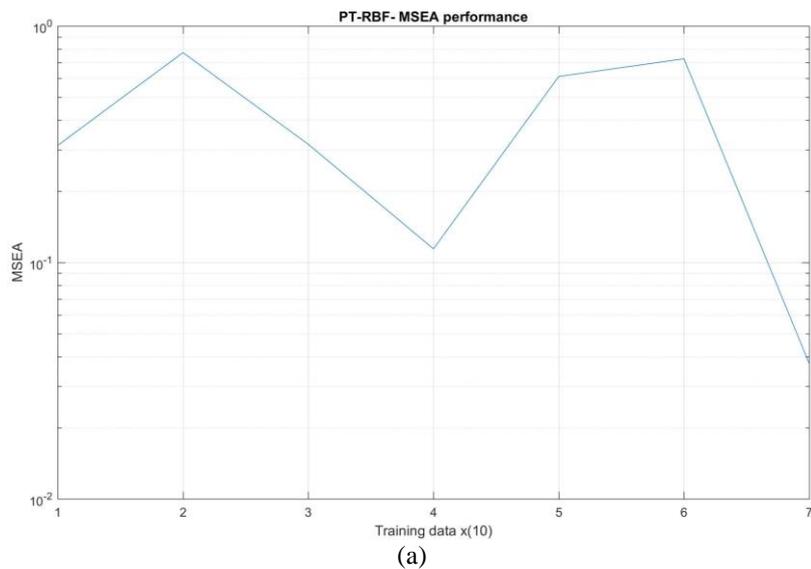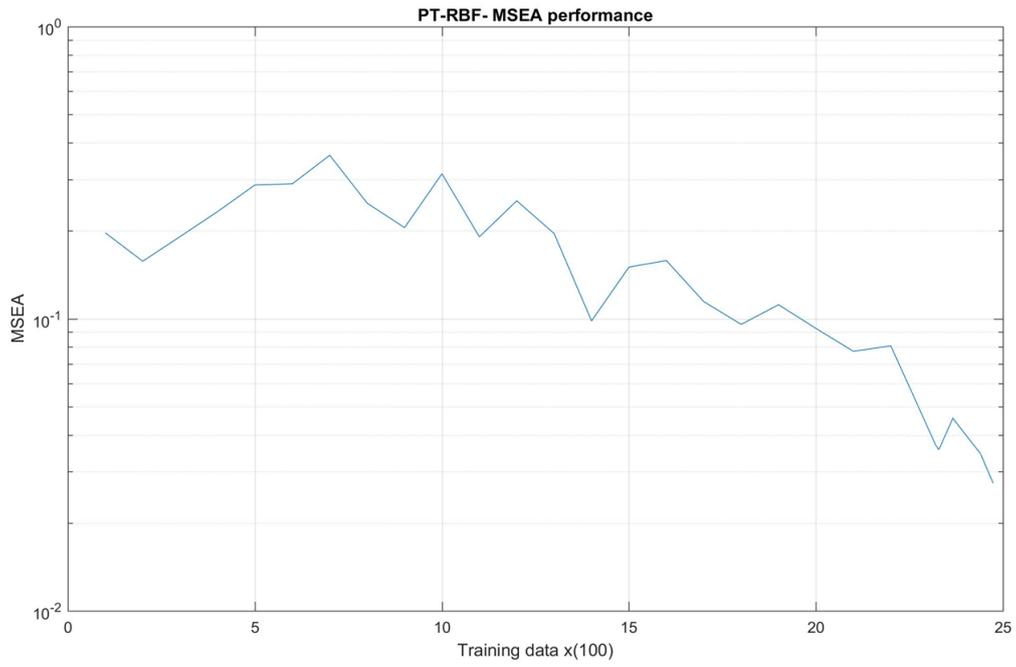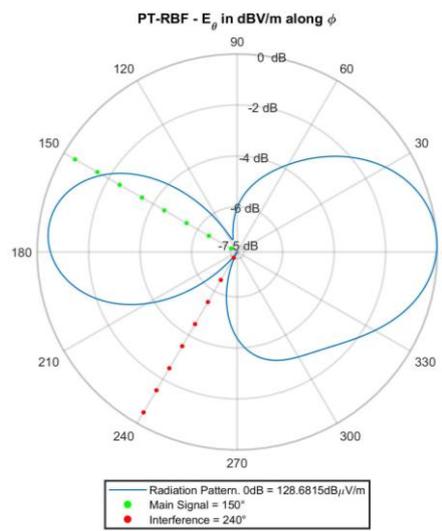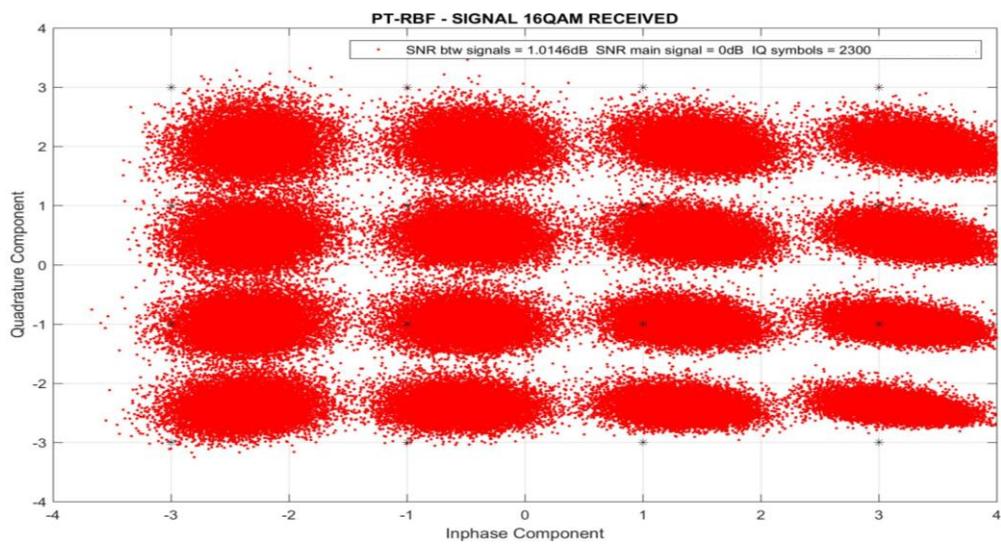
(f)

(g)

(h)

(i)
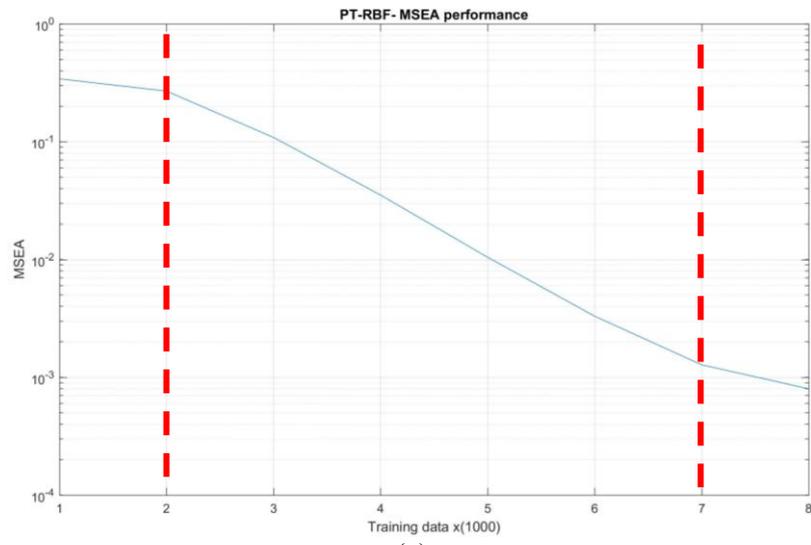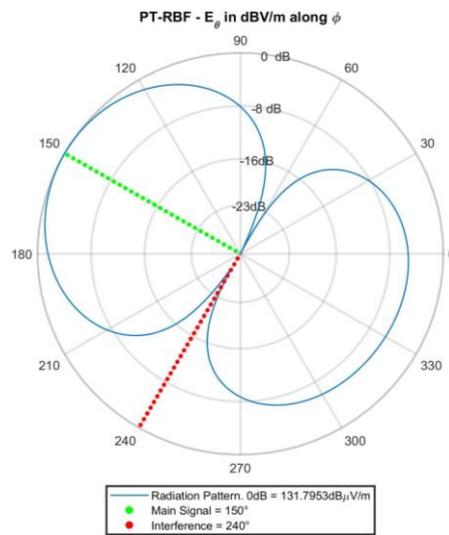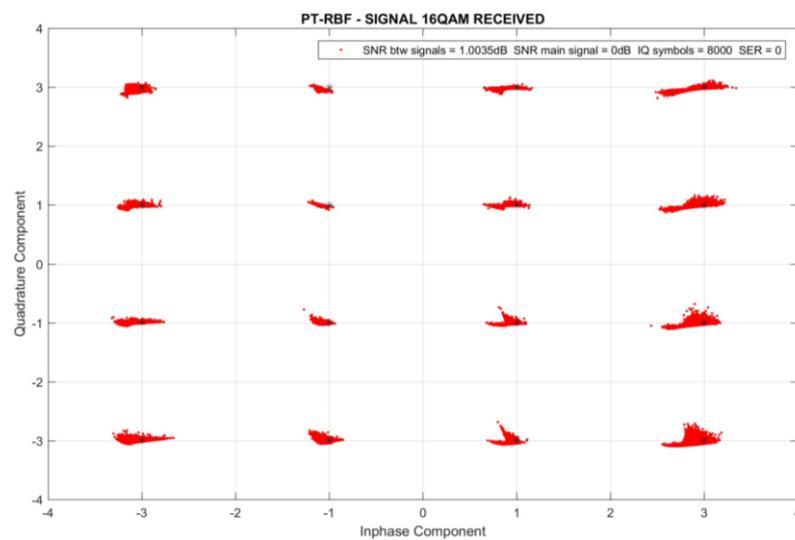
Source: (AUTHOR, 2018).

Note from the MSEA curve in Figure 26(g) that there are three distinct phases separated by the read lines. In the first phase, initially the MSEA is high and the slope of the MSEA curve is small. As the training advances and the cost function is minimized by the SD algorithm, the slope of the MSEA curve becomes negative so that the MSEA is reduced along time in a rate that is proportional to the absolute value of the curve negative slope. By the end of the training, the third phase, the slope of the MSEA curve becomes small again, indicating the ANN convergence (ready to start the execution). In the execution, new input data $\underline{u}$ (new IQ symbols which haven't been part of the TS) are applied to the already trained PT-RBF ANN, with the free parameters (t, w, σ^2) already stabilized around its values after the SD convergence. Note the ANN generalization capability on the execution, since new input IQ symbols which haven't been part of the TS are fed to the ANN inputs, and, even so, the ANN is able to infer a proper output based on the extrapolation of the underlying stochastic process learnt from the TS.

4 PERFORMANCE EVALUATION

This section presents the performance of the proposed PT-RBF beamforming solution. Simulation results are compared with the results of other beamforming techniques proposed in literature (FC-RBF, C-RBF, LMS). The proposed beamforming is evaluated in static and dynamic scenarios. In static scenarios, the signal of interest and the interfering signals do not change the DOA during the whole simulation. In dynamic scenarios, the interfering or desired signal moves around the receiving UCA.

For each scenario, the following performance parameters are evaluated:

I. The radiation pattern given a desired signal (D.S) and one or more interfering signal (I.S);

II. The demodulated D.S SER;

Also, for some scenarios:

III. The beamforming convergence rate through the analysis of the MSEA curve.

Apart from the simulations presented in this Chapter, further simulations are found in Appendix B.

4.1 STATIC SCENARIOS

At static scenarios, the proposed beamformer has been evaluated for different SIR levels with and without non-linearity (NL) at the front end. Also, different DOAs, symmetric and asymmetric around a reference bearing angle, have been analyzed. Table 4 presents the evaluated static scenarios.

Table 4 – Summary of static scenarios.

Scenario	N° I.S.	I.S. DOA	D.S. DOA	D.S. SNR	SIR	NL
1	1	270°	1°	35dB	-10dB	yes
2	1	270°	1°	∞	35dB	yes
3	2	150°,300°	30°	35dB	35dB	no
4	2	150°,300°	30°	35dB	10dB	yes
5	3	150°,210°,270°	90°	∞	-10dB	yes
6	3	150°,210°,270°	90°	∞	-10dB	no
7	9	60°,90°,120°,150°,180°, 210°,240°,270°,300°,330°	1°	∞	-10dB	no

Source: (AUTHOR, 2018).

The results of each scenario are organized in five graphics, which present:

I. The comparison of the radiation diagram obtained after the convergence of the four beamforming algorithms (PT-RBF, C-RBF, FC-RBS and LMS);

II. The individual radiation diagram obtained with each beamforming approach;

III. The received symbol constellation at the beamforming output;

IV. The MSEA curve of each beamforming approach;

V. A table summarizing the scenario results.

It is important to notice that all radiation diagrams in dB are normalized to a maximum received intensity of $0dB\mu V/m$, since the purpose is to express gain and not field strength. Besides, as the radiation pattern diagram has a logarithm scale, it might look different from the overall graphic with all networks to the individual performances because the minimum value varies. Nevertheless, they represent the same outcome.

4.1.1 Scenario 1

The first evaluated static scenario has the desired user signal coming from the bearing angle $\phi = 1°$ and one interference signal coming from $\phi = 270°$. WGN was added to the desired signal so that SNR $= 35dB$. The relation between the power of the desired signal and the power of the interfering signal is given by SIR $= -10dB$. Finally, the non-linearity expressed by Eq (2.140) is introduced at the RF front end in order to implement a more realistic simulation. It is possible to see, from Figure 27 (a), that the PT-RBF presented the highest attenuation of the interference signal. Note from Figure 27 (b) – (i) that, except the FC-RBF, all others beamformers yielded zero SER. Thus, from Figure 27, it is possible to infer that PT-RBF presented superior performance in this very low SIR (= high interference power) non-linear scenario in comparison to the other evaluated solutions.

Figure 27 – Static case 1. (a) Radiation pattern diagram comparing all beamformers gains, (b) radiation pattern diagram for PT-RBF beamformer and (c) 16QAM constellation presenting the gain and SER respectively for PT-RBF beamformer, (d) radiation pattern diagram and (e) 16QAM constellation presenting the gain and SER respectively for C-RBF beamformer, (f) radiation pattern diagram and (g) 16QAM constellation presenting the gain and SER respectively for FC-RBF beamformer, (h) radiation pattern diagram and (i) 16QAM constellation presenting the gain and SER respectively for LMS beamformer.

(a)

PT-RBF - E$_{\theta}$ in dBV/m along ϕ

Radiation Pattern. 0dB = 130.4506dBμV/m
Main Signal = 1°
Interference = 270°

(b)

PT-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -1.0076dB   SNR main signal = 35dB   IQ symbols = 160000   SER = 0

(c)

C-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 136.4468dBμV/m
Main Signal = 1°
Interference = 270°

(d)

C-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -1.0076dB SNR main signal = 35dB IQ symbols = 160000 SER = 0

(e)

FC-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 129.1367dBμV/m
Main Signal = 1°
Interference = 270°

(f)

FC-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -1.0076dB   SNR main signal = 35dB   IQ symbols = 160000   SER = 2.5e-05

(g)

LMS - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 130.0062dBμV/m
Main Signal = 1°
Interference = 270°

(h)

LMS - SIGNAL 16QAM RECEIVED

SNR btw signals = -1.0076dB   SNR main signal = 35dB   IQ symbols = 160000   SER = 0

(i)
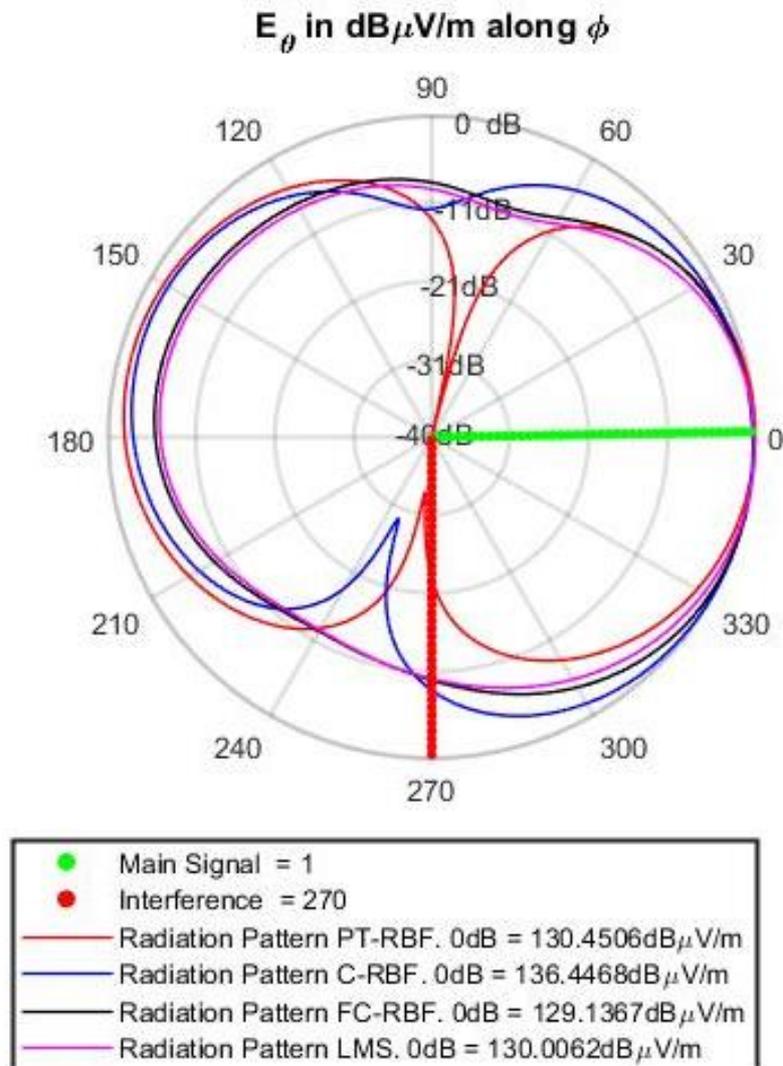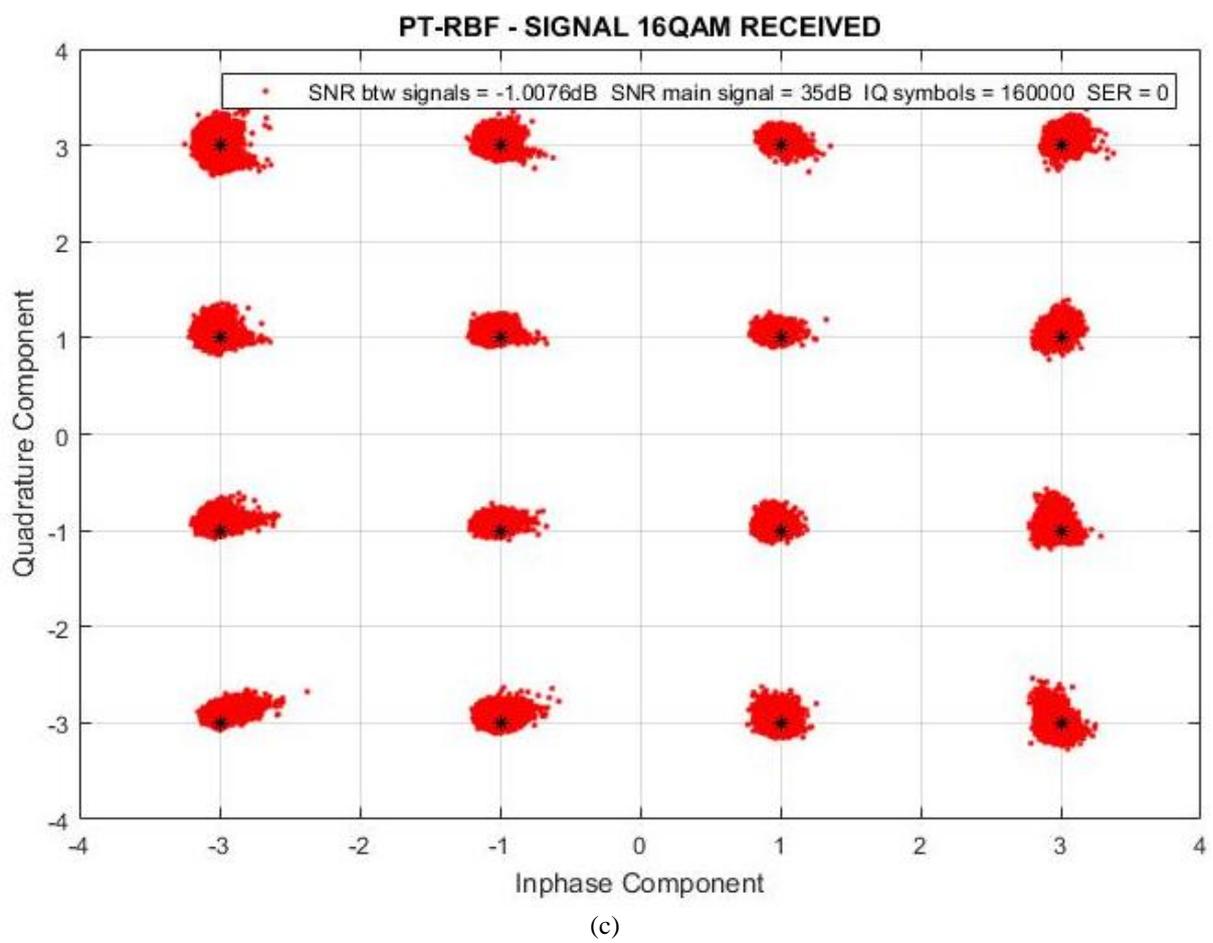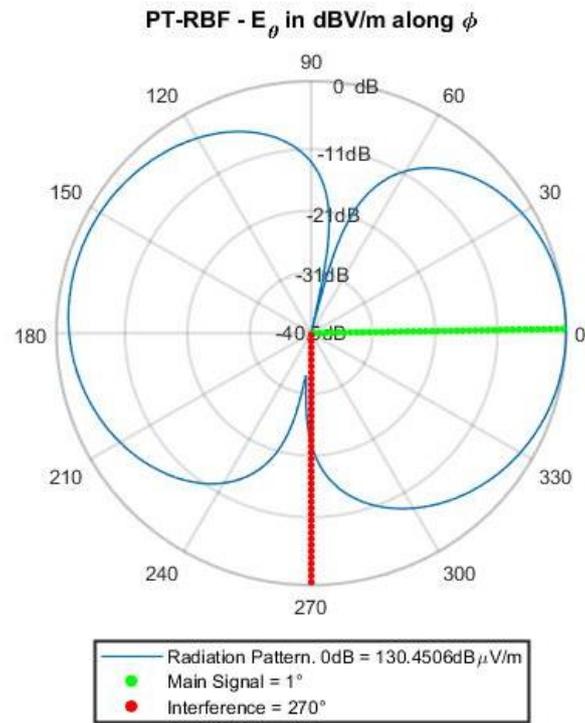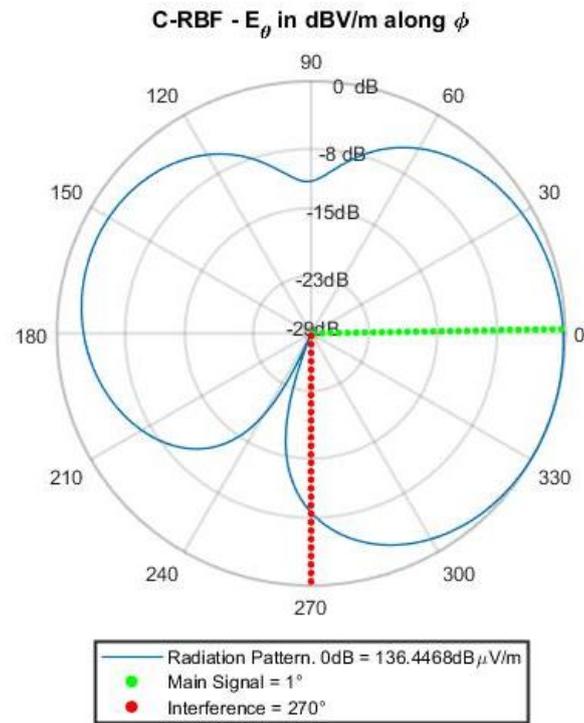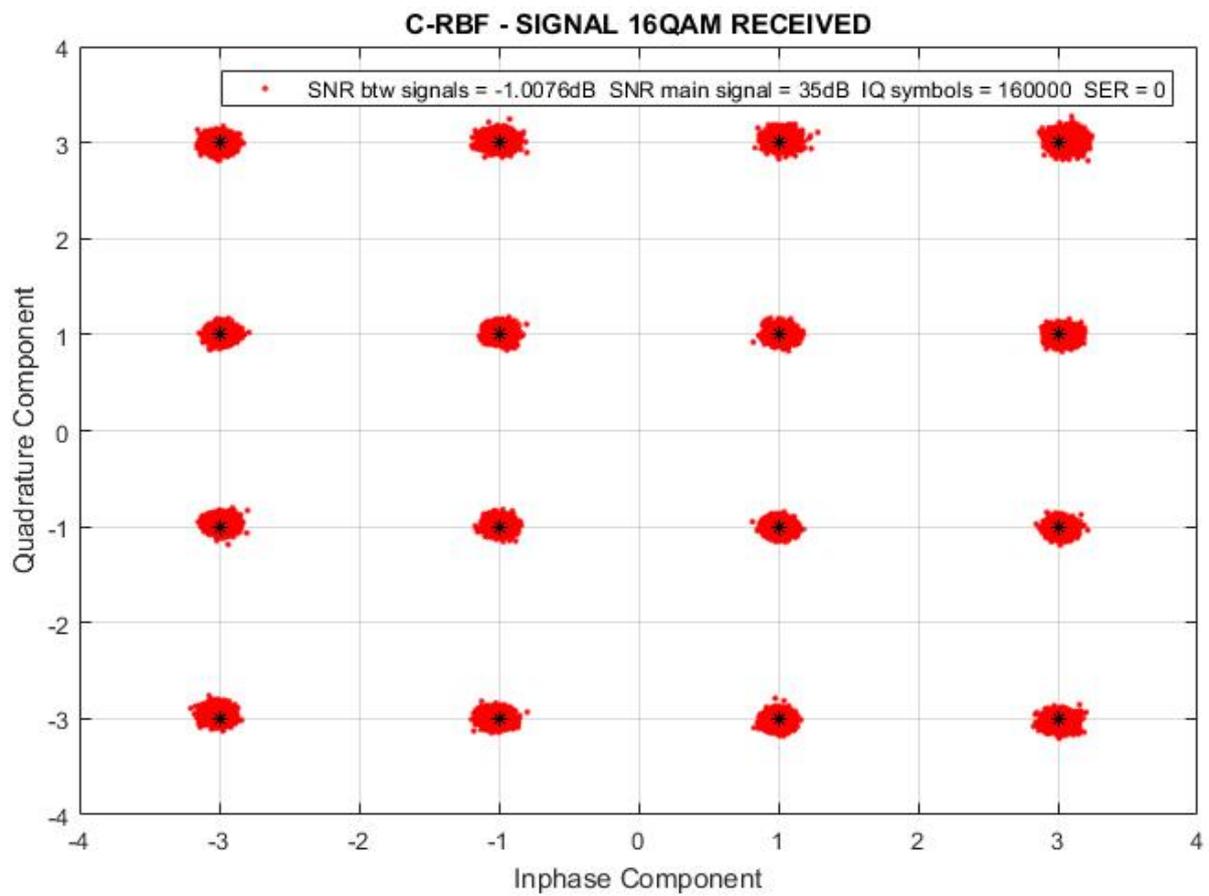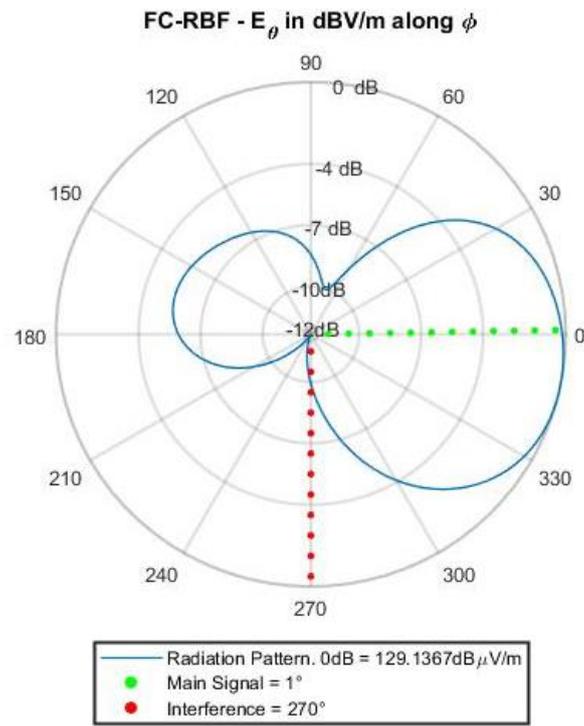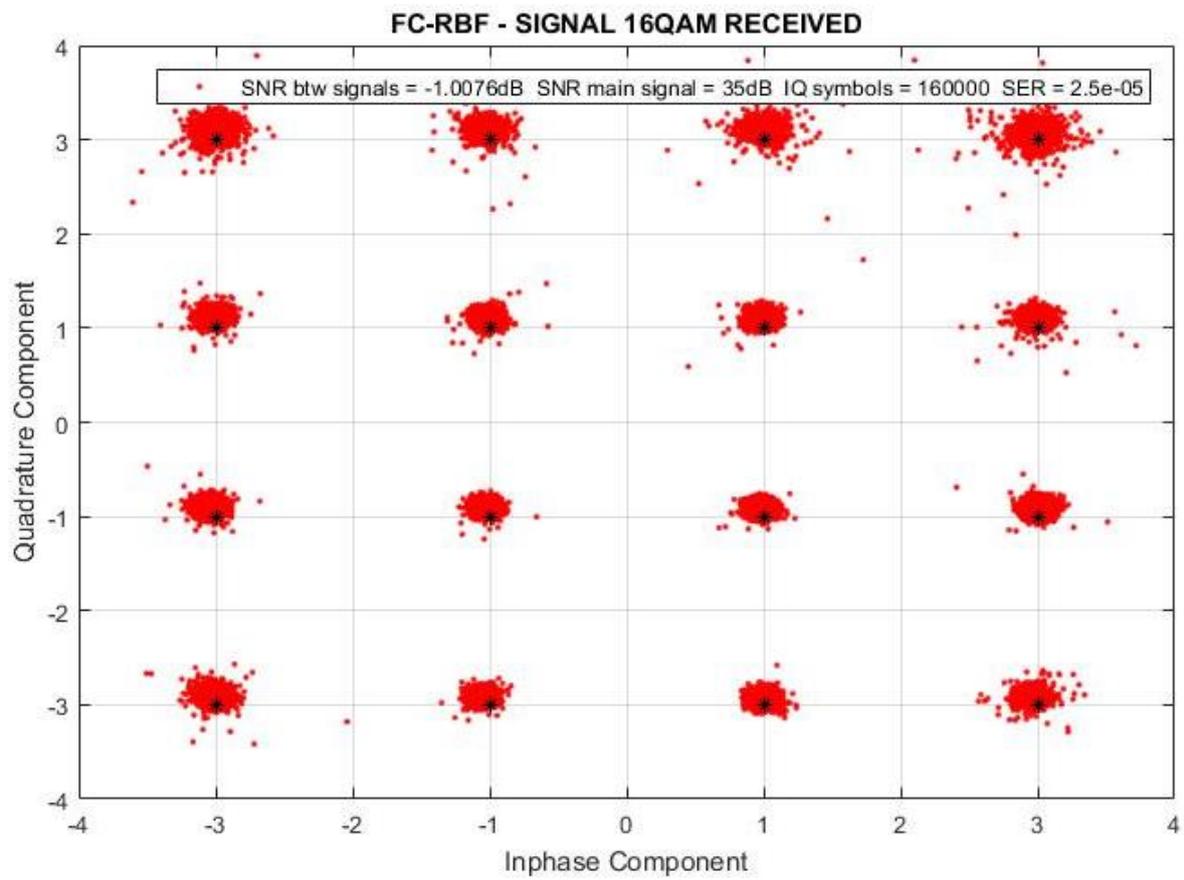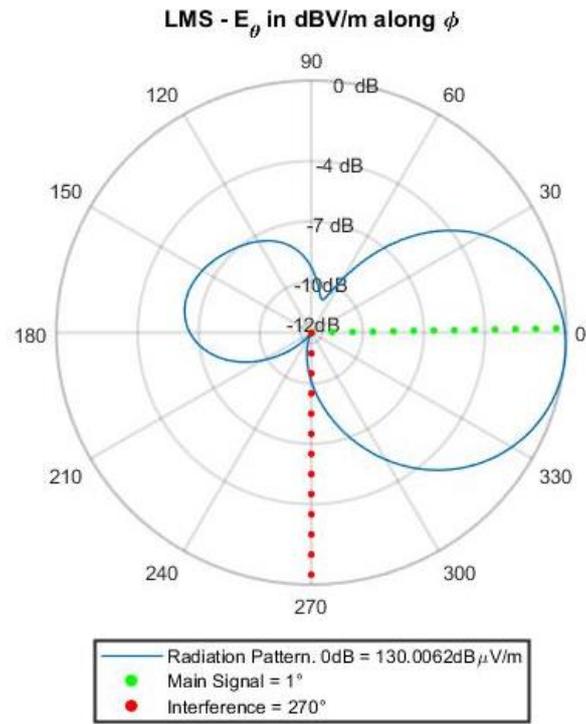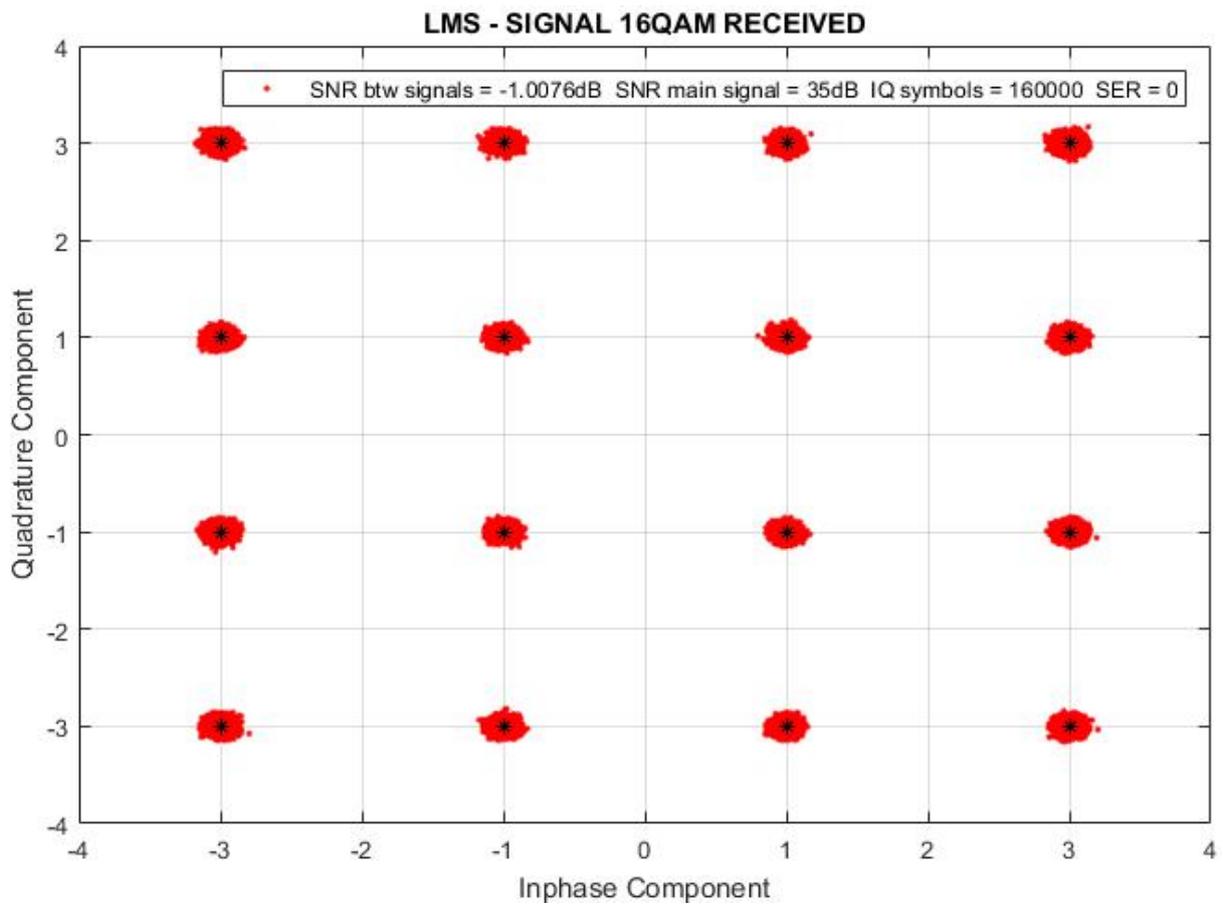
Source: (AUTHOR, 2018).

