Democratic and Popular Republic of Algeria

Ministry of Higher Education and Scientific Research

UNIVERSITY OF SAIDA Dr. TAHAR MOULAY SAIDA

Faculty of Technology

Department of Electronics

A Dissertation Submitted to the Department of Electronics

In Partial Fulfillment of the Requirement for the Degree of a Master in Telecommunications

Option: Systems of telecommunications

MODULATION DETECTION & CLASSIFICATION IN THZ COMMUNICATION SYSTEMS

Presented by:

HACHEMI Soumaya

KOURAT Ikram

Supervised by:

Dr. TAMI Abdelkader

Dr. OUESSAI Asmaa

Board of Examiners:

Dr. DAMOU Mahedi — *president*

Dr. OUARDI Aissa — *examiner*

Academic year: 2022-2023

Acknowledgments

First of all, all thanks and praise to our God " Allah" for granting us the courage and patience to complete this work.

I would like to express my sincere gratitude to my advisors, Dr. Tami Abdelkader and Dr. OUESSAI Asmaa , for their guidance, motivation and endless support.

My sincere thanks go to my parents for their continuous support, not to forget their prayers and kind words that helped me throughout my way.

We warmly thank our families for their moral and material assistance throughout the preparation period.

Last but not least I would especially show my faithful recognition to the jury members, the president and the examiner for their precious time reading and analyzing our work.

Dedication

Me and my friend dedicate this project with love, respect and appreciation:

To our tender mothers and the sweetheart of my heart, we cannot find words that can give her due. She is the epic of love, the joy of life, and an example of dedication and giving.

To our fathers , the ideal man , our role model and role model of my life, he is the one who taught our how to live with dignity and honor.

To our sisters and brothers.

To all the people in our large families.

ABSTRACT

The need to accommodate more users with a higher flow rate and better resolution with lower power consumption led to generation new to communications, THZ Wireless communications is a new technology implemented into future communication and harmless to health, It's a kind of channel nano communication we applied in KNN and SVM classification to determine the modulation type of a received signal.

The proposed system is tested for classifying BPSK, 4QAM, 16PSK , 16QAM, 4FSK, 32QAM, 8PSK,and 2 FSK modulation types, Moreover, the system uses High Order Cumulants (HOCs), variance, FOURIER transform of the received signal as discriminant features to distinguish between the different digital modulation types.

In addition, we have using support vector machine to classify the features and KNN classifier as a comparison, we have using linear, non-linear, one verses one and one verses all svm.

Keywords: Modulation Detection, Modulation classification, Nano communication, THZ communication, KNN ,SVM , NYUSIM simulator.

Resumé

La nécessité d'accueillir plus d'utilisateurs avec un débit plus élevé et une meilleure résolution avec une consommation d'énergie plus faible a conduit à la génération Nouveau dans les communications, Les communications sans fil THZ sont une nouvelle technologie mise en œuvre dans les communications futures et sans danger pour la santé C'est une sorte de canal de nano- communication que nous avons appliqué dans la classification KNN et SVM pour déterminer le type de modulation d'un signal reçu.

Le système proposé à tester pour classer les types de modulation BPSK, 4QAM, 16PSK, 16QAM, 4FSK, 32QAM, 8PSK et 2FSK. distinguer les différents types de modulation numérique.

De plus, nous avons utilisé une machine à vecteurs de support pour classer les caractéristiques et le classificateur KNN à titre de comparaison, nous avons utilisé linéaire, non linéaire, un vers un et un vers tous les svm.

Table of contents

Chapter III: classification method of digital modulation

List of figures

Chapter I : Overiew about Nano-Communication

Chapter II : Nano channel models

Chapter III : classification method of digital modulation

Chapter IV: Simulation and Results

Liste of Tables

APs: Access Points

AoD: Angle Of Departure

AoA: Angle Of Arrival

AWGN: Additive White Gaussian Noise

BPSK: Binary Phase Shift Keying

BS: Base Station

CIR: Channel Impulse Response

EM: Electromagnetic

FSK: frequency-shift keying

HITRAN: High-Resolution Transmission Molecular Absorption database

HOC: higher order cumulants

IoT : Internet of Things

KNN: K-Nearest Neighbor

LOS: line-of-sight

MC: Molecular Communications

MIMO: Multiple Input Multiple Output

MPCs: Multipath Components

NLOS: Non-line-of-sight

NYUSIM: New York University Simulator

PSK: Phase Shift Keying

QAM: quadrature amplitude modulation

Rx: Receiver

RBF: Radial Basis Function

SF: Shadow Fading

SNR: Signal-to-Noise

SVM: Support Vector Machine

Tx : Transmitter

THz: Terahertz

UEs: User Equipment

ULA: Uniform Linear Array

UT: user terminal

UMi : urban microcell

6G: sixth Generation

General introduction

Future wireless networks will have to deal with a substantial increase of data transmission due to a number of emerging applications that include machine-to-machine communications and video streaming This very large amount of data exchange is expected to continue and rise in the next decade so, The field of wired and wireless communications is considered one of the most important fields because it is in continuous development and prosperity and facilitates life in wireless communication systems so that sixth generation networks depend on technologies that can offer a major increase in transmission capacity as measured in bits/Hz.

Nano-technology has made significant developments and progresses well into what has very recently believed being science fiction. The research field of nano-technology is becoming a key area in science based on multi-disciplinary collaborations among medicine, engineering, physics, biology, computer science, and others. It turned out that many of the envisioned applications for nano-technology require or might exploit the potential advantages of communication and hence cooperative behavior of these nano-scale machines to achieve a common objective that exceeds the capabilities of a single device.

This document is structured around four chapters:

First chapter will present a Overview about Nano-Communication is the exchange of information at the nanoscale and it is at the basis of any wired/wireless interconnection of nanomachines in a nanonetwork, it can be protected from the attacker by confidentiality, integrity, and availability.

Second chapter will give Nano channel models of between it wireless thz channel Terahertz (0.1-10 THz) communications are envisioned as a key technology for sixth generation (6G) wireless systems. The study of underlying THz wireless propagation channels provides the foundations for the development of reliable THz communication systems and their applications.

Third chapter will deal with classification method of digital modulation.Complex classifiers are used for extended performance, such as K- nearest neighbor and support vector machine (SVM), or deep-learning methods .While Classification is an important problem in big data, data science and machine learning. The K-nearest neighbor (KNN) is one of the oldest, simplest and accurate algorithms for patterns classification and regression models. Many modern algorithms perform extraction of several different features (variance ,FFT,HOCs) simultaneously to know the types of modulation.

The last chapter will be dedicated to show of different simulations of reduced complexity algorithms using different parameters for three channels (selectif channel, AWGN, and TERAHERTZ channel) and show their impact on the performance of the system by measuring the **SNR.** We used several commonly digital modulations are phase-shift keying (PSK), frequency-shift keying (FSK) and quadrature amplitude modulation (QAM). In order to know the simulation results.

Chapter I

Overview about Nano Communication

I.1 INTRODUCTION

Nanotechnology was first developed in 1965[1], it consists of the processing of, separation, consolidation, and deformation of materials by one atom or by one molecule. It's fabrication of devices in a scale ranging from 1 to 100 nanometers [1]. Nanoscale science can be divided into three broad areas, nanostructures, nanofabrication and Nano characterization with typical applications in Nano electronics, life sciences and energy. There are several applications of nanotechnology nowadays as illustrated in **Figure I.1**, but with respect to electrical and electronics fields, a feasible applications of nanotechnology are; Communications, Bio-engineering, Medical Electronics and Robotics [1].

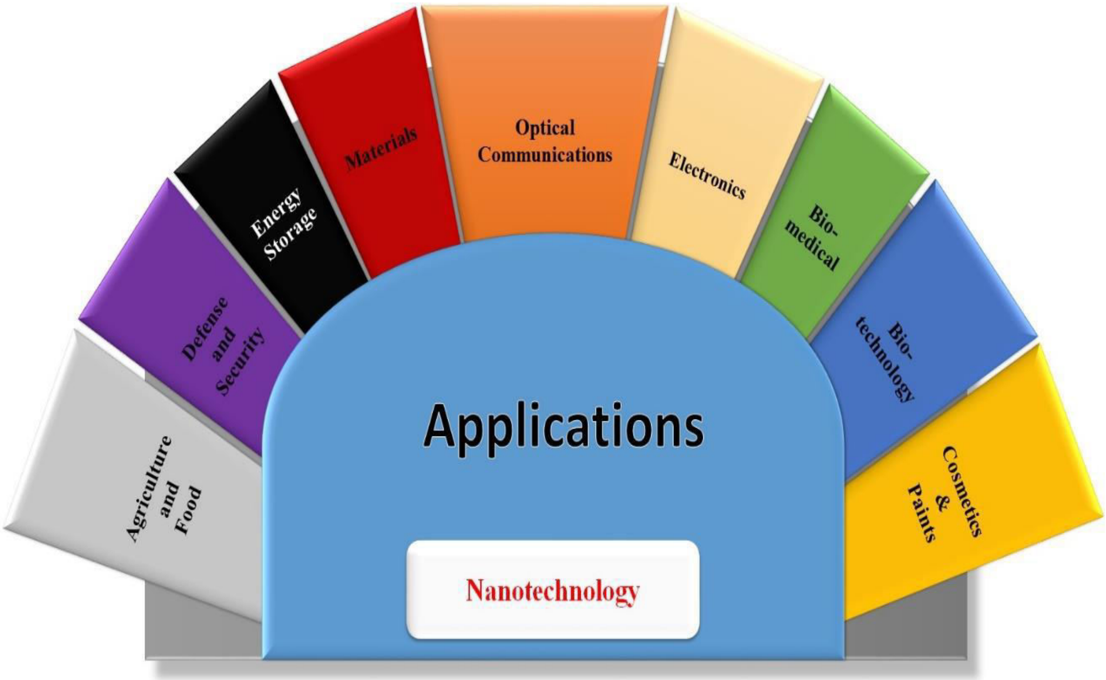


Figure I.1 : Nanotechnology Applications[1]

Nanotechnology will enable manufacturers to produce computer chips and sensors that are considerably smaller, faster, more energy efficient, and cheaper to manufacture than their present-day counterparts .Micromechanical sensors also became an elementary part of automotive technologies in mid-1990, roughly ten years later more miniaturized micromechanical sensors are enabling novel features for consumer electronics and mobile devices. Within the next ten years the development of truly embedded sensors based on nanostructures will become a part of our everyday intelligent environments [2]. Nanorobotics is the emerging technology field of creating machines or robots whose components are at or close to the microscopic scale of a nanometer (9-10 meters). Nanorobotics refers to the nanotechnology engineering discipline of designing and building Nano robots, with devices components [3].

Nano-technology has made significant developments and progresses well into what has very recently believed being science fiction. The research field of nano-technology is becoming a key area in science based on multi-disciplinary collaborations ranging in size from 0.1-10 micrometers and constructed of nanoscale or molecular among medicine, engineering, physics, biology, computer science, and others. It turned out that many of the envisioned applications for nano-technology require or might exploit the potential advantages of communication and hence cooperative behavior of these nano-scale machines to achieve a common objective that exceeds the capabilities of a single device. At this point, the term nano-networks is defined as a set of nano-scale devices, i.e., nano-machines, communicating with each other and sharing information to realize a common objective. Nano-networks allow nano-machines to communicate and share any kind of information required by wide range of applications including biomedical engineering, biological and chemical defense technologies, and environmental monitoring [3].

I.2 NANO COMMUNICATION :

Nano-communication has been a hot topic since its proposal, so is its concept. Below Are some explanations on both nano-network and nano-communication from the open Literatures:

- there are many nano-scale networks embedded within each device that might be otherwise more effectively utilized for communication. ...” (2006) [4]

- Nanonetworks. i.e., the interconnection of nano-machines are expected to expand the capabilities of single nano-machines by allowing them to cooperate and share information.” (2008)[5].

- Nanonetworks are communication networks that exist mostly or entirely at the nanometer scale.” (2010) [6].

- Nanocommunication is the exchange of information at the nanoscale and it is at the basis of any wired/wireless interconnection of nanomachines in a nanonetwork.”(2011) [7]
- As the name indicates, nano-communication encapsulates the communication between devices at the nano-scale applying novel and modified communication and radio propagation principles in comparison to conventional and existing solutions.” (2015) [8]
- Nanonetworking is an emerging field, communicating among nanomachines, expanding The capacity of single nanomachine.” (2015)[9].

Nano-communication, the exchange of information at nano-scale,is the only feature that enables nano-machines to work in a synchronous, supervised and cooperative manner to pursue a common objective. However, for the time being, it is still an unsolved challenge to enable the communication among nano-devices. Four main communication paradigms are proposed for nano-networks depending on the technology used to manufacture the nano-machines and the targeted application, namely, nano-electromagnetic, molecular, acoustic, and nano-mechanical [10]. Here we briefly discuss two popular communication methods i.e., electromagnetic and molecular communication two popular methods, and share some examples from the literature.

I.2.1 molecular communication :

Disseminating information from one place to another place has been investigated and deployed for hundreds of years and the modern electromagnetic (EM) wave [11] based technology has enabled a wide range of communication scenarios since Marconi invented the modern telecommunication system. The communication distance of the modern telecommunication Systems spans a wide range from thousands of kilometers to several meters while communications within the microscale remain undeveloped deeply. Although existing techniques, e.g, optical communications [12] and the terahertz (THz) techniques [13] can be applied to the microscale communications, these techniques will suffer from the path loss [14, 15] due to the conductive medium and the size of the components is non-negligible, forcing researchers to find another way to implement the communication systems in microscale.

The potential applications that necessitate the microscale communications include biosensing [14], nano-networks or nano-machines [16] etc. The scientists in biology have been working on figuring out how to fabricate nano-machines [16] in order to invent some robots that can be injected into blood vessels, identify the target organs and deliver specified drugs [17] to the organs. Meanwhile, environmental control and better understanding the biology are among the ultimate objectives of the microscale communications.

In nature, there are countless number of cases of communications that are not based on the EM wave. A typical example is in our bodies, e.g, dopamine molecules excite our nerve cells to make us feel happy after we hear some jokes. In this regard, the dopamine molecules convey the signals of molecular communications. The mechanism that the information is transferred by some molecules is referred to as molecular communications (MC) [18].

Molecular communication, the transmission, and reception of information encoded in molecules is a new and interdisciplinary field [19]. In molecular communication, a nano-transmitter releases small particles such as molecules or lipid vesicles into a fluidic or gaseous medium, where the particles propagate until they arrive at a receiver. The nano-receiver then detects and decodes the information encoded in these particles. Messages can be encoded in different properties such as concentration, number, type, release timing, and/or a ratio of molecules. A summary of the design and engineering of components for molecular communication systems from biology, chemistry, and nanotechnology is provided in [20] and [21]. The first approach for engineering bio-nano-sensor has been used and demonstrated in synthetic biology [22-24] by modifying a metabolic pathway of a biological cell, which then synthesizes and releases specific signal molecules to send information. Another approach to engineer sender and receiver bio-nano-machines is to create simplified cell-like structures using biological materials (e.g., by embedding proteins in a vesicle) [25, 26].

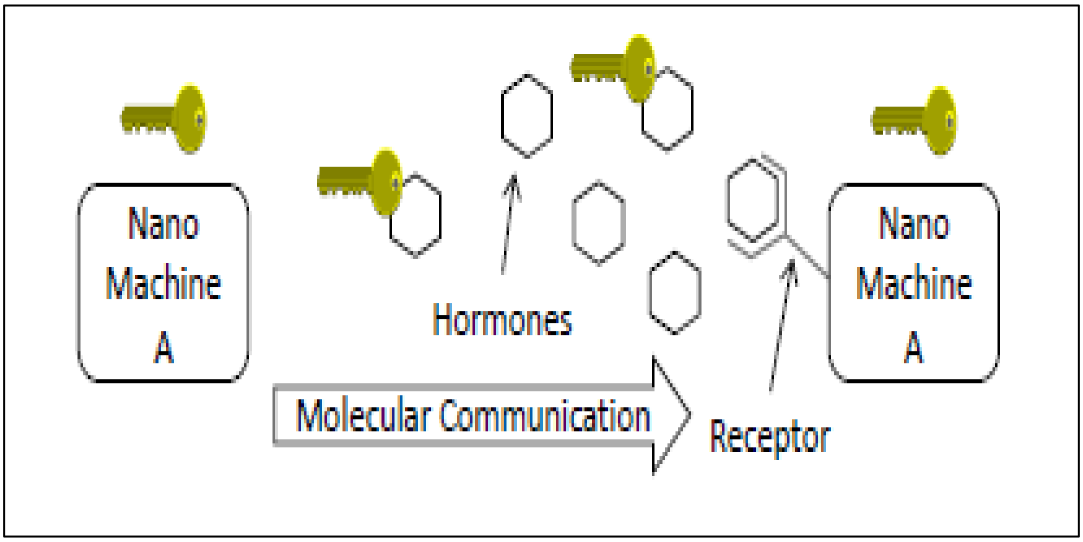


Figure I.2 : Bio-signaling using molecular communication

I.2.2 electromagnetic communication :

Nano-electromagnetic communication is defined as the transmission and reception of EM radiation from components based on novel nano-materials [27]. The latest advancements in graphene-based electronics have opened the door to EM communication among nano-devices in the terahertz (THz) band. The refinement of existing architectures and the utilization of new technologies brought THz communication paradigm closer to the reality .In order to enable communication among devices at the nano-scale, THz transmitters should be compact where their area size should reach to hundreds of square nanometers. It is found that electronic sources are capable of providing high levels of average output power at lower THz frequencies [28] and they can be feasible for THz biological research studies. Advancements in microelectronics led to miniature electronic components suitable for intra-body communication [29].

I.3 SECURITY IN NANO-COMMUNICATION:

Looking at security in nano communication, it is reasonable to start with the classical security goals confidentiality, integrity, and availability. The basic goals of security will not change when going from classical communication security to nano communication security. Facing an attacker that has a certain access to the nano communication system, we want to ensure:

- **Confidentiality:** an attacker should not be able to learn the content of a message exchanged between a sender and a receiver.
- **Integrity:** an attacker should not be able to modify the content of a message exchanged between a sender and a receiver.

- **Availability*:*** an attacker should not be able to disrupt of negatively affect communication.

Confidentiality and integrity imply authenticity, i.e., the receiver of a message should be able to verify the identity of the sender to prevent message spoofing. A further security goal that can be derived, e.g., from sensor or vehicular networks is data consistency, i.e., data transmitted should report true situations, measurements, or findings. Insiders should not be able to report arbitrary information. In classical networks, confidentiality, integrity, and authenticity are typically implemented based upon cryptographic primitives and protocols.

I.4 TERAHERTZ CHANNEL:

Terahertz (THz) band (0.1 − 10 THz) is well-known as a promising candidate for nano-scale communication [30,31], because of its non-ionization and robustness to the fading characteristics. Recently, the advent of novel nano-materials (e.g., graphene-based carbon nano tubes etc.) with remarkable electrical properties has led to a surge of interest in development of nano-scale devices [30, 31] for communication in THz band. The literature so far comprises of the studies which investigate the antenna design [32], propagation models [33,34], transceiver design [35-37], networking issues [38, 39], and data rates [33]. Nevertheless, the open literature on nano-scale communication in THz band still falls short of the problem of intelligent nano receiver design.

A promising application of THz communication is to enable active communication link between nano-sensors. In this regard, a pioneering work in proposing THz-based in-vivo networks at nano-scale is presented in (Abbasi et al., 2017)[40]. Although the communication at nano-scale can be performed in different ways, such as molecular, acoustic and electromagnetic (EM) communications at THz frequencies, which is considered as one of the most suitable approaches to enable communication amongst nano-sensors. This is because biological tissues remain non-ionised at THz frequencies. Furthermore, THz frequencies are more resistant to scattering phenomenon (Alomainy et al., 2019)[41]. Another reason of utilising THz frequencies for nano-scale communication is the property of molecular resonance at these frequencies. The underlying concepts of THz from physical to network layer are explained in (Abbasi et al ., 2016)[42], where a visionary architecture of THz based nano-communications is presented. Based on that work, **Figure I.3** envisions an architecture of THz enabled in-vivo networks incorporated in 6G communications (Abbasi et al ., 2017)[40]. Nano-sensors depicted in **Figure I.3** and **Figure I.5** are small devices with highly limited communication, memory and energy resources. They would perform the simplest tasks of communication and sensing. Contrarily, nano-routers are slightly larger than nano-sensors in terms of the resources they have and the actions they can perform. They act as the controller of nano-sensors by sending simple commands such as on, off, and sleep etc. Nano-micro interface are comparatively larger devices that can control the information passed to nano-routers. They have the ability to communicate over a nano-paradigm at the southbound interface and a classical communication paradigm at the northbound interface. The gateway is the central entity that connects the entire system to the Internet. The data collected through nano-sensors is transferred to a remote data analytics platform using 6G communications network. The data analytics platform sends the instruction back to the nano-network over the same 6G network.

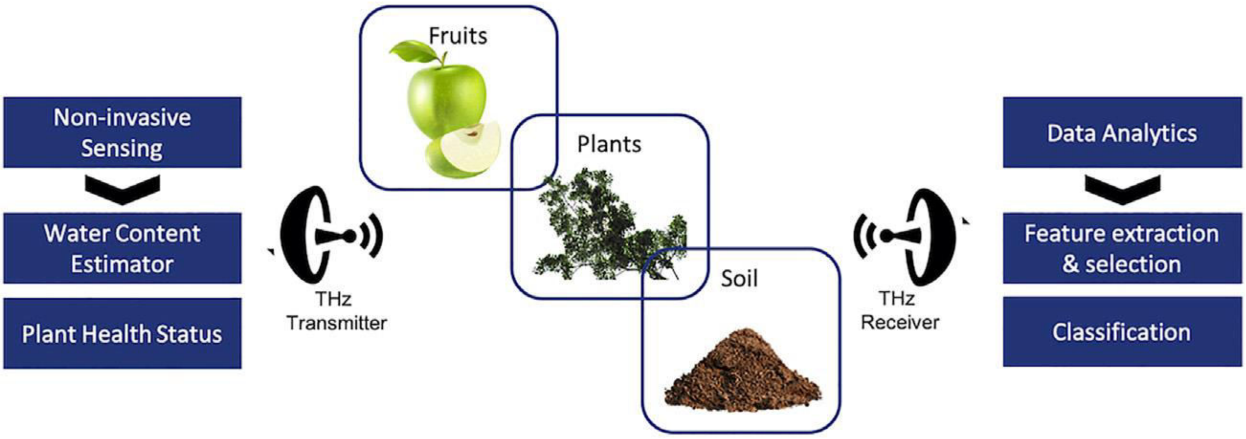


Figure I.3 : The envisioned architecture of in-vivo nano-networks connected with 6G communications [40]

consider an intelligent/cognitive nano receiver listening to a signal of interest in THz band. With the objective of modulation mode detection and modulation classification, two distinct application scenarios are foreseeable:

1) the nano receiver is the intended recipient of the signal transmitted by the nano transmitter;

2) the nano receiver is not the intended recipient; rather, it overhears the signal intended for some other nano receiver in the nearby vicinity (see **Fig I.4**). Note that the application scenario 1 (scenario 2) implies that the nano receiver has a trustworthy (untrustworthy) relationship with the nano transmitter. Scenario 2 depicts passive interception which arises in military, defense, and security applications.

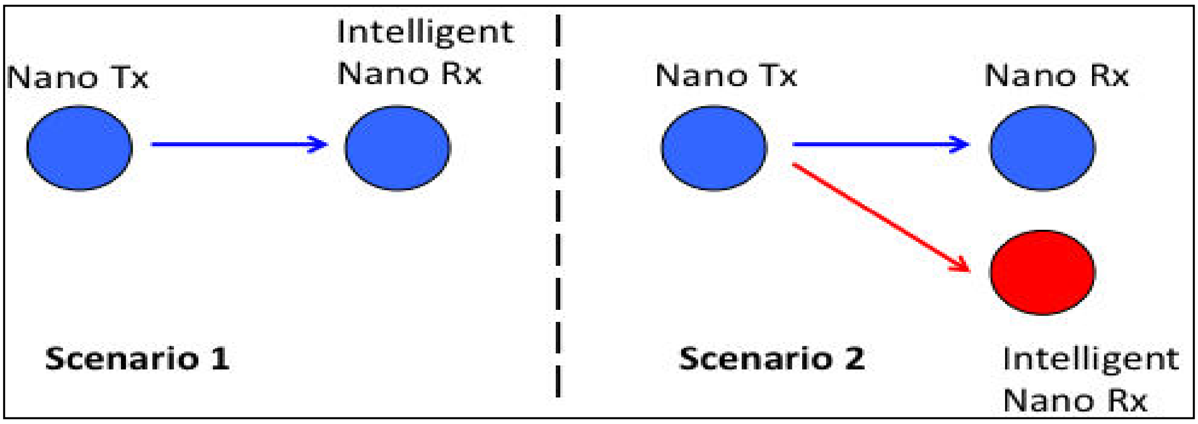


Figure I.4 : The system model: two distinct application scenarios for the considered intelligent nano receiver

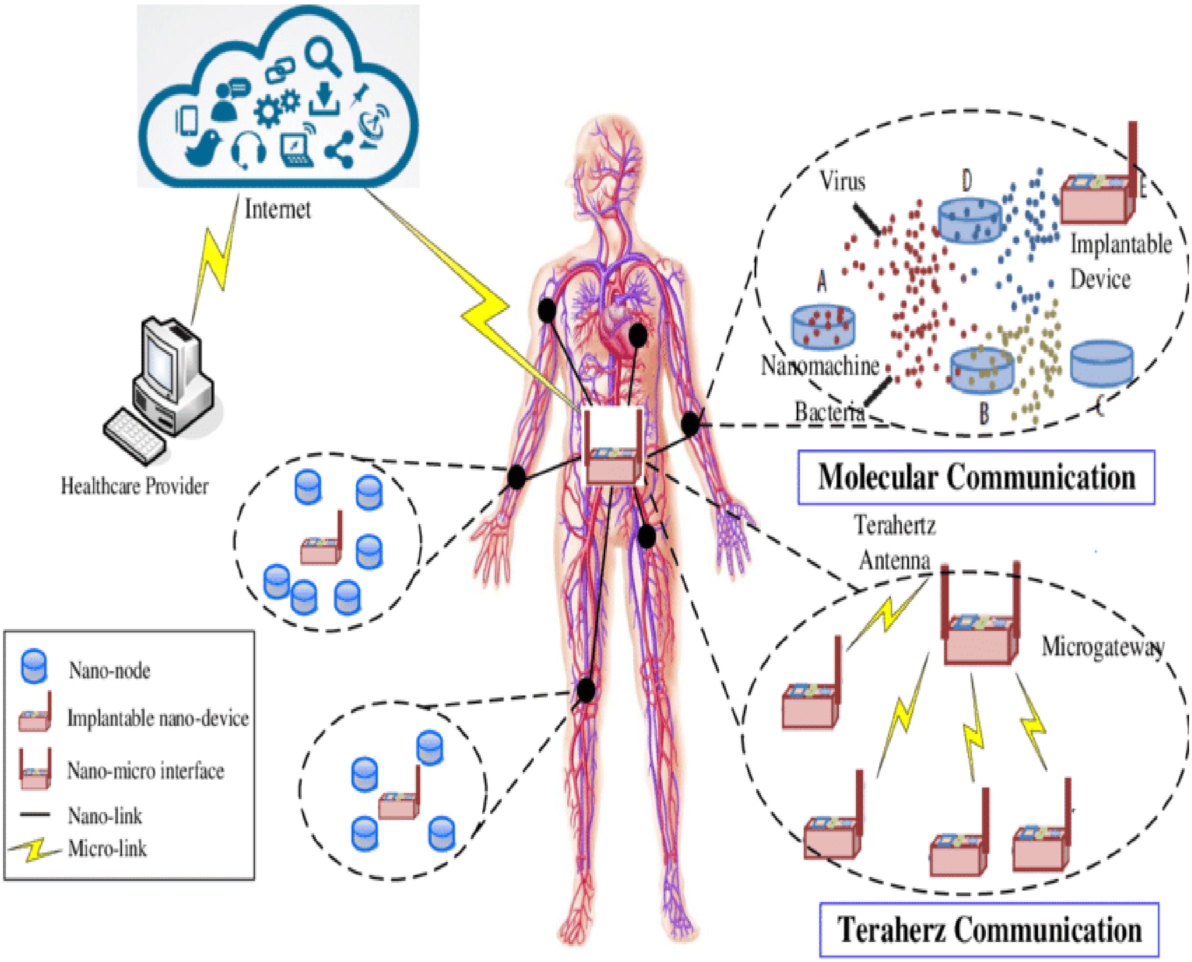


Figure I.5 : The sketch of the nano communication network

I.5 Nanonetworks applications :

The potential applications of nanonetworks are unlimited .We classify them in four groups: biomedical, environmental, industrial and military applications. However, since nanotechnologies have a key role in the manufacturing process of several devices, nanonetworks could be used extensively in many other fields such as consumer electronics, life style and home appliances among others.

I.5.1 Biomedical applications:

The most direct applications of nano-machines and nanonetworks are in the biomedical field. Biological models inspire and encourage the use of nanotechnologies to interact with organs and tissues. The advantages provided by nanonetworks are clearly in terms of size, biocompatibility and biostability, enabled by the control of system components at molecular level. These are some of the envisaged applications:

- Immune system support. The immune system is composed by several nano-machines that protect organism against diseases. These nano-machines are a collection of nano, micro and macro systems, including sensors and actuators, acting in a coordinated way to identify and control foreign and pathogen elements. Nanomachines can be used to help the detection and elimination procedures. They could realize tasks of localization and response to malicious agents and cells, such as cancer cells [43,44], resulting in a less aggressive and invasive treatments compared to the existing ones.

- Health monitoring. Oxygen and cholesterol level, hormonal disorders, and early diagnosis are some examples of possible applications that can take advantage of in body nano-sensor networks [45,46]. The information retrieved by these systems must be accessible by relevant actors. Thus, nanonetworks must provide the proper level of connectivity to deliver the sensed information.

- Genetic engineering. Manipulation and modification of nano-structures such as molecular sequences and genes can be achieved by nano-machines. The use of nanonetworks will allow expanding the potential applications in genetic engineering.

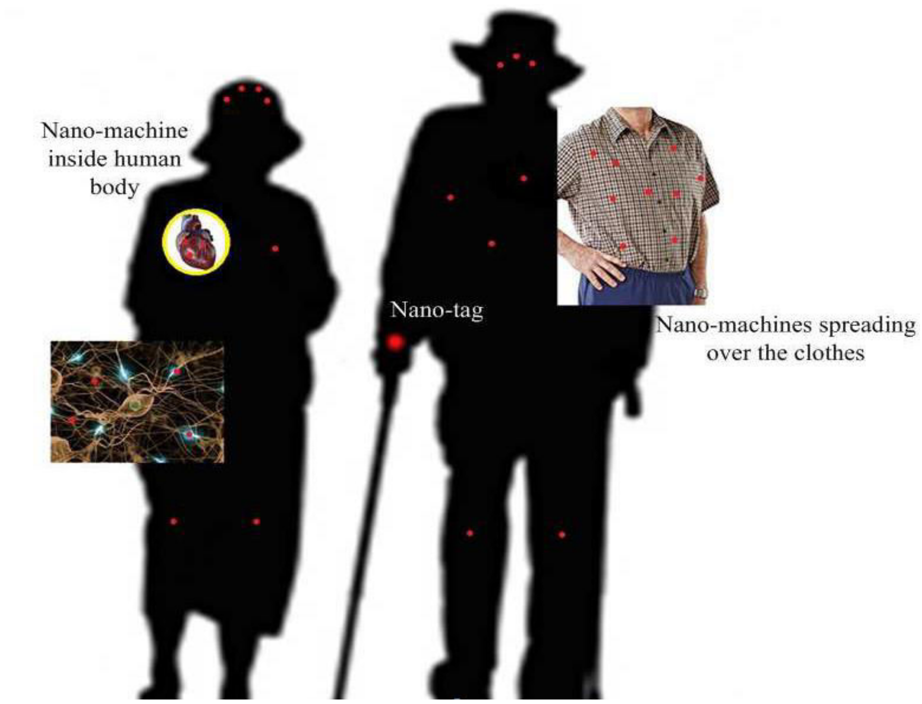


Figure I.6: Architecture of a health monitor nano-network

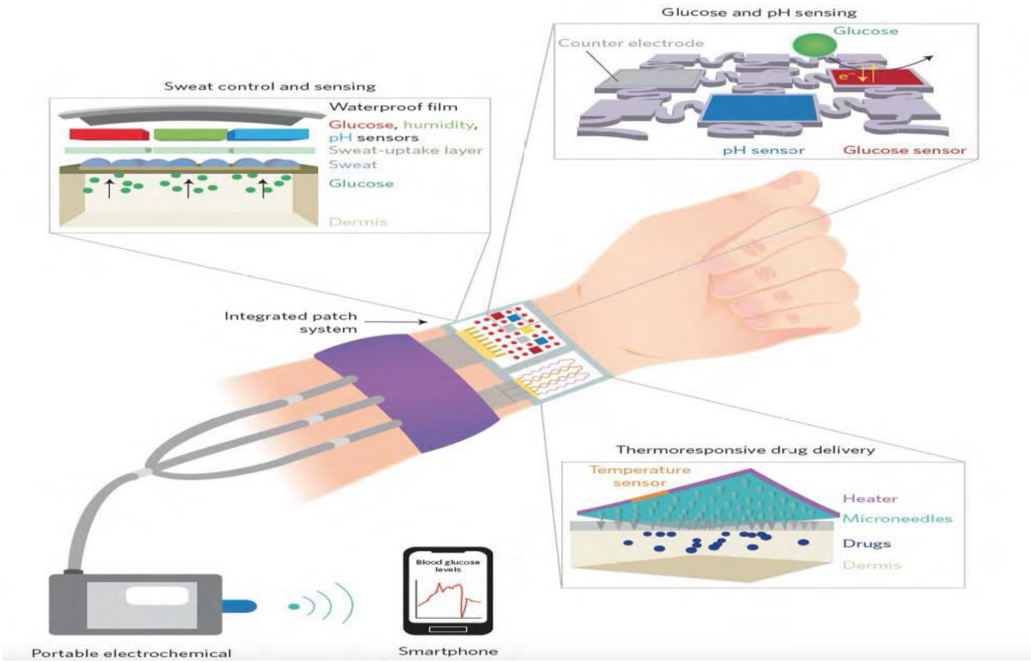


Figure I.7 : Nano-network: glucose graphene skin sweat sensor and drug delivery chip

I.5.2 Industrial and consumer goods applications:

Nanonetworks will be used not only the intra-body but also in industries. Nanonetworks can help with the development of new materials, manufacturing processes and quality control procedures. More specifically, these applications have already been proposed:

- Food and water quality control. Similar to health monitoring applications, food and fluids quality control can take advantage of nanonetworks. Nano-sensor networks can help detecting small bacteria and toxic components that can affect to the product quality and cannot be detected using traditional sensing technologies [47]. Advanced self-powered nano-sensor networks can be used to detect very small amount of chemical or biological agents installed in water supplies across the country.

I.5.3 Military applications:

Nanotechnologies can also have several applications in the military field. While in the applications pointed out before, the range covered by nanonetworks is short, in military field the deployment range of nanonetworks can be widely variable depending on the application. Battlefield monitoring and actuation demand a dense deployment of nanonetworks over large areas, while systems aimed at monitoring soldier performance are deployed in smaller areas, i.e., human body. Among the military applications,we can mention:

- Nuclear, biological and chemical (NBC) defenses. This is a classic application for large area deployment. Nanonetworks can be deployed over the battlefield or targeted areas to detect aggressive chemical and biological agents and coordinate the defensive response [48]. The overall system can be compounded of nano-sensors and nano-actuators, which would detect and control the hostile agents. Nano-sensor networks could also be deployed into cargo containers to detect the unauthorized entrance of chemical, biological or radiological materials.

I.5.4 Environmental applications:

Since nanoneworks are inspired in biological systems found in nature, they can also be applied in environmental fields achieving several goals that could not be solved with current technologies. Some environmental applications are:

- Biodegradation. There is an existing and growing problem with garbage handling around the world. In this line, nanonetworks could help with the biodegradation proc ess in the garbage dumps. Nanonetworks can help the biodegradation process by sensing and tagging different materials that can be later located and processed by smart nano actuators.

- Air pollution control. Similar to the quality control applications, air can be monitored using nanonetworks. Moreover, nano-filters can be developed to improve the air quality by removing harmful substances contained in it [49].

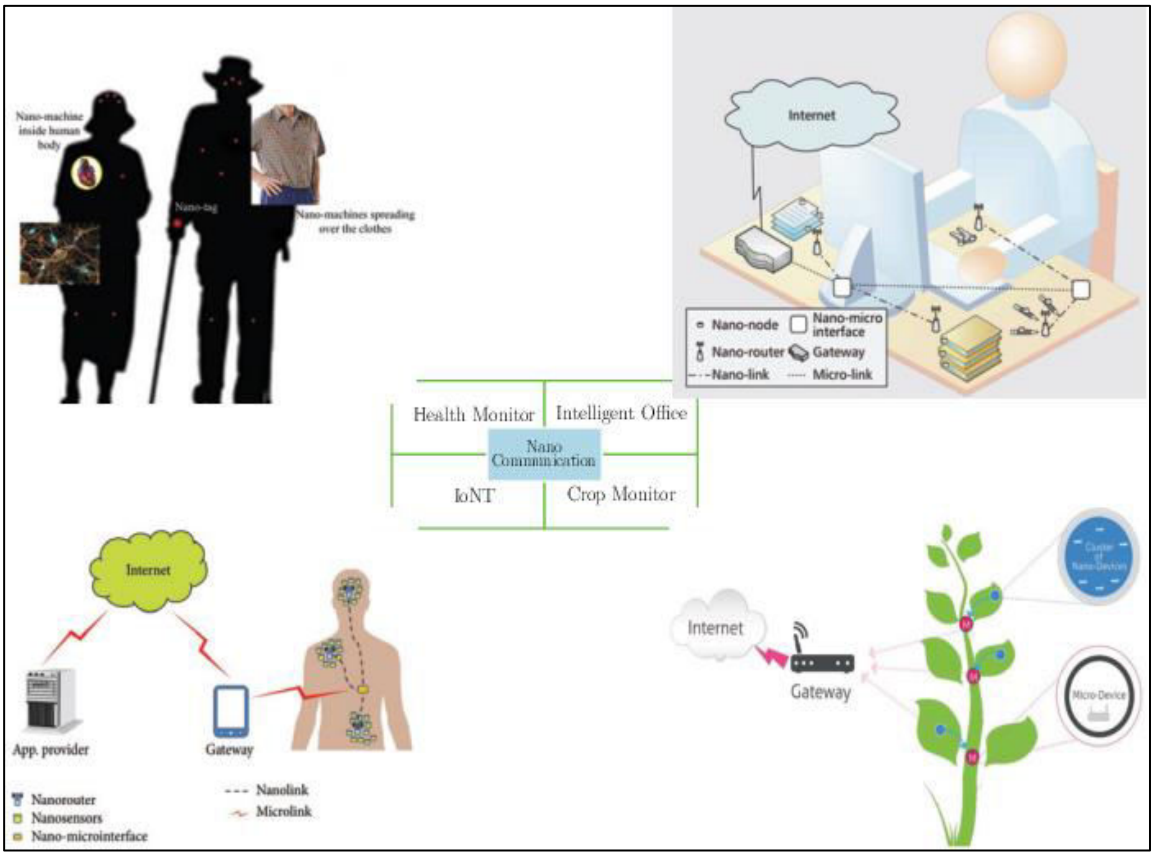


Figure I.8 : Envisioned applications for nano communication [50]

I.6 CONCLUSION:

Nanocommunications, understood as molecular communications (those performed in between groups of molecules or cells) is definitely one of the hottest research areas in IT science in the last few years. The progress in this discipline is strenuous, especially as it requires cooperation of research from both IT and medical fields, but the potential applications are enormous, extending from medical ones, like remote diagnostics and surgery, localised drug delivery, and aging care, via sports, entertainment and military, to the final vision of human-Internet integration realised by getting signals directly from human nerve cells and sending them to the Internet. This is a vision of truly personal communications, where parts of the Internet are related to particular people, as they are located inside their bodies.

In this chapter, we have provided an overview of nanotechnology and nanocommunications, the latter branching into several types, including molecular and electromagnetic communications. We use the terahertz channel band(0.1-10THZ) in nanocommunications. We also discussed applications ,We classify them in four groups: biomedical, environmental, industrial and military applications.We also addressed with the current development in nanocommunications, especially in the field of medicine in body humen.

In the next chapter we will talk about the channels models in nano communication system.

Chapter II

Nano channel models

II.1 INTRODUCTION

Today modern telecommunication systems transfer information through use of electrical or electromagnetic signals. There are, however, still many applications where these technologies are not convenient or appropriate. For example, use of electromagnetic wireless communication inside networks of tunnels, pipelines, or unpredictable underwater environments, can be very inefficient. As another example, at extremely small dimensions, such as between micro- or nano-scaled robots [51], electromagnetic communication is challenging because of constraints such as the ratio of the antenna size to the wavelength of the electromagnetic signal [52]. Inspired by nature, one possible solution to these problems is to use chemical signals as carriers of information, which is called molecular communication [53], [54]. In molecular communication, a transmitter releases small particles such as molecules or lipid vesicles into a fluidic or gaseous medium, where the particles propagate until they arrive at a receiver.

The recent interest in electromagnetic (EM) wireless nano-networks in the THz band stimulates the development of new models for performance assessment. Due to the advantages of size, bio-compatibility and bio-stability, this small-scale network is especially promising in biomedical fields [55]. For in-vivo nano-networks, EM communication is based on the transmission and reception of EM wave at the THz band to exchange information.

To date, some works have been performed on the channel characterisation and modelling of nano-communication at the THz band. Channel modelling of the THz wave propagating in the atmosphere with a different concentration of the water vapour was first studied in [56, 57]. Subsequently, both numerical and analytical EM channel modelling at the THz band for in-body nano-communication was investigated in [58] . In order to quantify the potential of the THz communication, channel capacity is used as an important performance metric and to proceed with realistic and accurate representation of the actual capacity and channel model in the THz band, consideration of channel noise model is necessary.

II.2 DEFINITION OF WIRELESS Thz CHANNEL

Terahertz (0.1-10 THz) communications are envisioned as a key technology for sixth generation (6G) wireless systems. The study of underlying THz wireless propagation channels provides the foundations for the development of reliable THz communication systems and their applications. The wireless propagation channel is the medium over which transportation of the signals from the transmitter (Tx) to the receiver (Rx) occurs, and as a consequence, channel properties determine the ultimate performance limits of wireless communications, as well as the performance of specific transmission schemes and transceiver architectures [59], [60]. Since wireless channels are the foundation for designing a wireless communication system in the new spectrum and new environments, it is imperative to study the THz radio propagation channels for 6G future wireless communications [61], [62], [63]. Studying wireless channel features relies on the physical channel measurements with channel sounders. The characteristics of the wireless channel are then analyzed based on the measurement results for developing channel models. Channel models capture the nature of wave propagation with reasonable complexity and allow the fair comparison of different algorithms, designs and performance in wireless networks.

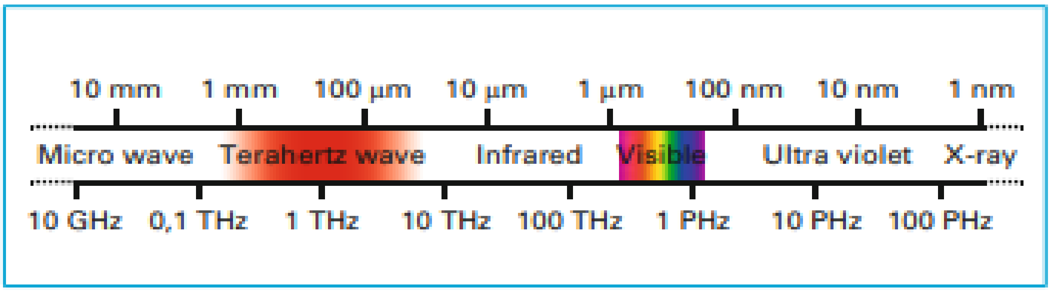


Figure II.1 : Spectrum of electromagnetic waves and identification of the terahertz frequency range

II.3 Properties of THZ:

THz waves are of considerable interest for a multitude of applications. They are non-ionizing, that is to say that they do not damage the objects that they pass through, and therefore THz waves are not harmful for health. We will present an example of THz applications.

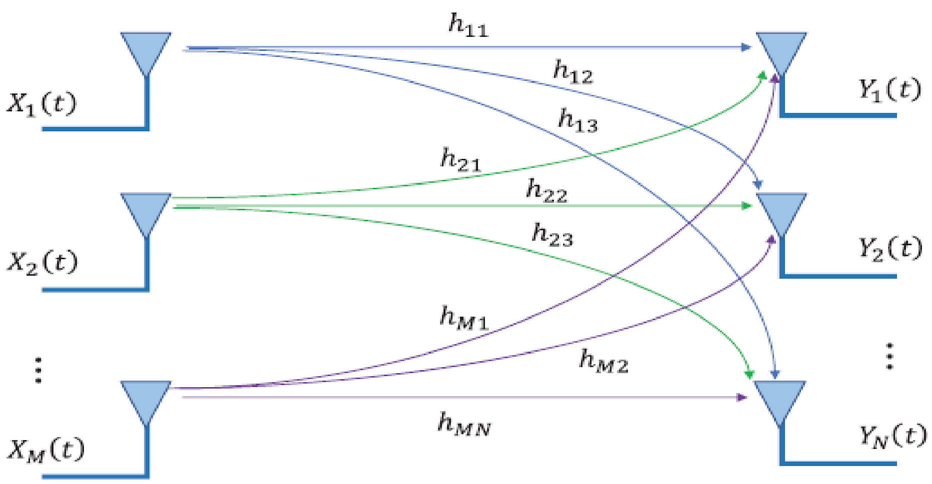
II.3.1 application in Telecommunications:

THz should play a decisive role in the development of wireless networks of the future that will have to satisfy a rapidly increasing demand for bandwidth. As THz frequencies are higher than those of the radio waves and Wi-Fi that are currently used, they should allow significantly higher data transmission speeds to be achieved (> 10 Gbps). THz waves will also facilitate lossless satellite communications over long distances.

II.4 Models of wireless Thz channel :

II.4.1 Gaussian Model of Thz channel :

The bandwidth of the terahertz band is very large which support the terabit transmission rate per second and the terahertz band itself has frequency selectivity [64]. In nano-communications, due to the limitations of the energy of nano-devices and other factors, the distance between the sending end and the receiving end is relatively close, but the antenna spacing is much smaller than the distance between the sending end and the receiving end. In the short-distance transmission, the anti-interference of the terahertz band is very good.
its MIMO channel is modeled as follows. The simplified MIMO channel model in the terahertz channel is shown in **FigureII.2**.



FigureII.2: Simplified MIMO channel in Terahertz communication

As shown in **FigureII.2**, it is assumed that there are M antennas at the transmitting end and N antennas at the receiving end. The signals transmitted on the antenna array can be expressed as

$$\mathbf{x(t)}=[x_1(t)\ x_2(t) \ldots x_M(t)]^T \quad (1)$$

among them, the symbol $[.]^T$ represents the transpose of the matrix and $x_j(t)$ is the signal of the **jth** antenna port at the transmitting end.