Table 5 – Summary of static case 1

D.S. DOA @ 1°	Gain [dB]			
	PT-RBF	C-RBF	FC-RBF	LMS
I.S. DOA @ 270°	-22.9	-8.6	-9.9	-10
SER	**0**	**0**	$\mathbf{2.5 \cdot 10^{-5}}$	**0**

Source: (AUTHOR, 2018).

4.1.2 Scenario 2

Scenario 2 has the desired user signal coming from the bearing angle $\phi = 1°$ and one interference signal coming from $\phi = 210°$. There is no WGN added to the desired signal, so that SNR $= \infty$. The relation between the power of the desired signal and the power of the interfering signal is given by SIR $= 35dB$. Finally, the non-linearity expressed by Eq (2.141) is introduced at the RF front end in order to implement a more realistic simulation. It is possible to see, from Figure 28 (a), that the C-RBF has presented the highest attenuation of the interference signal. All other beamformers had a very similar response. Note from Figure 28 (b) – (i) that the FC-RBF beamformer was the only one that presented symbol errors on demodulation. All other beamformers yielded zero SER. Note from Figure 28 that the C-RBF performed better in this moderate SIR (moderate interference power) non-linear scenario. The proposed PT-RBF performed similarly to the FC-RBF and to the LMS, presenting zero SER and gains at the I.S DOA of $-10.6dB, -10.4dB$ and $-11.3dB$ respectively.

Figure 28 – Static case 2. (a) Radiation pattern diagram comparing all beamformers gains, (b) radiation pattern diagram for PT-RBF beamformer and (c) 16QAM constellation presenting the gain and SER respectively for PT-RBF beamformer, (d) radiation pattern diagram and (e) 16QAM constellation presenting the gain and SER respectively for C-RBF beamformer, (f) radiation pattern diagram and (g) 16QAM constellation presenting the gain and SER respectively for FC-RBF beamformer, (h) radiation pattern diagram and (i) 16QAM constellation presenting the gain and SER respectively for LMS beamformer.

(a)

PT-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 127.9023dBμV/m
Main Signal = 1°
Interference = 210°

(b)

PT-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.99123dB   SNR main signal = 35dB   IQ symbols = 160000   SER = 0

Inphase Component

Quadrature Component

(c)

C-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 132.2288dBμV/m
Main Signal = 1°
Interference = 210°

(d)

C-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.99123dB   SNR main signal = 35dB   IQ symbols = 160000   SER = 0

(e)

FC-RBF - E$_\theta$ in dBV/m along ϕ

(f)

FC-RBF - SIGNAL 16QAM RECEIVED

(g)

(h)

(i)
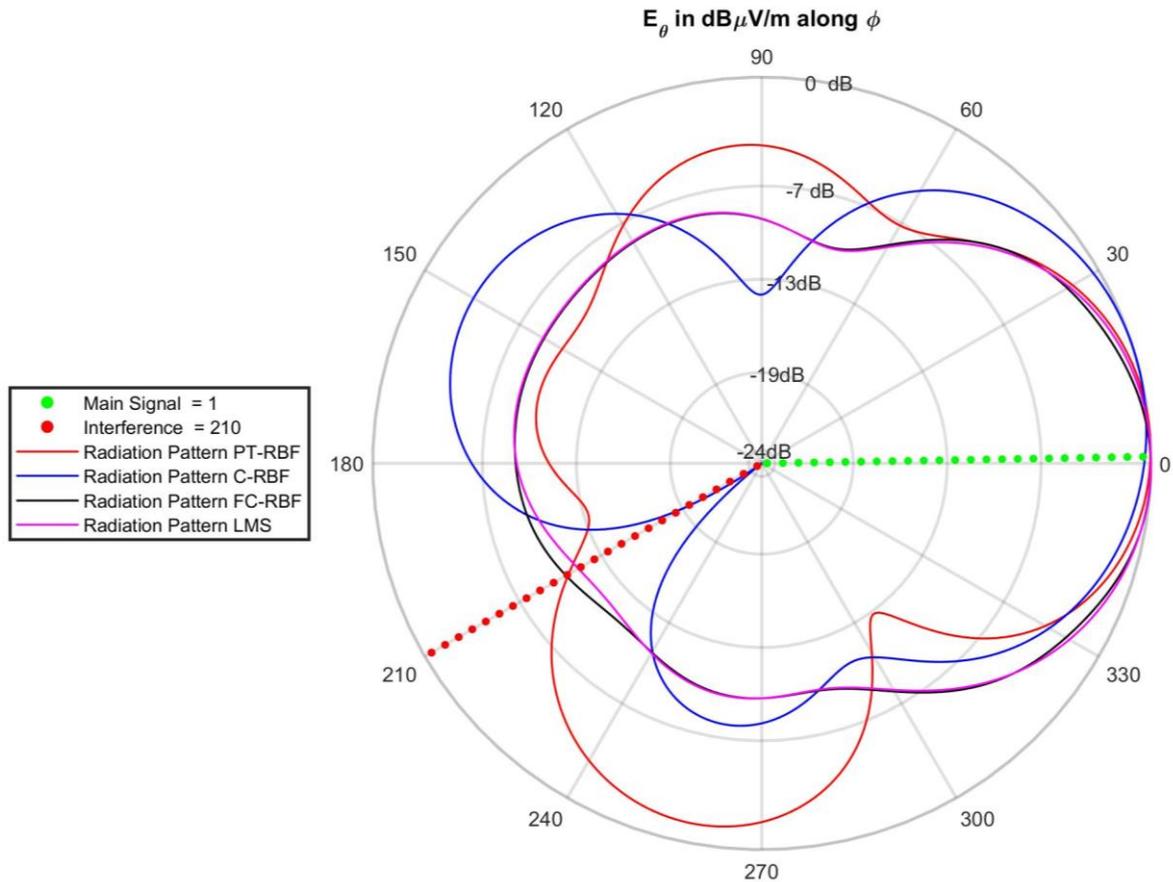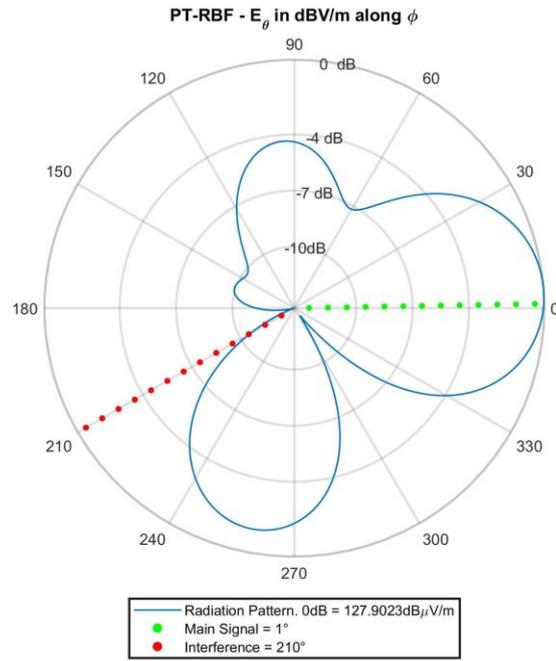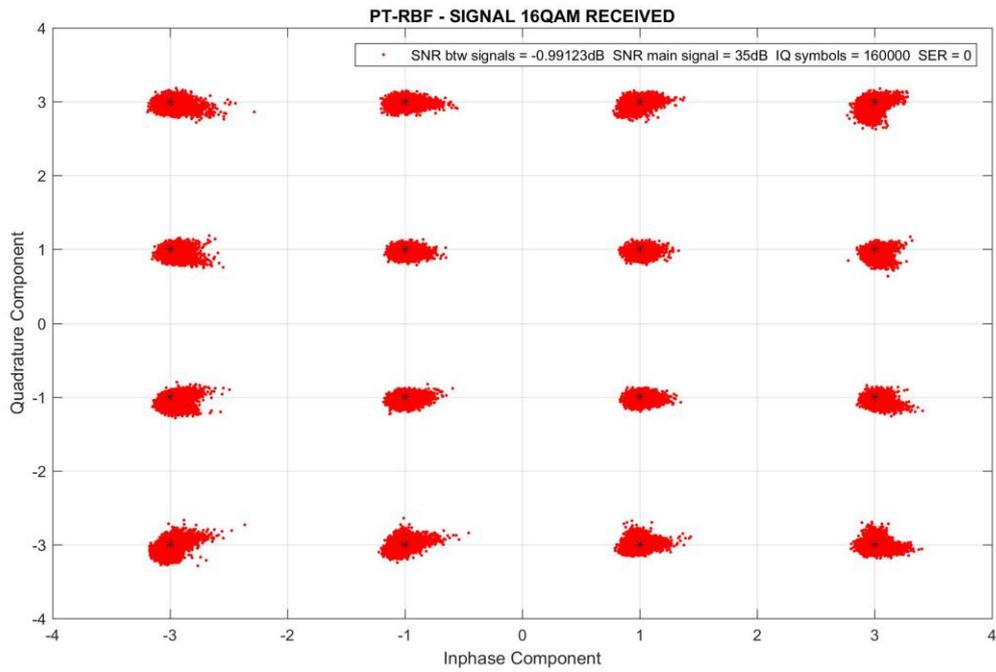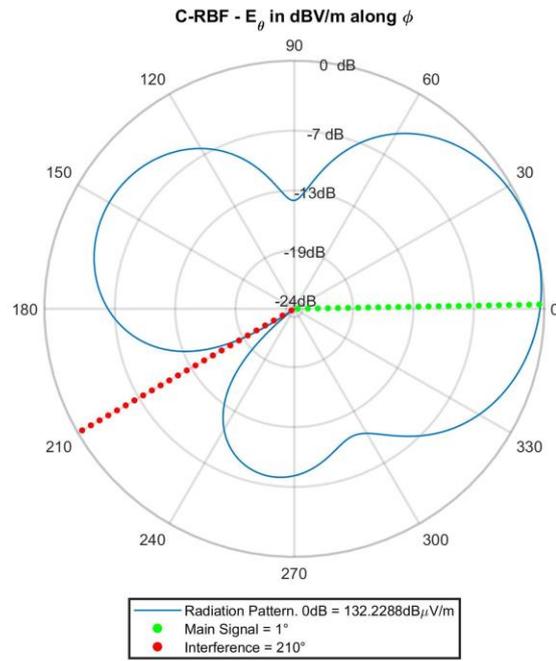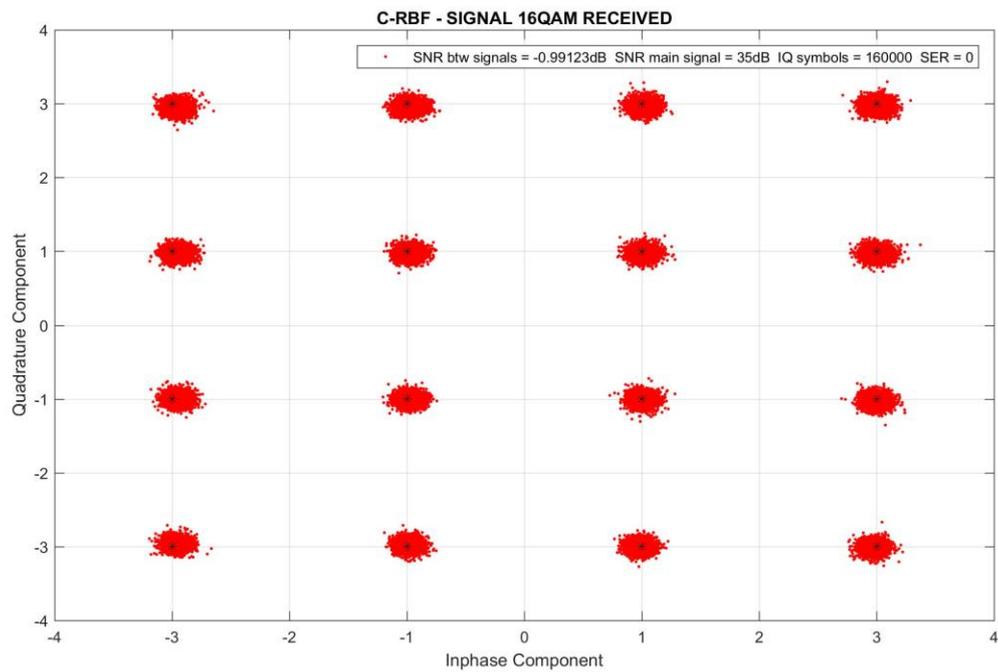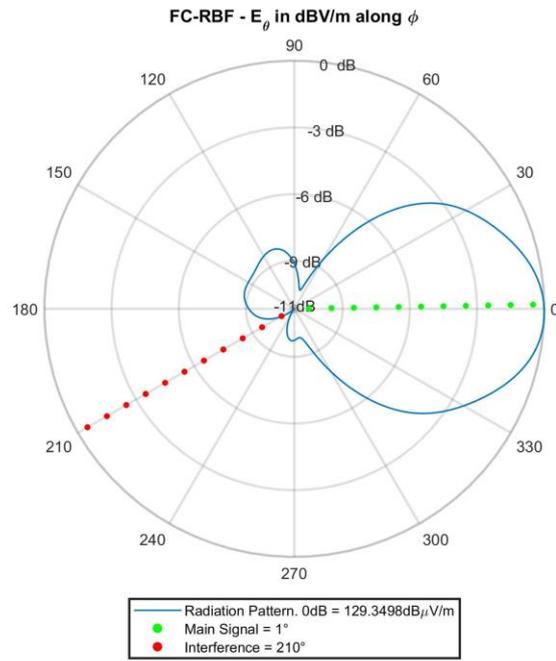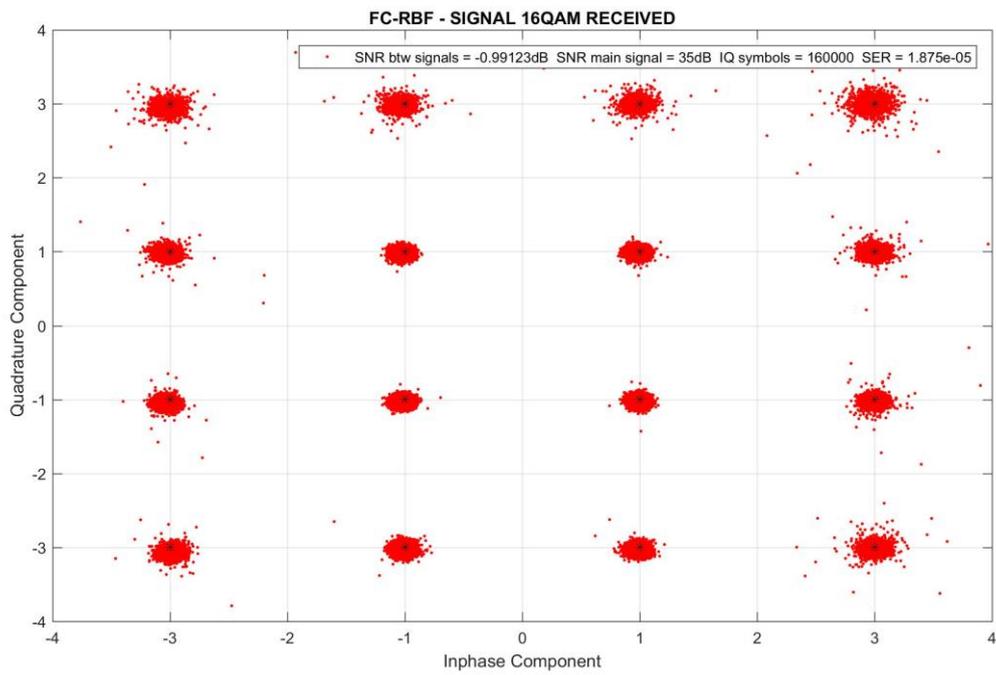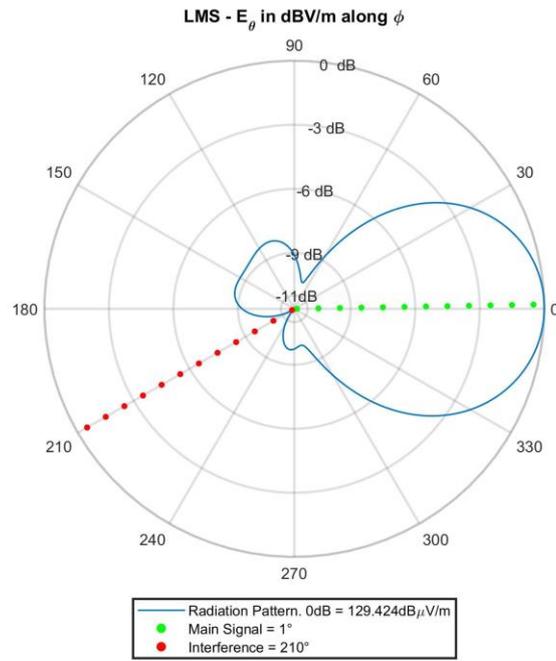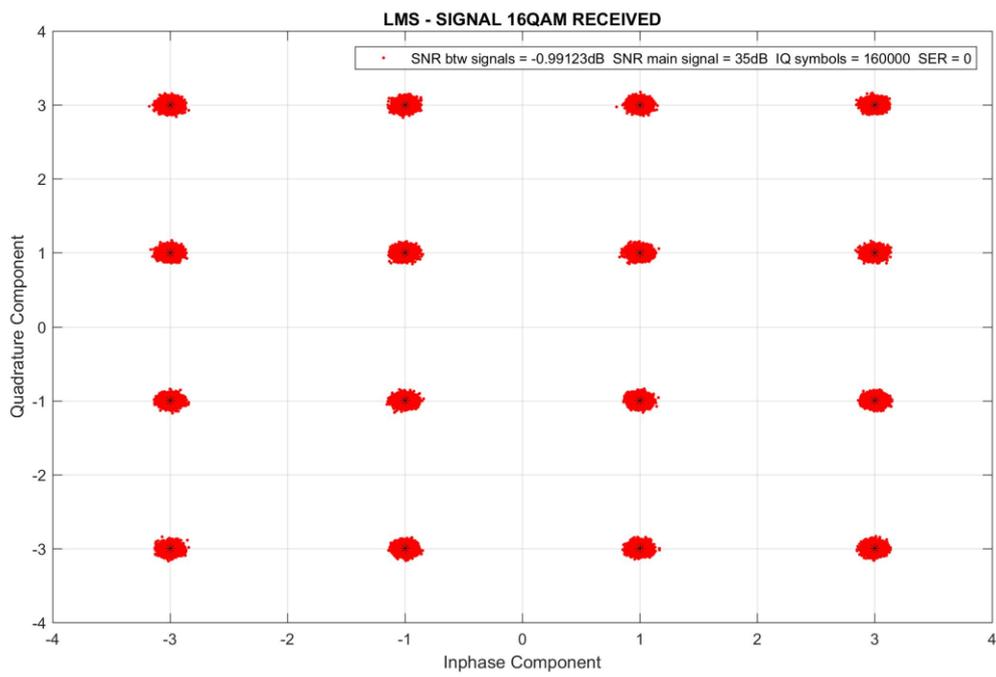
Source: (AUTHOR, 2018).

Table 6 - Summary of static case 2

D.S. DOA @ 1°	Gain [dB]			
	PT-RBF	C-RBF	FC-RBF	LMS
I.S. DOA @ 210°	−10.6	−18.5	−10.4	−11.3

SER	0	0	$1.87 \cdot 10^{-5}$	0

Source: (AUTHOR, 2018).

4.1.3 Scenario 3

Scenario 3 has the desired user signal coming from the bearing angle $\phi = 30°$ and two interfering signals coming from $\phi = 150°$ and $300°$. WGN is added to the desired signal so that SNR $= 35dB$. The relation between the power of the desired signal and the power of each interfering signal is given by SIR $= 35dB$. In this scenario, the non-linearity at the front end was not introduced. It is possible to see, from Figure 29 (a), that the C-RBF and the FC-RBF have similar performance yielding the highest attenuation of the interference signals. However, from Figure 29 (b) – (i), it is possible to see that, although C-RBF and FC-RBF yielded good attenuation of the interference signals, they fail on demodulating the received IQ symbols, presenting a SER $= 0.265$ and 0.00003 respectively. The PT-RBF and the LMS resulted zero SER, with the LMS presenting the received IQ symbols minimally dispersed around the 16-QAM reference symbols. Thus, from Figure 29, it is possible to infer that LMS performs better in a moderate SIR linear scenario.

Figure 29 – Static case 3. (a) Radiation pattern diagram comparing all beamformers gains, (b) radiation pattern diagram for PT-RBF beamformer and (c) 16QAM constellation presenting the gain and SER respectively for PT-RBF beamformer, (d) radiation pattern diagram and (e) 16QAM constellation presenting the gain and SER respectively for C-RBF beamformer, (f) radiation pattern diagram and (g) 16QAM constellation presenting the gain and SER respectively for FC-RBF beamformer, (h) radiation pattern diagram and (i) 16QAM constellation presenting the gain and SER respectively for LMS beamformer.