Similarly, the signals on the antenna array at the receiving end are

$$\mathbf{y(t)}=[y_1(t)\ y_2(t) \ldots y_N(t)]^T \quad (2)$$

Among them, $y_i(t)$ is the signal of the ith antenna port at the receiving end. In signal transmission, there is the influence of noise factors which can be expressed as

$$\mathbf{w(t)}=[w_1(t)\ w_2(t) \ldots w_N(t)]^T \quad (3)$$

Among them, $w_k(t)$ is the noise signal on the **k**-the channel. Besides, the channel gain moment $A_n$on the nth path is [65,66]

$$\mathbf{An}=\begin{bmatrix} a_{11}[n] & a_{12}[n] & \cdots & a_{1N}[n] \\ a_{21}[n] & a_{22}[n] & & a_{2M}[n] \\ \vdots & & \ddots & \vdots \\ a_{N1}[n] & a_{N2}[n] & \cdots & a_{NM}[n] \end{bmatrix} \quad (4)$$

Among them, $\boldsymbol{a_{ij}}[\boldsymbol{n}]$ represents the channel coefficient of the jth antenna at the sending end and the ith antenna at the receiving end on the nth path, which can be calculated by the following expression

$$\boldsymbol{a_{ij}}=\sqrt{\boldsymbol{P_j}}\ \mathrm{B}\boldsymbol{v_{ij}} \quad \textbf{(5)}$$

Among them, $\boldsymbol{P_j}$ is the average power of the signal on the transmitter $\boldsymbol{x_j(t)}$; **B** is a composite matrix of **i** * **j**; $\boldsymbol{v_{ij}}$ is the weight of the channel from the jth antenna at the sending end to the ith antenna at the receiving end. It is a complex Gaussian random variable, where**E** $\{|\boldsymbol{v_{ij}}|^{\boldsymbol{2}}\} = \boldsymbol{1}$, and the mean value is 0. We set the total number of MIMO channel paths in terahertz communication to **L** and the delay to $\boldsymbol{\sigma}$. In summary, the signal relationship between the discretized receiver and transmitter can be expressed as

$$\mathbf{y(t)}=\sum_{n=1}^{L} \boldsymbol{A_n}\, \boldsymbol{x(t-\sigma)}+ \mathbf{w(t)} \quad \textbf{(6)}$$

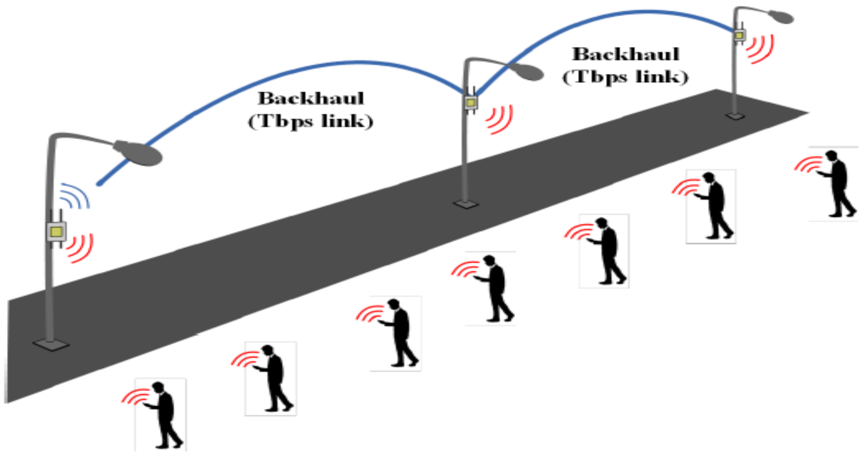
II.4.2 Geometrical Model of Thz channel :

The system model for the THz massive MIMO network is shown in **FigureII.3**. For the walkway scenario, we focus on the downlink where the evenly-spaced APs provide connectivity to the pedestrian users. We assume the APs are connected by Tbps wireless backhaul links and that each user is connected a single AP at each time instant. In the following subsections, we describe the channel model, the antenna array structure and the precoding techniques employed. We consider a three dimensional (3D) statistical spatial channel model following the implementation in [67] for the considered scenario. The large-scale fading is given by the effective omnidirectional path loss PL_{Leff} which combines the path loss (PL) and the shadow fading (SF).

$$\boldsymbol{P_{leff}} = \mathbf{20\ log_{10}}\left(\frac{\boldsymbol{4\pi f_c}}{\boldsymbol{c}}\right)+\mathbf{20\ \bar{a}\ \ log_{10}}(\boldsymbol{d})+\mathbf{X(0,\sigma)} \quad \textbf{(7)}$$

where f_c is the carrier frequency, c = 3*10^8 m/s is the speed of light, ā is the path loss exponent, d is the 3D distance between the AP and UE, and X is the log-normal random SF variable with zero mean and σ standard deviation. For the small-scale fading, the double-directional channel impulse response (CIR) h_{dir} with L multipath components (MPCs) for each transmission link is given as:

$$\boldsymbol{h_{dir}(t,\phi)} = \sum_{l=1}^{L} \boldsymbol{P_{RX,l}}.\ \boldsymbol{e^{j\varphi_l}}.\ \boldsymbol{\delta(t\text{-}\tau_l)}.\ \ \boldsymbol{G_{TX}(\phi-\phi_l^{TX})}.\ \boldsymbol{G_{RX}(\phi-\phi_l^{RX})} \quad \textbf{(8)}$$



FigureII.3: System deployment model

Where $P_{RX,l}$, φ_l and τ_l denote the received power magnitude, phase and propagation time delay of the MPCs, respectively; t is time and ϕ is the azimuth angle offset from the boresight direction. For each MPC, $\phi_l^{TX}$is the angle of departure (AoD) at the AP and ϕ_l^{RX} is the angle of arrival (AoA) for each UE. G_{TX} and G_{RX} are the transmit and receive antenna gains, respectively **[68]**, **[67]**. Given that G_{AE} is the gain of an antenna element, G_{TX} and G_{RX} can be given by (9) **[69]**, where X represents the TX or RX as appropriate.

$$G_X(\text{dB}) = G_{AE}\ (\text{dB}) + 10 \log_{10}(N_X) \quad \textbf{(9)}$$

II.4.2.1 Massive MIMO Antenna Model:

The AP and UEs are equipped with massive MIMO arrays with N_{TX} and N_{RX} antenna elements, respectively. We consider the uniform linear array (ULA) at both the AP and the UEs, with inter-element spacing $d_{TX} = d_{RX} = \lambda/2$. The $h_{dir}(t, \phi)$ in (8) is thus extended to the MIMO channel matrix H given by (4).

$$\mathbf{H}=\sqrt{\frac{N_{TX}\ N_{RX}}{L}} \sum_{l=1}^{L} \sqrt{PL_l(d)}\ .e^{j2\pi}\left(f_c\tau_l + \frac{\upsilon_{RX}\cos(\varphi_l)}{\lambda}\ \Delta t\right).a_{RX}\left(\phi_l^{RX}\right).\ a_{TX}^H\left(\phi_l^{TX}\right) \quad \textbf{(10)}$$

where υ_{RX} is the UE speed, λ= c/f_c is the wavelength, and φ_l is the phase, where φ_l is composed of the distancedependent phase change and the velocity-induced Doppler shift. The array response vectors a TX and a RX are given by (11) and (12), for the ULA TX and RX, respectively [70].

$$a_{TX}\left(\phi_l^{TX}\right) = \frac{1}{\sqrt{N_{TX}}}\left(e^{j\frac{2\pi}{\lambda}} d_{TX}(n-1)\sin(\phi_l^{TX})\right) \quad \textbf{(11)}$$

∀ n = 1 , 2, ….. , N_{TX}

$$a_{RX}\left(\phi_l^{RX}\right) = \frac{1}{\sqrt{N_{RX}}}\left(e^{j\frac{2\pi}{\lambda}} d_{RX}(m-1)\sin(\phi_l^{RX})\right) \quad \textbf{(12)}$$

∀m = 1, 2, …. , N_{RX}

II.5 Models of Molecular Thz channel :

After nano-device design and manufacturing, the terahertz channel is one of the main aspects that makes the realization of wireless nano networks a challenge. This still unlicensed band spans the frequencies between 100 GHz and 10 THz. Despite the major limitations of this band for short and medium range communications [71], the terahertz band offers new opportunities for communication in the nano scale.

In this section, the main characteristics of the terahertz channel in terms of total path-loss and noise are reviewed. For this, a novel propagation model based on radiative transfer theory [72], and that makes intensive use of the HITRAN (High resolution Transmission molecular absorption database) line catalog [73] is used. Despite not originally thought for the terahertz band or the nanoscale, the information contained in the HITRAN database is found to be a useful asset for the computation of the attenuation and noise in our frequency and transmission ranges of interest.

II.5.1 *Path-loss*

The total path-loss, **A**, for a traveling wave in the terahertz band is defined as the addition of the spreading loss, $\mathbf{A_{spread}}$, and the molecular absorption loss, $\mathbf{A_{abs}}$:

$$\mathbf{A(f,d)}=A_{spread}(\mathbf{f},\mathbf{d})+A_{abs}(\mathbf{f},\mathbf{d}) \qquad \mathbf{(13)}$$

where **f** stands for the wave frequency and **d** is the total path length. The spreading loss accounts for the attenuation due to the expansion of the wave as it propagates through the medium, and it is defined as:

$$\mathbf{A_{spread}}(\mathbf{f},\mathbf{d})=\left(\frac{\mathbf{4\pi fd}}{\mathbf{c}}\right)^{\mathbf{2}} \qquad \mathbf{(14)}$$

where **f** is the wave frequency, **d** is the total path length, and **c** stands for the speed of light in vacuum.

The absorption loss accounts for the attenuation that a propagating wave will suffer because of molecular absorption, i.e., the process by which part of the wave energy is converted into internal kinetic energy of the excited molecules in the medium. Indeed, several molecules present in a standard medium are excited by electromagnetic radiation at certain frequencies within the terahertz band, converting part of the radiation into internal vibrations. The absorption loss $\mathbf{A_{abs}}$ reflects this reduction in the wave energy, and it is defined as:

$$\mathbf{A_{abs}(f,d)} = \frac{\mathbf{1}}{\boldsymbol{\tau}\mathbf{(f,d)}} \qquad \textbf{(15)}$$

where **f** stands for the wave frequency, **d** is the total path length, and $\boldsymbol{\tau}$ is the transmittance of the medium. This parameter measures the fraction of incident radiation that is able to pass through the medium and can be calculated using the Beer-Lambert Law as [72]:

$$\boldsymbol{\tau}\ \mathbf{(f,d)=e^{-k(f)d}} \qquad \textbf{(16)}$$

where **f** is the frequency of the EM wave, **d** stands for the total path length, and **k** is the medium absorption coefficient. This last parameter depends on the composition of the medium, i.e., the particular mixture of molecules found along the path, and it is defined as:

$$\mathbf{k(f)=\sum_{\boldsymbol{i,g}} k^{\boldsymbol{i,g}}(f)} \qquad \textbf{(17)}$$

where $\mathbf{k^{i,g}}$ is the individual absorption coefficient for the isotopologue $\boldsymbol{i}$ of gas $\boldsymbol{g}$.
The absorption coefficient of an isotopologue $\boldsymbol{i}$ of gas $\boldsymbol{g}$, , $\mathbf{k}^{\boldsymbol{i,g}}$ in m^{-1}, for a molecular volumetric density, $\boldsymbol{Q^{i,g}}$, in molecules/m^3 at pressure **p** and temperature **T** can be written as:

$$\mathbf{k}^{\boldsymbol{i,g}}\mathbf{(f)} = \frac{\mathbf{p}}{\mathbf{p_0}}\frac{\mathbf{T_{STP}}}{\mathbf{T}}\boldsymbol{Q^{i,g}}\ \boldsymbol{\sigma^{i,g}}\mathbf{(f)} \qquad \textbf{(18)}$$

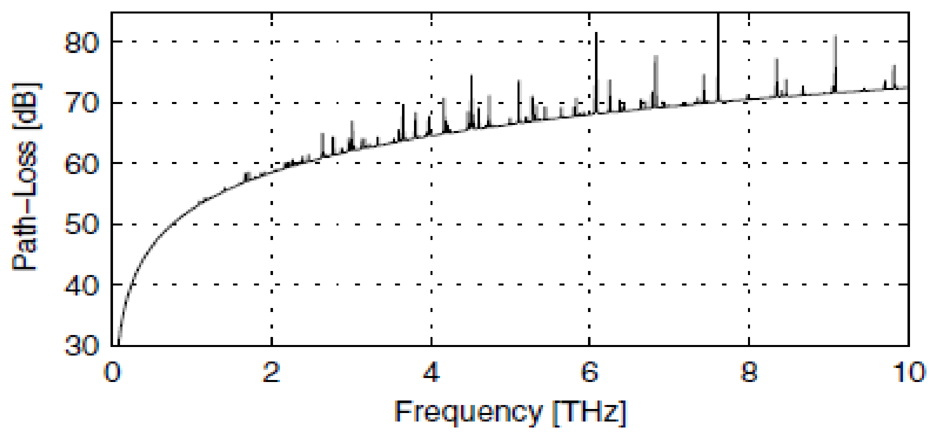


Figure II.4: Total path-loss as a function of frequency for a transmission distance equal to 10 mm, in a standard medium with 1% of water vapor molecules

where $\mathbf{p_0}$and $\mathbf{T_{STP}}$ are the Standard-Pressure-Temperature values and $\boldsymbol{\sigma^{i,g}}$ is the absorption cross section for the isotopologue $\boldsymbol{i}$ of gas $\boldsymbol{g}$ in m^2/molecule. Simply stated, the total absorption will depend on the number of molecules of a given gas that are found along the path. For a given gas mixture, the total number of molecules of the isotopologue $\boldsymbol{i}$ of gas $\boldsymbol{g}$ per volume unit, $\boldsymbol{Q^{i,g}}$, at pressure **p** and temperature **T**, can be obtained from the Ideal Gas Law [74] as:

$$\mathbf{Q}^{i,g} = \frac{\mathbf{n}}{\mathbf{v}}\mathbf{q}^{i,g}\mathbf{N_A} = \frac{\mathbf{p}}{\mathbf{RT}}\mathbf{q}^{i,g}\mathbf{N_A} \qquad \textbf{(19)}$$

where **n** is the total number of moles of the gas mixture that is being considered, **V** stands for the volume, $\mathbf{q}^{i,g}$ is the mixing ratio for the isotopologue $\boldsymbol{i}$ of gas $\boldsymbol{g}$,$\boldsymbol{N_A}$ stands for the Avogadro constant and **R** is the gas constant.

The absorption cross section $\boldsymbol{\sigma^{i,g}}$ in (18) can be further decomposed in terms of the line intensity $\boldsymbol{S^{i,g}}$for the absorption of the isotopologue $\boldsymbol{i}$ of gas $\boldsymbol{g}$ and the spectral line shape $\boldsymbol{G^{i,g}}$as:

$$\boldsymbol{\sigma^{i,g}}(\mathbf{f})=\boldsymbol{S^{i,g}}\boldsymbol{G^{i,g}}(\mathbf{f}) \qquad \textbf{(20)}$$

The line intensity $\boldsymbol{S^{i,g}}$ is a parameter directly obtained from the HITRAN database [73]. To obtain the line shape, $\boldsymbol{G^{i,g}}$, we first need to determine the position of the resonant frequency $\boldsymbol{f_c^{i,g}}$ for an isotopologue $\boldsymbol{i}$ of gas $\boldsymbol{g}$. This increases linearly with the pressure from its zero-pressure position as:

$$\boldsymbol{f_c^{i,g}} = \boldsymbol{f_{c_o}^{i,g}}+\boldsymbol{\delta^{i,g}}\,\mathbf{p}/\mathbf{p_0} \qquad \textbf{(21)}$$

where is $\boldsymbol{f_{c_o}^{i,g}}$ the zero-pressure position of the resonance, $\mathbf{p_0}$ is the reference pressure and $\boldsymbol{\delta^{i,g}}$ is the linear pressure shift. All these parameters are directly read from the HITRAN database [73].

The absorption from a particular molecule is not confined to a single frequency, but is spread over a range of frequencies. For a system in which the pressure is above 0.1 atm, the spreading will be mainly governed by the collisions between molecules of the same gas. The amount of broadening depends on the molecules involved in the collisions and it is usually referred as the Lorentz half-width $\boldsymbol{\alpha_L^{i,g}}$[72]. This can be obtained as a function of the air and self-broadened half widths, $\boldsymbol{\alpha_0^{air}}$ and $\boldsymbol{\alpha_0^{i,g}}$ respectively, as:

$$\boldsymbol{\alpha_L^{i,g}} = \left[(1-\mathrm{q^g})\,\alpha_0^{\mathrm{air}} + \mathrm{q_g}\alpha_0^{\mathrm{i,g}}\right]\left(\frac{\mathrm{p}}{\mathrm{p_0}}\right)\left(\frac{T_0}{T}\right)^{\gamma} \qquad \textbf{(22)}$$

in which the temperature broadening coefficient γ as well as $\boldsymbol{\alpha_0^{air}}$ and $\boldsymbol{\alpha_0^{i,g}}$ are obtained directly from the HITRAN database [73].

For the terahertz frequency band, an appropriate line shape to represent the molecular absorption is the Van Vleck- Weisskopf asymmetric line shape [75] :

$$\boldsymbol{F^{i,g}(f)} = \frac{\alpha_L^{i,g}}{\pi}\frac{f}{f_c^{i,g}}.\left[\frac{1}{\left(f-f_c^{i,g}\right)^2+\left(\alpha_0^{i,g}\right)^2}+\frac{1}{\left(f+f_c^{i,g}\right)^2+\left(\alpha_0^{i,g}\right)^2}\right] \qquad \textbf{(23)}$$

An additional adjustment to the far ends of the line shape can be done in order to account for the continuum absorption [76]:

$$\boldsymbol{G^{i,g}(f)} = \frac{\mathrm{f}}{f_c^{i,g}}\frac{\tanh\left(\frac{hc\mathrm{f}}{2K_BT}\right)}{\tanh\left(\frac{hcf_c^{i,g}}{2K_BT}\right)}.\boldsymbol{F^{i,g}(f)} \qquad \textbf{(24)}$$

where $\boldsymbol{h}$ is the Planck Constant, $\boldsymbol{c}$ is the speed of light in vacuum, $\boldsymbol{K_B}$ stands for the Boltzmann Constant and T is the system temperature.

With this, we are able to compute the contributions to the total molecular absorption from each isotopologue i of each gas g present in the medium, and consequently, the total molecular absorption loss (15) can be obtained. Molecular absorption makes the terahertz channel highly frequency selective. As an example, the total path-loss that a wave will suffer when traveling 10 mm in a standard medium with 1% of water vapor molecules, is shown in FigureII.4.

II.5.2 Noise

The ambient noise in the terahertz channel is mainly contributed by the molecular noise. The absorption from molecules present in the medium does not only attenuate the transmitted signal, but it also introduces noise. The parameter that measures this phenomenon is the emissivity of the channel, ε, and it is defined as :

$$\boldsymbol{\varepsilon(f, d) = 1 - \tau(f, d)} \qquad \textbf{(25)}$$

where $\boldsymbol{f}$ is the signal frequency, $\boldsymbol{d}$ stands for the total path length and $\boldsymbol{\tau}$ is the transmissivity of the medium (16).

The equivalent noise temperature due to molecular absorption T_{mol} in Kelvin that an omnidirectional antenna will detect from the medium is further obtained as:

$$T_{mol}(f, d) = T_0 \varepsilon (f, d) \quad \textbf{(26)}$$

where T_0 is the reference temperature. This type of noise will only be present around the frequencies in which the molecular absorption is considerably high. As an example, the molecular noise that a propagating wave will create when traveling 10 mm in a standard medium with 1% of water vapor molecules is shown in FigII.5.

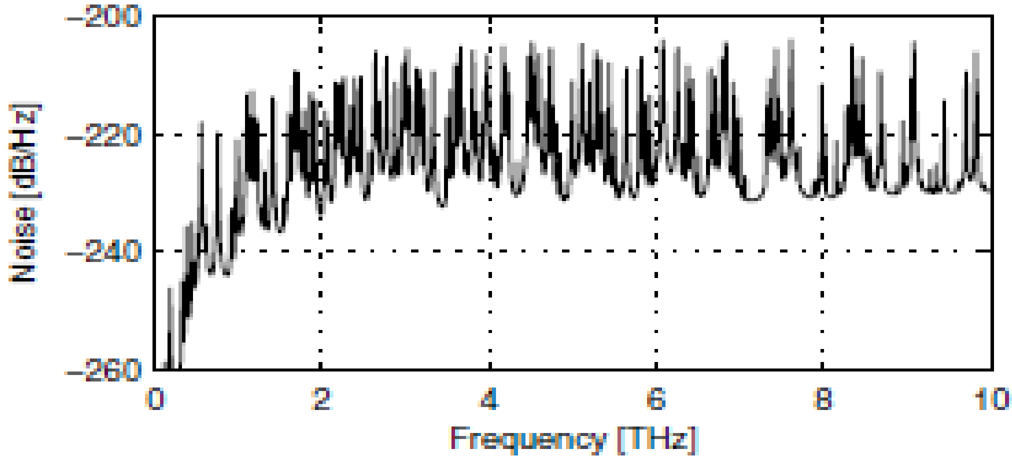


Figure II.5 : Molecular noise as a function of frequency, for a transmission distance equal to 10 mm, in a standard medium with 1% of water vapor molecules

To compute the equivalent noise power at the receiver, it is necessary to define the transmission bandwidth, which will on its turn depend on the transmission distance and the composition of the medium. For a given bandwidth, the total noise power p_n can be calculated as:

$$p_n(f, d) = K_B B \left(T_{mol}(f, d) + T_{other}(f)\right) \quad \textbf{(27)}$$

where f is the wave frequency, d is the total path length, K_B is the Boltzmann constant, B stands for the system bandwidth, T_{mol} refers to the molecular noise temperature and T_{other} is an additional term accounting for any other noise source present in the medium, e.g., electronic noise of the receiver.

II.6 CHANNEL CAPACITY:

The capacity, C given by (29), of the Terahertz channel is determined by the channel path loss, A given by (13), the noise p.s.d., N from (28), and the p.s.d. of the transmitted EM wave. In an intent to keep our numerical results realistic, and in light of the state of the art in molecular-electronics, the total signal energy is kept constant and equal to 500 pJ independently of the specific power spectral distribution. This number is chosen in light of the state of the art in novel energy harvesting systems for nano-devices, such as piezoelectric nano-generators [77], [78], [79]. These mechanisms are able to harvest energy from ambient vibrations, which is then stored in a nano-battery or a nano-ultra-capacitor. These mechanisms have been successfully used to power a laser diode [78] and a PH nanosensor [79]. The maximum amount of energy harvested by these mechanisms depends on several factors, such as the nano-battery size of the frequency of the harvested vibration, but values in the orders of hundreds of nJ have been illustrated. Because of this, we consider that the chosen energy is a reasonable value. However, note that the capacity results presented in this section strongly depend on this value.

To compute the equivalent noise power at the receiver, it is necessary to define the transmission bandwidth, which depends on its turn on the transmission distance and the composition of the medium. For a given bandwidth, B, the molecular absorption noise power at the receiver can be calculated as :

$$P_n(f, d) = \int_B N(f, d)\, df = k_B \int_B T_{noise}(f, d)\, df \tag{28}$$

where f stands for frequency, d is the transmission distance, N refers to the noise power spectral density (p.s.d.), k_B is the Boltzmann constant and T_{noise} refers to the equivalent noise temperature. Up to this point, we have given formulations to compute the total path loss and the molecular absorption noise power. In the following section, we investigate the capacity of the Terahertz channel by using the developed model.
Based on the proposed channel model, the Terahertz channel is highly frequency-selective and, in addition, the molecular absorption noise is non-white. Thus, the capacity can be obtained by dividing the total bandwidth into many narrow sub-bands and summing the individual capacities [80]. The i- th sub-band is centered around frequency f_i, i = 1, 2, ... and it has width Δf. If the sub-band width is small enough, the channel appears as frequency-nonselective and the noise p.s.d. can be considered locally flat. The resulting capacity in bits/s is then given by

$$\mathrm{C(d)} = \sum_i \Delta f \log_2 \left[1 + \frac{S(f_i)A(f_i,d)^{-1}}{N(f_i,d)}\right] \tag{29}$$

where d is the total path length, S is the transmitted signal p.s.d., A is the channel path loss and $\bar{\mathrm{U}}$ is the noise p.s.d..

II.7 Channel Model in Diffusion-Based Molecular Communications:

In molecular communications, the propagation process of the desired signal is mainly governed by the transport mechanism of the emitted messenger molecules. In diffusion-based molecular communication channels (i.e., where the propagation environment is stationary), the propagation process of emitted messenger molecules is accomplished by means of a thermally activated diffusion mechanism. In this transport mechanism, molecular flux is achieved from regions of high to low concentration via random collisions with the underlying medium. Hence, the diffusive molecular flux is proportional to the concentration gradient. Figure II.6 illustrates a diffusion-based molecular communications system.

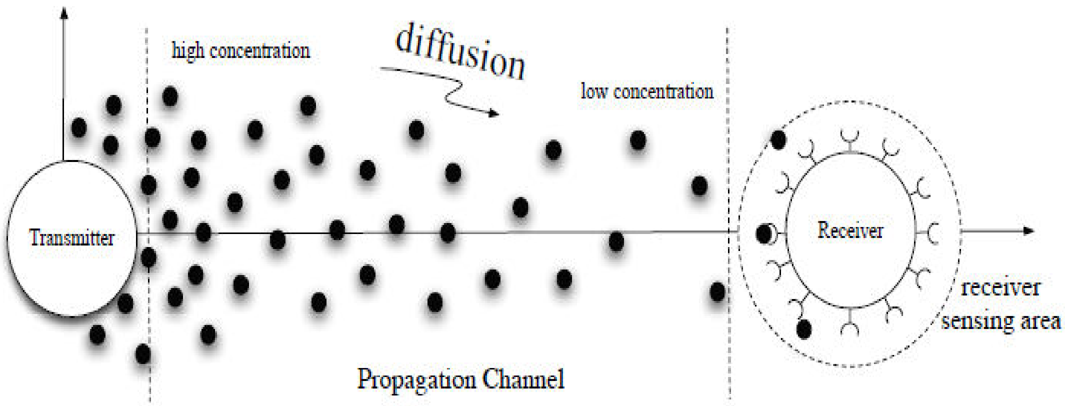


Figure II.6: Diffusion-based molecular exchange system

According to [81, 82], diffusion-based molecular propagation can be categorized into two groups:

- Normal diffusion, in which the interactions among the emitted particles (i.e., collisions and electrostatic forces) can be neglected and viscous forces of the fluid environment dominate the diffusion of the emitted molecules. Hence, the movement of each emitted particle is independent of one another. The diffusion process in this scenario is modelled by Fick's diffusion laws and Brownian motion [83, 82, 84].
- Anomalous diffusion, in which interactions among the emitted particles affect their diffusion process, such as collective diffusion [85]. In this diffusion mechanism, the

concentration of emitted particles is very high such that collisions among emitted particles affect their movement. The diffusion process in this scenario is modelled by a correlated random walk [83, 82].

Since in molecular communication channels the concentration of emitted molecules is much lower than the concentration of the medium molecules, the collisions and electrostatic forces among the emitted molecules are neglected [86, 87, 88 , 89]. Hence, the trajectories of different molecules are modelled as being independent of one another, and each released molecule moves through the fluid medium in a Brownian fashion.

In the following, using Brownian motion and Fick's diffusion laws, a signal propagation model is derived for the diffusion-based molecular channel in Subsection **II.7.1**

II.7.1 Signal Propagation Modelling :

In physics, the normal diffusion process is usually modelled by Fick's diffusion laws and a Wiener process [84, 90]. In mathematics, the Wiener process is a continuous-time stochastic process with independent Gaussian increments. After Robert Brown, who studied the motion of minute diffusing particles suspended in fluid for the first time, the Wiener process is also called Brownian motion.

According to Brownian motion, the displacement of each molecule during an infinitesimal interval of dt seconds is modelled with a zero-mean Gaussian distribution of variance 2Ddt [91, 92]:

$$\mathbf{X(t + dt) = X(t) + \sqrt{2Ddt}\ u(t)} \qquad \mathbf{(30)}$$

where D is the diffusion coefficient and X(t) = (x(t); y(t); z(t)) is the molecule position at time t. In the following, the notation of molecule position X(t) is simplified by omitting t.

The term u(t) is a vector of random variables that has a multivariate Normal distribution u(t) ~N(0_{1*3} ,I_{3*3}) where $0_{1*3}$and I_{3*3} are the null and identity matrices. The diffusion coefficient D in m^2/s is defined as follows [91]:

$$\mathbf{D=\frac{k_B\, T}{6\pi\, \eta\, r_m}} \qquad \mathbf{(31)}$$

where $k_B$is the Boltzmann constant, T is the temperature, η is the viscosity of the fluid medium, and $r_m$is the radius of the diffusing molecule.

Although the diffusion of messenger molecules is a stochastic process, the mean change in molecular concentration with time can be formalized by Fick's second law of diffusion [93]:

$$\frac{\partial \bar{U}(X,t)}{\partial t} = D\nabla^2 U(X,t) \qquad \mathbf{(32)}$$

Where Ū(X; t) is the mean molecular concentration in molecule/m^3 and t is the time passed after the transmission. By solving (32) when an impulse of Q messenger molecules (i.e., Qδ(t)) is transmitted at time zero, the mean molecular concentration Ū(X; t) can be obtained as follows:

$$\bar{U}(X,t)=\frac{Q}{(4\pi Dt)^{3/2}}\exp\left(\frac{-|X-X_{tx}|^2}{4Dt}\right) \quad (33)$$

where $X_{tx} = (x_{tx};\ y_{tx};z_{tx})$ is the location of the transmitting nano-machine. The mean concentration of diffused molecules is a function of both the location X and the time passed after.

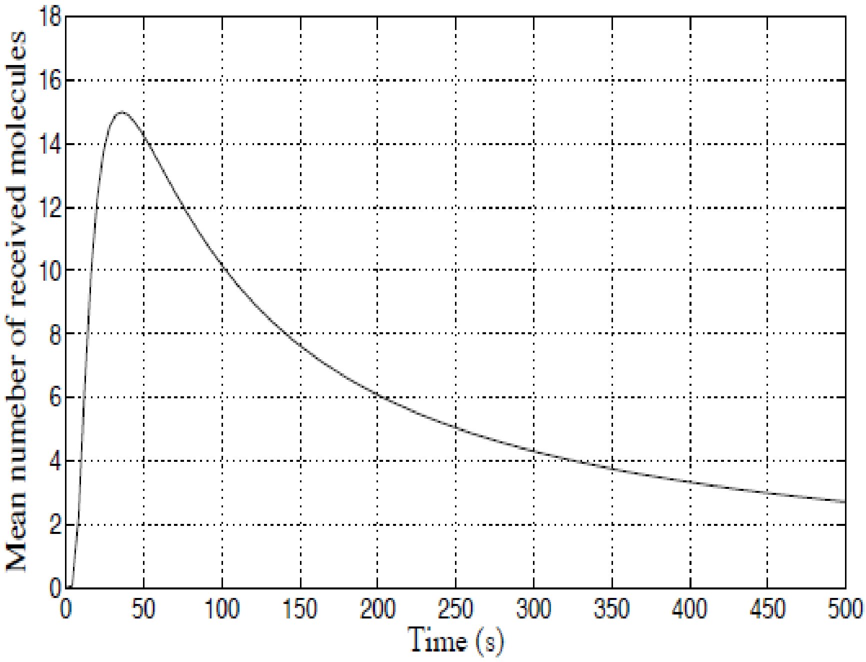


Figure II.7: The mean number of messenger molecules in the receiver sensing area, s(t), in terms of time

Fig II.7 shows the mean number of messenger molecules in the receiver sensing area in terms of time. For this figure, parameter assumptions are considered to be reasonable for the real-time brain activity mapping application which, is considered as the sample application scenario for all simulations in this dissertation. It is assumed that the transmitter and receiver nano-machines are located at (0; 0) µm and (500; 0) µm, respectively. The messenger molecules are considered to be calcium ions, Ca+2 [94, 95, 96] whose radius is 114 pm. The transmitter releases Q = 2*10ˆ4 molecules into the propagation medium which is considered to be human blood. This amount of released messenger molecules per symbol is non-toxic for biomedical applications since after communicating even 10ˆ5 symbols, the total volume of emitted messenger molecules is smaller than the volume of a human red blood cell (1.12*10ˆ2μm^3). The viscosity and temperature of human blood are η = 1.12*10ˆ-3 kg/s.m and T = 300 K [91]. As shown, the initial number of received molecules in the receiver sensing area is zero, but it experiences a sharp increase towards a maximum at 36 s. Afterwards, the number of received messenger molecules gradually degrades resulting in a long tail due to the effect of diffusion.

II.8 TERAHERTZ CHANNEL MODELING INSIDE THE SKIN :

A modified Friis equation has been proposed by Jornet et al. in [97] to calculate the path loss of the THz channel in water vapor, which can be divided into two parts: the spreading path loss, PL_{spr}which is due to the expansion of waves inside the tissue and the absorption path loss PL_{abs} (due to absorption of waves in tissue). Similarly, the path loss in human tissues can also be divided into two parts:

$$PL_{total}[dB] = PL_{spr}(f; d)[dB] + PL_{abs}(f; d)[dB] \quad (34)$$

$$=20 \log\frac{4\pi d}{\lambda_g} + 10\alpha d \log e \quad (35)$$

where, *f* stands for the frequency, *d* is the path length, $\alpha = 4\pi K/\lambda$, where *K* is extinction coefficient, which measures the amount of absorption loss. In this study, the electromagnetic power is considered to spread spherically with distance. This path loss model takes into account only distance and frequency and it doesn't explicitly discuss about inside the skin. This paper presents a novel model inside the human skin by considering additional parameters, which also plays a vital role in varying channel characteristics.

II.9 Conclusion:

In this chapter, we presented a comprehensive overview and analysis of studies of THz wireless and molecular channels, including channel measurement, channel modeling and channel characterization, and discuss open problems for future research directions concerned with THz wireless channels. After that, we gave a review of THz channel models, channel MIMO and Massive MIMO Antenna Model, We also discussed the channels models in the nanocommunication system , the main characteristics of the terahertz channel in terms of total path-loss and noise are reviewed , And we talked about Channel Capacity , Channel Model in Diffusion-Based Molecular Communications and Terahertz channel modeling inside the skin. This chapter provided a solid understanding of channel propagation at THz, setting a good knowledge base for future work to be conducted to realize various applications targeted at this band for operation in 6G networks. Further studies are also needed for preparing THz band standardization.

In the third chapter we will discuss about the classification method of digital modulation (nano communication).

Chapter III

classification method of digital modulation

III.1 INTRODUCTION

Classification is an area of machine learning that takes raw data and classifies it as belonging to a particular class based on the required parameter set. Bayesian theory is basically woks as a framework for making decision under uncertainty- a probabilistic approach to inference [98, 99]. Bayes theorized that the probability of future events could be calculated by determining their earlier frequency. The appeal of the Bayesian technique is its deceptive simplicity. The predictions are based completely on data culled from reality- the more data obtained, the better it works. Another advantage is that Bayesian models are self-correcting, meaning that when data changes, so do the result. One highly practical Bayesian learning method is the Naive Bayes Classifier. It is based on the so called Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Despite its, Naive Bayes can often outperform more sophisticated classification methods [100, 101]. In some domains its performance has been shown to be comparable to that of neural network and decision tree learning. In machine learning literature, nonparametric methods are called instance-based or memory based learning algorithms, since what they do is store the training instances in a lookup table and interpolate from these. The most basic instance-based method is the K-Nearest Neighbor algorithm.

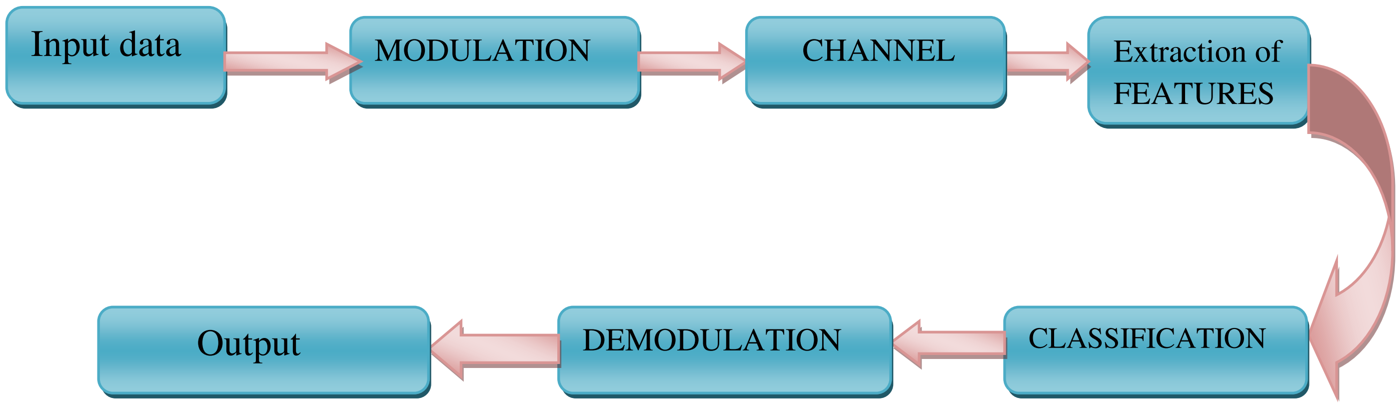


Figure III.1: classification steps in digital modulation

III.2 K-NEAREST NEIGHBORS:

III.2.1 Definition:

Classification is an important problem in big data, data science and machine learning. The K-nearest neighbor (KNN) is one of the oldest, simplest and accurate algorithms for patterns classification and regression models. KNN was proposed in 1951 by [102], and then modified by [103]. KNN has been identified as one of the top ten methods in data mining [104]. Consequently, KNN has been studied over the past few decades and widely applied in many fields [105]. Thus, KNN comprises the baseline classifier in many pattern classification problems such as pattern recognition [106], text categorization [107], ranking models [108], object recognition [109], and event recognition [110] applications. KNN is a non-parametric algorithm [111]. Non-Parametric means either there are no parameters or fixed number of parameters irrespective of size of data. Instead, parameters would be determined by the size of the training dataset. While there are no assumptions that need to be made to the underlying data distribution. Thus, KNN could be the best choice for any classification study that involves a little or no prior knowledge about the distribution of the data. In addition, KNN is one of the laziest learning methods. This implies storing all training data and waits until having the test data produced, without having to create a learning model [112].

Despite its slow characteristic, surprisingly, KNN is used extensively for Big data classification, this includes the works of, this is due to the other good characteristics of the KNN such as simplicity and reasonable accuracy, since the speed issue is normally solved using some kind of divided-and-conquer approaches. Slowness is not the only problem associated with the KNN classifier, in addition to choosing the best K-neighbors problem, choosing the best distance/similarity measure is an important problem, this is because the performance of the KNN classifier is dependent on the distance/similarity measure used.

III.2.2 Algorithm of KNN:

A KNN classifier calculates the distances between the tested sample and all the labeled training samples; then the identity of the tested sample is determined by the class of its majority neighbors.

The KNN algorithm is one of the most famous classification algorithms used for predicting the class of a record or (sample) with unspecified class based on the class of its neighbor records.

The algorithm is made of three steps:

1. Calculating the distance of input record from all training records.
2. Arranging training records based on the distance and selection of K-nearest neighbor.
3. Adopting the class which owns the majority among the k-nearest neighbors (this method considers the class as the class of input record which is observed more than all the other classes among the K-nearest neighbors).

Figure III.2 shows a KNN example, contains training samples with two classes, the first class is `blue square' and the second class is `red triangle'. The test sample is represented in green circle

These samples are placed into two dimensional feature spaces with one dimension for each feature. To classify the test sample that belongs to class `blue square' or to class `red triangle'; KNN adopts a distance function to find the K nearest neighbors to the test sample. Finding the majority of classes among the k nearest neighbors predicts the class of the test sample. In this case, when k = 3 the test sample is classified to the first class `red triangle' because there are two red triangles and only one blue square inside the inner circle, but when k = 5 it is classified to the "blue square" class because there are 2 red triangles and only 3 blue squares.

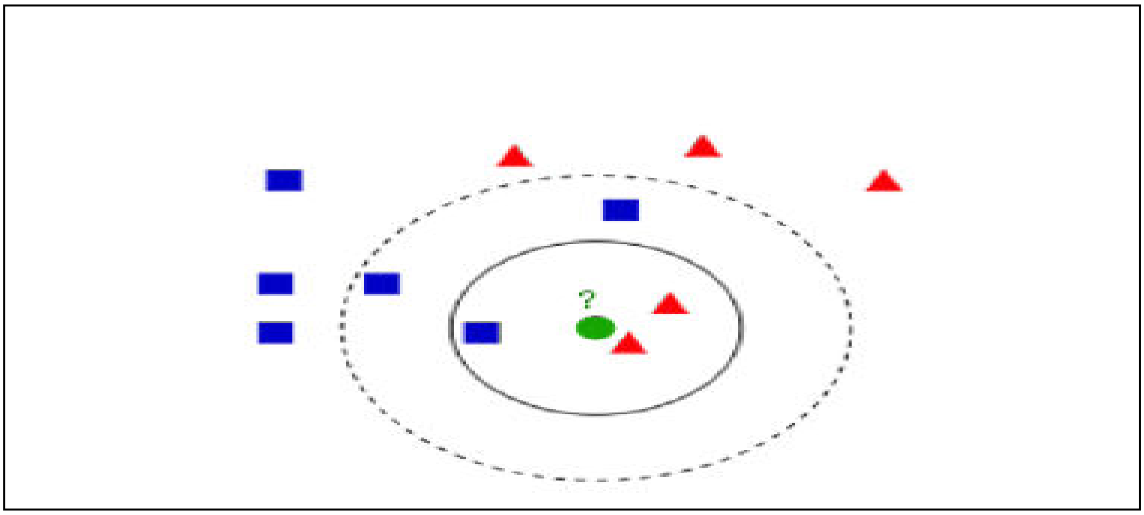


Figure III.2: An example of KNN classification with K neighbors K = 3 (solid line circle) and K = 5 (dashed line circle), distance measure is Euclidean distance

Different parameters may impact the performance of a KNN classifier such as:

A) Number of neighbors (K): the number of neighbors used by a KNN classifier to determine the class of an unlabeled sample.
The selection of an appropriate value of K is an important parameter to obtain high classification accuracy and maintain system simplicity. However, the best value of K differs from one application to another, and its selection depends on the problem at hand.
B) Distance measures: different types of distances can be used as a measure when using a KNN classifier.

III.2.3 Principle of KNN algorithm :

KNN algorithm is widely applied in pattern recognition and data mining for classification, which is famous for its simplicity and low error rate. The principle of the algorithm is that, if majority of the k most similar samples to a query point q_i in the feature space belong to a certain category, then a verdict can be made that the query point q_i fall in this category. Similarity can be measured by the distance in the feature space, so this algorithm is called K Nearest Neighbor algorithm. A train data set with accurate classification labels should be known at the beginning of the algorithm. Then for a query data q_i, whose label is not known and which is presented by a vector in the feature space, calculate the distances between it and every point in the train data set. After sorting the results of distances calculation, decision of the class label of the test point q_i can be made according to the label of the k nearest points in the train data set.

Each point in d-dimensional space can be expressed as a d-vector of coordinates, such as:

$$P=(p_1, p_2, \dots, p_n) \quad (36)$$

The distance between two points in the multi-dimensional feature space can be defined in many ways. Using Euclidean distance is usually to be the most ordinary method, that is:

$$\text{dist}(p, q)= \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \quad (37)$$

Alternatively, Manhattan distance can also be used as :

$$\text{dist}`(p, q)=\sum_{i=1}^{n}|p_i - q_i| \quad (38)$$

KNN algorithm is simple and easy to implement and it is versatile. It can be used for classification, regression, but KNN's main disadvantage of becoming significantly slower as the volume of data increases makes it an impractical choice in environments where predictions need to be made rapidly.

III.3 SUPPORT VECTORS MACHINES (SVM) :

III.3.1 Definition:

SVM are based on statistical learning theory and have the aim of determining the location of decision boundaries that produce the optimal separation of classes (Vapnik 1995). In the case of a two-class pattern recognition problem in which the classes are linearly separable the SVM selects from among the infinite number of linear decision boundaries the one that minimizes the generalization error. [113]Thus, the selected decision boundary will be one that leaves the greatest margin between the two classes, where margin is defined as the sum of the distances to the hyperplane from the closest points of the two classes (Vapnik 1995). This problem of maximizing the margin can be solved using standard quadratic programming (QP) optimization techniques. The data points that are closest to the hyperplane are used to measure the margin;

hence these data points are termed 'support vectors'. Consequently, the number of support vectors is small (Vapnik 1995).

III.3.1.1 Linear SVM:

The simplest type of classifier is a linear classifier. However, SVMs treat linear classification problems somewhat differently to the other common methods such as Linear Discriminant Analysis (LDA)[114–115] and Partial Least Squares Discriminant Analysis (PLS-DA).[116–117] Whereas for two linearly separable classes the method is probably overkill, if first understood this is an initial conceptual building block for understanding SVMs and the extension to more complex problems.

Each x n is p-dimensional real vector. The scaling is important to guard against variable (attributes) with larger variance. To view this Training data , by means of the dividing (or separating) hyperplane , which takes

$$\mathbf{w} \cdot \mathbf{x} + \mathbf{b} = \mathbf{o} \qquad \textbf{(39)}$$

Where b is scalar and w is p-dimensional Vector. The vector w points perpendicular to the separating hyperplane. Adding the offset parameter b allows us to increase the margin. Absent of b, the hyperplane is forced to pass through the origin, restricting the solution. As with the interest in the maximum margin, we are interested in SVM and the parallel hyperplanes. Parallel hyperplanes can be described by equation

$$\mathbf{w.x + b = 1}$$

$$\mathbf{w.x + b = -1}$$

If the training data are linearly separable, we can select these hyperplanes so that there are no points between them and then try to maximize their distance. [118]By geometry, We find the distance between the hyperplane is $2 / |w|$. So we want to minimize $|w|$. To excite data points, we need to ensure that for all I either

$$\mathbf{w.\, xi - b \geq 1 \text{ or } w.\, xi - b \leq -1}$$

This can be written as

$$\mathbf{yi\,(\, w.\, xi - b) \geq 1 \,,\, 1 \leq i \leq n} \qquad \textbf{(40)}$$

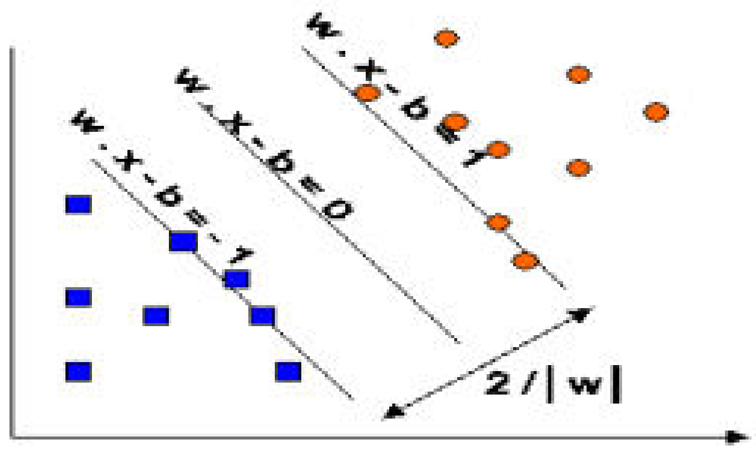


Figure III.3: Maximum margin hyperplanes for a SVM trained with samples from two classes

SVMs fall into the intersection of two research areas: kernel methods, and large margin classifiers. SVM has been applied to feature selection, time series analysis, reconstruction of a chaotic system, and non-linear principal components. Further advances in these areas are to be expected in the near future. SVMs and related methods are also being increasingly applied to real world data mining.