(a)

PT-RBF - E$_{\theta}$ in dBV/m along ϕ

(b)

PT-RBF - SIGNAL 16QAM RECEIVED

(c)

C-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 133.8374dBμV/m
Main Signal = 30°
Interference = 150° , 300

(d)

C-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.99848dB   SNR main signal = 0dB   IQ symbols = 200000   SER = 0.26312

(e)

FC-RBF - E_θ in dBV/m along ϕ

Radiation Pattern. 0dB = 132.2359dBμV/m
Main Signal = 30°
Interference = 150°, 300

(f)

FC-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.99848dB SNR main signal = 0dB IQ symbols = 200000 SER = 3e-05

(g)

LMS - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 130.3072dB$_\mu$V/m
Main Signal = 30°
Interference = 150", 300
SIR = 0.39769

(h)

LMS - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.99848dB SNR main signal = 0dB IQ symbols = 12888 SER = 0

(i)
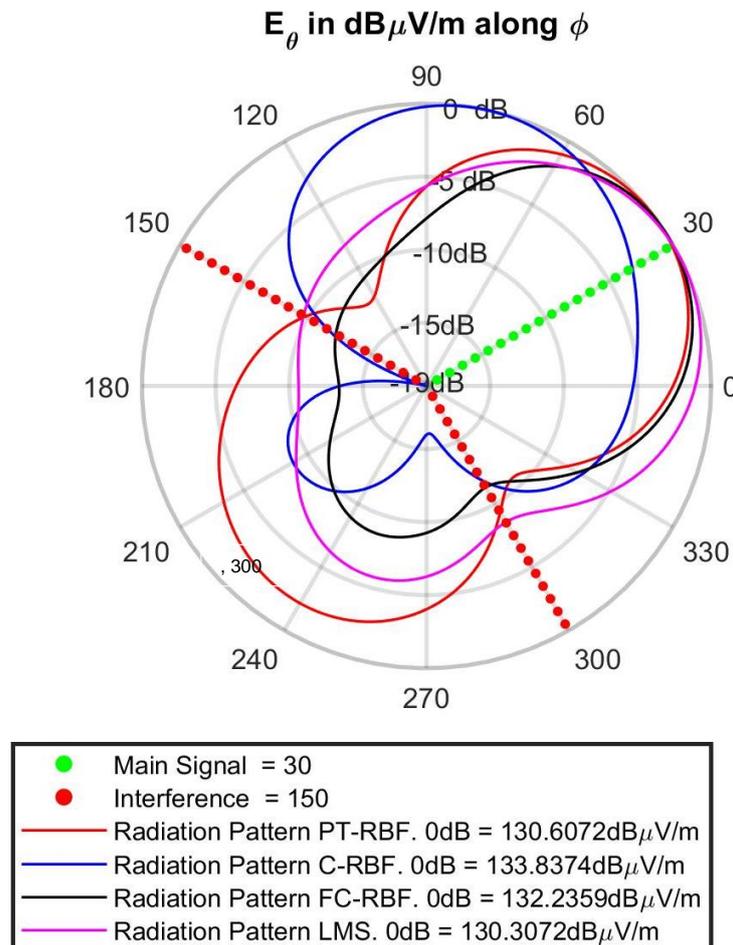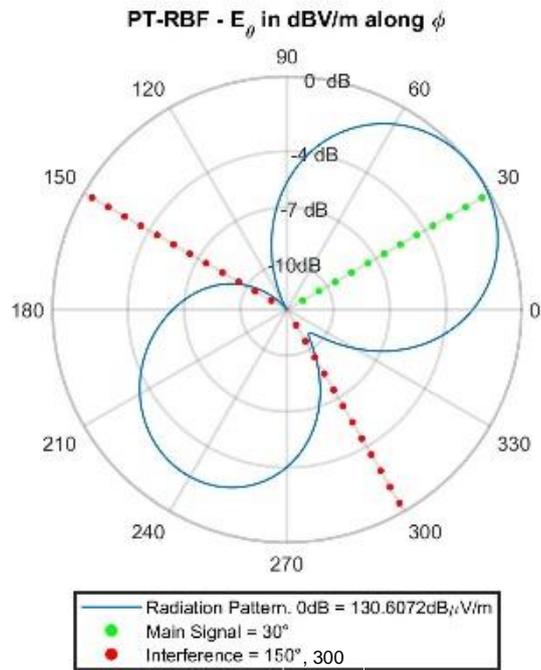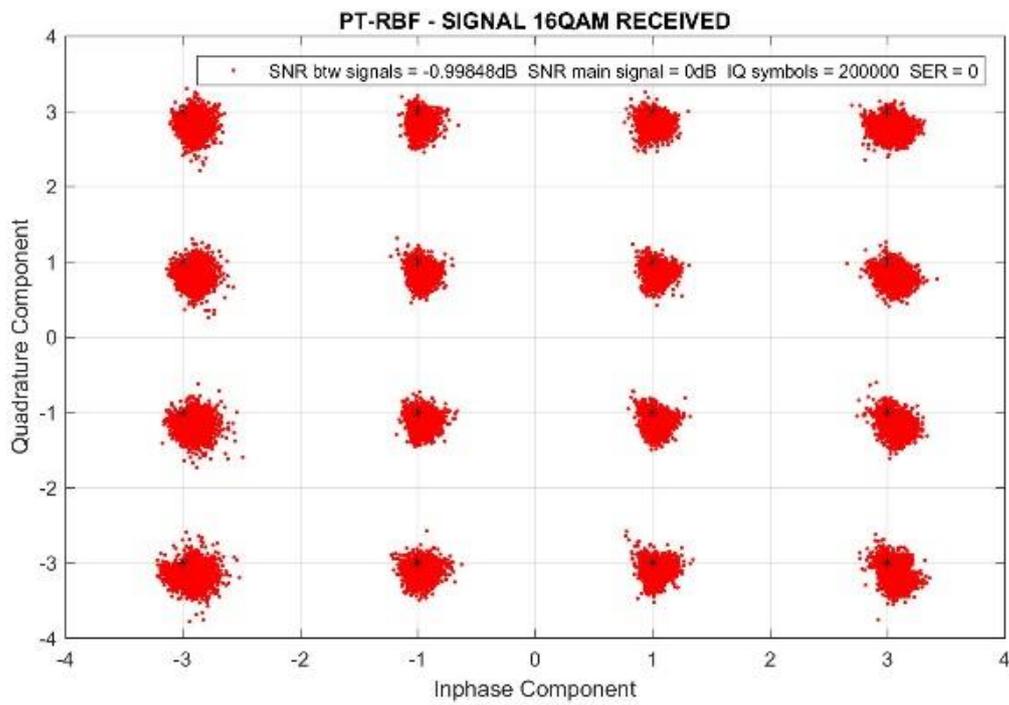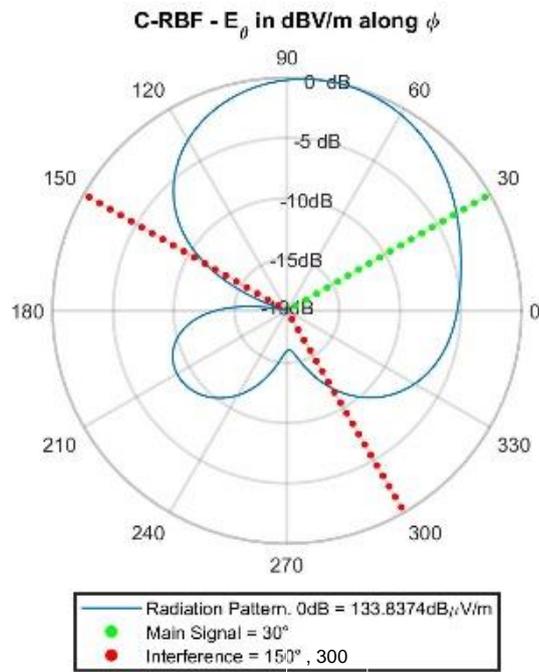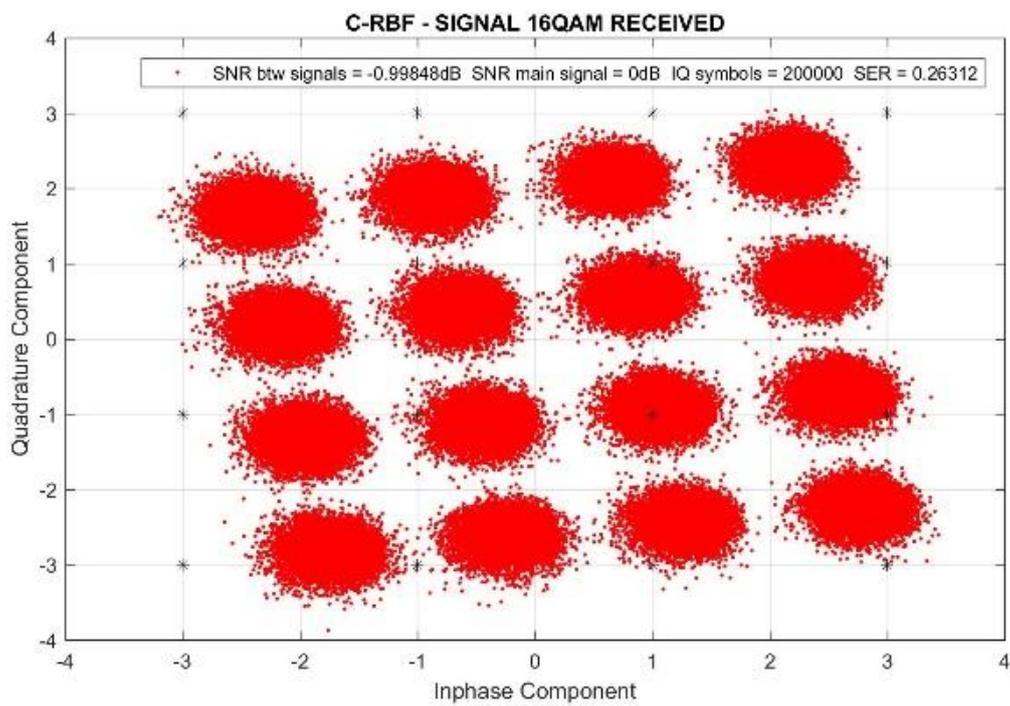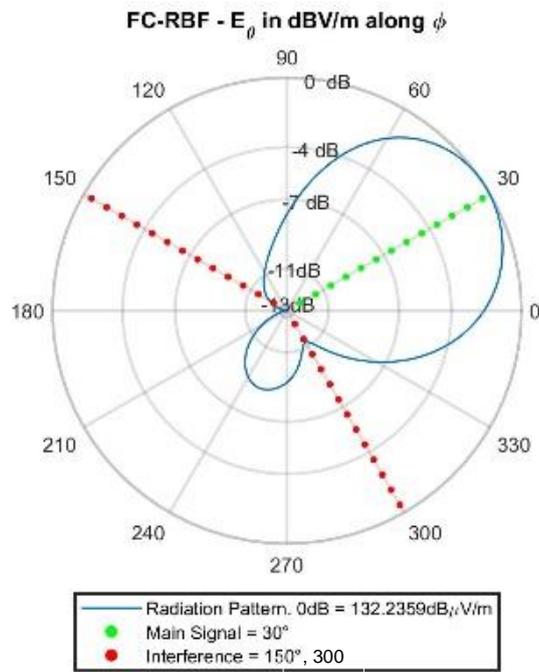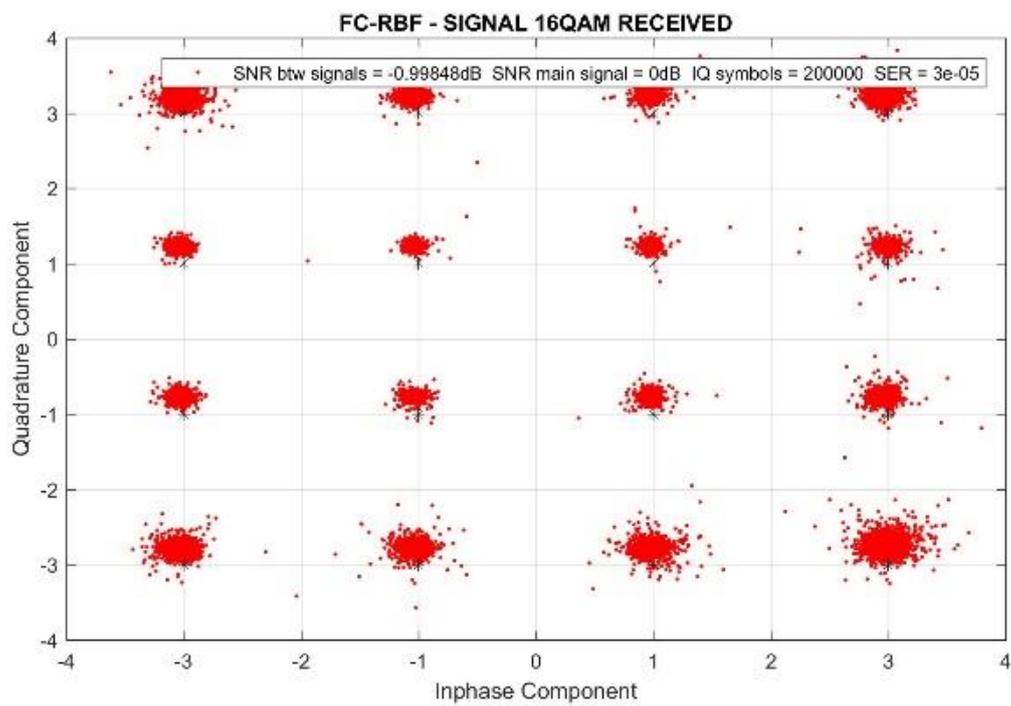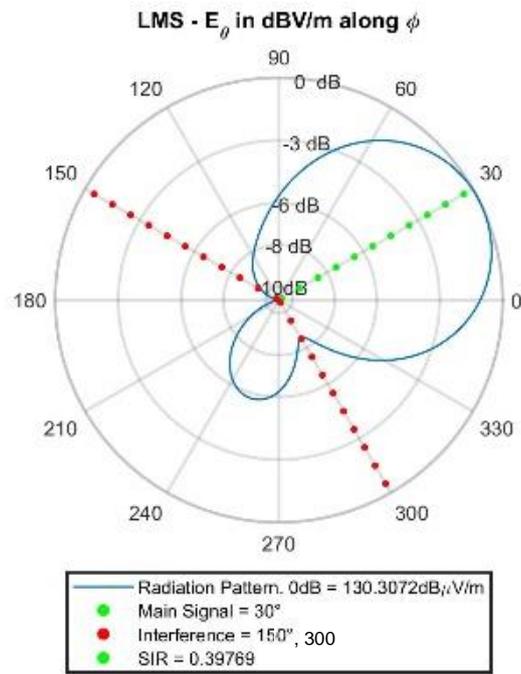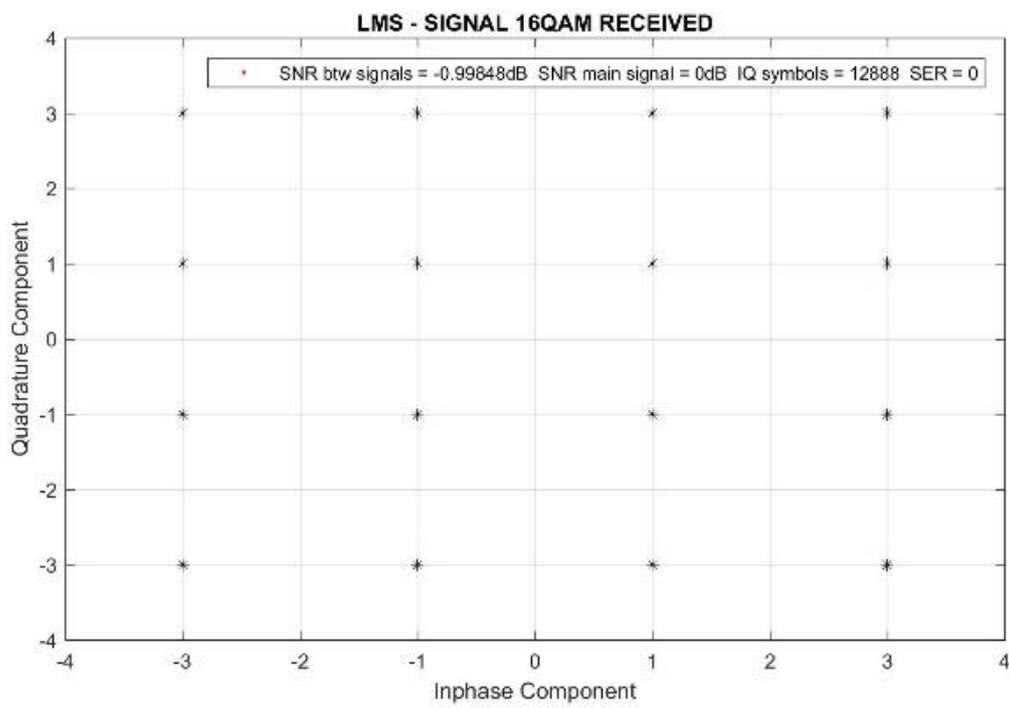
Source: (AUTHOR, 2018).

Table 7 – Summary of static case 3

D.S. DOA @ 30°	Gain [dB]			
	PT-RBF	C-RBF	FC-RBF	LMS
I.S. DOA @ 150°	−10.1	−19	−12	−9.5
I.S. DOA @ 300°	−9.5	−11.5	−11.3	−8.5

SER	**0**	**0.26312**	**3.10^{-5}**	**0**

Source: (AUTHOR, 2018).

4.1.4 Scenario 4

This scenario has the desired user signal coming from the bearing angle $\phi = 1°$ and two interfering signals coming from $\phi = 150°$ and $300°$. WGN is added to the desired signal so that $SNR = 35dB$. The relation between the power of the desired signal and the power of each interfering signal is given by $SIR = 10dB$. Finally, the non-linearity expressed by Eq (2.142) is introduced at the RF front end in order to implement a more realistic simulation. It is possible to see, from Figure 30 (a), that PT-RBF has presented the highest attenuation of the interference signals. Note from Figure 30 (b) – (i) that the C-RBF was the only beamformer that presented a non-zero SER. All other beamformers yielded zero SER. Thus, from Figure 30, it is possible to infer that the PT-RBF performs better in a low SIR and non-linear scenario.

Figure 30 – Static case 4. (a) Radiation pattern diagram comparing all beamformers gains, (b) radiation pattern diagram for PT-RBF beamformer and (c) 16QAM constellation presenting the gain and SER respectively for PT-RBF beamformer, (d) radiation pattern diagram and (e) 16QAM constellation presenting the gain and SER respectively for C-RBF beamformer, (f) radiation pattern diagram and (g) 16QAM constellation presenting the gain and SER respectively for FC-RBF beamformer, (h) radiation pattern diagram and (i) 16QAM constellation presenting the gain and SER respectively for LMS beamformer.

(a)

PT-RBF - E$_\theta$ in dBV/m along ϕ


Radiation Pattern. 0dB = 132.4293dBμV/m
Main Signal = 30°
Interference = 150°, 300

(b)

PT-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -1.013dB   SNR main signal = 35dB   IQ symbols = 200000   SER = 0

(c)

C-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 133.7233dBμV/m
Main Signal = 30°
Interference = 150°; 300

(d)

C-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -1.013dB SNR main signal = 35dB IQ symbols = 200000 SER = 0.58544

(e)

FC-RBF - E_θ in dBV/m along φ

(f)

FC-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -1.013dB   SNR main signal = 35dB   IQ symbols = 200000   SER = 0

(g)

LMS - E_θ in dBV/m along ϕ

Radiation Pattern. 0dB = 130.304dBμV/m
Main Signal = 30°
Interference = 150°; 300
SIR = 0.39674

(h)

LMS - SIGNAL 16QAM RECEIVED

SNR btw signals = -1.013dB SNR main signal = 35dB IQ symbols = 200000 SER = 0

(i)
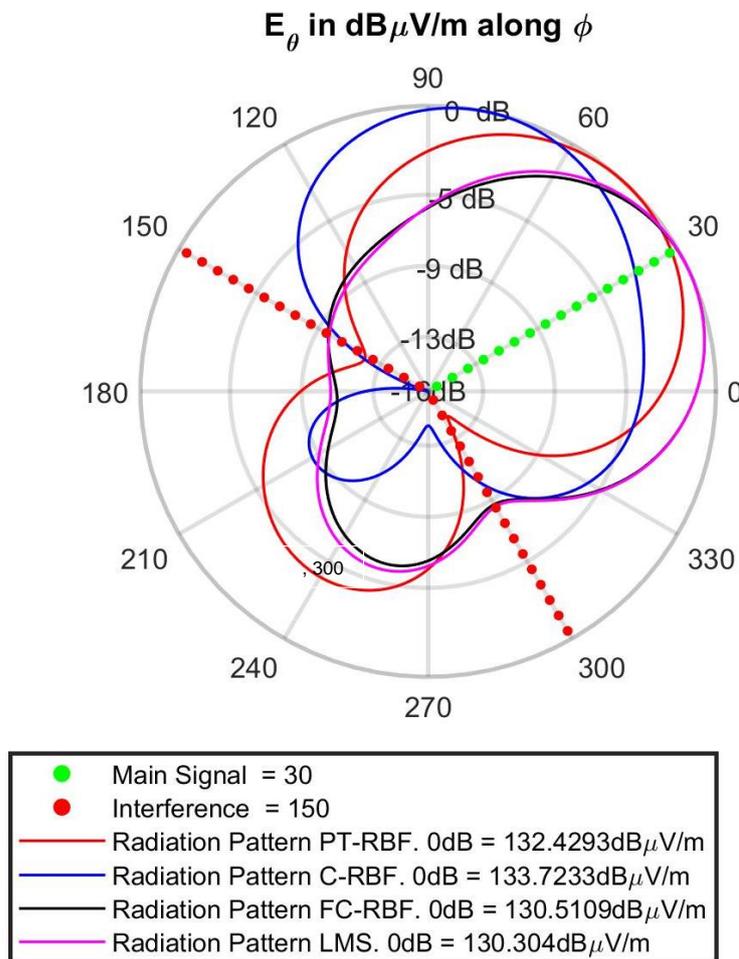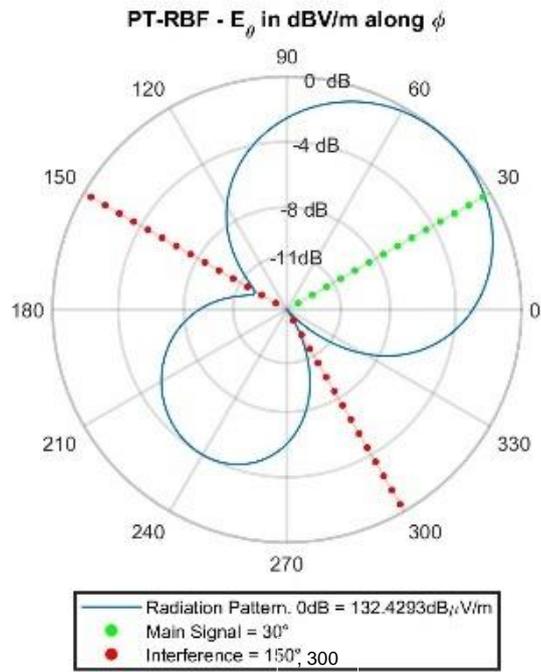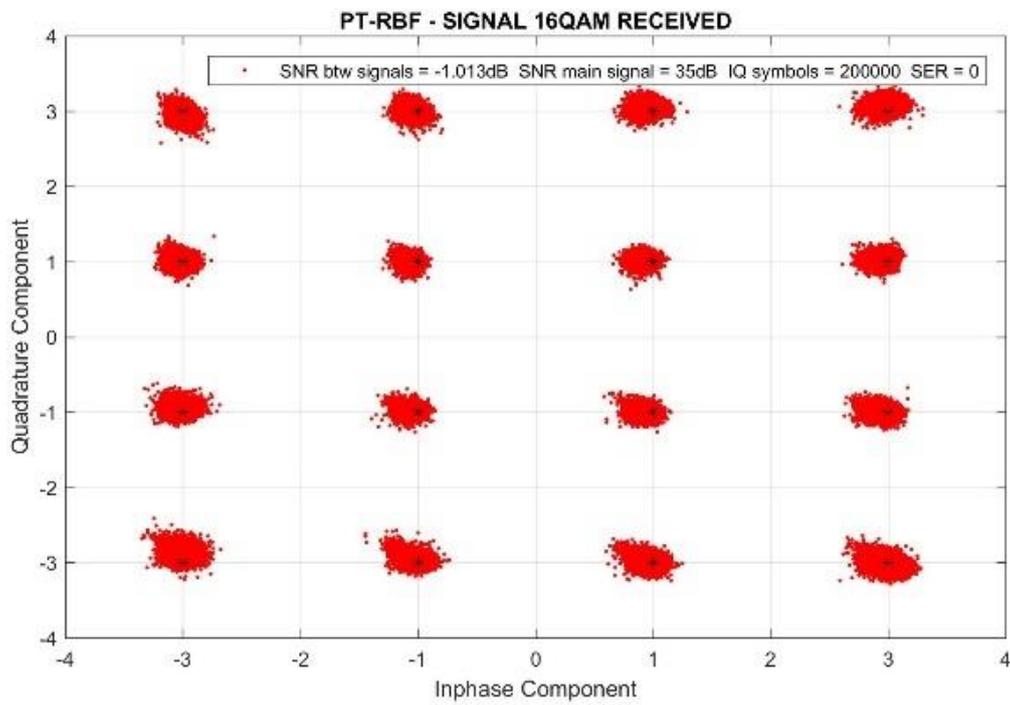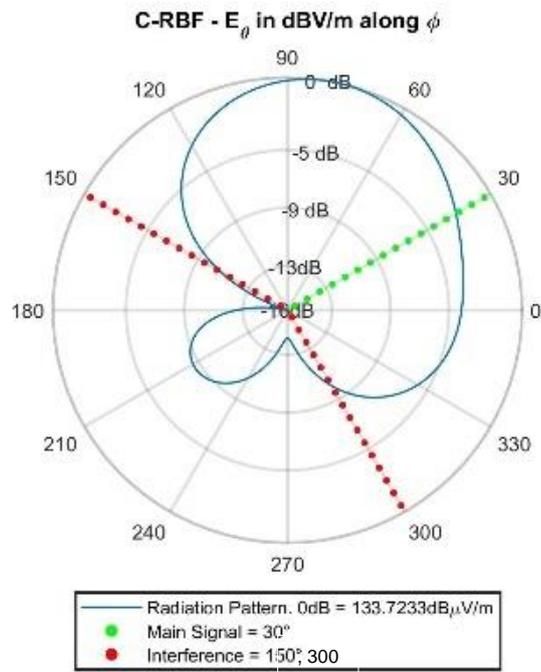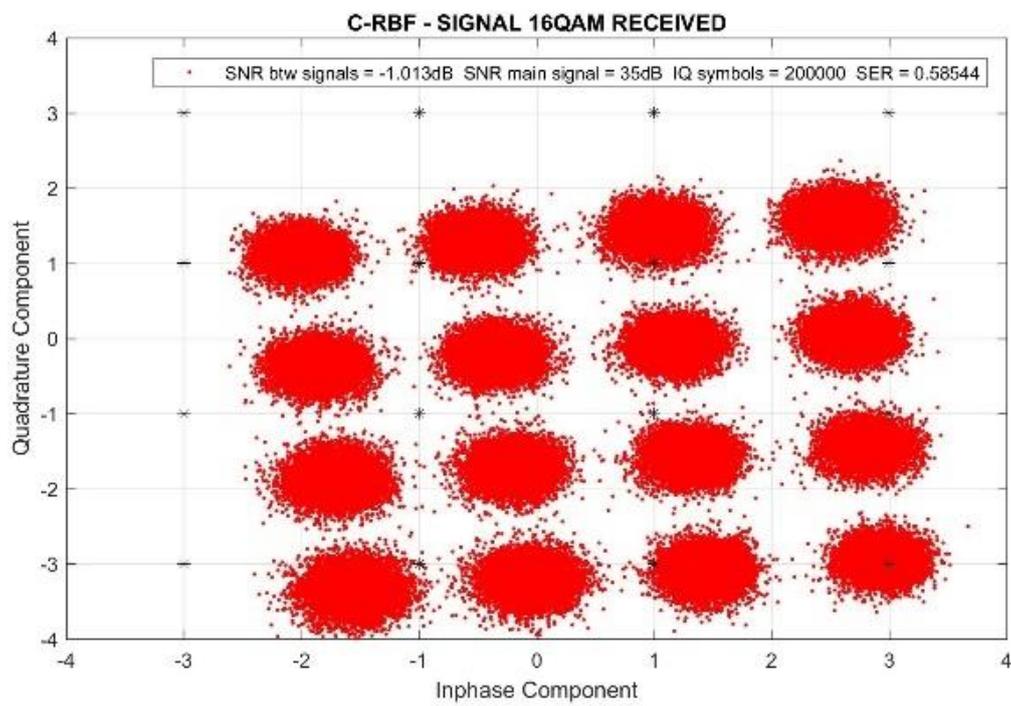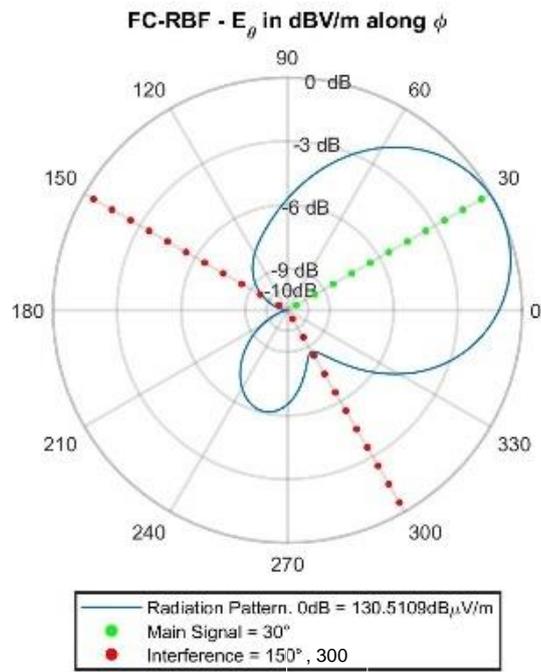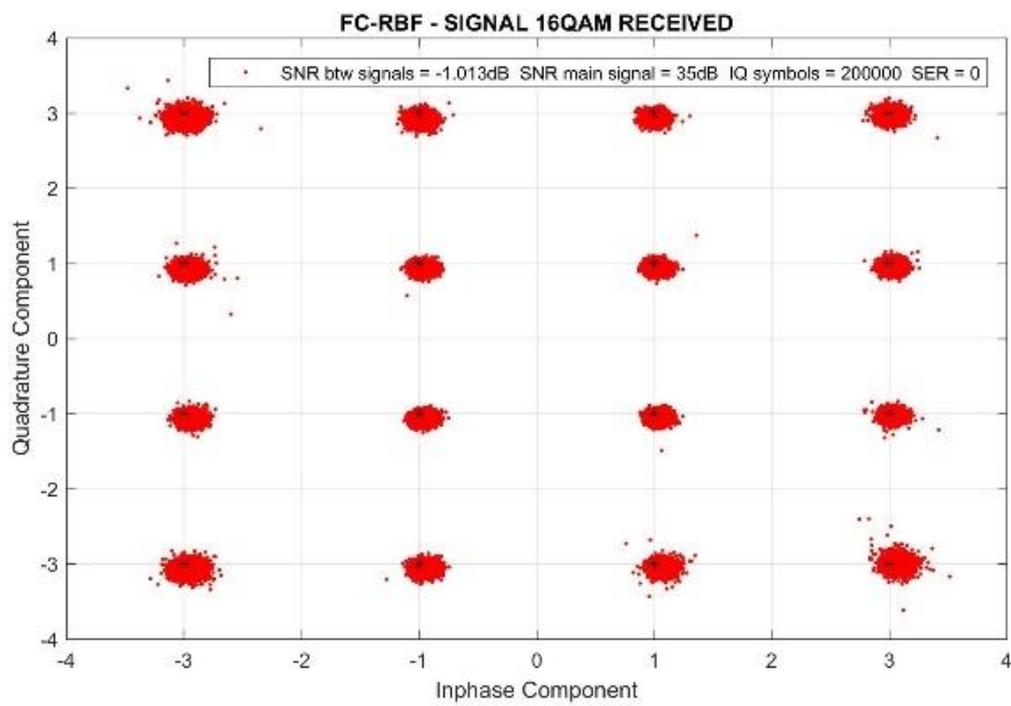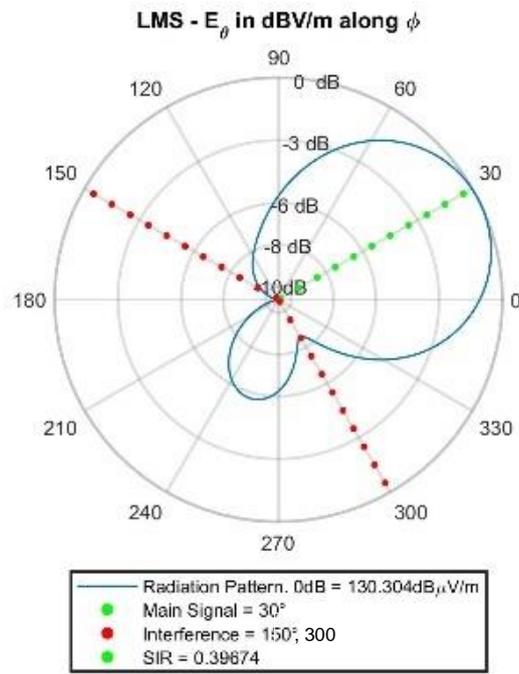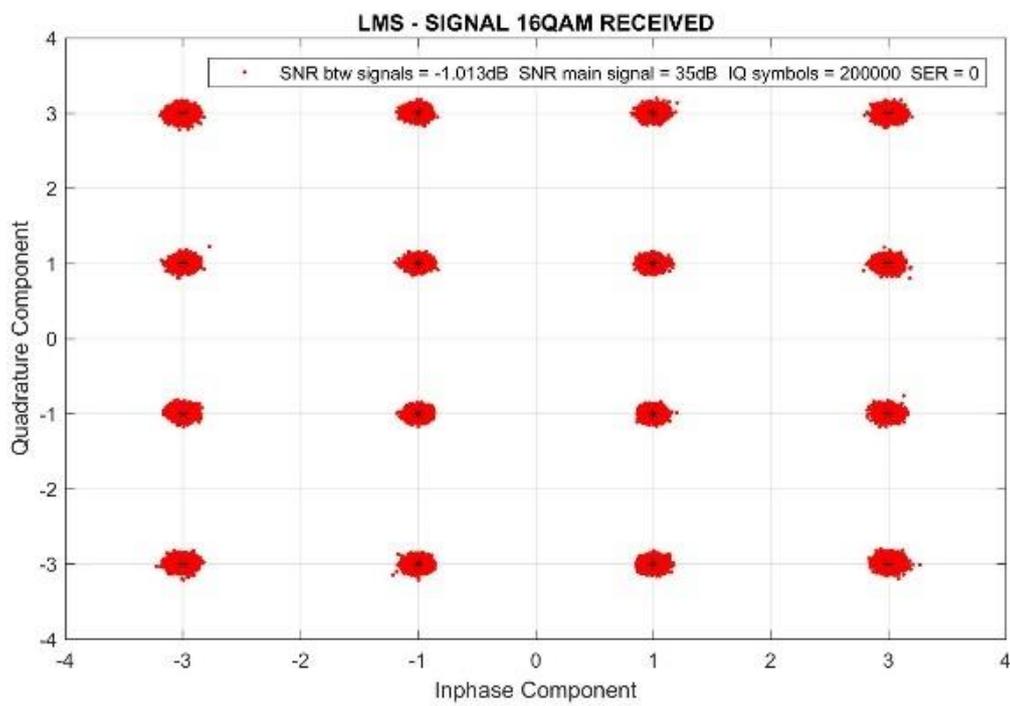
Source: (AUTHOR, 2018).