III.3.1.2 NON- Linear SVM:

In section **III.3.1.1** we discussed the separable case. We know that there are many cases in which the data is not linearly separable. In this case the perceptron learning algorithm simply does not terminate. However, we can set up an objective function that trades off misclassifications against minimizing ||w ||^2 to find an optimal compromise. To do this we add a “slack” variable$\varepsilon_i \geq 0$ for each training example and require that:

$$w.x_i + w_0 \geq 1 - \varepsilon_i \quad \text{for} \quad y_i = +1 \tag{41}$$

$$w.\ x_i + w_0 \leq -1 + \varepsilon_i \quad \text{for} \quad y_i = -1 \tag{42}$$

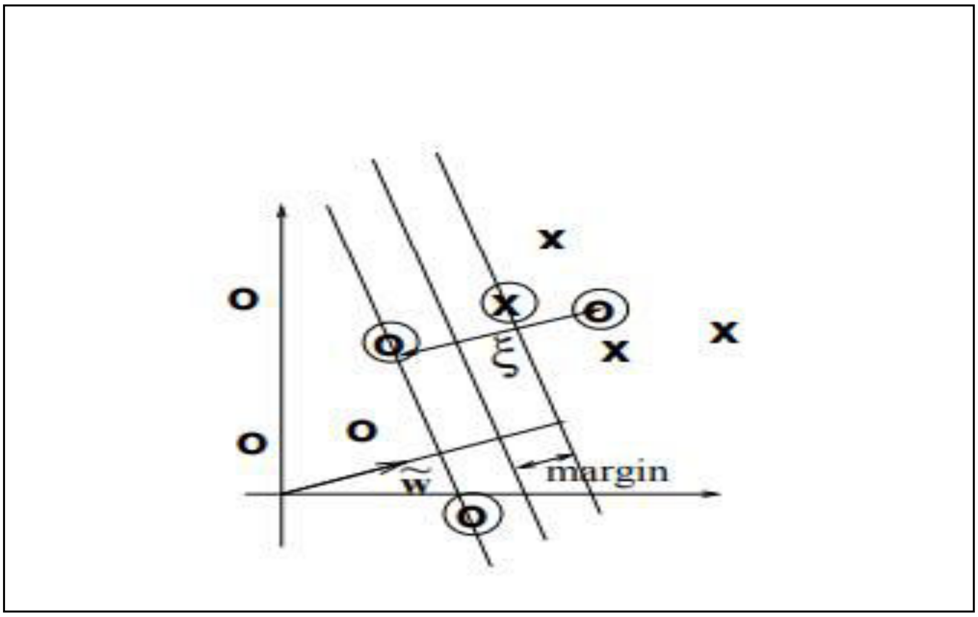


Figure III.4: A non-separable example. The data points ringed are the support vectors

In real SVM applications, classes cannot be separated linearly, so we are working with nonlinear SVM to work around this problem, That is to say, by applying a nonlinear transformation to the data to change dimension and easily find a hyperplane classification in this new space, and also to give the classifier more freedom to correctly classify the points even if they are initially points on the wrong side of the initial hyperplane (non separable categories).

where,
1. 'σ' is the variance. The RBF kernel function for two points $X_1$and X_2computes their similarity. This kernel can be represented as follows:

$$K(X_1, X_2) = exp\left(-\frac{\|X_1 - X_2\|^2}{2\sigma^2}\right) \quad \textbf{(43)}$$

2. $\|X_1$ - $X_2\|$ is the Euclidean distance between two points $X_1$and X_2.

- RBF is the main kernel is a popular method for non-linear modeling (svm) function because of following reasons:

1. The RBF kernel nonlinearly maps samples into a higher dimensional space unlike to linear kernel.
2. The RBF kernel has less hyper parameters than the polynomial kernel.
3. The RBF kernel has less numerical difficulties.

III.4 Multi class support vector machine models:

Multiclass models involve deciding which group a sample belongs to when there are more than two classes in a dataset.[119] Support Vector Machines were not originally formulated as a multiclass method, but there are extensions that are available when there are more than two groups in the data, e.g. case study R2 (NIR of food) where there are four classes and case study R3b (polymers where there are nine classes). In this and later sections we will define the group that is being modelled as the 'in group' and all other samples (which may arise from several classes) as the 'out group'. Two different examples of this approach are the One-vs-Rest (one-vs-all) and One-vs-One strategies (**Figure III.5**).

III.4.1 One versus one:

Given G classes, the one vs. one approach constructs **G(G _ 1)/2** two-class SVM classifiers, each classifier only separating two of the G classes. For example, we can form a model between classes A and B and ask which class each sample is assigned to, even if the origin of the samples is from outside these groups. consider a multi-class classification problem with four classes: '*A*,' '*B*,' and 'C,' '*D*.' This could be divided into six binary classification datasets as follows:

- **Binary Classification Problem 1**: A vs. B
- **Binary Classification Problem 2**: A vs. C
- **Binary Classification Problem 3**: A vs. D
- **Binary Classification Problem 4**: B vs. C
- **Binary Classification Problem 5**: B vs. D
- **Binary Classification Problem 6**: C vs. D

Each binary classification model may predict one class label and the model with the most predictions or votes is predicted by the one-vs-one strategy.

III.4.2. One versus all:

One vs. All [120] is the earliest method reported for extending two class SVMs to solve the multiclass problem and involves determining how well a sample is modelled by each class individually and choosing the class it is modelled by best. Given G classes under consideration, G binary SVM models can be constructed, samples either being considered as part of the class or outside it, so that each model consists of two groups, the first being class g and the second all other classes. The gth (g=1,.,G) SVM model is trained with all of the samples in the gth class being labelled by +1 and all other samples being labelled by-1 (note that the alternative approach of one-class SVDD) . Hence GSVM decision functions can be obtained, for each model.
For example , given a multi-class classification problem with examples for each class '*A*,' '*B*,' and '*C* '. This could be divided into three binary classification datasets as follows:

- **Binary Classification Problem 1**: A vs [B, C]
- **Binary Classification Problem 2**: B vs [A, C]
- **Binary Classification Problem 3**: C vs [A, B]

The bad side of this approach is that it requires one model to be created for each class. For example, three classes require three models. This could be an issue for large datasets (e.g. millions of rows), slow models (e.g. neural networks), or very large numbers of classes (e.g. hundreds of classes).

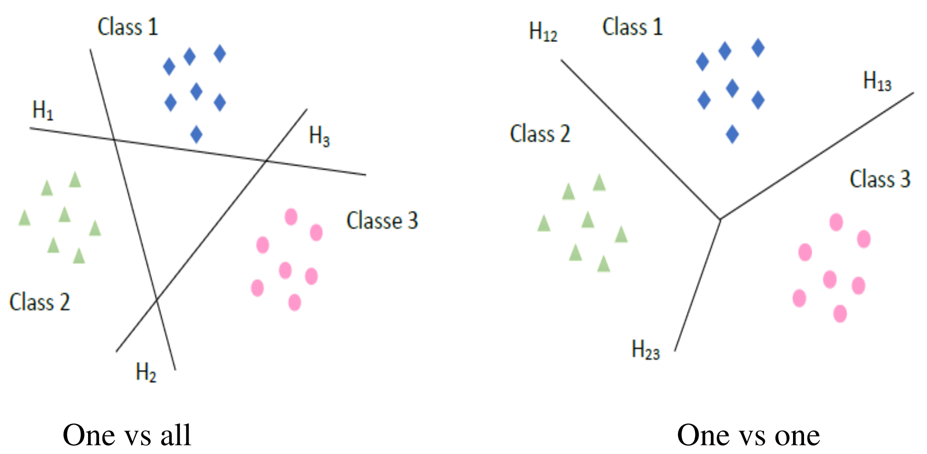


Figure III.5: Multi-class classification approaches

III.5 FEATURES EXTRACTION :

III.5.1 Higher order cumulants (HOC):

Higher order cumulants (HOC) is practically robust tool in modelling for many applications characterized by linear, non-linear systems and not affected by the noise environment. Numerous applications in the literature are based on the HOC [121–122] which use the Gaussian and non-Gaussian input signals [123–124]. HOC are encountered in various fields of system identification and signal processing such as modelling and prediction the real process.

The second order joint cumulants written as $C_{21,x}$ are defined as:

$C_{21,x} = M_{21}$

$= E(|x|^{2})$

where E ( ) is a mathematical expectation, and M is the k^{th} moments of received signal given by:

$Mpq = E[X(t)^{p-q} X^{*}(t)^{q}]$

$C_{20} = cum(r(n), r(n)) = M_{20}$ **(44)**

$C_{40} = cum(r(n), r(n), r(n), r(n))$

$C_{40} = M_{40} - 3M_{20}^{2}$ **(45)**

$C_{41} = cum(r(n), r(n), r(n), r^{*}(n)$

$C_{41} = M_{40} - 3M_{20}M_{21}$ **(46)**

$C_{42} = cum(r(n), r(n), r^{*}(n), r^{*}(n), r^{*}(n))$

$C_{42} = M_{42} - |M_{20}|^{2} - 2M_{21}$ **(47)**

$C_{60} = cum(r(n), r(n), r(n), r(n), r(n), r(n))$

$C_{60} = M_{60} - 15M_{20}M_{40} - 30M_{20}^{3}$ **(48)**

$C_{63} = cum(r(n), r(n), r(n), r^{*}(n), r^{*}(n), r^{*}(n))$

$C_{63} = M_{63} - 9M_{21}M_{42} + 21M_{21}^{3} - 3M_{20}M_{43} - 3M_{22}M_{41} + 18M_{20}M_{21}M_{22}$ **(49)**

III.5.2 Fourier transform:

Fourier Transform is a mathematical model which helps to transform the signals between two different domains, such as transforming signal from frequency domain to time domain or vice versa. Fourier transform has many applications in Engineering and Physics, such as signal processing, RADAR, and so on.[125]The Fourier transform helps to extend the Fourier series to the non-periodic functions, which helps us to view any functions in terms of the sum of simple sinusoids. the Fast Fourier Transform (FFT) of the received signal is used to identify its modulation type.

Fourier Series synthesis equation:

$$x(t) = \sum_{n=-\infty}^{+\infty} C_n e^{jnw_0 t} \quad (50)$$

The analysis equation for the Fourier Transform:

$$x(w) = \int_{-\infty}^{+\infty} \mathrm{x(t)}\, e^{-jwt}\, dt \quad (51)$$

III.5.3 Variance:

Variance is used in the field of statistics and probability as a measure used to characterize the dispersion of a distribution or a sample. [126] It can be interpreted as the dispersion of values from the mean. Concretely, the variance is defined as the mean of the squares of the deviations from the mean. Considering the square of these deviations prevents positive and negative deviations from canceling each other out. Visually, a distribution with a large variance will be more spread out, while a distribution with a small variance will be very tight around its mean.

$$VAR_x \quad = \sigma_x^2 = \mathrm{E}[(\mathrm{x} - \mu_x)^2] \quad (52)$$

Where x is random variable and μ_x is mean.

III.6 conclusion:

In this chapter, we have provided an overview of classification method of digital modulation .We have defined K-nearest neighbor (KNN) it as one of the oldest and simplest classification algorithm and an important problem in data science and machine learning. We also discussed the Principle of SVM algorithm (support Vectors Machines),it depends on the statistical learning theory, which is divided into two types: Linear SVM and NON- Linear SVM, We also discussed multi class support vector and we mentioned two different examples of this approach are the One-vs-Rest (one-vs-all) and One-vs-One strategies and features extraction Higher order cumulants (HOC), Fourier transform and Variance to know the types of modulation.

Chapter IV

Simulation and Results

IV.1 INTRODUCTION

In this chapter we will present the effect of channels (selectif channel, AWGN and TERAHERTZ channel) in classification. Based on the KNN and the SVM algorithms, this may allow the receiver to recognize the modulation types used by the transmitter, we have used 8 modulated signals (BPSK, QAM, 16PSK, 16QAM, 4FSK, 32QAM, 8PSK, 2FSK) .The features are extracted from the received signals and used in a classifier to decide upon the modulation level. The features used in this work are the HOCs, variance, and Fourier transform. Finally, we implement the classifier and calculate the accuracy.

In this chapter we present the simulation results. We use MATLAB as a main program for simulation and the NUYSIM software to simulate the Thz Channel.

IV.2 DATASET

In the following simulations 8 modulated signals are used, the data length is **N**=1024, the number of signal used is 8*100, the channel length was {**L**=1,**L**=3,**L**=5}, the number of transmitted and received antennas $\boldsymbol{Nt}$=1, $\boldsymbol{Nt}$=2 and $\boldsymbol{N}$r=1,$\boldsymbol{N}$r=2.
We show the impact on the performance of the system by using the **SNR** =[-5, 0, 5, 10]dB.

In the classification step we used 800 modulated signals. This dataset was divided into two parts the first is the learning dataset (Train dataset 75%),the second is the test base (Test dataset 25%).

IV.3 SELECTIVE CHANNEL

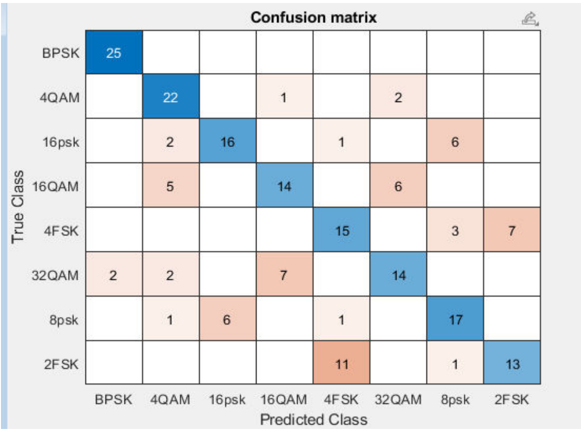
IV.3.1 Definition frequency selective channel

When we receive radio signal , the different spectral components fluctuates in unequal manner then this type of fading of signal occurs in case of frequency selective channel. Communication systems transmitting over frequency-selective channels generally employ an equalizer to recover the transmitted sequence corrupted by intersymbol interference (ISI). Frequency-selective channels are characterized by a constant gain and linear-phase response over a bandwidth that is smaller than the bandwidth of the signal to be transmitted. Equivalently, in the time domain, the length of the impulse response of the channel is equal to or longer than the width of the modulation signal. This is the case, for example, in high data rate wireless systems.

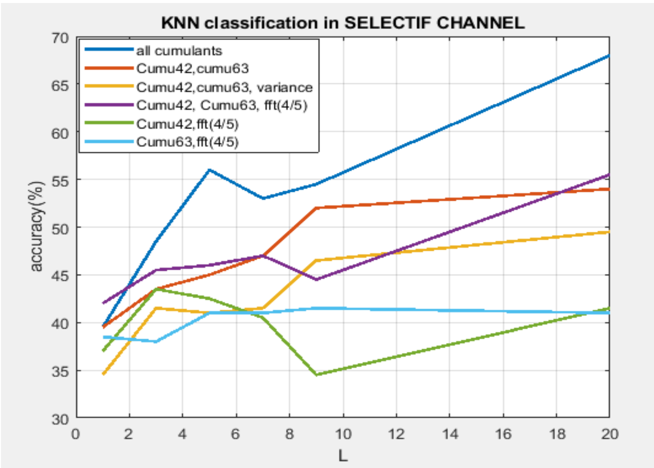
IV.3.2 classification results

Tab.1 KNN classification of modulated signal in Selectif channel for SNR=10

Parameter	Accuracy for L=1	Accuracy for L=3	Accuracy for L=5	Accuracy for L=7	Accuracy for L=9	Accuracy for L=20
All cumulants	39.5%	48.5%	56%	53%	54.5%	68%
Cumu42,cumu63	39.5%	43.5%	45%	47%	52%	54%
Cumu42,cumu63, variance	34.5%	41.5%	41%	41.5%	46.5%	49.5%
Cumu42, Cumu63, fft(4/5)	42%	45.5%	46%	47%	44.5%	55.5%
Cumu42,fft(4/5)	37%	43.5%	42.5%	40.5%	34.5%	41.5%
Cumu63,fft(4/5)	38.5%	38%	41%	41%	41.5%	41%



FigureIV.1: Confusion matrix of KNN classification using best parameters in selectif channel SNR=10, L=20



FigureIV.2: KNN modulation classification using variance, Cumulants and fft (4/5)

Tab.2 KNN classification of modulated signal in Selectif channel for SNR=5

Parameter	Accuracy for L=1	Accuracy for L=3	Accuracy for L=5	Accuracy for L=7	Accuracy for L=9	Accuracy for L=20
All cumulants	41.5%	48.5%	54%	52%	56%	58%
Cumu42,cumu63	41%	42.5%	55%	54%	54%	56.5%
Cumu42,cumu63, variance	38.5%	40.5%	46%	47%	48.5%	52.5%
Cumu42, Cumu63, fft(4/5)	38%	42.5%	53.5%	48%	50%	54.5%
Cumu42,fft(4/5)	34%	40.5%	45%	45.5%	45.5%	37.5%
Cumu63,fft(4/5)	36%	43.5%	44.5%	35.5%	36%	34%

Tab.3 KNN classification of modulated signal in Selectif channel for SNR=0

Parameter	Accuracy for L=1	Accuracy for L=3	Accuracy for L=5	Accuracy for L=7	Accuracy for L=9	Accuracy for L=20
All cumulants	37%	44.5%	47%	45%	55%	54%
Cumu42,cumu63	34%	44.5%	46%	39.5%	51.5%	52%
Cumu42,cumu63, variance	35.5%	47%	41.5%	44%	51.5%	56.5%
Cumu42, Cumu63, fft(4/5)	36%	45%	43.5%	41%	54.5%	51%
Cumu42,fft(4/5)	36.5%	38%	43.5%	43%	42%	45%
Cumu63,fft(4/5)	38%	41.5%	41.5%	34%	44%	35.5%

Tab.4 KNN classification of modulated signal in Selectif channel for SNR= -5

Parameter	Accuracy for L=1	Accuracy for L=3	Accuracy for L=5	Accuracy for L=7	Accuracy for L=9	Accuracy for L=20
All cumulants	28%	32%	36.5%	36%	42.5%	50%
Cumu42,cumu63	29.5%	31%	35.5%	37.5%	41%	54.5%
Cumu42,cumu63, variance	28.5%	31%	38%	37%	39.5%	51.5%
Cumu42, Cumu63, fft(4/5)	30%	30%	37.5%	37.5%	41%	54.5%
Cumu42,fft(4/5)	34%	37.5%	32%	35%	38%	39.5%
Cumu63,fft(4/5)	28%	31%	37%	34.5%	40%	42%

IV.4 Gaussien channel

IV.4.1 Definition Gaussian channel

This channel affected by Gaussian noise is a model among common communication channels, such as wired and wireless telephone channels and satellite links. Without further conditions, the capacity of this channel may be infinite. If the noise variance is zero, the receiver receives the transmitted symbol perfectly. The consideration of communication systems beyond Shannon's approach is useful in order to increase the efficiency of information transmission for certain applications.

Tab.5 KNN classification of modulated signal in AWGN channel

Parameter	**Accuracy (snr=10)**	**Accuracy (snr=5)**	**Accuracy (snr=0)**	**Accuracy (snr=-5)**
All cumulants	73%	70%	68%	55%
Cumu42,cumu63	65.5%	67.5%	65 .5%	57.5%
Cumu42,cumu63, variance	66.5%	67%	66.5%	57%
Cumu42, Cumu63, fft(4/5)	85%	86.5%	83%	60%
Cumu42,fft(4/5)	87.5%	85.5%	85%	70.5%
Cumu63,fft(4/5)	86%	86.5%	82.5%	60%

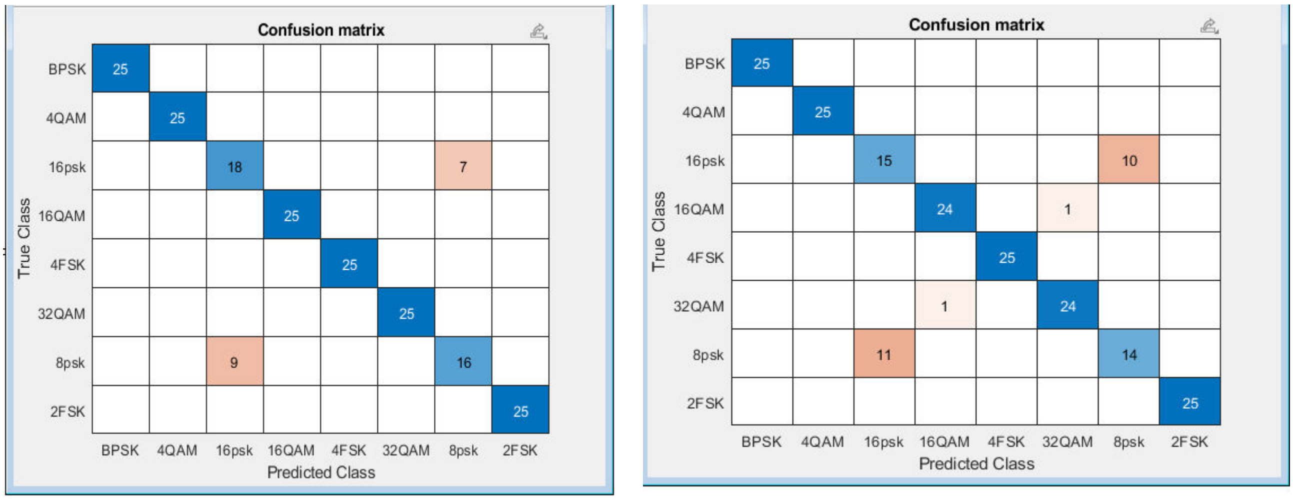
Tab.6 Best parameters for the AWGN

Parameter	**Accuracy (snr=10)**
Cum42,cum63,cum21,fft(4 /5)	90%
Cum21,cum63,fft(4/5)	**92%**
Cum21,cum63,fft(4/5),fft(2/3)	88 %

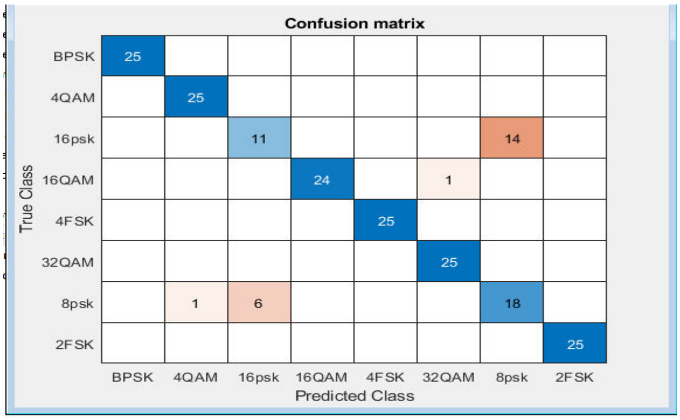
Parameter	**Accuracy (snr=5)**
Cum21,cum42,cum60,fft(4/5), cum41,cum40	87%
Cum21,cum42,fft(4/5)	**88.5 %**
Cum21,cum63,fft(4/5),fft(2/3)	87%

Parameter	**Accuracy (snr=-5)**
Cum40,fft(4/5),cum21	86.5%
Cum41,fft(4/5),cum21	87%
Cum41,fft(4/5),cum21,cum40	**89%**

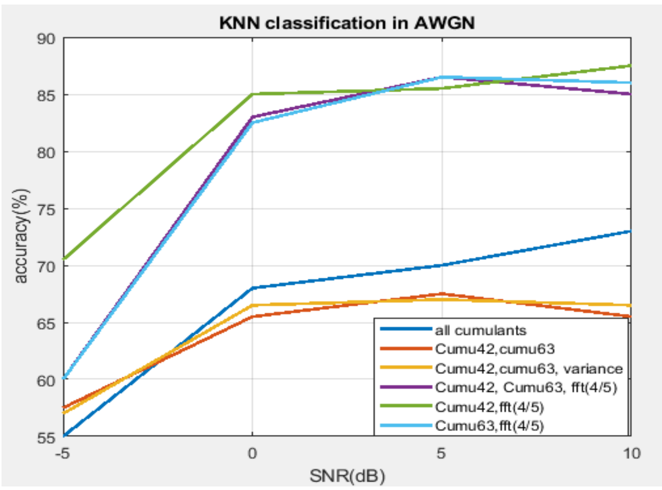
Parameter	**Accuracy (snr=0)**
Cum21,cum40,cum63,fft(4/5)	84.5%
Cum21,fft(4/5)	87.5**%**
Cum21,fft(4/5),cum42	**88.5%**