Table 8 – Summary of static case 4

D.S. DOA @ 30°	Gain [dB]			
	PT-RBF	C-RBF	FC-RBF	LMS
I.S. DOA @ 150°	-11.5	-10.4	-10.2	-9.5
I.S. DOA @ 300°	-15.1	-11.5	-8.6	-8.5

SER	**0**	**0.58544**	**0**	**0**

Source: (AUTHOR, 2018).

4.1.5 Scenario 5

Scenario 5 has the desired user signal coming from the bearing angle $\phi = 91°$ and three interfering signals coming from $\phi = 151°, 211°$ and $271°$. No WGN is added to the desired signal so that SNR $= \infty$. The relation between the power of the desired signal and the power of each interfering signal is given by SIR $= -10dB$. The non-linearity expressed by Eq (2.143) is introduced at the RF front end in order to implement a more realistic simulation. It is possible to see, from Figure 31 (a), that PT-RBF has presented the highest attenuation of the interference signals. From Figure 31 (b) – (m), it is possible to infer that, given the very low SIR, all beamformers yielded non-zero SER. Note also from the MSEA plots in Figure 31 that the LMS, the FC-RBF, the PT-RBF and the C-RBF beamformers took respectively $N = 2000, 7000, 10000$ and 25000 discrete time instants until convergence is achieved. This MSEA behavior has been observed in all simulations, with the LMS presenting the fastest convergence rate and the C-RBF presenting the slowest convergence rate. Note also from Figure 31 that the PT-RBF performs better in this very low SIR non-linear operating scenario, yielding the lowest SER $= 0.0009$.

Figure 31 – Static case 5. (a) Radiation pattern diagram comparing all beamformers gains, (b) radiation pattern diagram, (c) 16QAM constellation and (d) residual error curve presenting the gain, SER and MSEA respectively for PT-RBF beamformer. (e) radiation pattern diagram, (f) 16QAM constellation, (g) residual error curve presenting the gain, SER and MSEA respectively for C-RBF beamformer, (h) radiation pattern diagram, (i) 16QAM constellation and (j) residual error curve presenting the gain, SER and MSEA respectively for FC-RBF beamformer, (k) radiation pattern diagram, (l) 16QAM constellation and (m) residual error curve presenting the gain, SER and MSEA respectively for LMS beamformer.


(a)

PT-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 129.8246dBμV/m
Main Signal = 91°
Interference = 151, 210, 271

(b)

PT-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.012102dB   SNR main signal = 0dB   IQ symbols = 10000   SER = 0.0009

(c)

(d)

(e)

(f)

(g)

FC-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 130.8398dBμV/m
Main Signal = 91°
Interference = 151°, 210, 271

(h)

FC-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.012102dB   SNR main signal = 0dB   IQ symbols = 7000   SER = 0.026143

(i)

(j)

LMS - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 126.0255dBμV/m
Main Signal = 91°
Interference = 151°, 210, 271
SIR = 0.333

(k)

LMS - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.012102dB SNR main signal = 0dB IQ symbols = 1000 SER = 104.055

Quadrature Component

Inphase Component

(l)

(m)
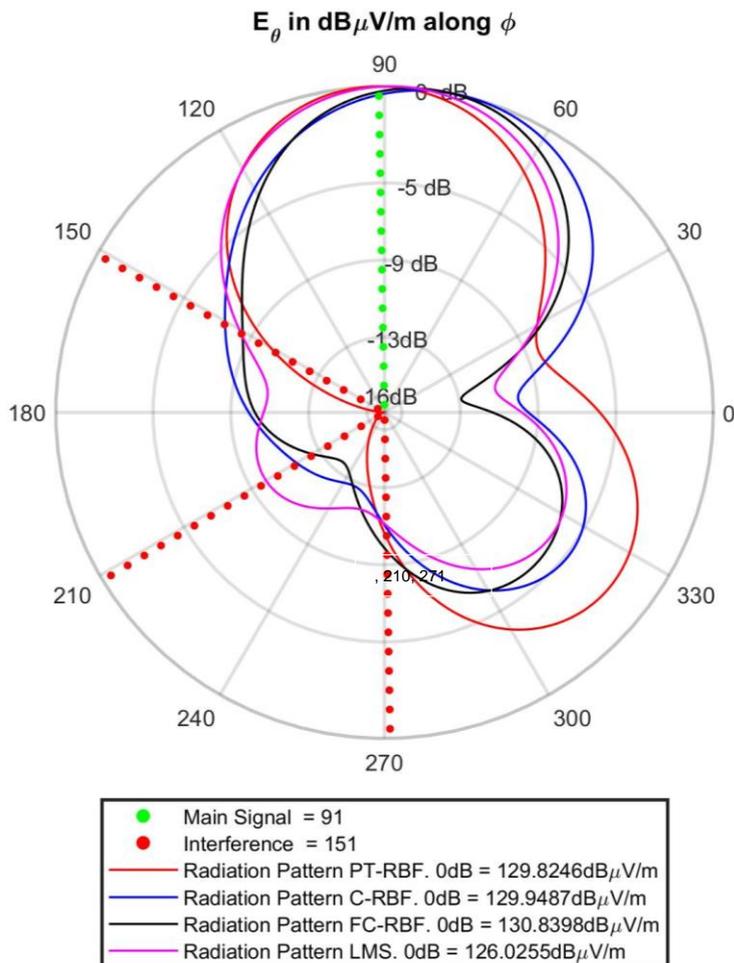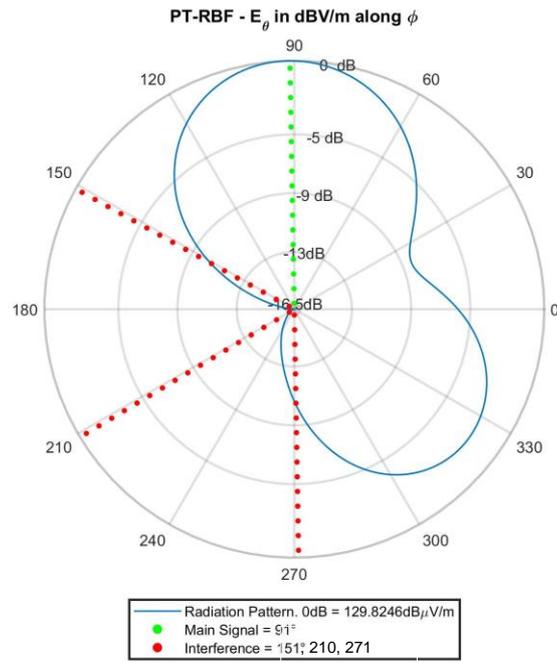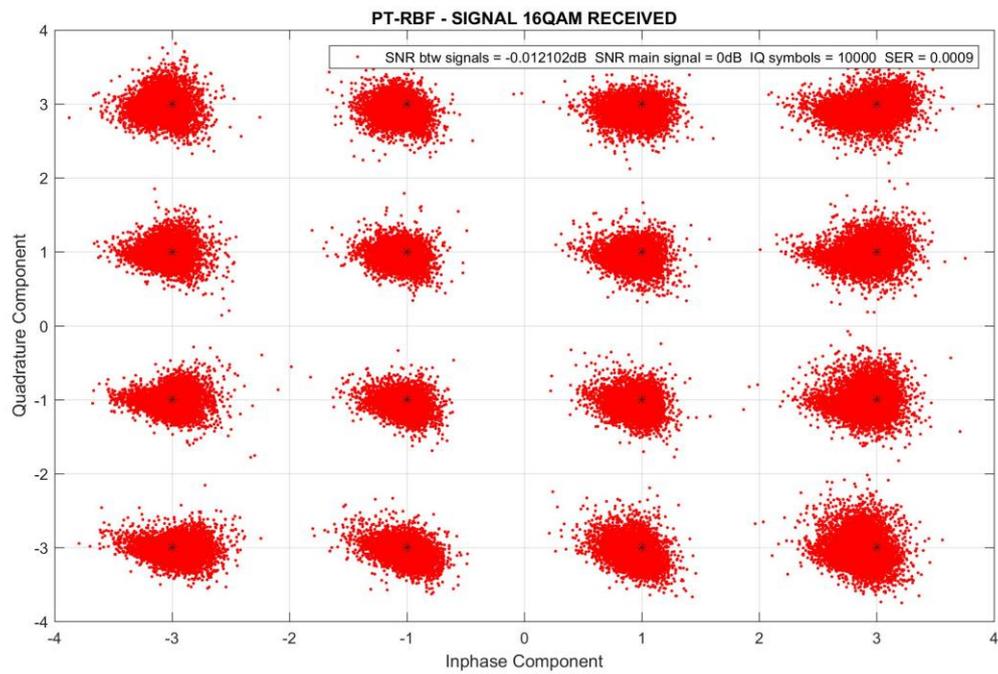
Source: (AUTHOR, 2018).

Table 9 – Summary of static case 5

D.S. DOA @ 91°	Gain [dB]			
	PT-RBF	C-RBF	FC-RBF	LMS
I.S. DOA @ 151°	−10.4	−7.5	−8.6	−8.5
I.S. DOA @ 211°	−16.6	−11.7	−12.7	−9.9
I.S. DOA @ 271°	−10.4	−10.9	−9.7	−11

SER	9.10^{-4}	0.1314	0.026143	0.686

Source: (AUTHOR, 2018).

4.1.6 Scenario 6

Scenario 6 has the desired user signal coming from the bearing angle $\phi = 91°$ and three interfering signals coming from $\phi = 151°, 211°$ and $271°$. T No WGN is added to the desired signal so that SNR $= \infty$. The relation between the power of the desired signal and the power of each interfering signal is given by SIR $= -10dB$. In this scenario, the non-linearity at the front end was not introduced. It is possible to see, from Figure 32 (a), that all beamformers presented similar attenuation of the interference signals. From Figure 32 (b) – (m), however, we see that the only beamformers capable of demodulating without symbol error is the LMS and the PT-RBF. Therefore, from Figure 32, it is possible to infer that the LMS and the PT-RBF performs better in this linear very low SIR scenario.

Figure 32 – Static case 6. (a) Radiation pattern diagram comparing all beamformers gains, (b) radiation pattern diagram, (c) 16QAM constellation and (d) residual error curve presenting the gain, SER and MSEA respectively for PT-RBF beamformer. (e) radiation pattern diagram, (f) 16QAM constellation, (g) residual error curve presenting the gain, SER and MSEA respectively for C-RBF beamformer, (h) radiation pattern diagram, (i) 16QAM constellation and (j) residual error curve presenting the gain, SER and MSEA respectively for FC-RBF beamformer, (k) radiation pattern diagram, (l) 16QAM constellation and (m) residual error curve presenting the gain, SER and MSEA respectively for LMS beamformer.

(a)

(b)

(c)

(d)

C-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 130.1894dBμV/m
Main Signal = 91°
Interference = 151° , 210, 271

(e)

C-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.0072dB SNR main signal = 0dB IQ symbols = 33000 SER = 3.0303e-05

(f)

(g)

FC-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 128.4552dBμV/m
Main Signal = 91°
Interference = 151, 210, 271

(h)

FC-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.0072517dB   SNR main signal = 0dB   IQ symbols = 21000   SER = 0.0046667

(i)

(j)

LMS - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 129.9541dBμV/m
Main Signal = 91°
Interference = 151°, 210, 271
SIR = 0.33395

(k)

LMS - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.0072517dB SNR main signal = 0dB IQ symbols = 6000 SER = 0

(l)

LMS - MSEA performance
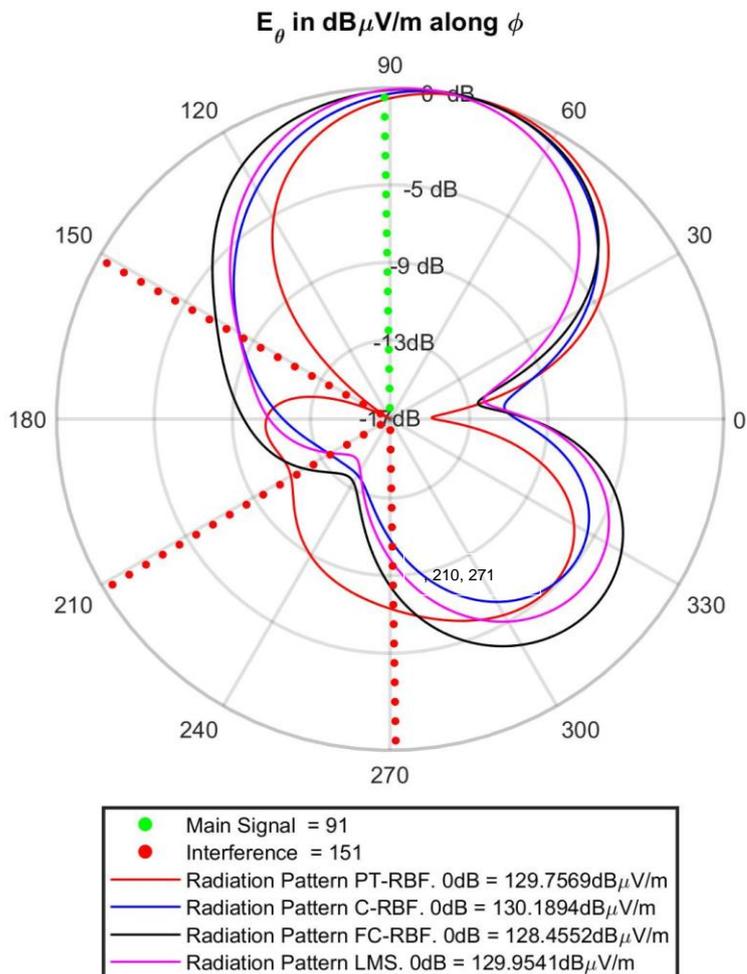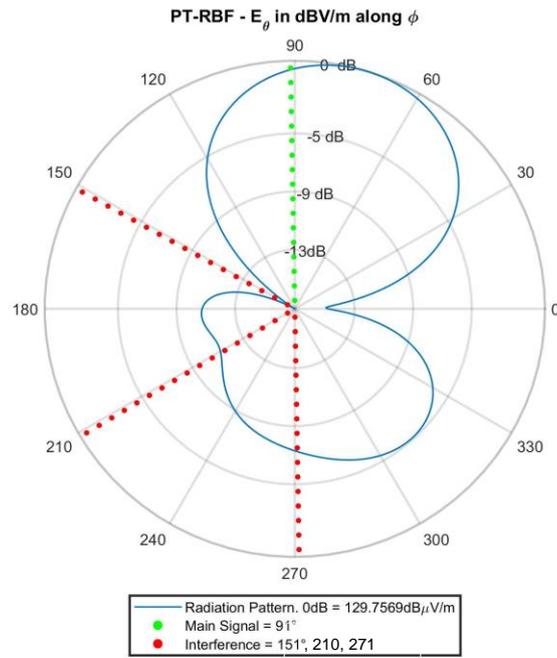
Source: (AUTHOR, 2018).

Table 10 – Summary of static case 6

D.S. DOA @ 91°	Gain [dB]			
	PT-RBF	C-RBF	FC-RBF	LMS
I.S. DOA @ 151°	-17.1	-8.3	-6.9	-8.2
I.S. DOA @ 211°	-11.1	-13.5	-12	-13.7
I.S. DOA @ 271°	-7.2	-10.7	-8.2	-9.7

SER	**0**	$3.03 \cdot 10^{-5}$	$4.67 \cdot 10^{-3}$	**0**

Source: (AUTHOR, 2018).

4.1.7 Scenario 7

Scenario 7 has the desired user signal coming from the bearing angle $\phi = 1°$ and nine interfering signals coming from $\phi = 60°, 90°, 120°, 150°, 180°, 210°, 240°, 270°, 300°$ and $330°$. No WGN is added to the desired signal so that $\text{SNR} = \infty$. The relation between the power of the desired signal and the power of each interfering signal is given by $\text{SIR} = -10dB$. In this scenario the nonlinearity at the front end was not introduced. It is possible to see, from Figure 33 (a), that PT-RBF has presented the highest attenuation of the interference signals. Note from Figure 33 (b) – (m) that the PT-RBF is the only beamformer capable of demodulating the received desired signal with zero SER. Thus, note from Figure 33, that the PT-RBF performs better in a linear scenario with very low SIR and a significant number of interfering signals.

Figure 33 – Static case 7. (a) Radiation pattern diagram comparing all beamformers gains, (b) radiation pattern diagram, (c) 16QAM constellation and (d) residual error curve presenting the gain, SER and MSEA respectively for PT-RBF beamformer. (e) radiation pattern diagram, (f) 16QAM constellation, (g) residual error curve presenting the gain, SER and MSEA respectively for C-RBF beamformer, (h) radiation pattern diagram, (i) 16QAM constellation and (j) residual error curve presenting the gain, SER and MSEA respectively for FC-RBF beamformer, (k) radiation pattern diagram, (l) 16QAM constellation and (m) residual error curve presenting the gain, SER and MSEA respectively for LMS beamformer.

(a)

PT-RBF - E$_\theta$ in dBV/m along ϕ

(b)

PT-RBF - SIGNAL 16QAM RECEIVED

(c)

(d)

C-RBF - E$_\theta$ in dBV/m along ϕ

Radiation Pattern. 0dB = 127.713dBμV/m
Main Signal = 1°
Interferenc

(e)

C-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = 10.0025dB   SNR main signal = 0dB   IQ symbols = 115000   SER = 0.59277

(f)

(g)

FC-RBF - E$_\theta$ in dBV/m along ϕ

(h)

FC-RBF - SIGNAL 16QAM RECEIVED

(i)

(j)

LMS - E$_\theta$ in dBV/m along φ

(k)

FC-RBF - SIGNAL 16QAM RECEIVED

(l)

(m)
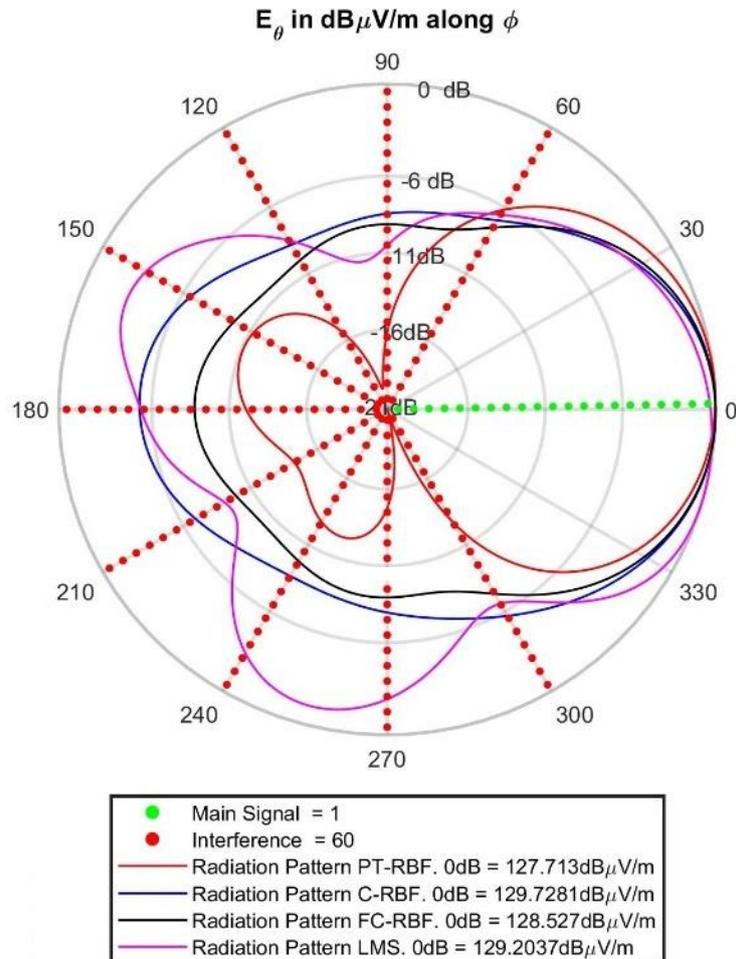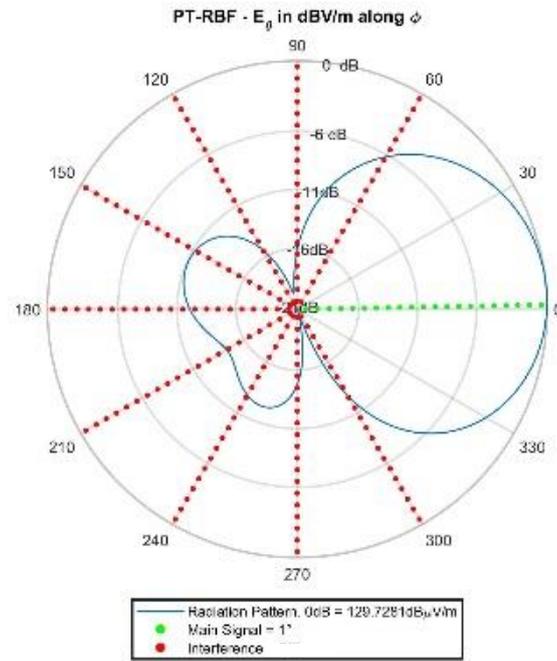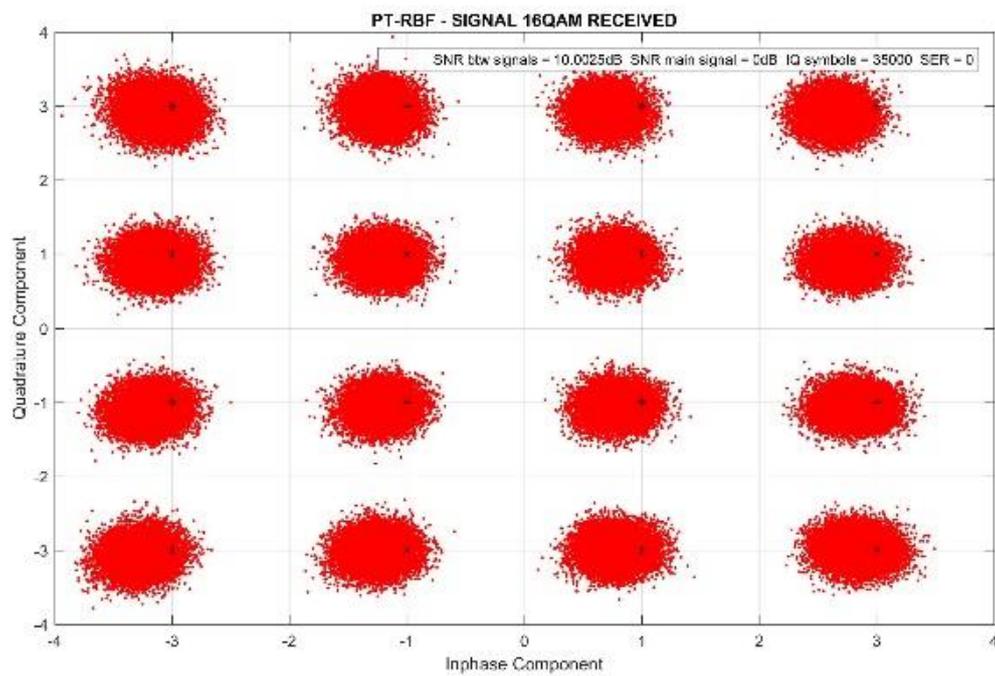
Source: (AUTHOR, 2018).

Table 11 – Summary of static case 7

D.S. DOA @ 1°	Gain [dB]			
	PT-RBF	C-RBF	FC-RBF	LMS
I.S. DOA @ 60°	−6	−7	−7.5	−6.4
I.S. DOA @ 90°	−16.2	−8.5	−9.1	−10.5
I.S. DOA @ 120°	−15.4	−8.8	−10	−8.3
I.S. DOA @ 150°	−11.2	−6.4	−9.8	−3.1
I.S. DOA @ 180°	−12	−5.2	−8.7	−5.3
I.S. DOA @ 210°	−14.3	−7	−10	−9.1
I.S. DOA @ 240°	−12.7	−8.6	−9.7	−2.7
I.S. DOA @ 270°	−14.8	−8	−9	−2.3
I.S. DOA @ 300°	−12.7	−5.7	−7.2	−6.6
SER	**0**	$\mathbf{0.5927}$	$\mathbf{4.4545 \cdot 10^{-3}}$	$\mathbf{13667 \cdot 10^{-2}}$

Source: (AUTHOR, 2018).

4.1.8 Summary of Static Scenarios

For non-linear low SIR (high interference power) operating scenarios, as the scenarios of Figure 27 and Figure 31, the PT-RBF has presented the best overall performance by suppressing the interfering signals, pointing the radiation diagram main lobe to the incoming desired signal and yielding the lowest SER over all other beamformers. Such scenarios approach the operational scenarios that new multiuser technologies (5G, IoT, etc) are expected

to operate with. Given the high number of users, these new technologies operate in a highly polluted electromagnetic environment, plenty of severe interfering signals. Thus, the proposed algorithm is an effective innovation for next generation beamforming problem.

As for the convergence speed, Figure 31 and Figure 32 also illustrate the MSEA for each beamformer. Note that the LMS is the fastest algorithm (as expected, given its linear approach), followed by the FC-RBF, by the PT-RBF and finally, by the C-RBF.

For linear low SIR scenarios, the PT-RBF also presented a better performance on energy focusing and symbol demodulation, as can be seen in Figure 33. For moderate SIR linear scenarios, Figure 29 shows that the LMS algorithm has a better performance, demodulating the signal with no error, presenting less dispersive constellations at the beamformer output. However, it is important to note that perfectly linear scenarios are really rare on real world.

In the presence of non-linearity, the LMS fast convergence is hampered. Figure 31 shows an example of a low SIR non-linear scenario and Figure 32 shows a low SIR linear scenario. In the non-linear case the LMS yields a significantly dispersed constellation. The nulls and main lobe on the radiation pattern are not well defined in comparison to the PT-RBF.

Figure 27 and Figure 28 shows a low SIR case and a moderate SIR case respectively. Note that, for moderate SIR cases, although the C-RBF has a reasonable SER performance, it yields a better suppression of the interfering signals and a better-defined main lobe on the radiation diagram. Nevertheless, it was observed in all simulations that the C-RBF has the slowest convergence rate.

On the other hand, for low SIR cases, the C-RBF cannot converge to the proposed MSEA threshold and its SER is higher than all other beamformers. For low SIR cases, the FC-RBF algorithm performs well on interference suppressing and on main lobe pointing, presenting a high convergence rate. However, the FC-RBF cannot handle the identification of the IQ symbols on the constellation diagram, yielding a higher SER. Despite the FC-RBF to converge faster than the PT-RBF, in non-linear scenarios with a high level of interference, the PT-RBF has an overall better performance on interference suppressing and on main lobe definition, presenting the lowest SER over all other beamformers.

In conclusion, the PT-RBF based beamformer was the only algorithm that presented satisfactory and uniform performance in all scenarios. The PT-RBF performance is overcome by some other solution in specific scenarios, however, none of the other solutions presented satisfactory and uniform performance in all evaluated scenarios.

The gains and SERs of each mentioned scenario can be seen on Tables 5, 6, 7, 8, 9 and 10.

4.2 DYNAMIC SCENARIOS

The dynamic scenario implementation follows the general procedure of Table 2 in Section 3.2 and considers that the signals that impinge over the receiving UCA are originated at a moving TX with respect to a static UCA. The movement of the TX is represented by a non-constant DOA of the desired signal, or a non-constant DOA of the interfering signal. If both the TX and the receiving UCA are moving, we still can consider the UCA as static and assume that the TX moves with respect the static UCA with a velocity given by the relative velocity between the TX and the UCA. For simplicity, it is considered that the TX moves circularly around the UCA in the *xy* plane with a uniform speed v, so that the DOA of the impinging signal varies in the range $1° < \phi < 360°$.

One of the most demanding challenges in the context of wireless networks are encountered in the dynamic operation scenario of wireless networks for CAS (Close Air Support) military operations (AWARE-; TRACKING, 2016; BRUCE R. PIRNIE et al., [s.d.]; CHEN et al., 2008; DARPA, [s.d.]; KOPP, 2010)

In this context, let's assume an operating scenario in which two aircrafts exchange tactical data via a wireless mesh network in a highly jammed environment and that the IQ symbol rate of the digital modulation is 1.0 Msymbol/s. Aircraft A transmits the desired signal to the Aircraft B, which houses the receiving UCA. The desired signal is jammed by the interfering signal of the TX of the opposing force (possibly the jammer of another aircraft or of a ground station). Both TX (Aircraft A) and the receiving UCA (Aircraft B) are moving in high speed, so the UCA can be considered static and the TX can be considered as moving with respect to the UCA with a velocity given by the relative velocity between the TX (Aircraft A) and the UCA (Aircraft B). Let's assume, as an approximation of a real scenario, the relative velocity between the TX (Aircraft A) and the UCA (Aircraft B) is approximately $v = 850$ knot and that the TX (Aircraft A) moves circularly around the UCA (Aircraft B) in the *xy* plane with a radius $r = 100$m, so that the DOA of the impinging signal varies in the range $1° < \phi < 360°$. This operational scenario defines a DOA angular rate $\frac{d\phi}{dt}$ of $1°$ per the time lapse of 4000 IQ symbols or, equivalently, a DOA angular rate of $\frac{d\phi}{dt} = 250 \frac{°}{s}$.

In this framework, the dynamic scenarios are summarized in Table 12. Note that a linear front end was only considered in the dynamic scenario 1. All other scenarios represent more realistic and severe conditions, which include non-linearity at the receiver front end. As seen

in the previous section, the LMS can't handle non-linearity at the RF front end. Thus, the LMS beamformer results are presented only for the dynamic scenario 1, which is a linear scenario. For dynamic non-linear scenarios 2, 3 and 4, the LMS yielded a received symbol constellation with a high dispersion level (similar to the Figure 31(i) of Scenario 5 in previous section 4.1.5), resulting in SER > 0.5. Such high SER indicates that the LMS failed for all these dynamic non-linear scenarios.

Table 12 - Summary of the dynamic scenarios.

Scenario	I.S. DOA	D.S. DOA	SIR	NL
1	170°	1° to 360°	-10dB	no
2a	270°	1° to 360°	35dB	yes
2b	1° to 360°	90°	35dB	yes
3	270°	1° to 360°	10dB	yes
4a	270°	1° to 360°	-10dB	yes
4b	1° to 360°	90°	-10dB	yes

Source: (AUTHOR, 2018).

The results are shown in terms of the polar plot of the SER curve for each beamforming as a function of the DOA angle ϕ of the desired/interfering signal, with ϕ varying in the range $1° < \phi < 360°$, as shown by Figure 34 (a), and for a DOA angular rate of $\frac{d\phi}{dt} = 250\frac{°}{s}$.

For the sake of clarity, the results are also shown in terms of the rectangular plot of the SER curve for each beamforming as a function of the DOA angle ϕ of the D.S., as shown by Figure 34 (b).

4.2.1 Scenario 1

The first evaluated dynamic scenario has the D.S. DOA varying in the range $1° < \phi < 360°$ and one static I.S. with DOA angle $\phi = 170°$. The relation between the power of the D.S. and the power of the I.S. is given by $SIR = -10dB$. The receiver front end is linear. It is possible to see, from Figure 34, that the PT-RBF and the LMS beamformer yield a better isolation from the I.S, since their respective SERs are high only when the D.S. and the I.S. are located in the same DOA. Notice from Figure 34 (a), that the LMS presents a high SER over a smaller angular range of the D.S. DOA in comparison to the PT-RBF which presents a slightly larger angular range of the D.S. DOA that corresponds to the same high SER. From Figure 34 (b), it is possible to notice that LMS has a higher residual SER than the PT-RBF, which makes the PT-RBF more reliable. In this scenario, the other networks failed to converge and to follow the varying D.S. DOA as a consequence of the very low SIR. Thus, the LMS performs better

in linear very low SIR scenarios. Same behavior is observed for the LMS in moderate SIR linear scenarios

Figure 34 – Dynamic case 1

(a)
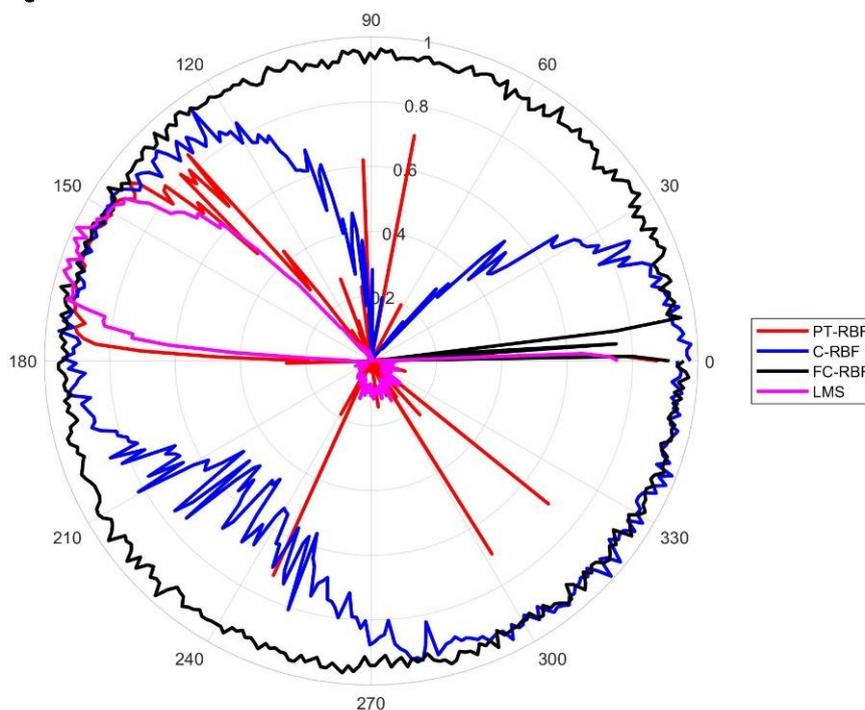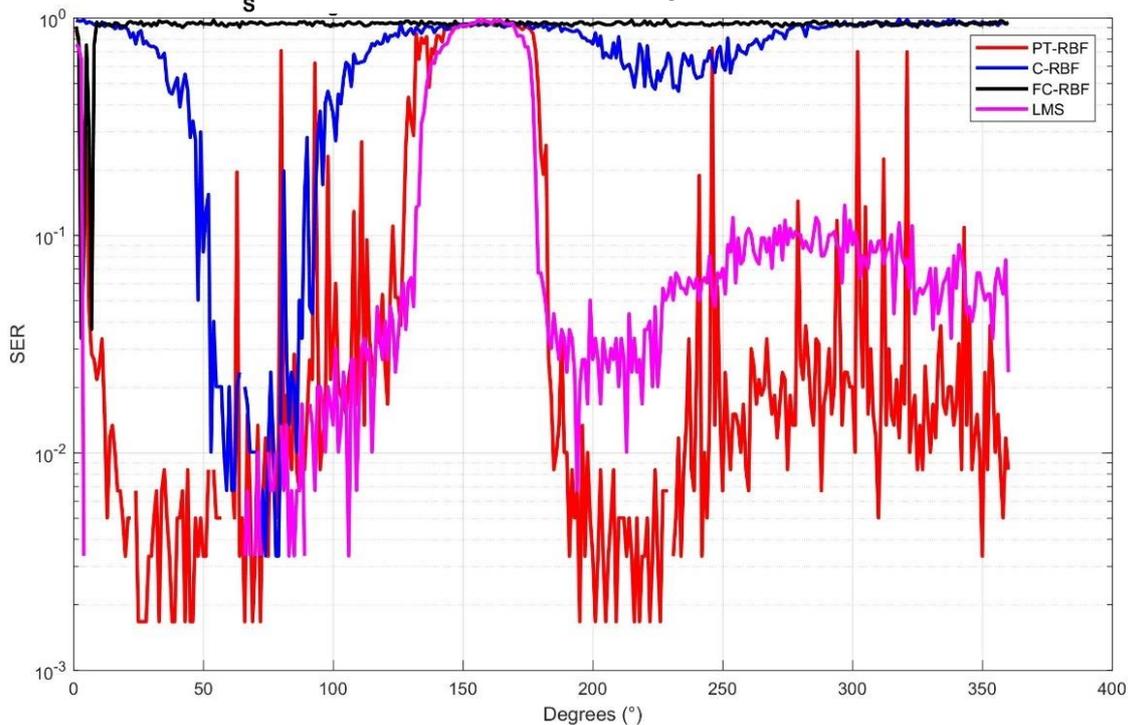

(b)

4.2.2 Scenario 2

In this dynamic scenario the relation between the power of the D.S. and the power of the I.S. is given by SIR = $35dB$. The non-linearity expressed by Eq (2.144) is introduced at the RF front end in order to implement a more realistic simulation.

The operational conditions in terms of D.S. and I.S. are:

I. Figure 35 and Figure 36 (see row 2a of Table 12): D.S. DOA varying in the range $1° < \phi < 360°$ and a stationary I.S. DOA at $\phi = 270°$;

II. Figure 37 (see row 2b of Table 12): D.S. DOA stationary at $\phi = 90°$ and I.S. DOA varying in the range $1° < \phi < 360°$.

Note in Figure 35 and in Figure 37 that the C-RBF beamformer presents SER > 0.5 along a larger angular range of the varying DOA, not being able to isolate the D.S from the I.S in this angular range. Note also that the FC-RBF beamformer presents a SER > 0.5 only over a smaller angular range of the varying DOA, thus performing better in this very low SIR non-linear dynamic scenario.

Figure 35 – Dynamic case example performance

(a)

SER curve for each beamformer for 16-QAM modulation and AWGN

(b)
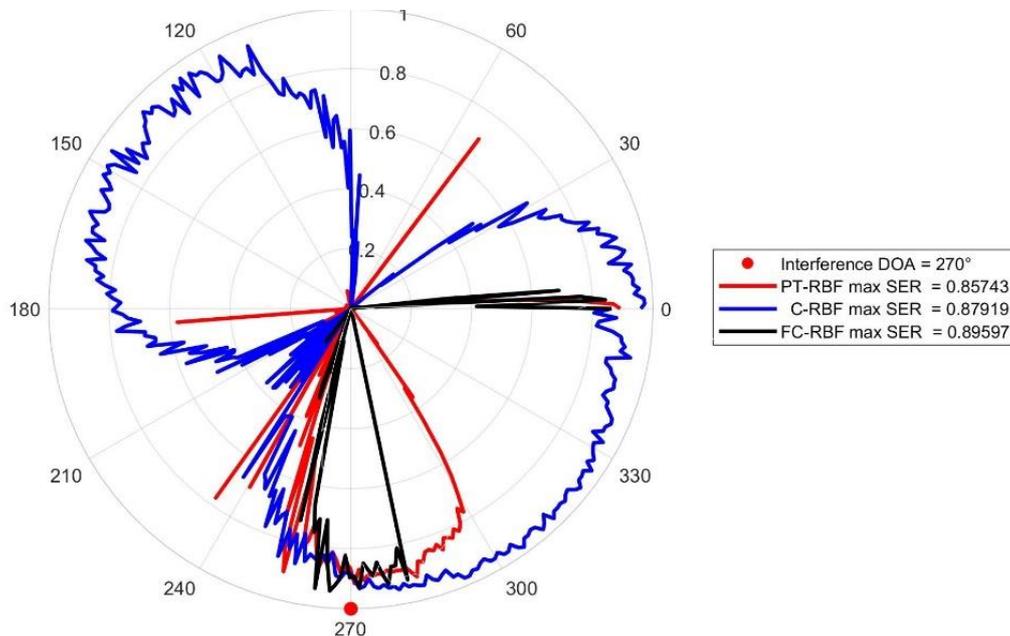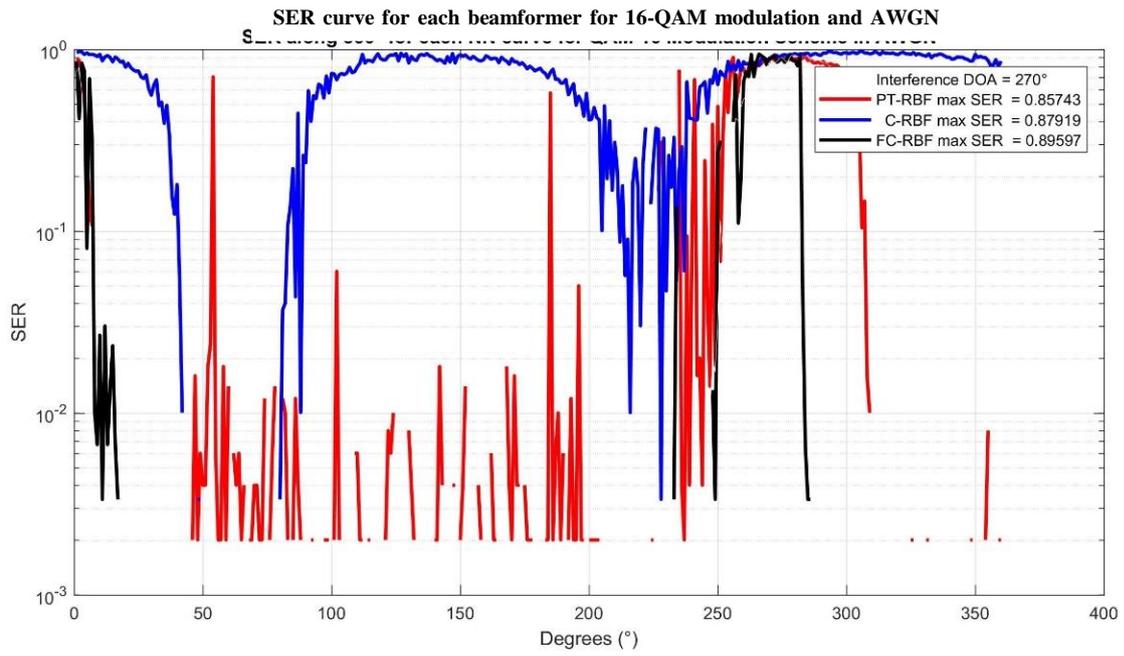
Source: (AUTHOR, 2018).

Figure 36 – Dynamic case 2a

SER curve for each beamformer for 16-QAM modulation and AWGN channel

(a)

SER curve for each beamformer for 16-QAM modulation and AWGN

(b)
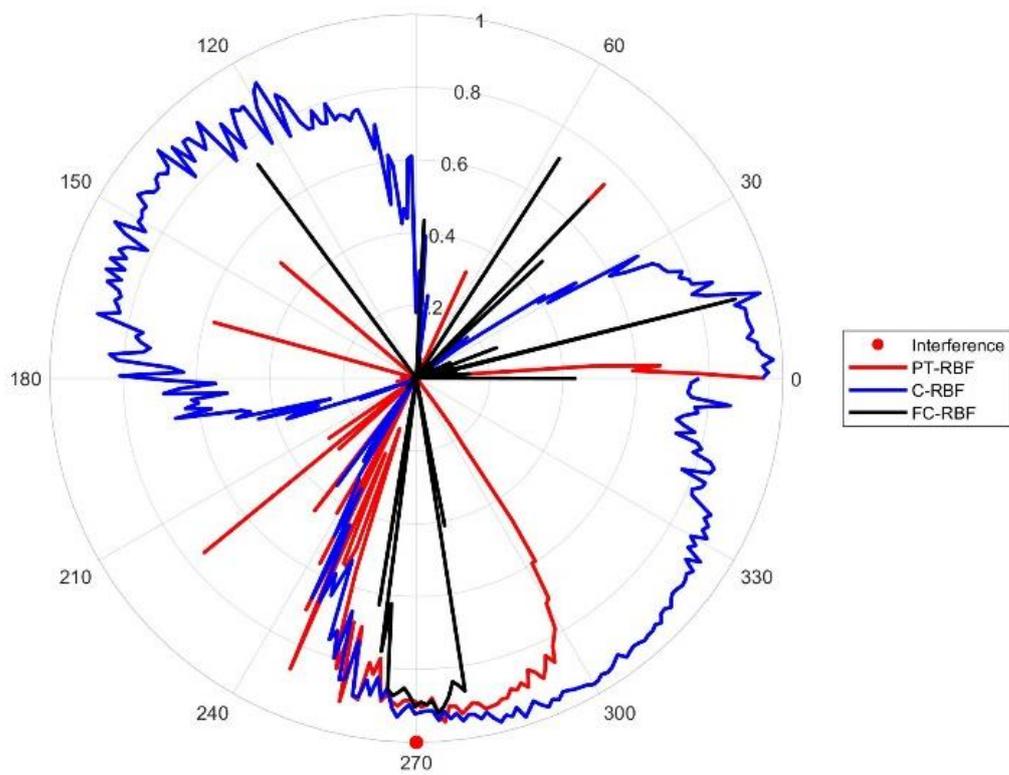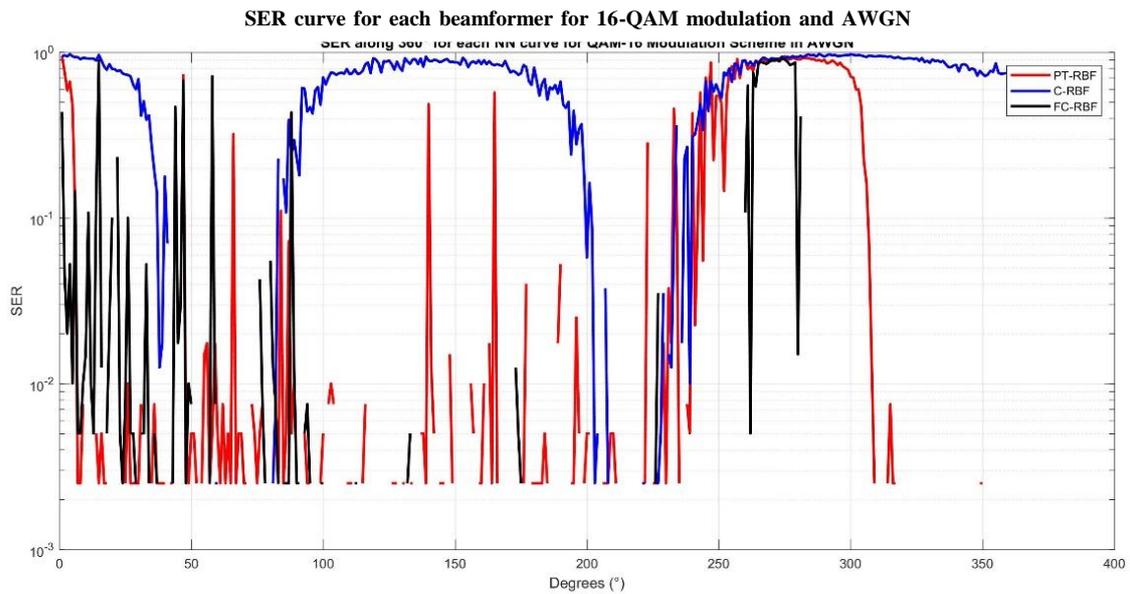
Source: (AUTHOR, 2018).

Table 13 – Summary of dynamic case 1 for SER performance relative to the varying D.S. DOA and the stationary I.S. DOA at $\phi = 270°$

DOA	SER		
	PT-RBF	**C-RBF**	**FC-RBF**
D.S. @ 20°	0	7.939698e-01	1.005025e-01
D.S. @ 40°	2.512563e-03	1.783920e-01	0
D.S. @ 60°	5.025126e-03	2.512563e-03	0
D.S. @ 80°	0	0	5.527638e-02
D.S. @ 100°	5.025126e-03	7.060302e-01	2.512563e-03
D.S. @ 120°	5.025126e-03	9.045226e-01	0
D.S. @ 140°	4.874372e-01	8.919598e-01	0
D.S. @ 160°	2.512563e-03	8.492462e-01	0
D.S. @ 180°	2.512563e-03	8.090452e-01	5.025126e-03
D.S. @ 200°	5.025126e-03	5.778894e-02	0
D.S. @ 220°	7.512563e-01	0	0
D.S. @ 240°	4.296482e-01	3.090452e-01	0
D.S. @ 260°	8.969849e-01	8.894472e-01	1.080402e-01
D.S. @ 280°	9.045226e-01	9.371859e-01	1.507538e-02
D.S. @ 300°	7.135678e-01	9.698492e-01	0
D.S. @ 320°	0	9.145729e-01	0
D.S. @ 340°	0	8.391960e-01	0
D.S. @ 360°	2.512563e-03	7.713568e-01	0

Source: (AUTHOR, 2018).

Table 14 – Summary of dynamic case 1 for gain performance relative to the varying D.S. DOA and the stationary I.S. DOA at $\phi = 270°$

DOA	Electric Field Strength [dB]		
	PT-RBF	C-RBF	FC-RBF
D.S. @ 20°	-6.877313e+00	-1.367491e+01	-8.373341e+00
D.S. @ 40°	-6.181504e+00	-1.810273e+01	-9.956139e+00
D.S. @ 60°	-6.181504e+00	-2.366600e+01	-8.959150e+00
D.S. @ 80°	-7.697863e+00	-2.826316e+01	-1.069834e+01
D.S. @ 100°	-8.067933e+00	-1.180232e+01	-1.119299e+01
D.S. @ 120°	-7.223531e+00	-8.441693e+00	-7.684838e+00
D.S. @ 140°	-6.685709e+00	-7.128674e+00	-4.246069e+00
D.S. @ 160°	-4.501440e+00	-4.271749e+00	-1.569562e+00
D.S. @ 180°	-3.986757e+00	-5.484926e+00	-1.315858e+00
D.S. @ 200°	-2.942710e+00	-5.674915e+00	-7.593159e-01
D.S. @ 220°	2.091105e+00	-5.641567e+00	-9.312603e-01
D.S. @ 240°	-2.411940e+00	-6.422948e+00	-1.083138e+00
D.S. @ 260°	-5.418204e+00	-7.727672e+00	-1.629417e+00
D.S. @ 280°	-3.389932e+00	-1.061945e+01	-2.098566e+01
D.S. @ 300°	-3.509452e+00	-4.114691e+00	-2.795604e+01
D.S. @ 320°	-4.676069e+00	-5.615469e+00	-2.391013e+01
D.S. @ 340°	-4.280646e+00	-7.248593e+00	-2.420425e+01
D.S. @ 360°	-2.915954e+00	-6.933852e+00	-3.455894e+01
Average	-4.71E+00	-1.00E+01	-1.11E+01

Source: (AUTHOR, 2018).

Figure 37 – Dynamic case 2b

SER curve for each beamformer for 16-QAM modulation and AWGN

(a)

SER curve for each beamformer for 16-QAM modulation and AWGN

(b)
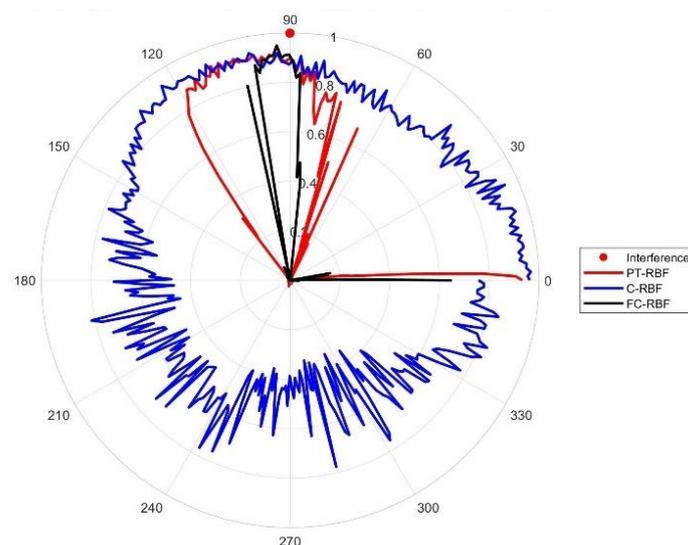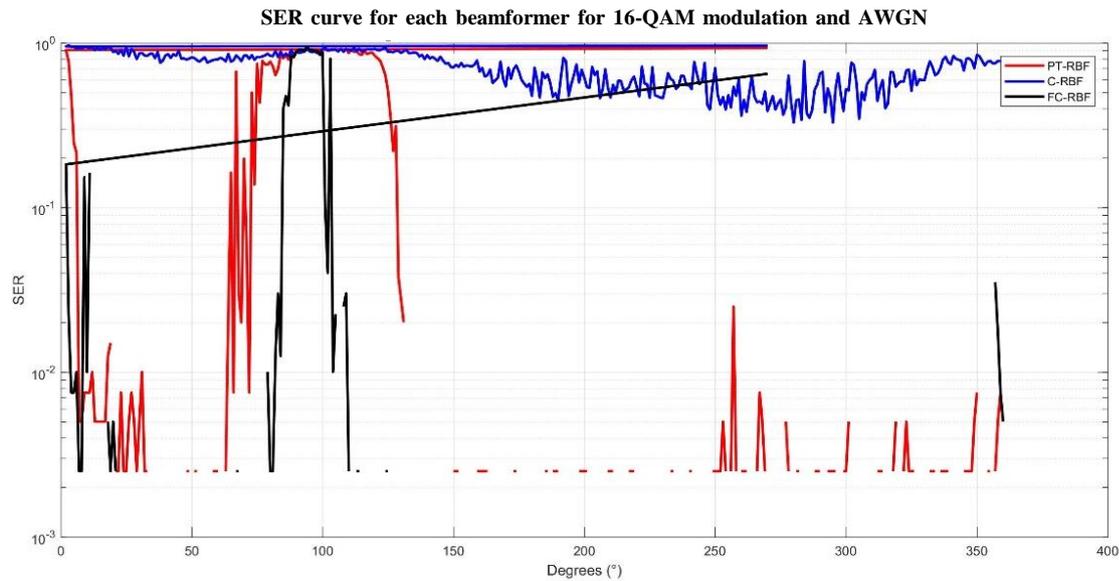
Source: (AUTHOR, 2018).

Table 15 – Summary of dynamic case 4 for SER performance relative to the stationary D.S. DOA at $\phi = 90°$ and the varying I.S. DOA

DOA	SER		
	PT-RBF	**C-RBF**	**FC-RBF**
I.S. @ 20°	2.512563e-03	7.738693e-01	0
I.S. @ 40°	7.537688e-03	9.296482e-01	1.633166e-01
I.S. @ 60°	1.005025e-02	8.718593e-01	0
I.S. @ 80°	2.512563e-03	8.115578e-01	0
I.S. @ 100°	6.030151e-02	8.417085e-01	0
I.S. @ 120°	8.944724e-01	8.718593e-01	9.120603e-01
I.S. @ 140°	8.844221e-01	9.346734e-01	0
I.S. @ 160°	0	9.120603e-01	0
I.S. @ 180°	2.512563e-03	7.160804e-01	0
I.S. @ 200°	0	6.080402e-01	0
I.S. @ 220°	0	8.140704e-01	0
I.S. @ 240°	0	4.974874e-01	0
I.S. @ 260°	0	6.557789e-01	0
I.S. @ 280°	2.512563e-03	4.221106e-01	0
I.S. @ 300°	7.537688e-03	4.773869e-01	0
I.S. @ 320°	2.512563e-03	4.447236e-01	0
I.S. @ 340°	0	7.788945e-01	0
I.S. @ 360°	0	3.793970e-01	0

Source: (AUTHOR, 2018).