FigureIV.3: Confusion matrix of KNN classification using best parameters in AWGN channel (with (SNR=10 (left), SNR=5 (right))



FigureIV.4: Confusion matrix of KNN classification using best parameters in AWGN channel SNR=-5



FigureIV.5: KNN modulation classification using variance, Cumulants and fft (4/5)

IV.5 Discussion:

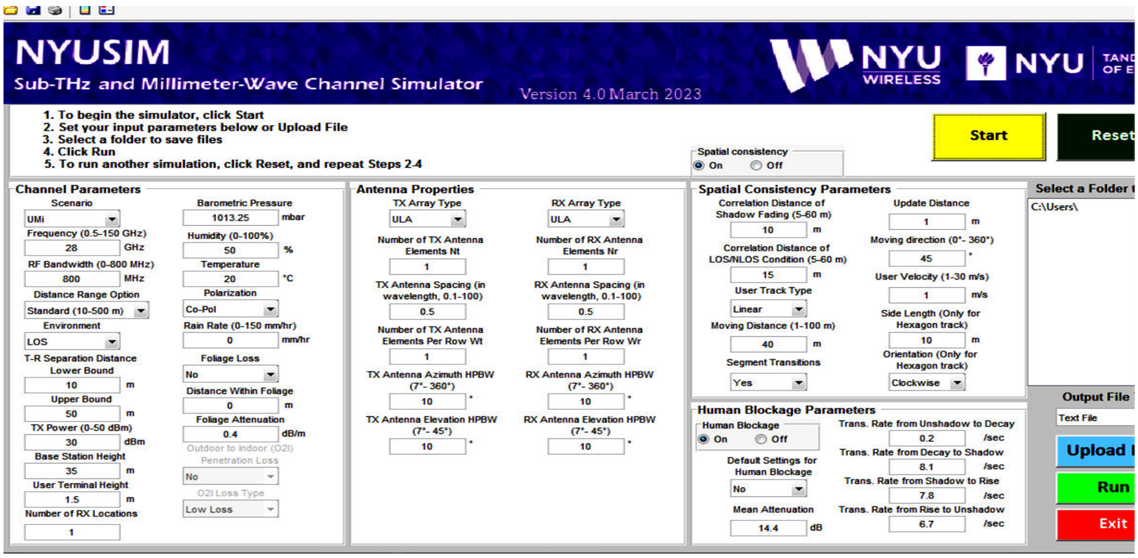
In our work some related features were used in order to improve classifier performance and their applicability to some specific modulation types. We have studied two channels Gaussien channel(AWGN) and Selectif channel in which AWGN gives good performance compared to the selectif channel, and through table 6, the combination of cumulants C21,C63 with FFT (4/5) gives higher value (92%) in the case of SNR=10, while the cumulants C21,C42 with FFT (4/5) gives 88.5% in the cases of SNR=0 and SNR=5. The combination of C21,C40, C41 with FFT (4/5) gives 89% for SNR=-5. The FFT (Fourier transform) is considered as better feature which may insensitive to noise and ideal to classification. While the C42, C63 and the variance gives a low accuracy at SNR=[10,5,0,-5], the variance provide low accuracy when it was combined with cumulants.

In the case of selectif channel, we notice that when the value of L increases the accuracy goes up and when the value of SNR decreases the accuracy rate decreases. The combination of All cumulants at L=20 gives a high value (68%) at SNR=10, furthermore, the FFT is considered the better feature which can be insensitive to noise and ideal for the classification.

IV.6 TERAHERTZ CHANNEL:

IV.6.1 Description of Simulator NYUSIM:

NYUSIM is an open source simulator based on a 3-D spatial statistical channel model (SSCM) [127], which adopts a time-cluster spatial-lobe (TCSL) approach to characterize temporal and angular properties of multipath components. NYUSIM provides physical-level simulations that realistically simulate the downlink channel between a base station (BS) and a UT. The screenshot in Figure VI.6 shows the graphical user interface (GUI) of NYUSIM.



FigureIV.6: Graphical User Interface (GUI) of the NYUSIM channel model simulator

Tab.7: parameters of channel 1(1x1)

Parameter	Value	Parameter	Value
Frequency	150	RxAzHPBW	10
Bandwidth	1000	RxElHPBW	10
TXPower	30	SCenable	Off
Environment	LOS	CorrDistanceSF	10
Scenario	UMi	CorrDistanceLOS	15
TXHeight	35	TrackType	Linear
RXHeight	1.5	MovDistance	40
Pressure	1013.25	MovDirection	45
Humidity	0	UpdateDist	1
Temperature	20	Velocity	1
RainRate	0	HBenable	Off
Polarization	Co-Pol	HBdefault	No
Foliage	No	SEMean	14 .4
DistFol	0	U2Drate	0.20
FoliageAttenuation	0.4	D2Srate	8.1
TxArrayType	ULA	S2Rrate	7.8
RxArrayType	ULA	R2Urate	6.7
NumberOfTxAntenna	1	distType	Standard (10-500 m)
NumberOfRxAntenna	1	dmin	10
NumberOfTxAntennaPerRow	1	dmax	50
NumberOfRxAntennaPerRow	1	Nsim	1
TxAntennaSpacing	0.5	o2iLoss	No
RxAntennaSpacing	0.5	o2iType	Low Loss
TxAzHPBW	10	side_length	10
TxElHPBW	10	orient	Clockwise
transExist	Yes	/	/

Tab.8: parameters of channel 2(2x2)

Parameter	Value	Parameter	Value
Frequency	100	RxAzHPBW	10
Bandwidth	800	RxElHPBW	10
TXPower	30	SCenable	On
Environment	LOS	CorrDistanceSF	16
Scenario	UMi	CorrDistanceLOS	13
TXHeight	35	TrackType	Linear
RXHeight	1.5	MovDistance	46
Pressure	1013.25	MovDirection	45
Humidity	2	UpdateDist	1
Temperature	20	Velocity	3.2
RainRate	0	HBenable	On
Polarization	Co-Pol	HBdefault	No
Foliage	No	SEMean	15.8
DistFol	0	U2Drate	0.21
FoliageAttenuation	0.4	D2Srate	7.88
TxArrayType	ULA	S2Rrate	7.7
RxArrayType	ULA	R2Urate	7.67
NumberOfTxAntenna	2	distType	Standard (10-500 m)
NumberOfRxAntenna	2	dmin	10
NumberOfTxAntennaPerRow	2	dmax	50
NumberOfRxAntennaPerRow	2	Nsim	1
TxAntennaSpacing	0.5	o2iLoss	No
RxAntennaSpacing	0.5	o2iType	Low Loss
TxAzHPBW	10	side_length	12
TxElHPBW	10	orient	Clockwise
transExist	Yes	/	/

IV.7 Classification of modulated signals in Thz channel (1x1)

Firstly, we used the same set of parameters from previous experiments to compare the classification accuracy in both the AWGN channel, channel selectif and the Thz channel. Secondly, we conducted a search to find the optimal set of parameters that would yield high classification accuracy using the KNN classifier.

Tab.9: Classification rates without signals amplification

Parameters	**Accuracy (snr=10)**	**Accuracy (snr=5)**	**Accuracy (snr=0)**	**Accuracy (snr=-5)**
All cumulants	28%	27%	21%	15%
Cumu42,cumu63	15.5%	15%	13.5%	15%
Cumu42,cumu63, variance	20.5%	18.5%	14.5%	17.5%
Cumu42, Cumu63, fft(4/5)	12.5%	17%	16%	17.5%
Cumu42,fft(4/5)	11%	13.5%	16.5%	11.5%
Cumu63,fft(4/5)	11.5%	20.5%	15%	15%

Tab.10: Amplification of signals at the emission step (*1000)

Parameters	Accuracy (snr=10)	Accuracy (snr=5)	Accuracy (snr=0)	Accuracy (snr=-5)
All cumulants	74%	65.5%	33.5%	24.5%
Cumu42,cumu63	70.5%	54.5%	28.5%	15%
Cumu42,cumu63, variance	70%	62.5%	31.5%	24%
Cumu42, Cumu63, fft(4/5)	75.5%	68%	44%	24%
Cumu42,fft(4/5)	77.5%	70%	38.5%	23%
Cumu63,fft(4/5)	70.5%	63.5%	38%	15%

Tab.11: Amplification of signals at the emission step (*10000)

Parameters	Accuracy (snr=10)	Accuracy (snr=5)	Accuracy (snr=0)	Accuracy (snr=-5)
All cumulants	72.5%	71%	64.5%	66%
Cumu42, Cumu63	78%	80%	71.5%	58%
Cumu42, Cumu63, variance	70%	66%	64.5%	58%
Cumu42, Cumu63, fft(4/5)	86.5%	84%	77.5%	59%
Cumu42,fft(4/5)	82%	82.5%	86%	76%
Cumu63,fft(4/5)	81%	80%	71.5%	58%

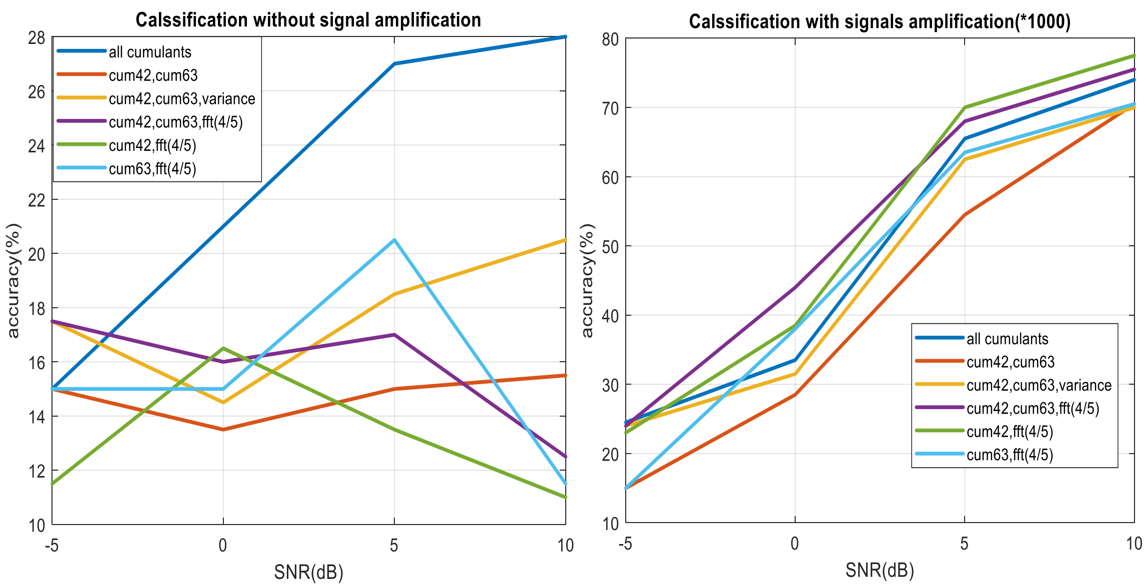


Figure IV.7: KNN modulation classification without signals amplification and with signals amplification

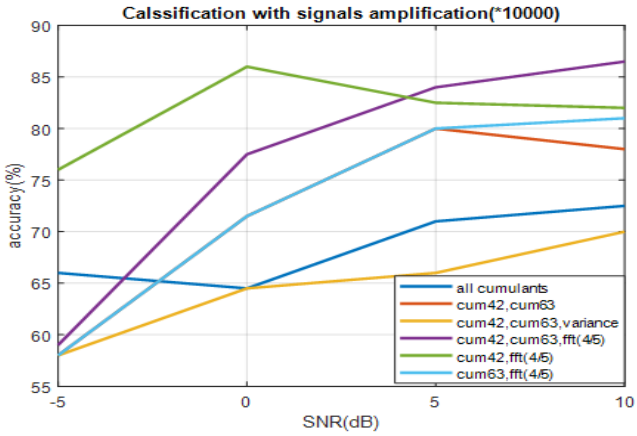


Figure IV.8: KNN modulation classification with signals amplification

The classification accuracy is higher when the signals are multiplied by 10,000 compared to when they were multiplied by 1,000, especially in the case of SNR=-5. Therefore, in subsequent experiments, we utilized the 10,000 factor.

- ## Cross-Correlation

This is a kind of correlation, in which the signal in-hand is correlated with another signal so as to know how much resemblance exists between them. Mathematical expression for the cross-correlation of continuous time signals x (t) and y (t) is given by

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y^*(t-\tau)dt \quad \textbf{(53)}$$

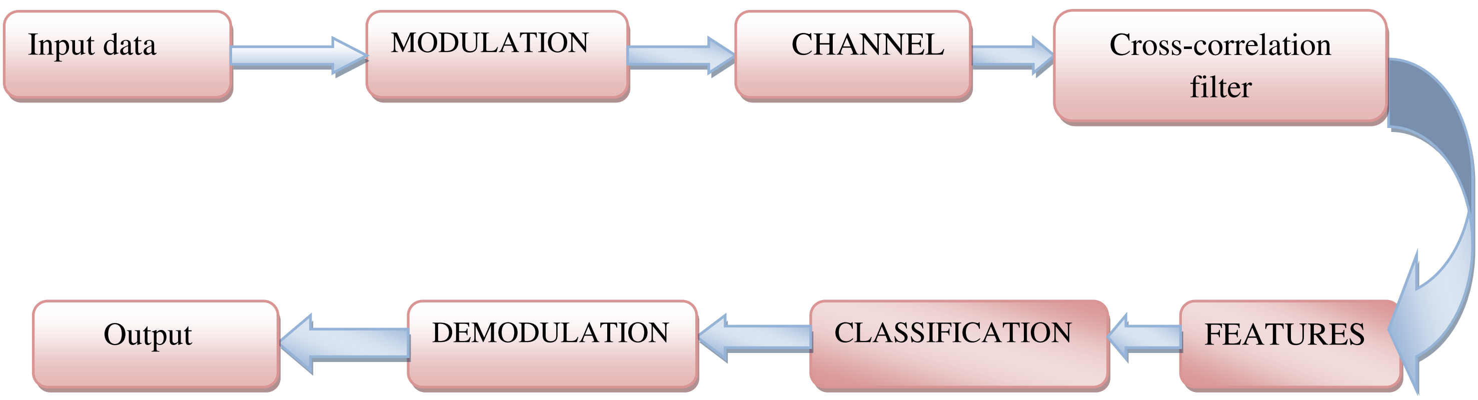


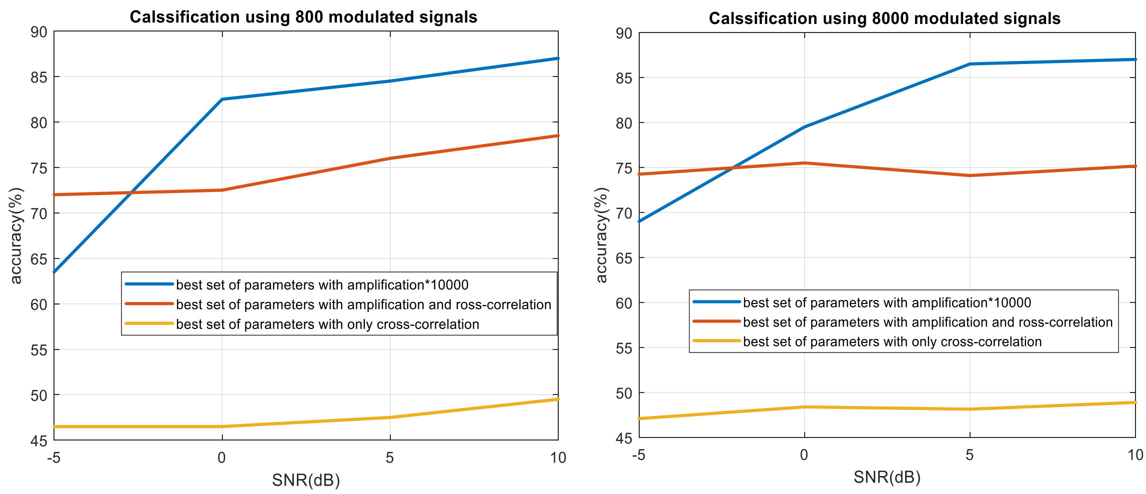
Figure IV. 9: scheme of classification in digital modulation

Tab.12: Best set of parameters (number of signals (8*100))

Parameters	Accuracy (snr=10)	Accuracy (snr=5)	Accuracy (snr=0)	Accuracy (snr=-5)
Cumu41, Cumu42, Cumu60, Cumu63, fft(4/5) with amplification*10000	87%	84.5%	82.5%	63.5%
Same parameters with amplification and Cross-correlation	78.5%	76%	72.5%	72%
Same parameters and Cross-correlation without amplification	49.5%	47.5%	46.5%	46.5%

Tab.13: Best set of parameters (number of signals (8*1000))

Parameters	Accuracy (snr=10)	Accuracy (snr=5)	Accuracy (snr=0)	Accuracy (snr=-5)
Cumu41, Cumu42,Cumu60, Cumu63, fft(4/5)with amplification*10000	87%	86.5%	79.5%	69%
Same parameters/ amplification and Cross-correlation	75.15%	74.10%	75.5%	74.25%
Same parameters and Cross-correlation without amplification	48.9%	48.15%	48.4%	47.10%

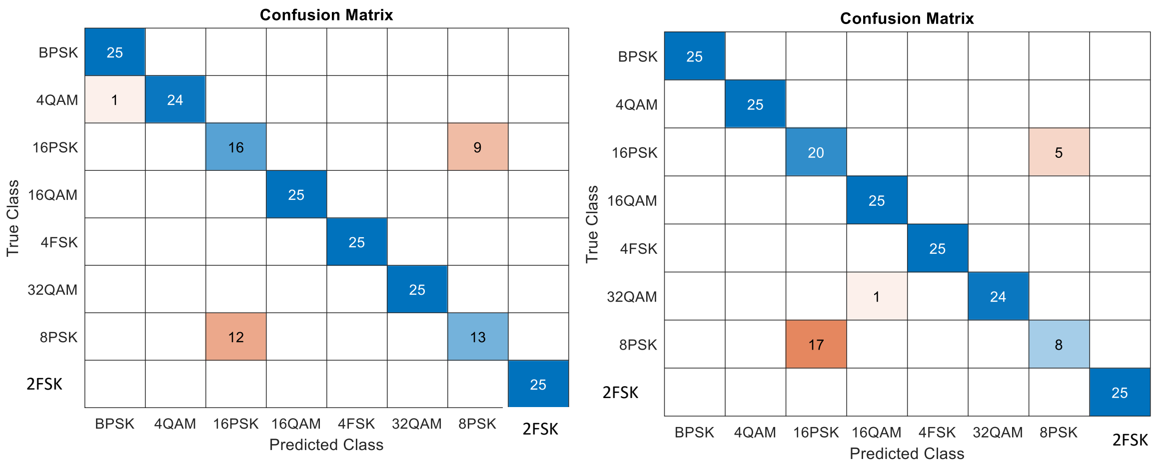


FigureIV.10: KNN modulation classification Best set of parameters

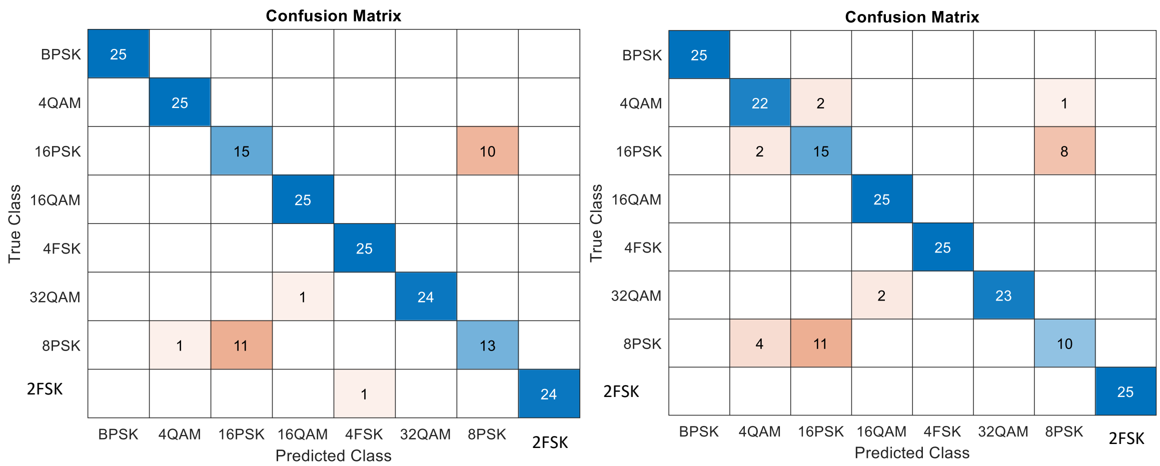
In the case of SNR equal to 10, the classification result of a dataset consisting of 8,000 signals is comparable to that of a dataset comprising 800 signals.