Table 16 – Summary of dynamic case 4 for gain performance relative to the stationary D.S. DOA at $\phi = 90°$ and the varying I.S. DOA

DOA	Electric Field Strength [dB]		
	PT-RBF	**C-RBF**	**FC-RBF**
I.S. @ 20°	-7.725741e+00	-9.679569e+00	-1.606977e+00
I.S. @ 40°	-9.522216e+00	-8.532491e+00	-8.233090e+00
I.S. @ 60°	-9.538507e+00	-1.263406e+01	-9.998363e+00
I.S. @ 80°	-5.637271e+00	-5.594768e+00	-1.089866e+01
I.S. @ 100°	-1.221045e+01	-1.164342e+01	-1.368665e+01
I.S. @ 120°	-1.041854e+01	-1.474614e+01	-1.100067e+01
I.S. @ 140°	-1.026047e+01	-7.021735e+00	-1.379547e+01
I.S. @ 160°	-9.950636e+00	-6.438903e+00	-1.782669e+01
I.S. @ 180°	-1.561892e+01	-6.560060e+00	-1.038638e+01
I.S. @ 200°	-1.925637e+00	-2.440484e+00	-2.907011e+01
I.S. @ 220°	-8.689752e+00	-3.422606e-01	-1.806978e+01
I.S. @ 240°	-9.228414e+00	-5.056951e+00	-1.482441e+01
I.S. @ 260°	-1.196496e+01	-8.364223e+00	-1.299227e+01
I.S. @ 280°	-3.227717e+01	-7.114682e+00	-1.145319e+01
I.S. @ 300°	-1.256369e+01	-1.353599e+01	-1.148091e+01
I.S. @ 320°	-8.369845e+00	-7.864740e+00	-1.293155e+01
I.S. @ 340°	-7.055841e+00	-6.628043e+00	-1.631702e+01
I.S. @ 360°	-7.139694e+00	-6.558158e+00	-1.380562e+01
Average	-1.06e+01	-7.82e+00	-1.32e+01

Source: (AUTHOR, 2018).

4.2.3 Scenario 3

In this dynamic scenario the relation between the power of the D.S. and the power of the I.S. is given by SIR $= 10dB$. The non-linearity expressed by Eq *(2.145)* is introduced at the RF front end in order to implement a realistic simulation.

I. The operational condition in terms of D.S. and I.S. is: Figure 38: D.S. DOA varying in the range $1° < \phi < 360°$ and a stationary I.S. DOA at $\phi = 270°$.

Note in Figure 38 that the C-RBF beamformer presents SER > 0.5 along a larger angular range of the varying DOA, not being able to isolate the D.S from the I.S in this angular range. Note also that both PT-RBF and FC-RBF beamformer present a SER > 0.5 only over a smaller angular range of the varying DOA, thus performing better in this low SIR non-linear dynamic scenario. Specifically note in Figure 38 (b) that the PT-RBF has a higher residual SER in the

range $1° < \phi < 360°$. However, the FC-RBF shows more spikes of high SER outside the interference angular zone

Figure 38 – Dynamic case 3

SER curve for each beamformer for 16-QAM modulation and AWGN

(a)

**SER curve for each beamformer for 16-QAM modulation and AWGN**


(b)
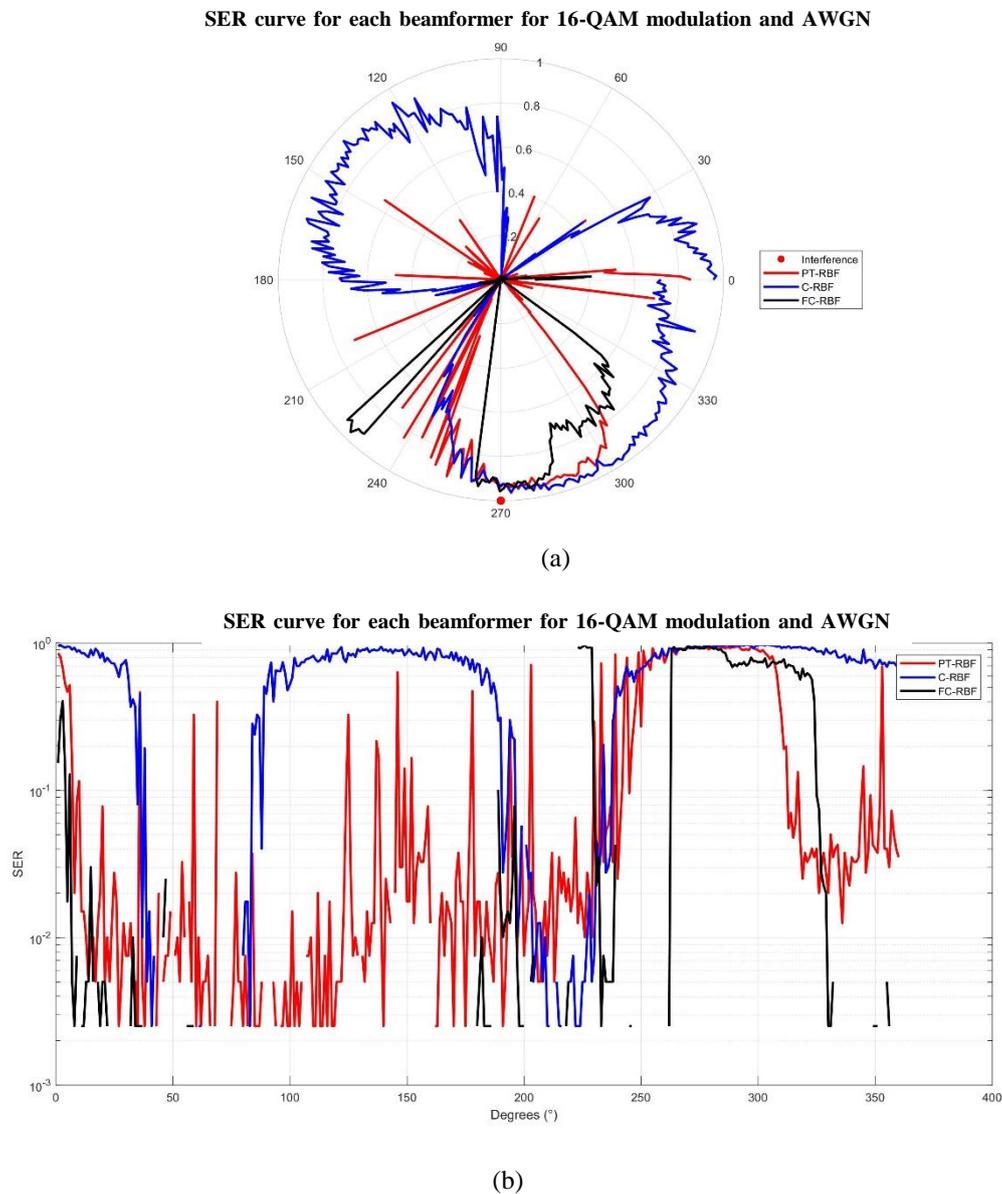
Source: (AUTHOR, 2018).

Table 17 – Summary of dynamic case 2 for SER performance relative to the varying D.S. DOA and the stationary I.S. DOA at $\phi = 270°$

DOA	SER		
	PT-RBF	**C-RBF**	**FC-RBF**
D.S. @ 20°	7.788945e-02	8.567839e-01	5.025126e-03
D.S. @ 40°	7.537688e-03	1.507538e-02	0
D.S. @ 60°	2.512563e-03	0	0
D.S. @ 80°	2.512563e-03	7.537688e-03	5.025126e-03

D.S. @ 100°	2.512563e-03	5.226131e-01	0
D.S. @ 120°	5.025126e-03	8.467337e-01	1.256281e-02
D.S. @ 140°	2.512563e-03	8.894472e-01	0
D.S. @ 160°	1.256281e-02	8.743719e-01	0
D.S. @ 180°	1.758794e-02	7.236181e-01	2.512563e-03
D.S. @ 200°	7.537688e-03	0	2.512563e-03
D.S. @ 220°	1.507538e-02	5.025126e-03	0
D.S. @ 240°	2.512563e-02	5.075377e-01	0
D.S. @ 260°	7.361809e-01	8.291457e-01	0
D.S. @ 280°	9.346734e-01	9.321608e-01	9.346734e-01
D.S. @ 300°	8.919598e-01	9.673367e-01	7.386935e-01
D.S. @ 320°	3.768844e-02	9.296482e-01	5.954774e-01
D.S. @ 340°	3.517588e-02	7.638191e-01	0
D.S. @ 360°	3.517588e-02	7.185930e-01	2.512563e-03

Source: (AUTHOR, 2018).

Table 18 – Summary of dynamic case 2 for gain performance relative to the varying D.S. DOA and the stationary I.S. DOA at $\phi = 270°$

DOA	**Electric Field Strength [dB]**		
	PT-RBF	**C-RBF**	**FC-RBF**
D.S. @ 20°	-9.395371e+00	-2.402690e+01	-1.038062e+01
D.S. @ 40°	-9.597881e+00	-2.066167e+01	-1.194914e+01
D.S. @ 60°	-8.130699e+00	-2.502196e+01	-1.579953e+01
D.S. @ 80°	-6.947941e+00	-2.232902e+01	-2.301815e+01
D.S. @ 100°	-4.908692e+00	-1.606819e+01	-4.413900e+01
D.S. @ 120°	-3.561117e+00	-8.272529e+00	-1.307676e+01
D.S. @ 140°	-2.327531e+00	-5.677650e+00	-1.247065e+01
D.S. @ 160°	-1.663094e+00	-8.719291e+00	-4.901814e+01
D.S. @ 180°	-1.848336e+00	-5.206581e+00	-1.209449e+00
D.S. @ 200°	-1.858923e+00	-5.586549e+00	-1.459911e-01
D.S. @ 220°	-2.309702e+00	-5.869049e+00	-5.699030e+00
D.S. @ 240°	-2.873681e+00	-5.867065e+00	-3.953419e+01
D.S. @ 260°	-4.071306e+00	-6.234586e+00	-3.830678e+01
D.S. @ 280°	-4.317314e+00	-1.034033e+01	-1.742199e+01
D.S. @ 300°	-2.200940e+00	-5.840143e+00	-3.659269e+01
D.S. @ 320°	-8.663505e+00	-8.967178e+00	-2.125631e+01
D.S. @ 340°	-6.121685e+00	-7.565683e+00	-3.468598e+01
D.S. @ 360°	-7.171596e+00	-7.916176e+00	-1.370737e+01
Average	-4.89E+00	-1.11E+01	-2.16E+01

Source: (AUTHOR, 2018).

4.2.4 Scenario 4

In this dynamic scenario the relation between the power of the D.S. and the power of the I.S. is given by $SIR = -10dB$. The non-linearity expressed by Eq (2.146) is introduced at the RF front end in order to implement a more realistic simulation.

The operational conditions in terms of D.S. and I.S. are:

    I.   Figure 39: D.S. DOA varying in the range $1° < \phi < 360°$ and a stationary I.S. DOA at $\phi = 270°$;

    II.   Figure 40: D.S. DOA varying in the range $1° < \phi < 360°$ and a stationary I.S. DOA at $\phi = 270°$;

    III.   Figure 41: I.S. DOA varying in the range $1° < \phi < 360°$ and a stationary D.S. DOA at $\phi = 90°$.

Note in Figures 39, 40 and 41 that the C-RBF and the FC-RBF beamformers presents $SER > 0.5$ along a larger angular range of the varying DOA, not being able to isolate the D.S from the I.S in this angular range. Note also that the PT-RBF beamformer presents a $SER > 0.5$ only over a smaller angular range of the varying DOA comparing to the other beamformers, thus performing better in this very low SIR non-linear dynamic scenario.

Figure 39 – Dynamic case 4 example

SER performance over 360° variation on ϕ on a non-linear scenario with low SIR
Main signal moving around the array and interference impinging from $\phi = 270°$

SER curve for each beamformer for 16-QAM modulation and AWGN channel

(a)
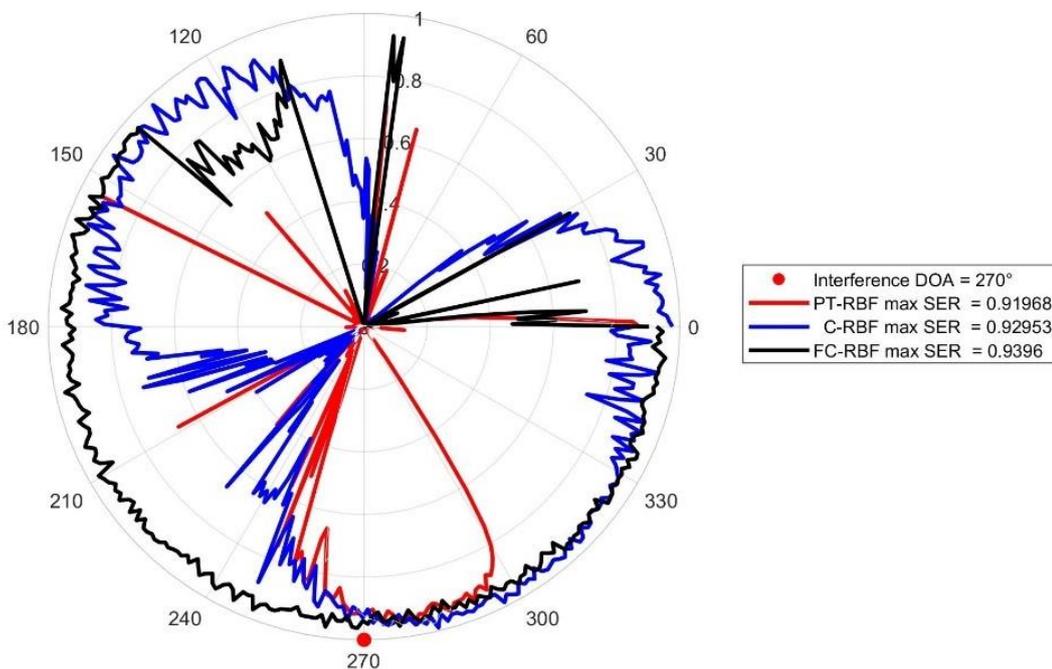
SER curve for each beamformer for 16-QAM modulation and AWGN channel

(b)
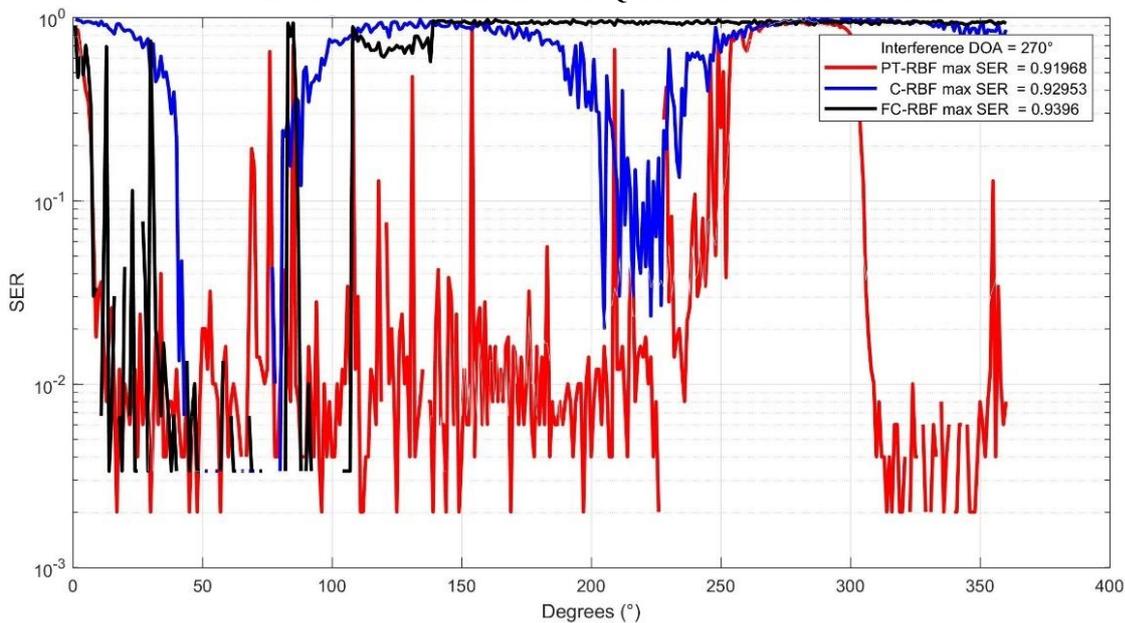
Source: (AUTHOR, 2018)

Figure 40 – Dynamic case 4: 1 interference at 270° & 1 moving desired signal

Non-linearity added on the front end.

SER curve for each beamformer for 16-QAM modulation and AWGN

(a)

SER curve for each beamformer for 16-QAM modulation and AWGN

(b)
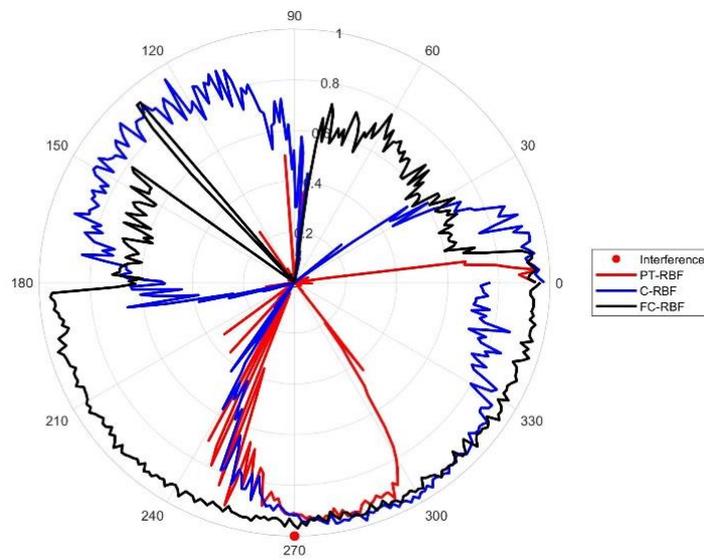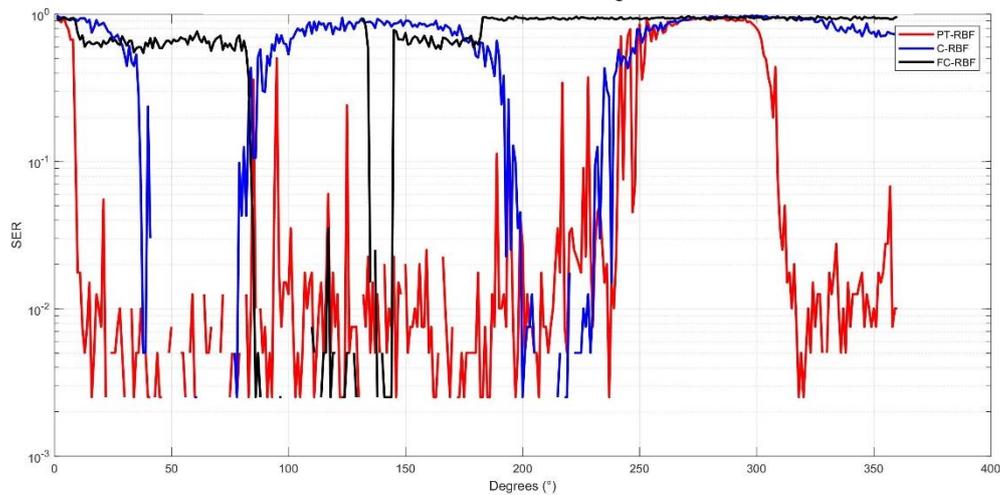
Source: (AUTHOR, 2018).

Table 19 – Summary of dynamic case 3 for SER performance relative to the varying D.S. DOA and the stationary I.S. DOA at $\phi = 270°$

DOA	SER		
	PT-RBF	**C-RBF**	**FC-RBF**
D.S. @ 20°	7.537688e-03	8.542714e-01	5.979899e-01
D.S. @ 40°	2.512563e-03	2.361809e-01	6.055276e-01
D.S. @ 60°	2.512563e-03	2.512563e-03	7.060302e-01
D.S. @ 80°	0	4.271357e-02	5.653266e-01
D.S. @ 100°	1.256281e-02	5.326633e-01	0

D.S. @ 120°	7.537688e-03	8.190955e-01	5.025126e-03
D.S. @ 140°	0	8.819095e-01	5.025126e-03
D.S. @ 160°	5.025126e-03	8.366834e-01	6.005025e-01
D.S. @ 180°	5.025126e-03	5.226131e-01	6.482412e-01
D.S. @ 200°	2.763819e-02	2.512563e-03	9.396985e-01
D.S. @ 220°	3.266332e-02	1.758794e-02	9.396985e-01
D.S. @ 240°	2.261307e-02	4.120603e-01	9.246231e-01
D.S. @ 260°	7.437186e-01	7.763819e-01	: 9.547739e-01
D.S. @ 280°	9.045226e-01	9.648241e-01	9.346734e-01
D.S. @ 300°	7.989950e-01	9.798995e-01	9.522613e-01
D.S. @ 320°	2.512563e-03	9.221106e-01	9.321608e-01
D.S. @ 340°	7.537688e-03	7.939698e-01	9.095477e-01
D.S. @ 360°	1.005025e-02	7.663317e-01	9.623116e-01

Source: (AUTHOR, 2018).

Table 20 – Summary of dynamic case 3 for gain performance relative to the varying D.S. DOA and the stationary I.S. DOA at $\phi = 270°$

DOA	**Electric Field Strength [dB]**		
	PT-RBF	**C-RBF**	**FC-RBF**
D.S. @ 20°	-9.293696e+00	-2.990623e+01	-7.288000e+00
D.S. @ 40°	-7.659862e+00	-2.577442e+01	-3.707812e+00
D.S. @ 60°	-6.694569e+00	-2.659082e+01	-4.975148e+00
D.S. @ 80°	-6.047070e+00	-2.853064e+01	-1.942162e+01
D.S. @ 100°	-4.056004e+00	-1.770678e+01	-2.155988e+01
D.S. @ 120°	-2.440175e+00	-1.166392e+01	-3.335533e+01
D.S. @ 140°	-1.876166e+00	-7.736798e+00	-3.983197e+01
D.S. @ 160°	-1.024554e+00	-4.187546e+00	-2.008144e+01
D.S. @ 180°	-1.797481e+00	-5.291249e+00	-4.416581e+01
D.S. @ 200°	-3.789623e+00	-5.579686e+00	-1.068180e+00
D.S. @ 220°	-5.132653e+00	-6.074186e+00	-1.095994e+00
D.S. @ 240°	-5.494058e+00	-5.777771e+00	-1.149550e+00
D.S. @ 260°	-7.654779e+00	-5.709359e+00	-1.277008e+00
D.S. @ 280°	-6.656713e+00	-8.825583e+00	-1.240542e+00
D.S. @ 300°	-1.127285e+01	-1.255308e+01	-1.260519e+00
D.S. @ 320°	-7.002743e+00	-7.293321e+00	-1.296944e+00
D.S. @ 340°	-7.859370e+00	-7.191911e+00	-1.264195e+00
D.S. @ 360°	-8.758862e+00	-9.254261e+00	-1.319732e+00
Average	-5.81E+00	-1.25E+01	-1.14E+01

Source: (AUTHOR, 2018).

Figure 41 – Dynamic case 4: 1 moving interferences & 1 desired signal at 90°
Non-linearity added on the front end.

SER curve for each beamformer for 16-QAM modulation and AWGN

(a)

SER curve for each beamformer for 16-QAM modulation and AWGN

(b)
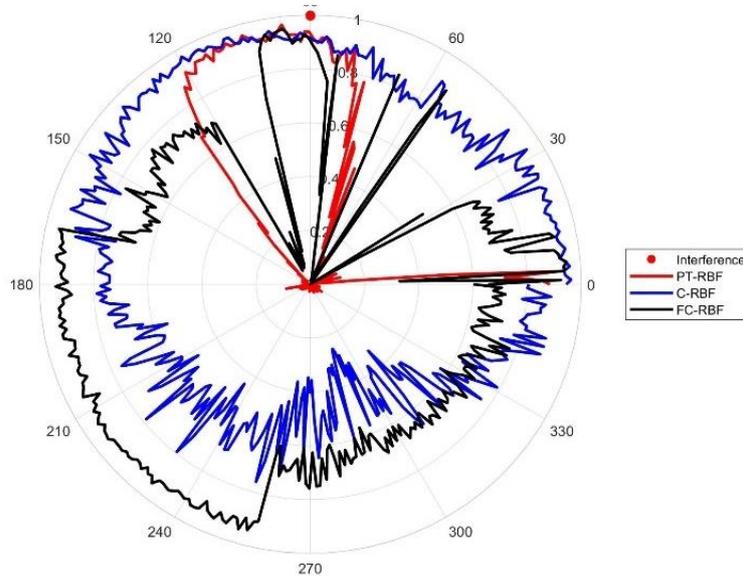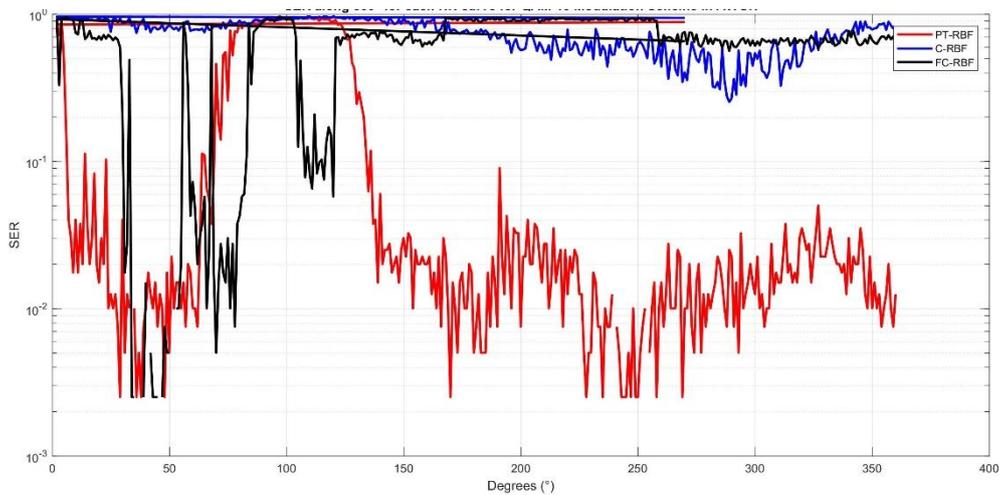
Source: (AUTHOR, 2018).

Table 21 – Summary of dynamic case 5 for SER performance relative to the stationary D.S. DOA at $\phi = 90°$ and the varying I.S. DOA

DOA	SER		
	PT-RBF	**C-RBF**	**FC-RBF**
I.S. @ 20°	1.758794e-02	-1.514478e+01	7.286432e-01
I.S. @ 40°	5.025126e-03	-7.360742e+00	1.758794e-02
I.S. @ 60°	2.261307e-02	-7.155615e+00	0
I.S. @ 80°	1.859296e-01	-9.358594e+00	1.507538e-02
I.S. @ 100°	9.371859e-01	-1.185892e+01	9.120603e-01

I.S. @ 120°	9.271357e-01	-4.426984e+00	2.085427e-01
I.S. @ 140°	2.437186e-01	-6.447938e+00	7.336683e-01
I.S. @ 160°	3.266332e-02	-3.298882e+00	6.934673e-01
I.S. @ 180°	1.005025e-02	-5.567844e+00	9.346734e-01
I.S. @ 200°	2.261307e-02	-8.514257e+00	9.321608e-01
I.S. @ 220°	1.005025e-02	-1.783105e+01	9.120603e-01
I.S. @ 240°	1.507538e-02	-1.651852e+01	9.422111e-01
I.S. @ 260°	7.537688e-03	6.180905e-01	9.070352e-01
I.S. @ 280°	2.010050e-02	5.100503e-01	7.512563e-01
I.S. @ 300°	1.507538e-02	2.839196e-01	5.979899e-01
I.S. @ 320°	2.261307e-02	4.246231e-01	6.582915e-01
I.S. @ 340°	3.517588e-02	6.934673e-01	6.482412e-01
I.S. @ 360°	1.005025e-02	8.090452e-01	6.281407e-01

Source: (AUTHOR, 2018).

Table 22 – Summary of dynamic case 5 for gain performance relative to the stationary D.S. DOA at $\phi = 90°$ and the varying I.S. DOA

DOA	**Electric Field Strength [dB]**		
	PT-RBF	**C-RBF**	**FC-RBF**
I.S. @ 20°	-1.169052e+01	-1.246813e+01	-4.026603e+00
I.S. @ 40°	-1.080302e+01	-1.344394e+01	-7.768834e+00
I.S. @ 60°	-7.107378e+00	-9.210507e+00	-1.194063e+01
I.S. @ 80°	-4.887604e+00	-6.816877e+00	-2.043980e+01
I.S. @ 100°	-1.864843e+01	9.271357e-01	-8.027247e+00
I.S. @ 120°	-1.403257e+01	8.743719e-01	-7.304055e+00
I.S. @ 140°	-1.705383e+01	8.140704e-01	-7.949298e+00
I.S. @ 160°	-1.979587e+01	8.316583e-01	-7.594112e+00
I.S. @ 180°	-1.528136e+01	9.170854e-01	-1.459609e+01
I.S. @ 200°	-7.300080e+00	9.422111e-01	-3.253233e+01
I.S. @ 220°	-6.216366e+00	9.422111e-01	-2.343960e+01
I.S. @ 240°	-7.872151e+00	8.969849e-01	-2.393313e+01
I.S. @ 260°	-1.118034e+01	7.814070e-01	-2.343762e+01
I.S. @ 280°	-2.136617e+01	7.361809e-01	-2.162928e+01
I.S. @ 300°	-1.580750e+01	6.155779e-01	-2.122316e+01
I.S. @ 320°	-1.177968e+01	7.939698e-01	-1.107852e+01
I.S. @ 340°	-9.390392e+00	-8.929682e+00	-3.583003e+00
I.S. @ 360°	-9.509514e+00	-8.267338e+00	-4.292567e-01
Average	-1.22e+01	-2.73e+00	-1.39e+01

Source: (AUTHOR, 2018).

4.2.5 Summary of Dynamic Scenarios

As expected, on dynamic operation, the beamforming algorithms performed similarly to the static system. The PT-RBF based beamforming performed better under non-linear very low SIR scenarios, yielding a small residual SER and precisely pointing the boresight lobe of the UCA to the D.S. DOA for both situations of varying D.S. DOA and varying I.S. (jammer signal) DOA.

On linear operational scenario – which is a rare operational condition of nowadays crowded spectrum, full of interfering or even jamming signals – the LMS based beamforming had a better overall performance as for the boresight lobe pointing, as for the I.S. isolation and as for the SER. As the LMS has a faster convergence rate, it is more suitable than all other beamformers for dynamic linear operational conditions, regardless the SIR level. On the other hand, the LMS performance significantly decreases on non-linear scenarios. As consequence, the isolation between desired signal and interferences is low and the algorithm finds some trouble to follow the dynamic operation and presents a delayed behavior on adapting the radiation pattern. Figure 34 illustrates the SER performance under linear operation for all beamformers with the DOA varying in the range $1° < \phi < 360°$. It is possible to see the LMS's better performance for the referred linear scenario.

The C-RBF based beamforming, which performed well on low power interference static scenarios, proved not suitable to perform over dynamic operations. This is because the C-RBF is the ANN with slowest convergence, which needs a much higher number of training samples N to reach a fair MSEA threshold.

For the FC-RBF based beamforming, it has been seen that, over a moderate SIR scenario, it presents the best overall performance. As it has the fastest convergence rate among the RBFs, the FC-RBF can follow the moving signal and reaching a satisfactory MSEA threshold rapidly enough so the isolation between desired signal and interferences are high enough to allow the QAM de-mapper to correctly identify all transmitted IQ symbols. However, once again, over a very low SIR scenario, the FC-RBF is still able to null interferences and to point the UCA boresight lobe to the interest point, but it is unable to recover the transmitted IQ symbols. Figure 35 and Figure 39 illustrate the SER performance between the three non-linear Radial Basis beamformers for a non-linear moderate SIR and very low SIR environment respectively.

In conclusion, just like for static scenarios, the PT-RBF based beamformer was the only algorithm that presented satisfactory performance in all moving scenarios, keeping uniform

performance, even if overcome by some other solution in specific scenarios. However, none of the other solutions presented satisfactory performance in all evaluated scenarios.

5 FINAL REMARKS

5.1 CONCLUSIONS

This thesis proposed a novel beamforming technique based on the PT-RBF ANN which, different from most ANN found in literature, has a characteristic of avoiding the phase invariance issue. It started by a small gastropod performing every single process lately explored in the state of the art section. Hopefully, this analogy served the purpose of instigating the reader to go through all sections of this work and realize that all cutting-edge technology on telecommunications are nothing more than an attempt to mimic the communication system of the simplest life forms around us.

A beamformer aims the beam to the desired signal direction and mitigates impairments from interfering signal sources, which may adversely affect the communication system performance. The main functional features of the PT-RBF based beamforming method can be summarized as follows: (1) it increases coverage range, (2) it increases throughput at network nodes, (3) it reduces transmission power, (4) it allows multipath links, (5) it allows multiple concurrent transmissions using the same frequency, (6) it allows an efficient use of the spectrum resource.

A six half-wavelength ($\lambda/2$) dipole UCA has been used to evaluate the proposed PT-RBF beamforming. It provides a good control of the boresight lobe according to the spherical coordinate ϕ, making it possible to aim the radiation pattern boresight of the array to any direction of azimuth ϕ. The number of dipoles was chosen in order to improve the overall system performance without drastically increasing the computational complexity. The adopted frequency $f = 850MHz$ lies near on center of the UHF band, where a great number of wireless systems operate.

The beamformer architecture in Section 3 and Appendix A was built in order to assess the PT-RBF ANN and to train it over a realistic scenario.

The performance of the PT-RBF based beamforming has been compared to the state-of-the-art solutions over several static and dynamic scenarios through simulations using 16QAM modulation and comparing the gain performance, the symbol error rate and the ANN convergence rate. The algorithms were tested under linear and non-linear scenarios, with moderate (35dB), low (10dB) and very low (-10dB) SIR levels for both static and dynamic operating conditions, with SIR= -10dB representing a tactical jamming scenario.

For static non-linear scenarios, although the C-RBF presented the slowest convergence rate, it had a better performance on moderate SIR cases. For dynamic non-linear scenarios, the

C-RBF proved to be useless and the FC-RBF, which has the fastest convergence rate between the non-linear networks, had the best performance on moderate SIR cases.

Nevertheless, for both static and dynamic systems, operating on a very low SIR environment, the FC-RBF cannot reach a satisfactory MSEA threshold. The proposed PT-RBF beamformer has presented significant results when compared with the state-of-the-art solutions, making possible to operate communication links under static scenarios on self-organizing networks and in dynamic scenarios with access in motion, both with multiple interferences, thus maximizing the throughput and the spectrum efficiency.

The proposed approach is robust over static and dynamic low SIR scenarios, in the sense that the beamforming architecture presents a higher gain between desired and interfering signals and can demodulate with a lower SER, given the state-of-the-art architectures. Nevertheless, the proposed algorithm performed reasonably well for all other simpler scenarios. Finally, it is important to note that sometimes the proposed beamforming doesn't create a high gain between boresight and null, but always guarantee a zero or a low SER.

In conclusion, the novel non-blind PT-RBF beamformer proposed in this thesis achieved both a lower SER and a higher isolation between desired and interfering signals with a fast convergence rate for the most critical scenarios, including tactical jamming scenarios with SIR= -10dB. Also, different from the state-of-the-art beamformers, the PT-RBF based beamformer was the only algorithm that presented satisfactory performance in all scenarios, keeping uniform performance, even if overcome by some other solution in specific scenarios. These results validate the effectiveness of the proposed approach architecture for static and dynamic beamforming.

5.2 FUTURE WORK

For those who felt challenged by this work, there are several ideas that might contribute for a better performance of a PT-RBF beamforming. Some additional techniques found in literature like a mathematical approach for an optimum learning rate μ definition proposed by (KIM, 2010) could improve the presented performance.

Also, a fuzzy controller proposed by (CHEN et al., 2010) or a neuro-fuzzy system proposed by (ABIYEV; AL-SHANABLEH, 2006) could be implemented in order to control the PT-RBF beamforming leaning rates (μ_w, μ_t and μ_σ). The current implementation needs the human intervention and decision to set these parameters and does not allow them to be changed along the training process. A controlled leaning rate could therefore improve even more the presented performance.

Finally, a metacognitive learning framework, proposed by (SAVITHA; SURESH; SUNDARARAJAN, 2012), which could regulate the PT-RBF learning process might also enhance its performance. In every epoch, when a sample is presented, the metacognitive component decides what to learn, when to learn, and how to learn based on the knowledge acquired by the ANN and the new information contained in the sample.

REFERENCES

ABIYEV, R.; AL-SHANABLEH, T. **Neuro-Fuzzy Network for Adaptive Channel Equalization**. 2006 Fifth Mexican International Conference on Artificial Intelligence. **Anais**...IEEE, nov. 2006Disponível em: <http://ieeexplore.ieee.org/document/4022157/>

ALPAY, D. Cauchy–Riemann Equations and $\mathbb{C}$-differentiable Functions. In: **A Complex Analysis Problem Book**. 1. ed. Cham: Springer International Publishing, 2016. p. 151–207.

AWARE-, S.; TRACKING, B. F. **Digitally Aided Close Air Support**Black Diamond Advanced Technology, , 2016. Disponível em: <https://www.bdatech.com/wp-content/uploads/2015/08/Whitepaper_DACAS.pdf>

BALANIS, C. A. **Antenna Theory: Analysis and Design**. 4. ed. [s.l.] John Wiley & Sons, Inc., 2016.

BISHOP, C. **Pattern Recognition and Machine Learning**. 1. ed. [s.l.] Springer-Verlag New York, 2006.

BLACK, B. E. J. Holographic Beam Forming and MIMO. **Commware, Pivotal**, 2016.

BLAUNSTEIN, N.; CHRISTODOULOU, C. G. Adaptive Antennas for Wireless Networks. In: **Radio Propagation and Adaptive Antennas for Wireless Communication Networks**. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2014. p. 216–279.

BRUCE R. PIRNIE et al. **Beyond Close Air Support**. [s.l: s.n.].

CASTRO, F. DE; CASTRO, M. C. DE. **Redes Neurais Artificiais**. [s.l: s.n.].

CHEN, S. et al. Fully complex-valued radial basis function networks: Orthogonal least squares regression and classification. **Neurocomputing**, v. 71, n. 16–18, p. 3421–3433, 2008.

CHEN, S. et al. Fuzzy-logic tuned constant modulus algorithm and soft decision-directed scheme for blind equalisation. **Digital Signal Processing**, v. 20, n. 3, p. 846–859, maio 2010.

COVER, T. M. Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition. **IEEE Transactions on Electronic Computers**, v. EC-14, n. 3, p. 326–334, jun. 1965.

DARPA. **From Close Air Support to Fire Suppression**. Disponível em: <https://www.darpa.mil/news-events/2014-06-09>. Acesso em: 10 maio. 2018.

DE CASTRO, F. DE; FRANCO, P. R. . **Antenas**. Porto Alegre: [s.n.].

DIAZ-ROZO, J.; BIELZA, C.; LARRANAGA, P. Clustering of Data Streams with Dynamic Gaussian Mixture Models. An IoT Application in Industrial Processes. **IEEE Internet of Things Journal**, p. 1–1, 2018.

DU, K.-L. et al. Neural methods for antenna array signal processing: a review. **Signal Processing**, v. 82, n. 4, p. 547–561, 2002.

FARUQUE, S. **Radio Frequency Modulation Made Easy**. 1. ed. [s.l.] Springer, 2017.

FORTUNER, B. **Can neural networks solve any problem?**

GODARA, L. C. Application of antenna arrays to mobile communications, part II: Beam-forming and direction-of-arrival considerations. **Adaptive Antennas for Wireless Communications**, v. 85, n. 8, p. 95–145, 2009.

GUTIÉRREZ, A.; MARCO, S. **Biologically Inspired Signal Processing for Chemical Sensing**. 1. ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. v. 188

HAGHIGHI, G.; HOSSEIN, M. Beamforming in Wireless Networks. In: **Towards 5G Wireless Networks - A Physical Layer Perspective**. [s.l: s.n.]. p. 163–192.

HANZO, L.; WEBB, W. T.; KELLER, T. **Single– and Multi–carrier Quadrature Amplitude Modulation: Principles and Applications for Personal Communications, WLANs and Broadcasting**. 2. ed. [s.l.] Wiley-Blackwell, 2000.

HAYKIN, S. **Neural Networks and Machine Learning**. 3. ed. Hamilton, Ontario, Canada: Pearson Education, Inc., 2009.

HAYKIN, S. O. **Adaptive Filter Theory**. 5. ed. [s.l.] Pearson, 2013.

IOANNIDES, P.; BALANIS, C. A. Uniform circular arrays for smart antennas. **IEEE Antennas and Propagation Magazine**, v. 47, n. 4, p. 192–206, 2005.

KANUNGO, T. et al. An efficient k-means clustering algorithm: analysis and implementation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 24, n. 7, p. 881–892, jul. 2002.

KIM, N. Step-size control for width adaptation in radial basis function networks for nonlinear channel equalization. **Journal of Communications and Networks**, v. 12, n. 6, p. 600–604, dez. 2010.

KOPP, C. Close Air Support in COIN operations. **Defence Today**, p. 30–32, 2010.

KUN-CHOU LEE; TAH-HSIUNG CHU. **A circuit model for antenna array mutual coupling effects**. IEEE Antennas and Propagation Society International Symposium. 1995 Digest. **Anais**...IEEE, 1995

LI, X. et al. Obtaining polynomial coefficients from intercept points of RF power amplifiers. v. 48, n. 19, p. 1–2, 2012.

LI, Y. et al. Clustering Analysis in the Wireless Propagation Channel with a variational Gaussian Mixture Model. **IEEE Transactions on Big Data**, p. 1–1, 2018.

LOSS, D. V. et al. Phase transmittance RBF neural networks. **Electronics Letters**, v. 43, n. 16, p. 882, 2007.

MATHWORKS. **MATLAB The Language of Technical ComputingComponents**, 2016.

MHASKAR, H. N. Neural Networks for Optimal Approximation of Smooth and Analytic Functions. **Neural Computation**, v. 8, n. 1, p. 164–177, jan. 1996.

NIELSEN, M. A. **Neural Networks and Deep Learning**. Disponível em: <neuralnetworksanddeeplearning.com/>.

NIYOGI, P.; GIROSI, F. On the Relationship between Generalization Error, Hypothesis Complexity, and Sample Complexity for Radial Basis Functions. **Neural Computation**, v. 8, n. 4, p. 819–842, 15 maio 1996.

POCKLINGTON, H. C. Electrical Oscillations in Wire. **CAMBRIDGE PHILOSOPHICAL SOCIETY - PROCEEDINGS**, 1897.

POZAR, D. M. **Microwave Engineering**. 4. ed. [s.l: s.n.].

PROAKIS, J. G.; SALEHI, M. **Communication Systems Engineering**. 2. ed. [s.l: s.n.].

SADIKU, M. **Elements of Electromagnetics**. 7. ed. [s.l.] The Oxford Series in Electrical and Computer Engineering, 2018.

SAVITHA, R.; SURESH, S.; SUNDARARAJAN, N. A FULLY COMPLEX-VALUED RADIAL BASIS FUNCTION NETWORK AND ITS LEARNING ALGORITHM. **International Journal of Neural Systems**, v. 19, n. 04, p. 253–267, ago. 2009.

SAVITHA, R.; SURESH, S.; SUNDARARAJAN, N. Metacognitive Learning in a Fully Complex-Valued Radial Basis Function Neural Network. **Neural Computation**, v. 24, n. 5, p. 1297–1328, maio 2012.

SCHELKUNOFF, S. A. Conversion of Maxwell's Equations into Generalized Telegraphist's Equations*. **Bell System Technical Journal**, v. 34, n. 5, p. 995–1043, set. 1955.

SOCIETY, I. A. AND P. **IEEE Standard for Definitions of Terms for Antennas**. [s.l: s.n.]. v. 2013

SOLOMON, E. P. **Introduction to Human Anatomy and Physiology**. 4. ed. [s.l: s.n.].

STUTZMAN, W.; GARY, T. **Antenna Theory and Design**. 3. ed. [s.l.] John Wiley & Sons, Inc., 2013.

SURESH, S.; SUNDARARAJAN, N.; SAVITHA, R. **Supervised Learning with**

Complex-valued Neural Networks. 1. ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. v. 421

SVANTESSON, T. **Direction Finding In The Presence Of Mutual Coupling**. [s.l.] chalmers university of technology, 1999.

SVANTESSON, T. **Antennas and Propagation from a Signal Processing Perspective**. [s.l.] chalmers university of technology, 2001.

VAN TREES, H. L. Optimum Array Processing. In: 1. ed. New York, USA: John Wiley & Sons, Inc., 2002.

WIDROW, B.; HOFF, M. E. Adaptive Switching Circuits. **IRE WESCON Convention Record**, n. 4, p. 96–104, 1960.

WINKLER, R. Elon Musk Launches Neuralink to Connect Brains With Computers. **The Wall Street Journal**, 2017.

WU, S. et al. Frequency and quadrature amplitude modulation for 5G networks. **EUCNC 2016 - European Conference on Networks and Communications**, p. 1–5, 2016.

YANG, W.; ZHANG, D.; FU, Y. **Research of a Diagonal Recurrent Neural Network and Artificial Neural Networks**. 2016 International Symposium on Computer, Consumer and Control (IS3C). **Anais**...IEEE, jul. 2016

YUANJIAN, Z.; XIAOHUI, Y. A Novel Adaptive Beamforming Algorithm for Smart Antenna System. **2016 12th International Conference on Computational Intelligence and Security (CIS)**, p. 522–525, 2016.

ZOOGHBY, A. H. E.; CHRISTODOULOU, C. G.; GEORGIOPOULOS, M. Neural network-based adaptive beamforming for one- and two-dimensional antenna arrays. **IEEE Transactions on Antennas and Propagation**, v. 46, n. 12, p. 1891–1893, 1998.

APPENDIX A – BEAMFORMING ALGORITHM FULL FLOW CHART

Figure 42 – Full flow chart algorithm for static scenarios
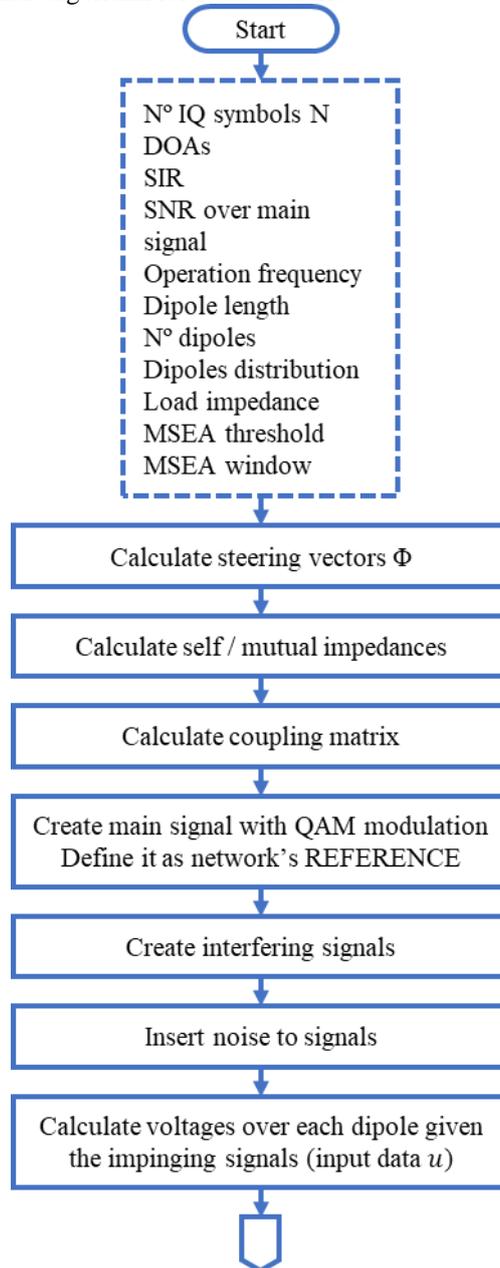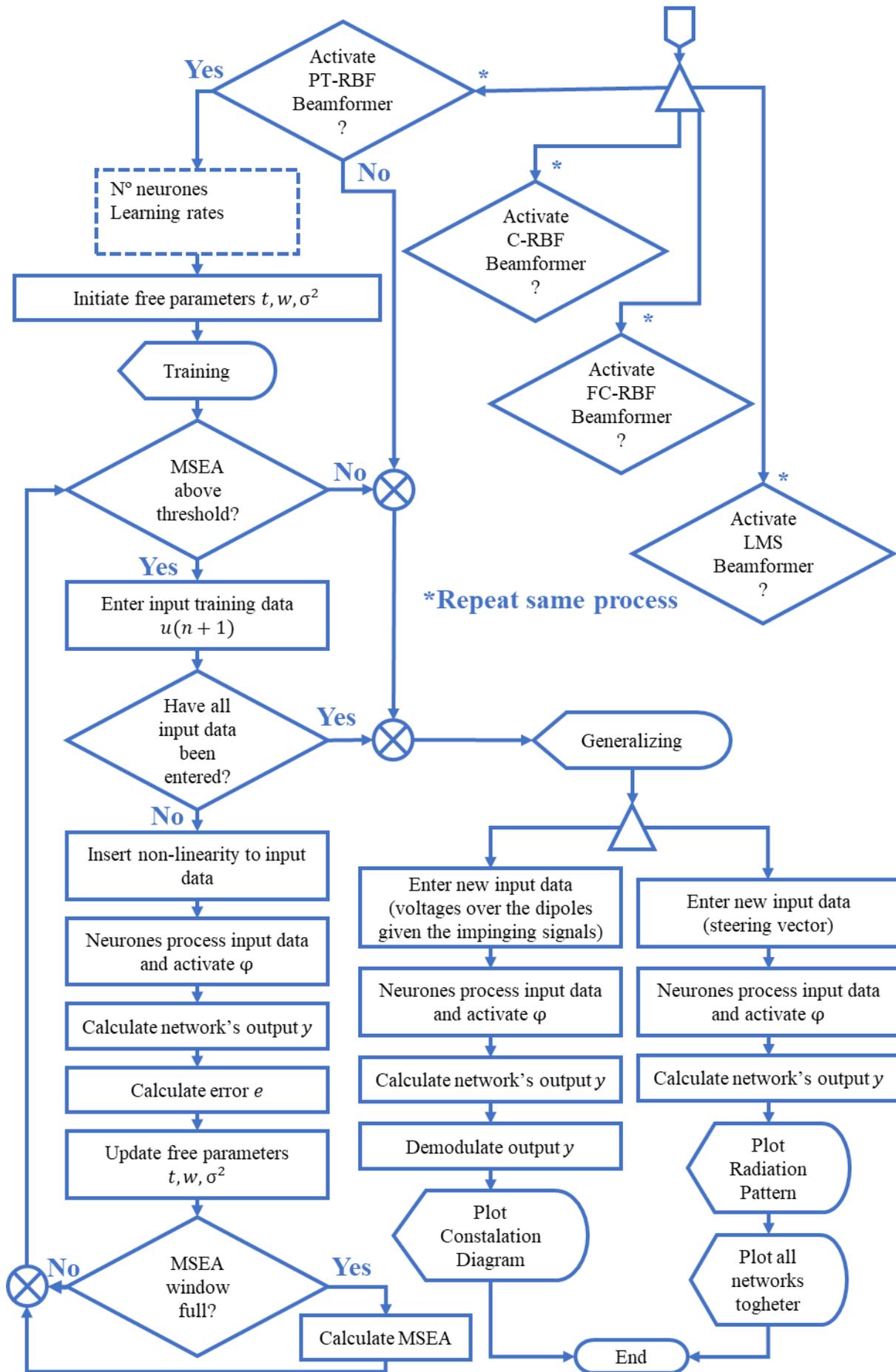
Figure 43 – Full flow chart algorithm for dynamic scenarios
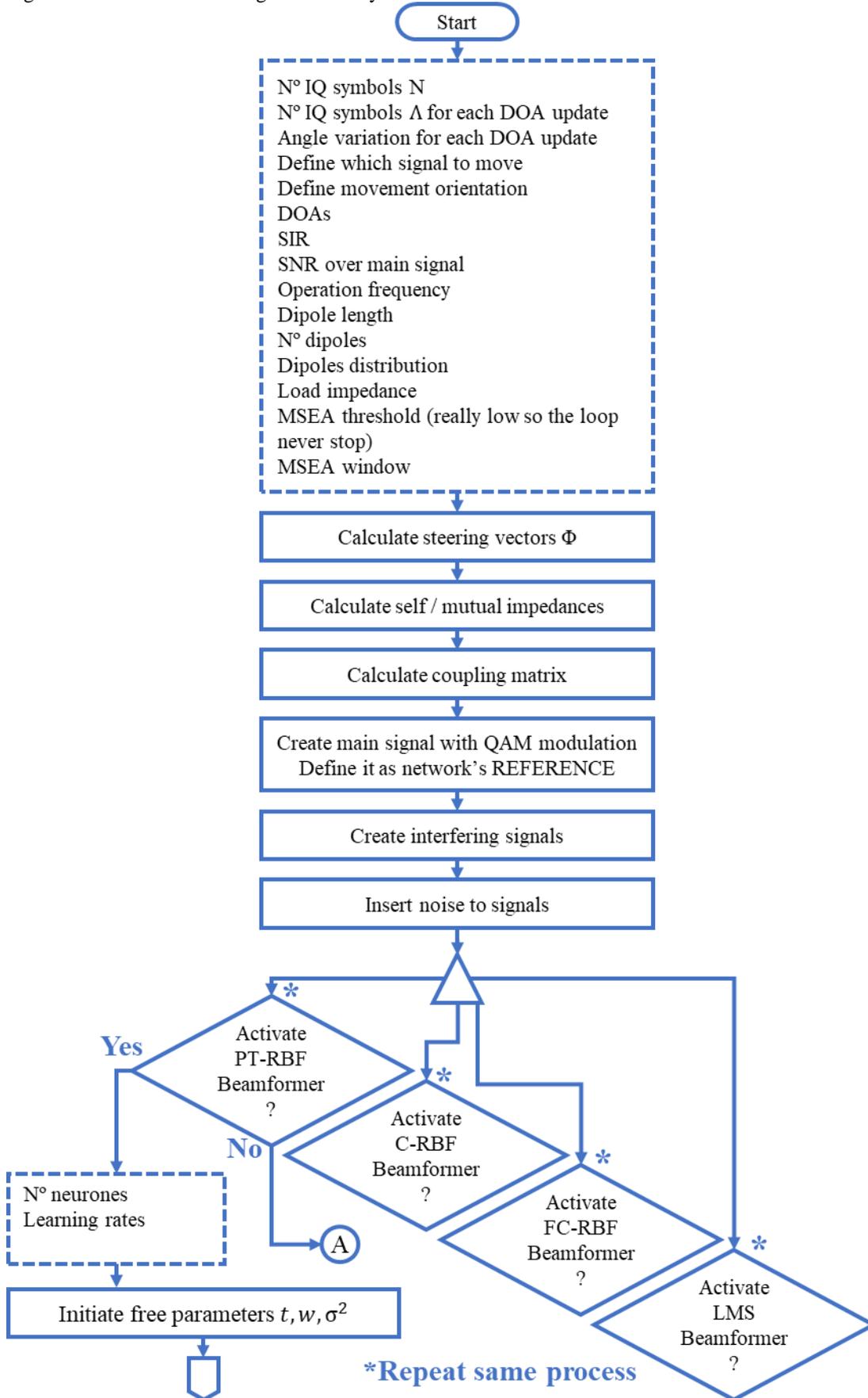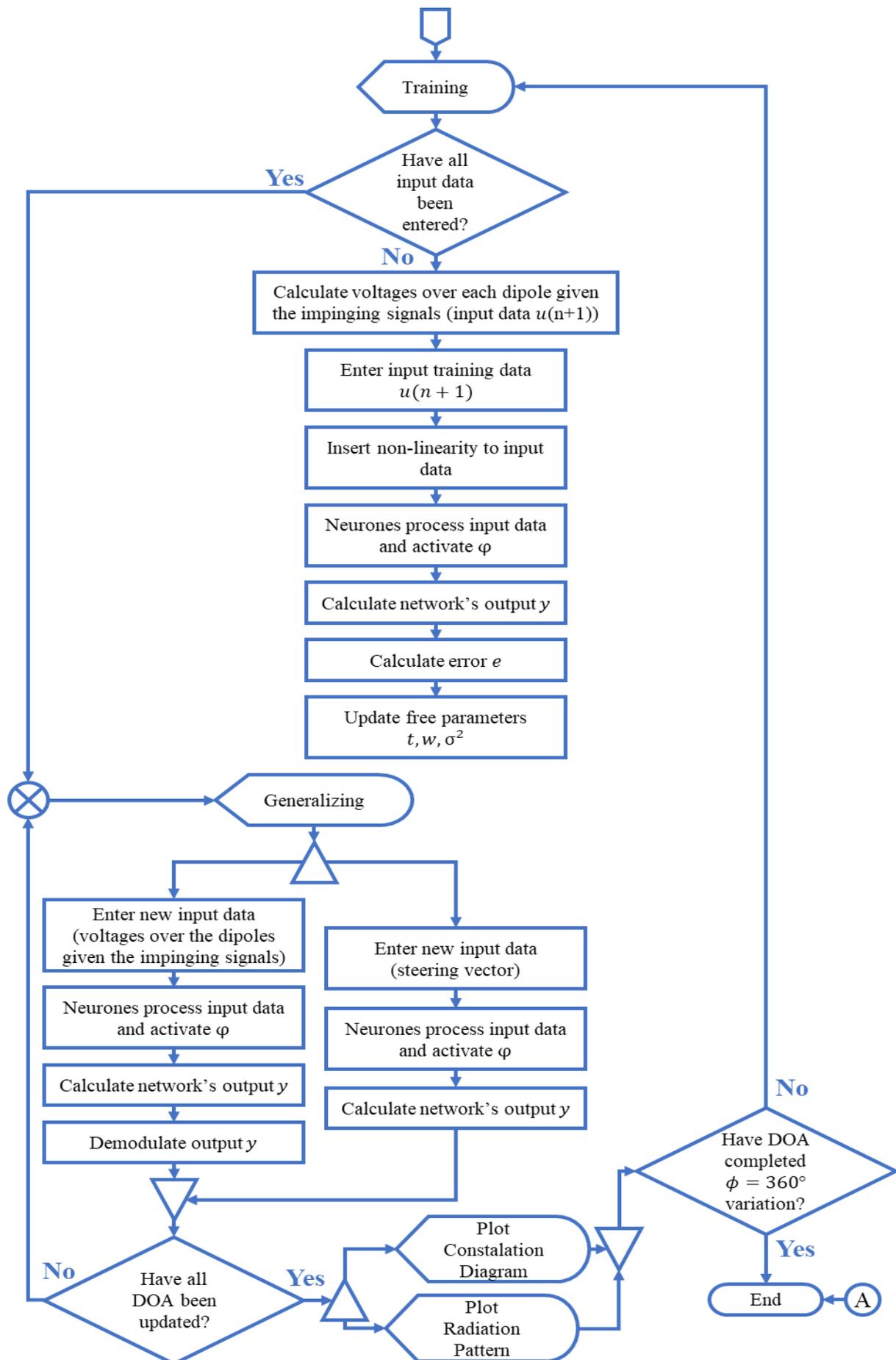
APPENDIX B – FURTHER SIMULATIONS RESULTS

Figure 44 – Static scenario, D.S. DOA @ $\phi = 1°$, I.S. DOA @ $\phi = 300°$, no NL, no noise on D.S, SIR $= 10dB$