Tab.14: Classification of the modulated signals using the best set of parameters and the SVM classifier (number of signals (8*100))

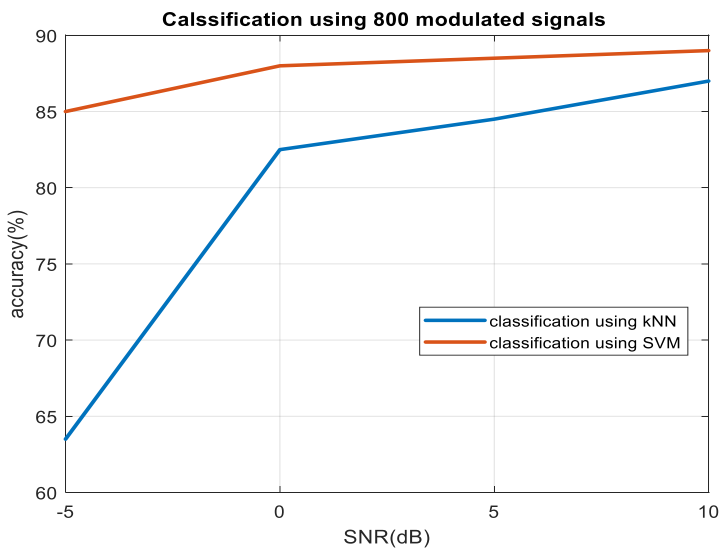
	SNR=10	SNR=5	SNR=0	SNR=-5
KNN	87%	84.5%	82.5%	63.5%
SVM	89%	88.5%	88%	85%



FigureIV.11: Confusion matrix of SVM classification using best parameters in Thz channel(1x1) with (SNR=10 (left), SNR=5 (right))



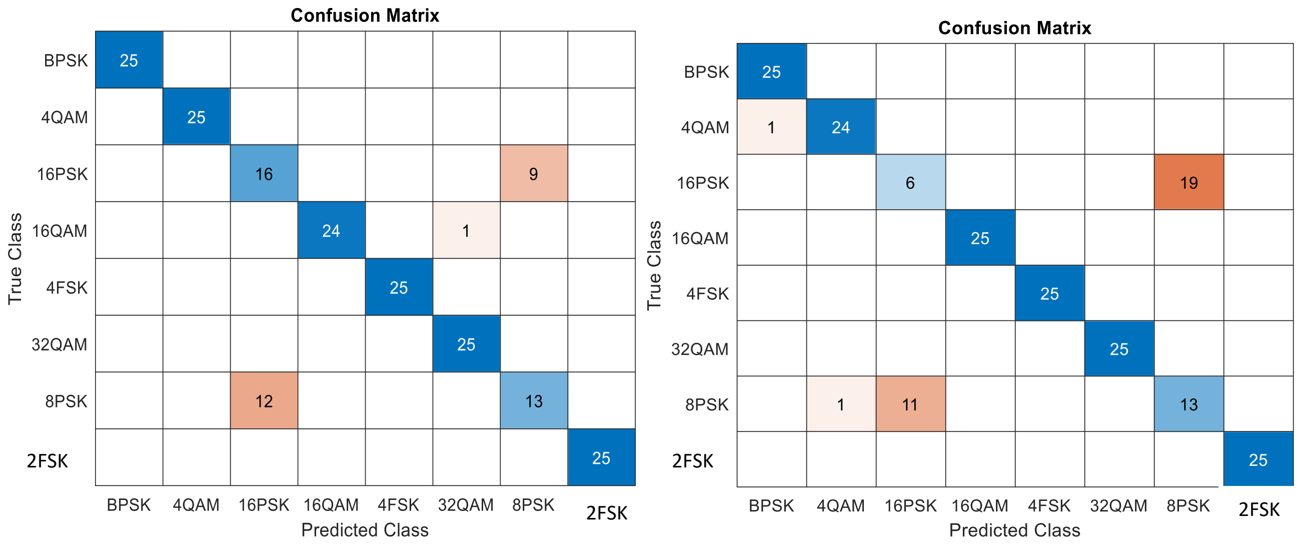
FigureIV.12: Confusion matrix of SVM classification using best parameters in Thz channel (1x1) with (SNR=0 (left), SNR=-5 (right))



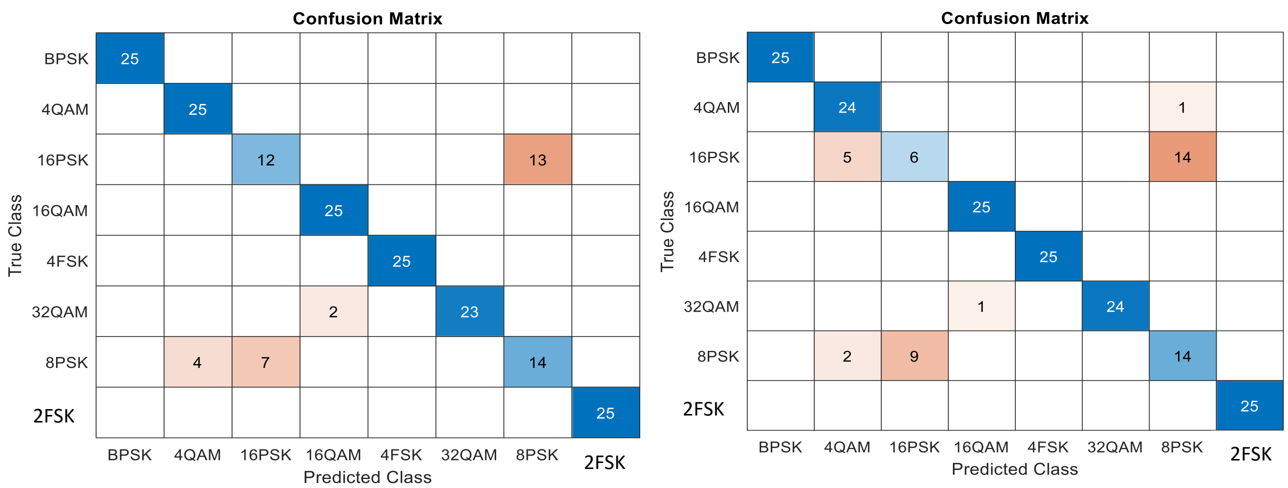
FigureIV.13: Comparison between KNN and SVM

Tab.15: Classification of the modulated signals using the best set of parameters and the SVM classifier (number of signals (8*1000))

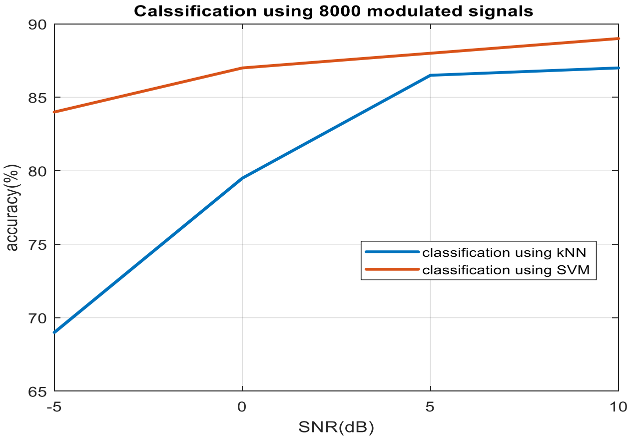
	SNR=10	SNR=5	SNR=0	SNR=-5
KNN	87%	86.5%	79.5%	69%
SVM	89%	88%	87%	84%



FigureIV.14: Confusion matrix of SVM classification using best parameters in Thz channel (1x1) with (SNR=10 (left), SNR=5 (right))



FigureIV.15: Confusion matrix of SVM classification using best parameters in Thz channel(1x1) with (SNR=0(left), SNR=-5 (right))



FigureIV.16: Comparison between KNN and SVM

IV.8 Classification of modulated signals in Thz channel (2x2)

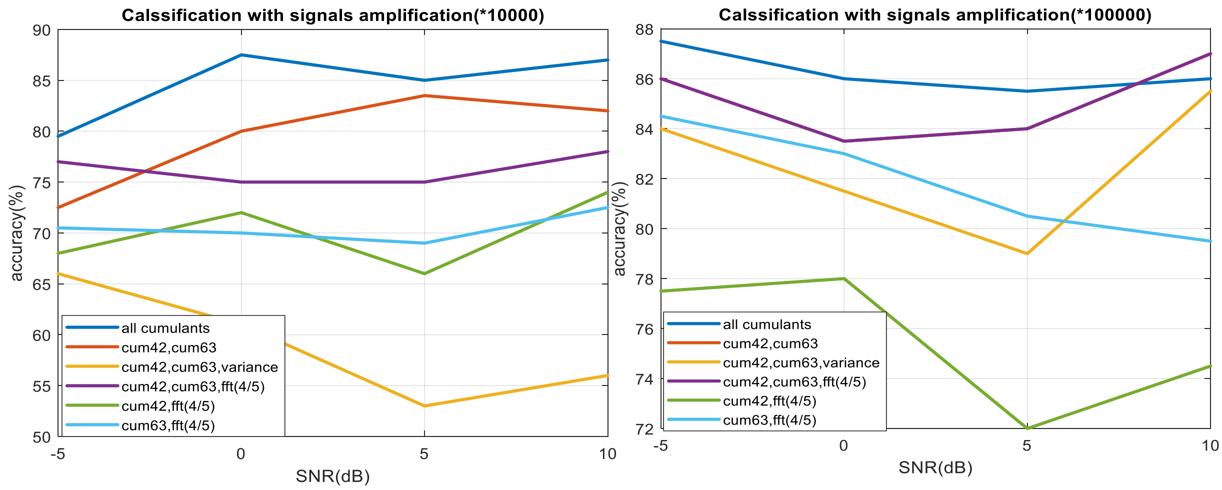
The following tables show the classification accuracies using (2x2) antennas and the same previous set of parameters.

Tab.16: Amplification of signals at the emission *10,000 (number of signals (8*100))

Parameters	Accuracy (snr=10)	Accuracy (snr=5)	Accuracy (snr=0)	Accuracy (snr=-5)
All cumulants	87%	85%	87.5%	79.5%
Cumu42,cumu63	82%	83.5%	80%	72.5%
Cumu42,cumu63, variance	56%	53%	61%	66%
Cumu42, Cumu63, fft(4/5)	78%	75%	75%	77%
Cumu42,fft(4/5)	74%	66%	72%	68%
Cumu63,fft(4/5)	72.5%	69%	70%	70.5%

Tab.17: Amplification of signals at the emission *100,000 (number of signals (8*100))

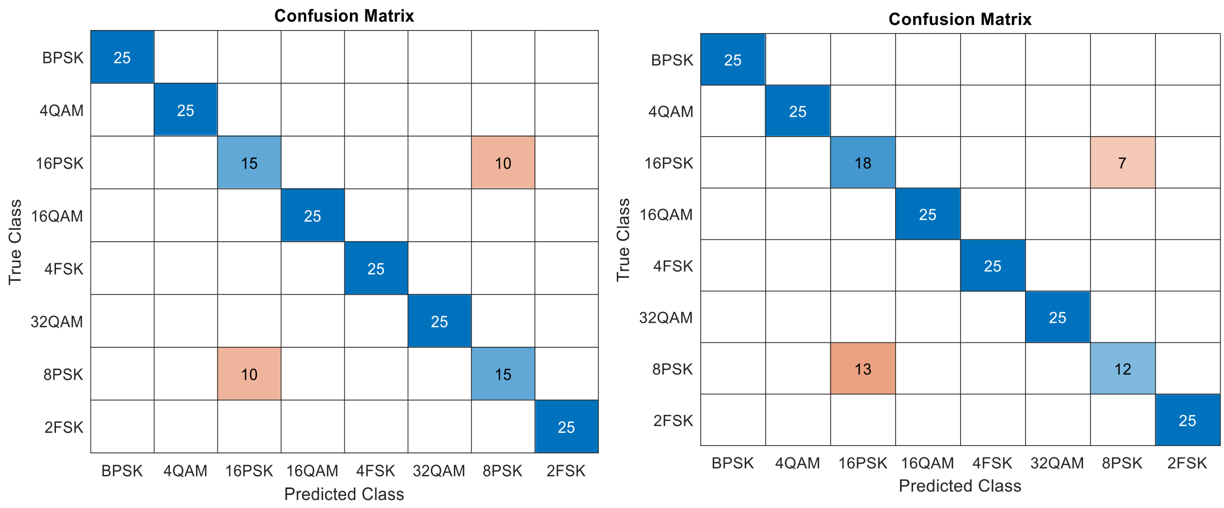
Parameters	Accuracy (snr=10)	Accuracy (snr=5)	Accuracy (snr=0)	Accuracy (snr=-5)
All cumulants	86%	85.5%	86%	87.5%
Cumu42,cumu63	87%	84%	83.5%	86%
Cumu42,cumu63, variance	85.5%	79%	81.5%	84%
Cumu42, Cumu63, fft(4/5)	87%	84%	83.5%	86%
Cumu42,fft(4/5)	74.5%	72%	78%	77.5%
Cumu63,fft(4/5)	79.5%	80.5%	83%	84.5%



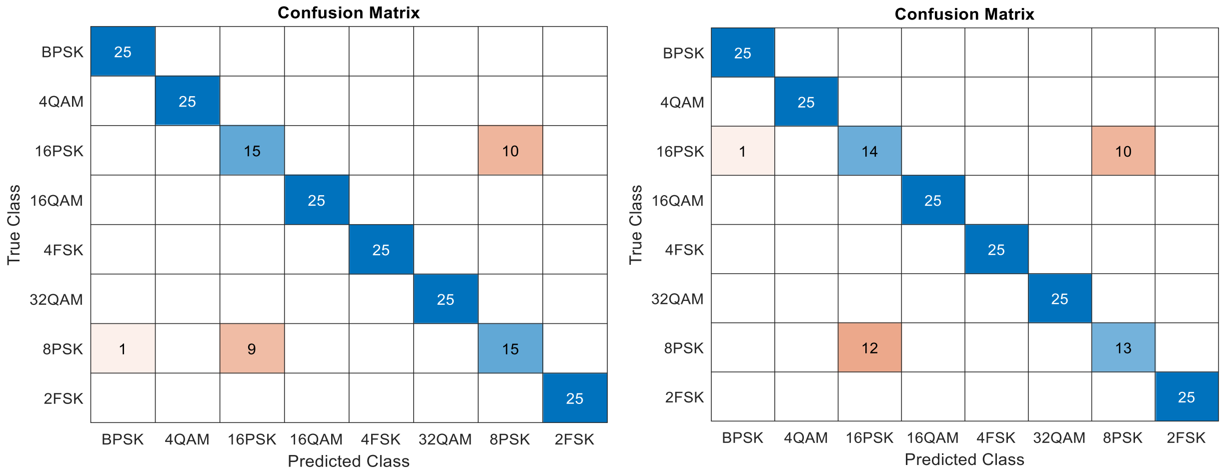
FigureIV.17: KNN modulation classification with signals amplification

Tab.18: Best set of parameters (number of signals (8*100)) using an amplification*100,000

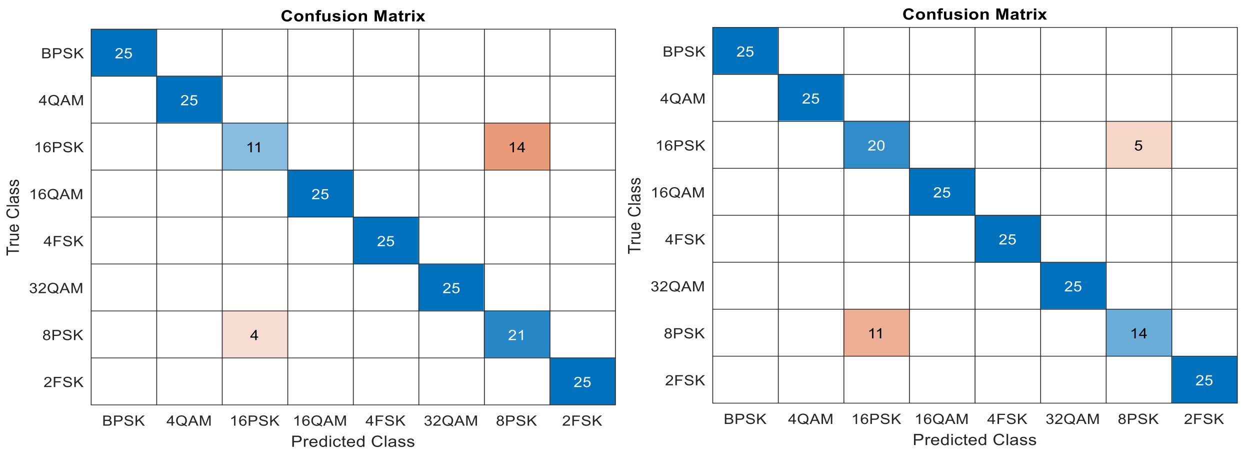
Parameters/Classifier	**Accuracy (snr=10)**	**Accuracy (snr=5)**	**Accuracy (snr=0)**	**Accuracy (snr=-5)**
Cumu21, fft(4/5) with KNN	89%	90%	90%	88.5%
Cumu21, fft(4/5) with SVM	91%	92%	89.5%	89%



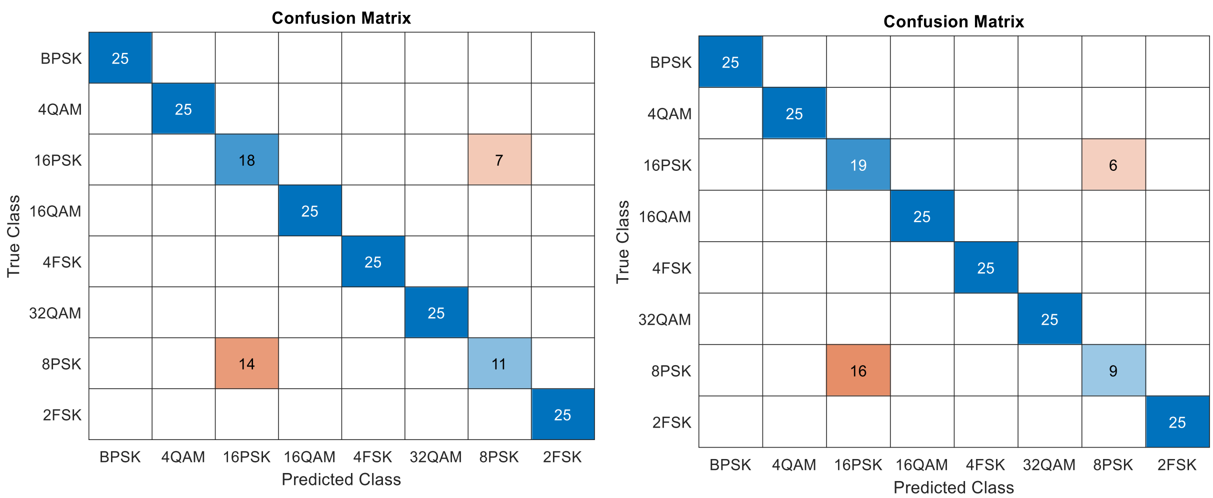
FigureIV.18: Confusion matrix of classification using best parameters in Thz channel (2x2) using KNN with SNR=10 (left) and SNR=5 (right)



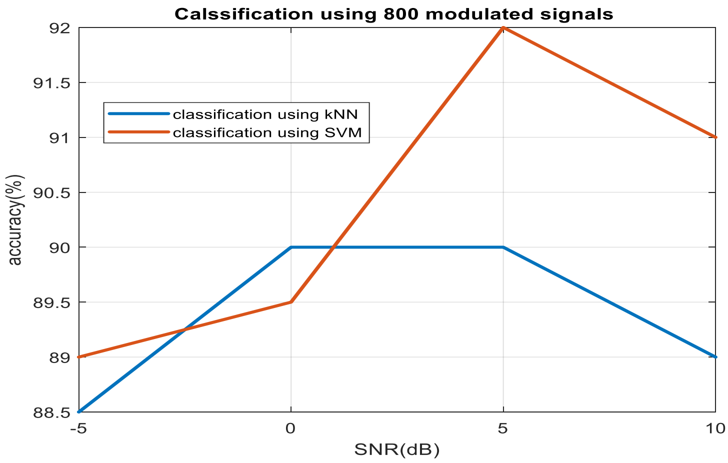
FigureIV.19: Confusion matrix of classification using best parameters in Thz channel (2x2) using KNN with SNR=0 (left) and SNR= -5 (right)



FigureIV.20: Confusion matrix of classification using best parameters in Thz channel(2x2) using SVM with SNR=10 (left) and SNR= 5 (right)



FigureIV.21: Confusion matrix of classification using best parameters in Thz channel(2x2) using SVM with SNR=0 (left) and SNR= -5 (right)



FigureIV.22: Comparison between KNN and SVM

IV.9 Discussion

We presented in the table (9) the classification rates of the modulated signals in the THz channel (1x1) without signal amplification which gives low accuracy values, however, the classification accuracies are higher when multiplying signals by 10,000 compared to the multiplication by 1,000, especially in the case of SNR=-5.In table (10) the amplification of the signals at the emission step (*1000) gave average values.

We note that in the case of the amplification of signals at the emission step (*10000), good accuracy was obtained (86.5%) using the combination of cumulants C42, C63 with FFT (4/5) in the case of SNR=10.

In tables (12) and (13) we present the best set of parameters which gives the higher accuracy. When we increased the number of signals in the datasets from 800 to 8000,the results exhibit negligible differences.

We used the Cross-correlation filter which helps to reduce the channel noise. This filter was used in two cases: with and without amplification. However, the obtained results in the second case were very low in the contrast of the first case.

In table (15) we compared between the KNN and the SVM classifiers, we concluded that the SVM is more accurate than the KNN.

We notice that for the SVM classifier we used the RBF kernel with one vs one and one vs all schemes for multiclass classification.

In the second part of simulations, we used the THZ channel (2x2). Firstly, we increased the amplification of signals at the emission from ***10,000** to***100,000** without changing in the number of signals (800). The results presented tables (16) and (17) indicate that the accuracy in the cases of SNR=[10, 5, 0] using the two amplifications factors are comparables. But, in the case of SNR=-5 the amplification with ***100,000** gives the best results.

Additionally, we compared between Cumu21, FFT (4/5) with KNN and Cumu21, FFT (4/5) with SVM with an amplification using *100,000. The obtained results indicated that the SVM classifier is more accurate than the KNN.

The corresponding confusion matrix shown that the misclassification occurs mainly in the cases of 8PSK and 16PSK.

IV.10 Conclusion

In this chapter we presented the simulations results of the modulation classification in the cases of AWGN channel, selectif channel and THZ channels (using (1x1) and (2x2) antennas on the emitter and the receiver).
The classification results of the 8 modulated signals in the case of AWGN channel exhibited a high degree of accuracy and performance compared to the selectif channel. However, the goal of our work was to perform a modulation classification in the THZ channel, therefore we conducted firstly, a classification experiment over the (1x1) THZ channel then, on the (2x2) THZ channel.Unfortunately, the obtained results were unsatisfactory, which forced us to find alternative approach. Therefore, we suggested amplifying the transmitted signals. While this solution leads to high accuracy, it may be practically unfeasible. However, the obtained results served as an indication to the original problem, which is related to the attenuation of the signal in the THZ channel and do not depend to the quality of the selected parameters types or to the accuracy of the KNN and the SVM classifiers. The use of the matched filter using the cross-correlation operation without amplification slightly improved the classification results.

The overall obtained results were theoretically encouraging. Hence, we suggest that the matched filter can be practically utilized in future works if the results demonstrate improvement.

General Conclusion

THz plays a decisive role in the development of the future wireless networks which will have to satisfy a rapid increasing demand for bandwidth. As THz frequencies are higher than those of the radio waves and Wi-Fi that are currently used, they should allow significantly higher data transmission speeds to be achieved (> 10 Gbps). This has a great importance in many applications of life science, telecommunication and the environment, which encouraged research.

In this work, a modulation classification system using a support vector machines (SVM) classifier ,K-Nearest Neighbor (KNN) and a set of features was simulated using MATLAB ,and we used NYUSIM Simulator for the simulation of the Terahertz channel.

As part of this work, we established, firstly a general Overview about Nano-Communication Secondly, we presented a channels model in Nano-Communication system finally, we described some classification method of digital modulation which were used in our simulations of features extraction methods from modulated signal.

Support vector machines recently achieved high accuracy in automatic modulation classification, thus we used it in our work and compared its performance against the K-nearest neighbor classifier.

In the order to find the best possible features and the best relationship among them, we conducted different tests over several features in high and low SNR values. The simulation results revelated that the features group formed by theCumu21, FFT(4/5) with SVM achieves higher classification accuracy for the eight proposed modulation types in the (2x2) THZ channel. Furthermore, we noticed that the SVM classifiers outperforms the KNN classifier in the most cases.

For the future works:

-We think that it will be interesting to find another parameters which will be able to increase the classification accuracy for the 8PSK and 16 PSK.

- An alternative solution to the amplification operation needs to be explored in order to solve the problem of signal attenuation.

- Deep Neural Networks can be suggested as a solution to improve the classification accuracy of modulated signals in the THZ channel without features extraction.

References

[1] Ian F. Akyildiz, Fernando Brunetti , Cristina Blazquez. Nano networks: A new communication paradigm. *Computer Networks* 52 (2008) 2260-2279 doi:10.1016/j.comnet.2008.04.001.

[2] Martin Zikmunda. Communication technology for nano sensors and nano devices applications in the health industry. NANOCON 2011, Brno, Czech Republic, EU.

[3] Vimal Upadhayayl and Dr. Sonali Agarwal. Application of wireless nano sensor networks for wild lives. *International Journal of Distributed and Parallel Systems* (IJDPS) Vol. 3, No. 4, July 2012.

[4] S. F. Bush and Y. Li, "Nano-communications: A new field? an exploration into a carbon nanotube communication network," *GE Global Research Technical Report*, 2006.

[5] I. F. Akyildiz, F. Brunetti, and C. Bl´azquez, "Nanonetworks: A new communication paradigm," *Computer Networks*, vol. 52, no. 12, pp. 2260–2279, 2008.

[6] S. F. Bush, *Nanoscale Communication Networks*. Artech House, 2010.

[7] I. F. Akyildiz, J. M. Jornet, and M. Pierobon, "Nanonetworks: A new frontier in communications,"*Communications of the ACM*, vol. 54, no. 11, pp. 84–89, 2011.