(a)

(b)

(c)

(d)

(e)
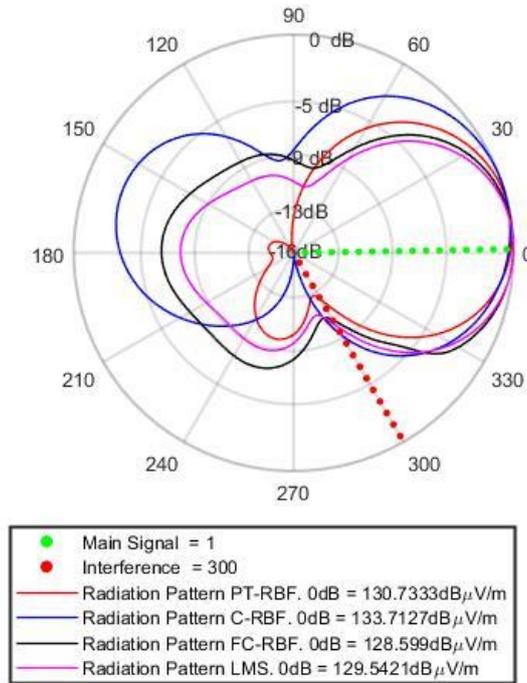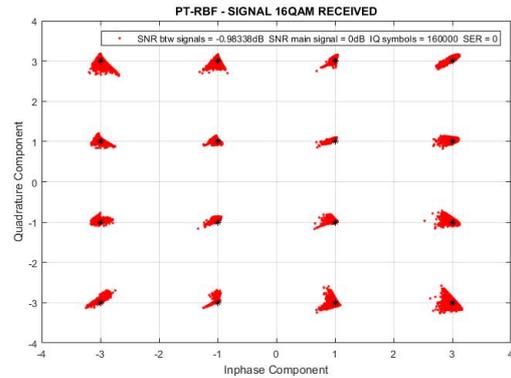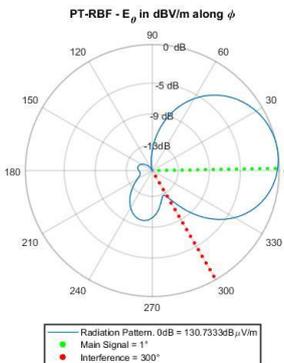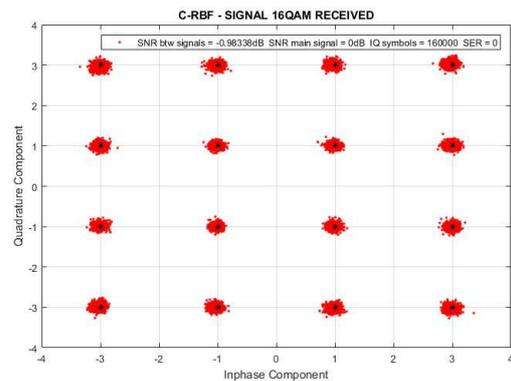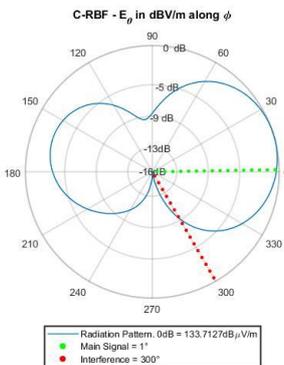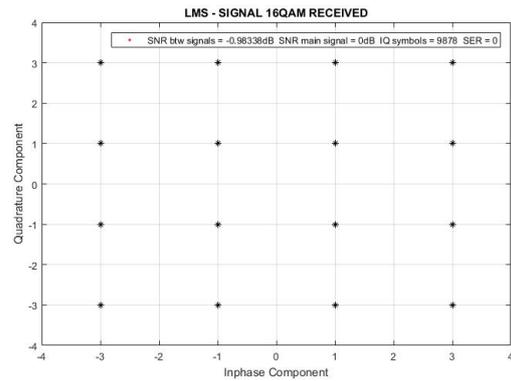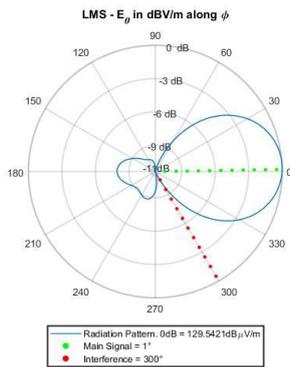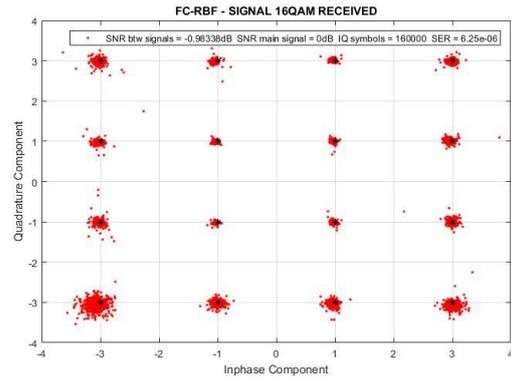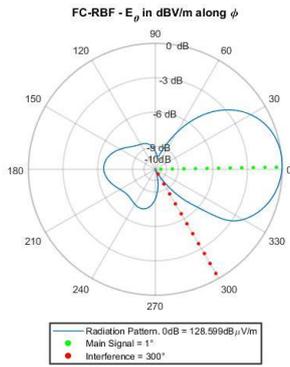
Source: (AUTHOR, 2018).

Figure 45 – Static scenario, D.S. DOA @ $\phi = 1°$, I.S. DOA @ $\phi = 30°$, no NL, no noise on D.S, SIR $= 10dB$

(a)

(b)

(c)

(d)

(e)
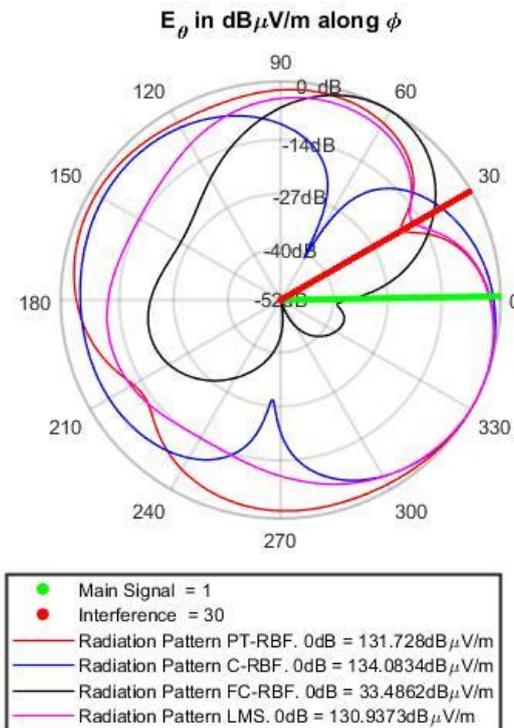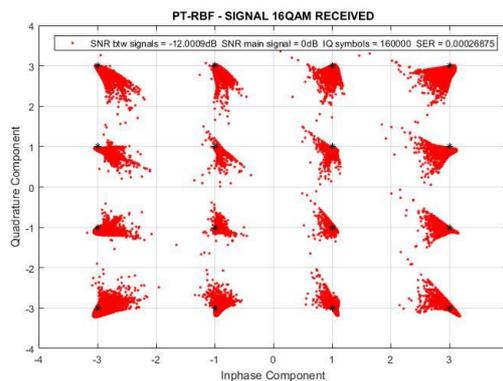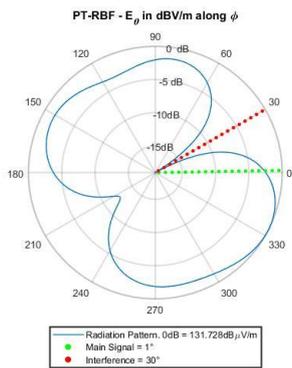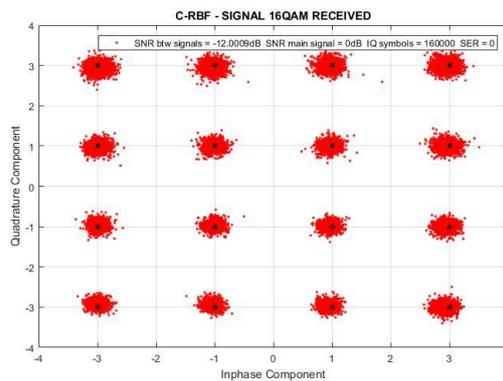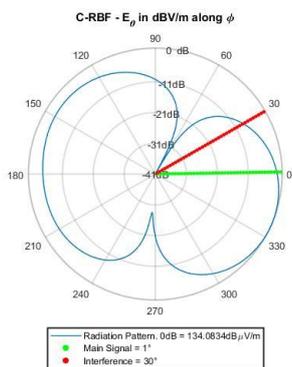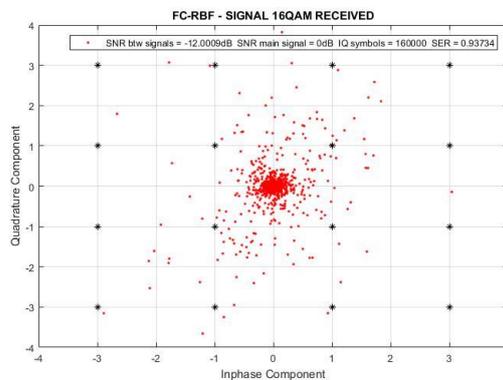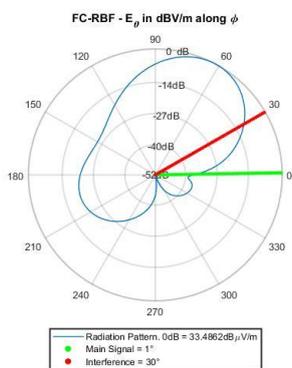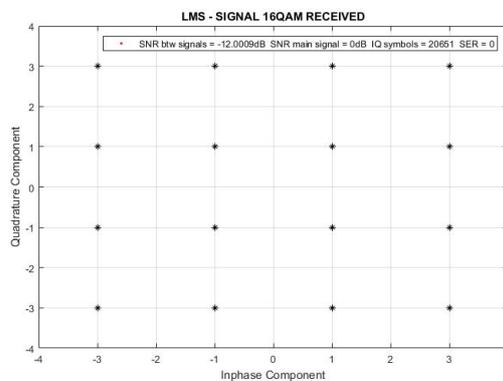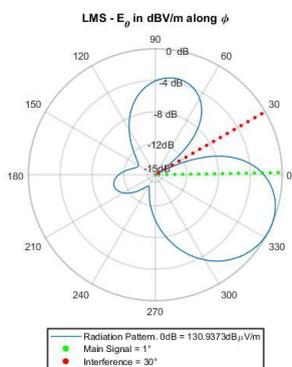
Source: (AUTHOR, 2018).

Figure 46 – Static scenario, D.S. DOA @ $\phi = 1°$, I.S. DOA @ $\phi = 300°$, with NL, noise at D.S $35dB$, SIR $= 35dB$

E_θ in dBμV/m along ϕ

Main Signal = 1
Interference = 300
Radiation Pattern PT-RBF. 0dB = 128.4468dBμV/m
Radiation Pattern C-RBF. 0dB = 133.7722dBμV/m
Radiation Pattern FC-RBF. 0dB = 130.872dBμV/m
Radiation Pattern LMS. 0dB = 129.5468dBμV/m

(a)

PT-RBF - E_θ in dBV/m along ϕ

Radiation Pattern. 0dB = 128.4468dBμV/m
Main Signal = 1°
Interference = 300°

PT-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.99234dB   SNR main signal = 35dB   IQ symbols = 160000   SER = 0

(b)

C-RBF - E_θ in dBV/m along ϕ

Radiation Pattern. 0dB = 133.7722dBμV/m
Main Signal = 1°
Interference = 300°

C-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -0.99234dB   SNR main signal = 35dB   IQ symbols = 160000   SER = 0

(c)

(d)

(e)

Source: (AUTHOR, 2018).

Figure 47 – Static scenario, D.S. DOA @ $\phi = 1°$, I.S. DOA @ $\phi = 270°$, no NL, noise at D.S $35dB$, SIR $= 35dB$

(a)

(b)

(c)

(d)

(e)
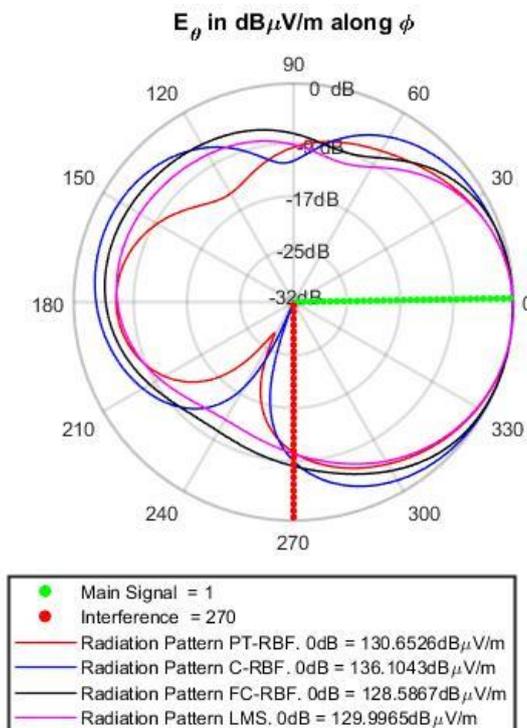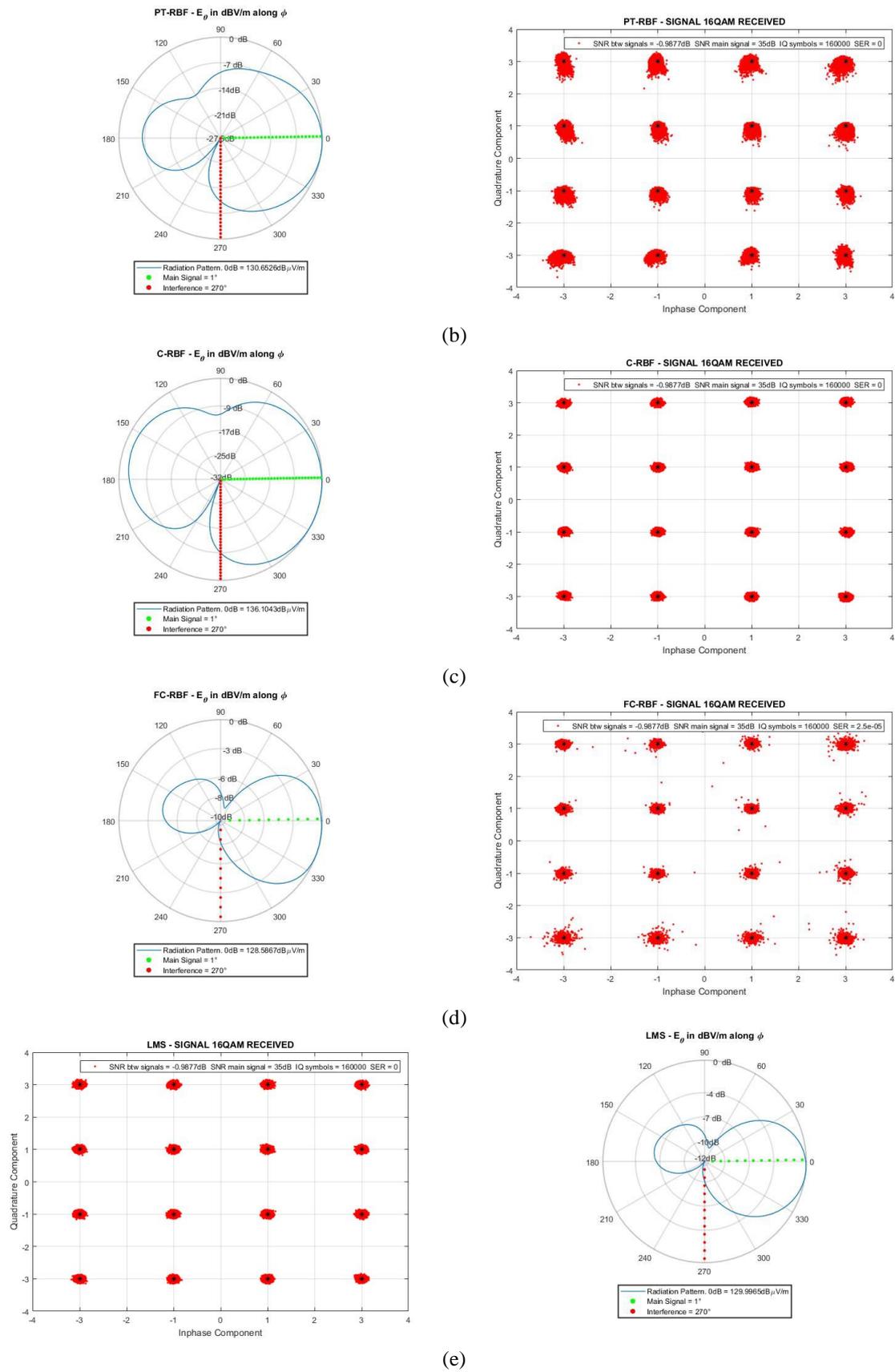
Figure 48 – Static scenario, D.S. DOA @ $\phi = 1°$, I.S. DOA @ $\phi = 270°$, with NL, noise at D.S $= 35dB$, SIR $= 10dB$

E_θ in dBμV/m along ϕ

Main Signal = 1
Interference = 270
Radiation Pattern PT-RBF. 0dB = 130.9203dBμV/m
Radiation Pattern C-RBF. 0dB = 136.6941dBμV/m
Radiation Pattern FC-RBF. 0dB = 124.4258dBμV/m
Radiation Pattern LMS. 0dB = 131.2967dBμV/m

(a)

PT-RBF - E_θ in dBV/m along ϕ

Radiation Pattern. 0dB = 130.9203dBμV/m
Main Signal = 1°
Interference = 270°

PT-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -10.0191dB SNR main signal = 35dB IQ symbols = 160000 SER = 0.00059375

(b)

C-RBF - E_θ in dBV/m along ϕ

Radiation Pattern. 0dB = 136.6941dBμV/m
Main Signal = 1°
Interference = 270°

C-RBF - SIGNAL 16QAM RECEIVED

SNR btw signals = -10.0191dB SNR main signal = 35dB IQ symbols = 160000 SER = 0.0049375

(c)

(d)

(e)

Source: (AUTHOR, 2018).

Figure 49 – Static scenario, D.S. DOA @ $\phi = 1°$, I.S. DOA @ $\phi = 60°, 90°, 120°, 150°, 180°, 210°, 240°, 300°$, no NL, no noise at D.S, SIR $= -10dB$


(a)

(b)

(c)

(d)

(e)
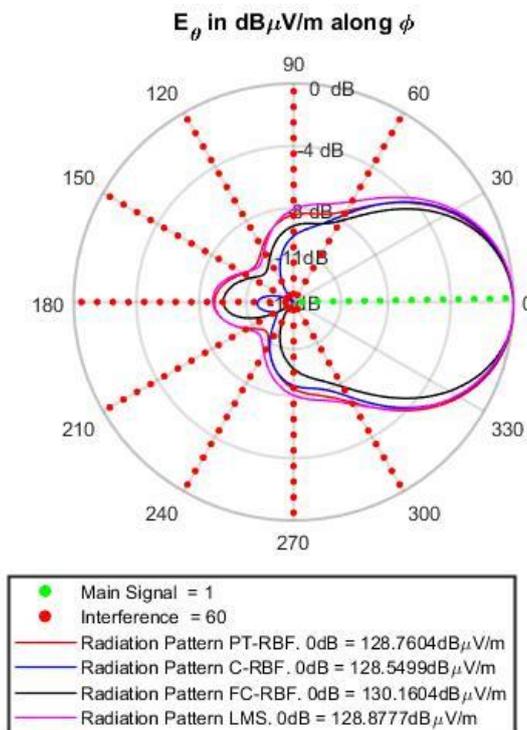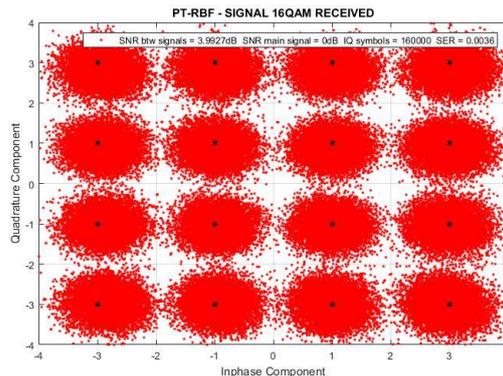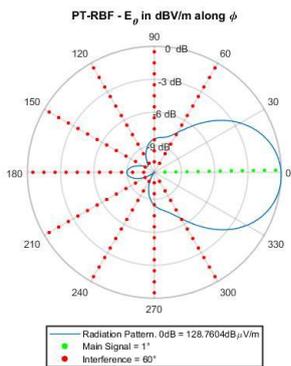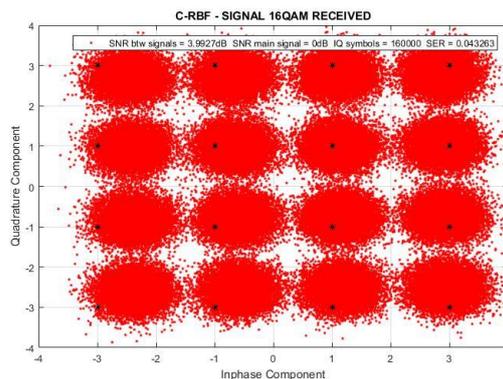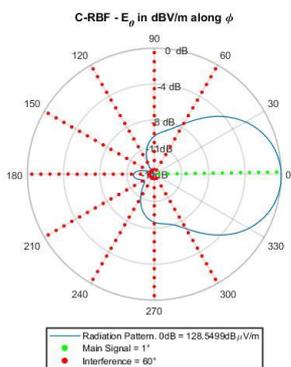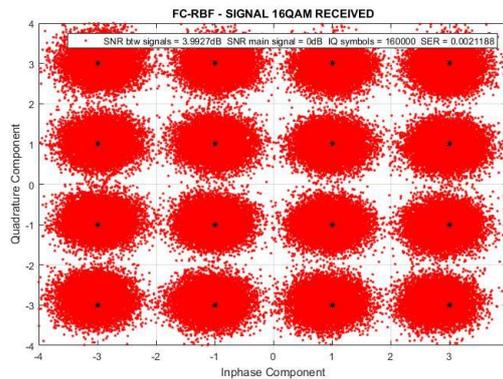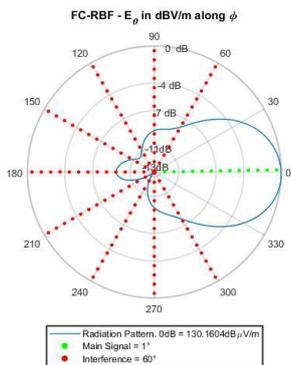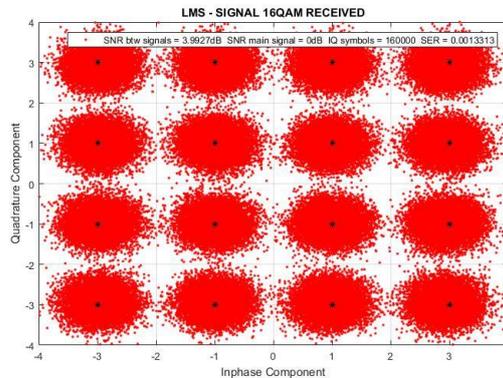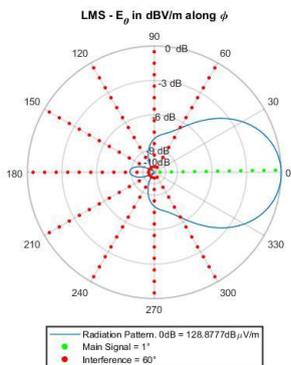
Source: (AUTHOR, 2018).

Figure 50 – Static scenario, D.S. DOA @ $\phi = 1°$, I.S. DOA @ $\phi = 60°, 90°, 120°, 150°, 180°, 210°, 240°, 300°$, with NL, no noise at D.S, SIR $= -10dB$

E_θ in dBμV/m along ϕ

(a)

(b)

(c)

(d)

(e)
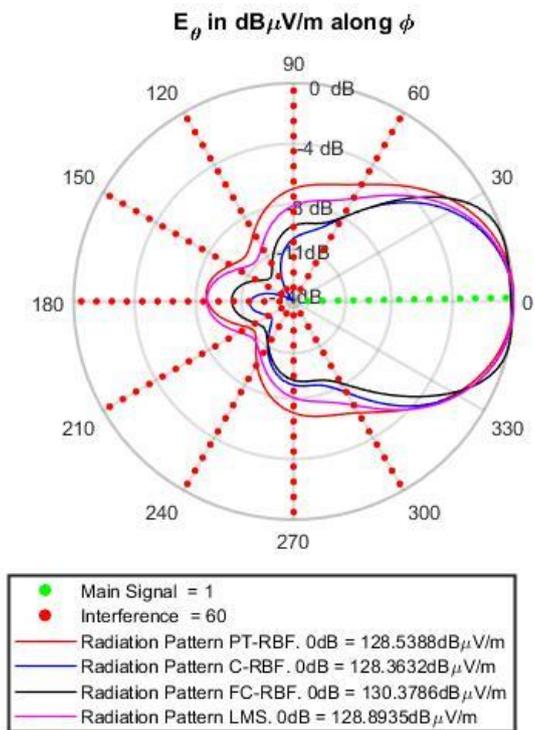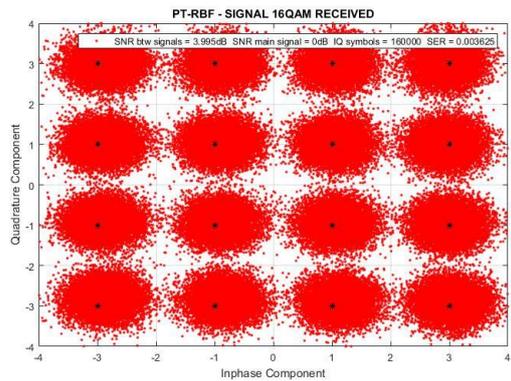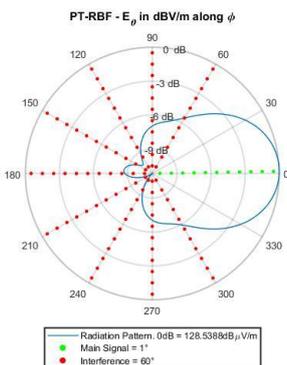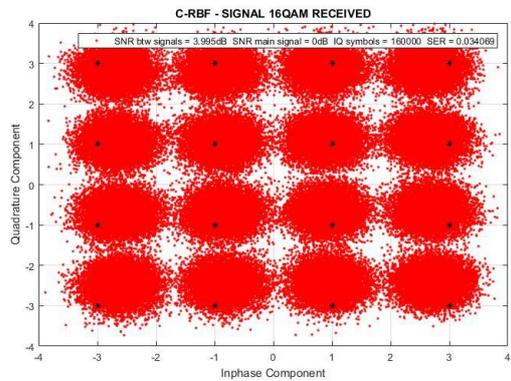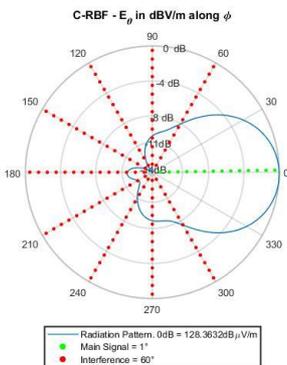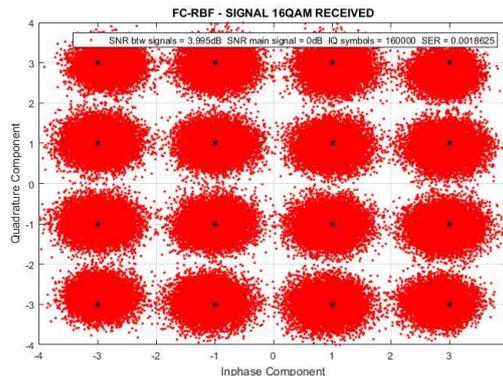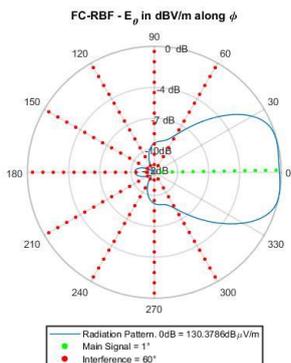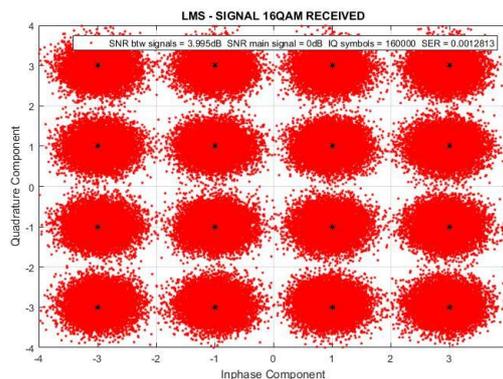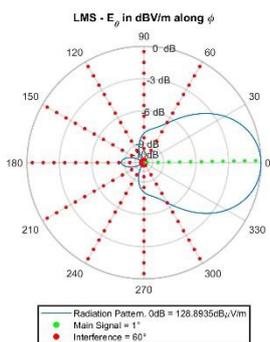
Source: (AUTHOR, 2018).

Figure 51 – Static scenario, D.S. DOA @ $\phi = 1°$, I.S. DOA @ $\phi = 60°, 90°, 120°, 150°, 180°, 210°, 240°, 300°$, no NL, no noise at D.S, SIR = $10dB$

(a)

(b)

(c)

(d)

(e)

Source: (AUTHOR, 2018).