[8] K. Yang, A. Pellegrini, M. Munoz, A. Brizzi, A. Alomainy, and Y. Hao, "Numerical analysis and characterization of thz propagation channel for body-centric nanocommunications," *Terahertz Science and Technology, IEEE Transactions on*, vol. 5, no. 3, pp. 419–426,May 2015.

[9] S. J. Lee, C. A. Jung, K. Choi, and S. Kim, "Design of wireless nanosensor networks for intrabody application," *International Journal of Distributed Sensor Networks*, vol. 2015, 2015.

[10] R. A. Freitas, Jr., ``Nanotechnology, nanomedicine and nanosurgery," *Int. J. Surg.*, vol. 3, no. 4, pp. 243_246, 2005.

[11] Yg Li and Gl Stuber. *Orthogonal frequency division multiplexing for wireless communications*. 2006.

[12] Chen Gong and Zhengyuan Xu. Channel estimation and signal detection for optical wireless scattering communication with inter-symbol interference. *IEEE Transactions on Wireless Communications*, 14(10):5326–5337, 2015.

[13] Josep Miquel Jornet and Ian F Akyildiz. Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band. *IEEE Transactions on Wireless Communications*, 10(10):3211–3221, 2011.

[14] Sarah Hussein, Youmni Ziade, and RaedMShubair. On propagation losses due to in-vivo electromagnetic nanoscale communication. In *2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, pages 1309–1310. IEEE, 2019.

[15] Weisi Guo, Christos Mias, Nariman Farsad, and Jiang-Lun Wu. Molecular versus electromagnetic wave propagation loss in macro-scale environments. *IEEE Transactions on Molecular, Biological andMulti-Scale Communications*, 1(1):18–25, 2015.

[16] GeorgeMWhitesides. The once and future nanomachine. *Scientific American*, 285(3):78–83, 2001.

[17] Uche AK Chude-Okonkwo, Reza Malekian, Bodhaswar T Maharaj, and Athanasios V Vasilakos. Molecular communication and nanonetwork for targeted drug delivery: A survey. *IEEE Communications Surveys & Tutorials*, 19(4):3046–3096, 2017.

[18] Nariman Farsad, H Birkan Yilmaz, Andrew Eckford, Chan-Byoung Chae, andWeisi Guo. A comprehensive survey of recent advancements in molecular communication. *IEEE Communications Surveys & Tutorials*, 18(3):1887–1919, 2016.

[19] C. Rutherglen and P. Burke, ``Nanoelectromagnetics: Circuit and electromagnetic properties of carbon nanotubes,'' *Small*, vol. 5, no. 8, pp. 884_906, 2009.

[20] T. Nakano, M. J. Moore, F.Wei, A. V. Vasilakos, and J. Shuai, ``Molecular communication and networking: Opportunities and challenges,'' *IEEE Trans. Nanobiosci.*, vol. 11, no. 2, pp. 135_148, Jun. 2012.

[21] O. B. Akan, H. Ramezani, T. Khan, N. A. Abbasi, and M. Kuscu, ``Fundamentals of molecular information and communication science,'' *Proc. IEEE*, vol. 105, no. 2, pp. 306_318, Feb. 2017.

[22] L. You, R. S. Cox, R. Weiss, and F. H. Arnold, ``Programmed population control by cell_cell communication and regulated killing,'' *Nature*, vol. 428, pp. 868_871, Apr. 2004.

[24] M.-T. Chen and R. Weiss, ``Arti_cial cell-cell communication in yeast *saccharomyces cerevisiae* using signaling elements from Arabidopsis thaliana,'' *Nature Biotechnol.*, vol. 23, pp. 1551_1555, Nov. 2005.

[25] M. J. Doktycz and M. L. Simpson, ``Nano-enabled synthetic biology,'' *Mol. Syst. Biol.*, vol. 3, no. 1, p. 125, 2007.

[26] Y. Sasaki *et al.*, ``A nanosensory device fabricated on a liposome for detection of chemical signals,'' *Biotechnol. Bioeng.*, vol. 105, no. 1, pp. 37_43, 2010.

[27] C. Rutherglen and P. Burke, ``Nanoelectromagnetics: Circuit and electromagnetic properties of carbon nanotubes,'' *Small*, vol. 5, no. 8, pp. 884_906, 2009.

[28] M. Feiginov, R. Gonzalo, I. Maestrojun, O. Cojocari, M. Hoe_e, and E. Limiti, *THz Electronics*. Hoboken, NJ, USA: Wiley, 2015, pp. 254_303.

[29] H. Elayan, R. M. Shubair, J. M. Jornet, and P. Johari, ``Terahertz channel model and link budget analysis for intrabody nanoscale communication," *IEEE Trans. Nanobiosci.*, vol. 16, no. 6, pp. 491_503, Sep. 2017.

[30] I. F. Akyildiz and J. M. Jornet, “Electromagnetic wireless nanosensor networks,” Nano Communication Networks (Elsevier) Journal, vol. 1, no. 1, pp. 3–19, Mar. 2010.

[31] I. F. Akyildiz, F. Brunetti, and C. Bl´azquez, “Nanonetworks: A new communication paradigm,” Computer Networks, vol. 52, no. 12, pp. 2260–2279, 2008.

[32] M. Nafari and J. M. Jornet, “Metallic plasmonic nano-antenna for wireless optical communication in intra-body nanonetworks,” in Proceedings of the 10th EAI International Conference on Body Area Networks, ser. BodyNets ’15. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2015, pp. 287–293. [Online]. Available: http://dx.doi.org/10.4108/eai.28-9-2015.2261410

[33] J. M. Jornet and I. F. Akyildiz, “Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band,” IEEE Transactions on Wireless Communications, vol. 10, no. 10, pp. 3211– 3221, October 2011.

[34] Q. H. Abbasi, H. E. Sallabi, N. Chopra, K. Yang, K. Qaraqe, and A. Alomainy, “Terahertz channel characterisation inside the human skin at the nano-scale,” IEEE Transactions on THz Science and Technology, vol. 6, no. 3, pp. 427 – 434, May, 2016.

[35] A. Gupta, M. Medley, and J. M. Jornet, “Joint synchronization and symbol detection design for pulse-based communications in the thz band,” in Global Communications Conference (GLOBECOM), 2015 IEEE. IEEE, 2015, pp. 1–7.

[36] R. G. Cid-Fuentes, J. M. Jornet, I. F. Akyildiz, and E. Alarc´on, “A receiver architecture for pulse-based electromagnetic nanonetworks in the terahertz band,” in Communications (ICC), 2012 IEEE International Conference on. IEEE, 2012, pp. 4937–4942.

[37] J. M. Jornet and I. F. Akyildiz, “Femtosecond-long pulse-based modulation for terahertz band communication in nanonetworks,” IEEE Transactions on Communications, vol. 62, no. 5, pp. 1742–1754, May 2014.

[38] J. M. Jornet, J. Capdevila-Pujol, and J. Sole-Pareta, “Phlame: A physical layer aware mac protocol for electromagnetic nanonetworks in the terahertz band,” Nano Communication Networks (Elsevier) Journal,vol. 3, no. 1, pp. 74 – 81, 2012.

[39] C. Han, J. M. Jornet, E. Fadel, and I. F. Akyildiz, “A cross-layer communication module for the internet of things,” Computer Networks, vol. 57, no. 3, pp. 622–633, 2013.

[40]Abbasi, Q. H., Nasir, A. A., Yang, K., Qaraqe, K. A., and Alomainy, A. (2017). Cooperative In-Vivo Nano-Network Communication at Terahertz Frequencies. IEEE Access 5, 8642–8647. doi:10.1109/ACCESS.2017.2677498.

[41]Alomainy, A., Yang, K., Imran, M., Yao, X.-W., and Abbasi, Q. H. (2019). Nano Electromagnetic Communication at Terahertz and Optical Frequencies: Principles and Applications. Institution of Engineering and Technology.

[42]Abbasi, Q. H., Yang, K., Chopra, N., Jornet, J. M., Abuali, N. A., Qaraqe, K. A., et al. (2016). Nano-communication for Biomedical Applications: A Review on the State-Of-The-Art from Physical Layers to Novel Networking Concepts. IEEE Access 4, 3920–3935. doi:10.1109/ACCESS.2016.2593582.

[43] C. Chen, Y. Haik, J. Chatterjee, Development of nanotechnology for biomedical applications, in: Proceedings of the Emerging Information Technology Conference, August 2005.

[44] R.A. Freitas, Nanotechnology, nanomedicine and nanosurgery, International Journal of Surgery 3 (November) (2005) 243–246.

[45] T. Donaldson, 24th century medicine, Cryonics (December) (1988) 16–34.

[46] R. Freitas, What is nanomedicine?, Nanomedicine: Nanotechnology, Biology and Medicine 1 (November) (2004) 2–9.

[47] J.W. Aylott, Optical nanosensors – an enabling technology for intracelular measurements, Analyst 128 (2003) 309–312.

[48] R. Smalley, M.S. Dresselhaus, G. Dresselhaus, P. Avouris, in: M.S. Dresselhaus (Ed.), Carbon Nanotubes: Synthesis, Structure, Properties and Applications, Springer, 2001.

[49] J. Han, J. Fu, R.B. Schoch, Molecular sieving using nanofilters: past, present and future, Lab on a Chip 8 (January) (2008) 23–33.

[50] I. F. Akyildiz and J. M. Jornet. Electromagnetic wireless nanosensor networks. *Nano Communication Networks (Elsevier) Journal*, 1(1):3– 19, March 2010.

[51] S. Sengupta, M. E. Ibele, and A. Sen, “Fantastic voyage: Designing self-powered nanorobots,” Angewandte Chemie International Edition, vol. 51, no. 34, pp. 8434–8445, 2012

[52] I. F. Akyildiz, F. Brunetti, and C. Blazquez, “Nanonetworks: A new communication paradigm,” Computer Networks, vol. 52, no. 12, pp. 2260–2279, Aug. 2008.

[53] S. Hiyama, Y. Moritani, T. Suda, R. Egashira, A. Enomoto, M. Moore, and T. Nakano, “Molecular communication,” in Proc. of NSTI Nanotechnology Conference, 2005, pp. 391–394.

[54] T. Nakano, A. W. Eckford, and T. Haraguchi, Molecular communication, Cambridge University Press, 2013.

[55] I. F. Akyildiz and J. M. Jornet, "Electromagnetic wireless nanosensor networks," *Nano Communication Networks,* vol. 1, pp. 3-19, 2010.

[56] J. M. Jornet and I. F. Akyildiz, "Channel capacity of electromagnetic nanonetworks in the terahertz band," in *Communications (ICC), 2010 IEEE International Conference on*, 2010, pp. 1-6.

[57] J. M. Jornet and I. F. Akyildiz, "Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band," *Wireless Communications, IEEE Transactions on,* vol. 10, pp. 3211-3221, 2011.

[58] K. Yang, A. Pellegrini, M. O. Munoz, A. Brizzi, A. Alomainy, and Y. Hao, "Numerical Analysis and Characterization of THz Propagation Channel for Body-Centric Nano-Communications," *IEEE Transactions on Terahertz Science and Technology,* vol. 5, pp. 419 - 426, 2015.

[59] A. F. Molisch, Wireless Communications, 2nd ed. Chichester, U.K.: IEEE Press/Wiley, 2011.

[60] J. Zhang, P. Tang, L. Yu, T. Jiang, and L. Tian, “Channel measurements and models for 6G: Current status and future outlook,” Front. Inf. Technol. Electron. Eng., vol. 21, pp. 39–61, Mar. 2020.

[61] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, “6G wireless systems: Vision, requirements, challenges, insights, and opportunities,” Proc. IEEE, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.

[62] C. X. Wang, J. Huang, H. Wang, X. Gao, X. You, and Y. Hao, “6G wireless channel measurements and models: Trends and challenges,” IEEE Veh. Technol. Mag., vol. 15, no. 4, pp. 22–32, Dec. 2020.

[63] T. S. Rappaport, K. A. Remley, C. Gentile, A. F. Molisch, and A. Zajic, Radio Propagation Measurements and Channel Modeling: Best Practices for Millimeter-Wave and Sub-Terahertz Frequencies. Cambridge, U.K.: Cambridge Univ. Press, 2021.

[64] S. A. Busari, K. M. S. Huq, S. Mumtaz and J. Rodriguez, "Terahertz massive MIMO for beyond-5G wireless communications, " in IEEE Int. Conf. on Communications. Shanghai, China, pp. 1–6, 2019.

[65] A. M. Magableh, T. Aldalgamouni, S. Mater and O. S. Badarneh, "Ergodic capacity and SNR analysis for dual hop amplify and forward cooperative communication systems over $\alpha - \eta - \mu$ channels," IEEE Transactions on Vehicular Technology, vol. 67, no. 12, pp. 12384–12388, 2018.

[66] A. K. Vavouris, F. D. Dervisi, V. K. Papanikolaou, P. D. Diamantoulakis, G. K. Karagiannnidis et al., "An energy efficient modulation scheme for body-centric terahertz (THz) nanonetworks," Technologies, vol. 7, no. 1, pp. 1–9, 2019.

[67] M. K. Samimi and T. S. Rappaport, "3-D Millimeter-Wave Statistical Channel Model for 5G Wireless System Design," IEEE Transactions on Microwave Theory and Techniques, vol. 64, no. 7, pp. 2207–2225, July 2016.

[68] S. A. Busari, K. M. S. Huq, G. Felfel, and J. Rodriguez, "Adaptive Resource Allocation for Energy-Efficient Millimeter-Wave Massive MIMO Networks," in 2018 IEEE Global Communications Conference, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

[69] S. Sun, T. S. Rappaport, M. Shafi, P. Tang, J. Zhang, and P. J. Smith, "Propagation Models and Performance Evaluation for 5G Millimeter- Wave Bands," IEEE Transactions on Vehicular Technology, vol. 67, no. 9, pp. 8422–8439, Sep. 2018.

[70] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-Wave Massive MIMO Communication for Future Wireless Systems: A Survey," IEEE Communications Surveys Tutorials, vol. 20, no. 2, pp. 836–869, Secondquarter 2018.

[71] R. Piesiewicz, T. Kleine-Ostmann, N. Krumbholz, D. Mittleman, M. Koch, J. Schoebei, and T. Kurner, "Short-range ultra-broadband terahertz communications: Concepts and perspectives," *IEEE Antennas and Propagation Magazine*, vol. 49, no. 6, pp. 24–39, Dec 2007.

[72] R. M. Goody and Y. L. Yung, *Atmospheric Radiation: Theoretical basis*, 2nd ed. Oxford University Press, 1989.

[73] L. Rothman, I. Gordon, A. Barbe, D. C. Benner, P. Bernath, M.Birk, V. Boudon, L. Brown, and A.Campargue, "The hitran 2008 molecular spectroscopic database," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 110, no. 9-10, pp. 533–572, Jun 2009.

[74] N. J. Tro, *Principles of Chemistry: a molecular approach*. Prentice- Hall, Inc., 2009.

[75] J. H. Van Vleck and V. F. Weisskopf, "On the shape of collisionbroadened lines," *Rev. Mod. Phys.*, vol. 17, no. 2-3, pp. 227–236, Apr1945.

[76] S. Clough, F. Kneizys, and R. Davies, “Line shape and the water vapor continuum,” *Atmos. Res.*, vol. 23, no. 3-4, pp. 229 – 241, Oct 1989.

[77] Z. L. Wang, “Towards self-powered nanosystems: from nanogenerators to nanopiezotronics,” *Advanced Functional Materials*, vol. 18, no. 22, pp. 3553–3567, 2008.

[78] S. Xu, B. J. Hansen, and Z. L. Wang, “Piezoelectric-nanowire-enabled power source for driving wireless microelectronics,” *Nature Commun.*, vol. 1, no. 7, pp. 1–5, Oct. 2010.

[79] S. Xu, Y. Qin, C. Xu, Y. Wei, R. Yang, and Z. L. Wang, “Self-powered nanowire devices,” *Nature Nanotechnol.*, vol. 5, no. 5, pp. 366–373, May 2010.

[80] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[81] L. Vlahos, H. Isliker, Y. Kominis, and K. Hizanidis, \Normal and anomalous diffusion: A tutorial," Order and Chao, vol. 10, pp. 1-40, March 2008.

[82] I. Llatser, I. Pascual, N. Garralda, A. C. Aparicioa, M. Pierobon, E. Alarcon, and J. Sole-Pareta, \Exploring the physical channel of diffusion-based molecular communication by simulation," in Global Telecommunications Conference (GLOBECOM 2011), December 2011, pp. 1-5.

[83] N. Garralda, I. Llatser, A. Aparicio, and M. Pierobon, \Simulation-based evaluation of the diffusion-based physical channel in molecular nano networks," in 1st IEEE International Workshop on Molecular and Nano Scale Communication (MoNaCom 2011), April 2011, pp. 443 - 448.

[84] J. Philibert, \One and a half century of diffusion: Fick, einstein, before and beyond," Diffusion Fundamentals, vol. 4, pp. 6.1-6.19, 2006.

[85] A. Ladd, H. Gang, J. Zhu, and D. Weit, \Time-dependent collective diffusion of colloidal particles," Physical Review Letters, vol. 74, no. 2, pp. 318-321, January 1995.

[86] A. Eckford, \Achievable information rates for molecular communication with distinct molecules," in Bio-Inspired Models of Network, Information and Computing Systems, December 2007, pp. 313-315.

[87] A. W. Eckford, \Nanoscale communication with brownian motion," in 41st Annual Conference on Information Sciences and Systems, March 2007, pp. 160-165.

[88] S. Kadloor, R. Adve, and A. Eckford, \Molecular communication using brownian motion with drift," IEEE Trans. on NanoBioscience, July 2011.

[89] M. Gregori, I. Llatsera, A. C. Aparicioa, and E. Alarcon, \Physical channel characterization for medium-range nano-networks using agellated bacteria," Journal Computer Networks: The International Journal of Computer and Telecommunications Networking, vol. 55, pp. 779-791, February 2011.

[90] C. Shalizi, \Stochastic processes (advanced probability ii)," 2007, Carnegie Mellon University.

[91] R. Freitas, Nanomedicine, Volume I: Basic Capabilities. Georgetown, TX: Landes Bioscience, 1999.

[92] A. Einstein, Investigations on the Theory of the Brownian Movement. Dover Publications, 1956.

[93] L. Parcerisa and I. F. Akyildiz, \Molecular communication options for long range nanonetworks," Computer Networks (Elsevier) Journal, vol. 53, pp. 2753-2766, November 2009.

[94] I. F. Akyildiz, F. Brunetti, and C. Blazquez, \ Nanonetworks: A new communication paradigm," Computer Networks (Elsevier) Journal, vol. 52, pp. 2260-2279, August 2008.

[95] E. Carafoli, \Calcium signaling: A tale for all seasons," Proceedings of the National Academy of Sciences of the United States of America, vol. 99, no. 3, pp. 1115-1122, 2002.

[96] T. S. T. Nakano, M. Moore, R. Egashira, A. Enomoto, and K. Arima, \Molecular communication for nanomachines using intercellular calcium signaling," in Proceedings of 2005 5th IEEE Conference on Nanotechnology Nagoya, Japan, July 2005.

[97] J. M. Jornet and I. F. Akyildiz, “Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band,” *Wireless Communications, IEEE Transactions on*, vol. 10, no. 10, pp. 3211–3221, 2011.

[98] K Weise and W Woger, “Comparison of Two Measurement Results Using the Bayesian Theory of Measurement Uncertainty”, Measurement Science and Technology, vol. 5, pp. 879-882, 1994.

[99] D.D. Lewis, “Naïve (Bayes) at Forty: The Independence Assumption in Information Retrieval”, AT&T lab Research, NJ, USA.

[100] I Rish, “An Empirical Study of the Naive Bayes Classifier”, Proceedings of IJCAI-01, New Zealand, 2001.

[101] S Eyheramendy, D Lewis, D Madigan, “On the Naive Bayes Model for Text Categorization”, Proceedings Artificial Intelligence & Statistics, 2003.

[102] Fix, E. and Hodges, J.L. (1951). Discriminatory analysis. Nonparametric discrimination; consistency properties. Technical Report 4, USAF School of Aviation Medicine, Randolph Field, TX, USA, 1951.

[103] Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13 (1), 21–27.

[104] Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., et al. (2008). Top 10 algorithms in data mining. Knowledge and information systems, 14 (1), 1–37.

[105] Bhatia, N., & Vandana. (2010). Survey of Nearest Neighbor Techniques. International Journal of Computer Science and Information Security, 8 (2), 302–305.

[106] Xu, S., & Wu, Y. (2008). An algorithm for remote sensing image classification based on artificial immune B-cell network. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37, 107–112.

[107] Manne, S., Kotha, S., & Fatima, S. S. (2012). Text categorization with K-nearest neighbor approach. In Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2012 (INDIA 2012) held in Visakhapatnam, India, January 2012 (Vol. 132, pp. 413–420). Springer.

[108] Xiubo, G., Tie-Yan, L., Qin, T., Andrew, A., Li, H., & Shum, H.-Y. (2008). Query dependent ranking using k-nearest neighbor. In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval (pp. 115–122).

[109] Bajramovic, F., Mattern, F., Butko, N., & Denzler, J. (2006). A comparison of nearest neighbor search algorithms for generic object recognition. In Advanced Concepts for Intelligent Vision Systems (pp. 1186–1197). Springer.

[110] Yang, Y., Ault, T., Pierce, T., & Lattimer, C. W. (2000). Improving text categorization methods for event tracking. In Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval (pp. 65–72).

[111] Kataria, A., & Singh, M. D. (2013). A Review of data classification Using K-nearest neighbour Algorithm. International Journal of Emerging Technology and Advanced Engineering, 3 (6), 354–360.

[112] Wettschereck, D., Aha, D. W., & Mohri, T. (1997). A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. Artificial Intelligence Review, 11 (1-5), 273–314.

[113] M. PAL and P. M. MATHER School of Geography, University of Nottingham, UK (Received 8 April 2002; in final form 7 August 2004)

[114] R. A. Johnson and D. W. Wishern, Applied Multivariate Statistical Analysis, Prentice Hall, London, 1988.

[115] I. E. Frank and J. H. Friedman, J. Chemom., 1989, 3, 463–475.

[116] S. J. Dixon, Y. Xu, R. G. Brereton, A. Soini, M. V. Novotny, E. Oberzaucher, K. Grammer and D. J. Penn, Chemom. Intell. Lab. Syst., 2007, 87, 161–172.

[117] M. Baker and W. Rayens, J. Chemom., 2003, 17, 166–173.

[118] (Himani Bhavsar, Mahesh H. Panchal International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 10, December 2012.

[119] R. G. Brereton, Chemometrics for Pattern Recognition, Wiley, Chichester, 2009

[120] L. Bottou, C. Cortes, J. Denker, H. Drucker, I. Guyon, L. Jackel, Y. LeCun, U. Muller, E. Sackinger, P. Simard and V. Vapnik, in International Conference on Pattern Recognition, IEEE Computer Society Press, 1994, vol. 2, pp. 77–82.

[121] G.B. Giannakis, On the identifiability of non-Gaussian ARMA models using cumulants, IEEE Transactions on Automatic Control 35 (1) (1990) 18–26.

[122] G.B. Giannakis, Wavelet phase reconstruction using cumulants, IEEE Transactions on Geoscience and Remote Sensing 27 (4) (1989) 452–455.

[123] S. Safi, A. Zeroual, Blind identification in noisy environment of nonminimum phase finite impulse response (FIR) systems using higher order statistics, Systems Analysis Modelling Simulation 43 (5) (2003) 671–681.

[124] J. Antari, R. Iqdour, A. Zeroual, Forecasting the wind speed process using higher order statistics and fuzzy systems, Energies Renouvelables 9 (4) (2006) 237–251.

[125] https://kobia.fr/quest-ce-que-la-variance/ 19 janv. 2022

[126] https://byjus.com/maths/fourier-transform/ 2023, BYJU'S. All rights reserved.

[127] S. Sun, G. R. MacCartney, Jr., and T. S. Rappaport, "A novel millimeterwave channel simulator and applications for 5G wireless communications," in 2017 IEEE International Conference on Communications (ICC), May 2017, pp. 1–7.