INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Molecular Communications techniques for the Internet of Bio-Nano Things

Sofia Alexandra Duarte de Figueiredo

Master Degree in Telecommunications and Computer Engineering

Supervisor
PhD Nuno Manuel Branco Souto, Associate Professor (with Habilitation)
Iscte-IUL

Co-Supervisor
PhD Francisco António Bucho Cercas, Full Professor
Iscte-IUL

November, 2021

TECNOLOGIAS
E ARQUITETURA

Department of Information Science and Technology

Molecular Communications techniques for the Internet of Bio-Nano Things

Sofia Alexandra Duarte de Figueiredo

Master Degree in Telecommunications and Computer Engineering

Supervisor
PhD Nuno Manuel Branco Souto, Associate Professor (with Habilitation)
Iscte-IUL

Co-Supervisor
PhD Francisco António Bucho Cercas, Full Professor
Iscte-IUL

November, 2021

Acknowledgments

I would like to express my deepest appreciation and great gratitude for my supervisors. To PhD Nuno Souto who introduced me to this field of research and helped me to through all my difficulties. And, to PhD Francisco Cercas who continuously provide me with precious and essential comments and feedback. They were both extremely essential to keep me always motivated and making my delight for research grow. I did this dissertation with their help and it was a privileged to work with this two extraordinary professors.

My gratitude is extend to ISCTE - Instituto Universitário de Lisboa, which was a second home for me, giving me always the best space to learn and grow my knowledge. Also, to Instituto de Telecomunicações (IT) for the research program and for supplying me with the resources necessary to my work.

I would like to give a special thanks to my friends for hearing me and advising me to always try my best. They are very important and I really appreciate all the help and friendship they gave me.

Finally, I want to dedicate this thesis to my family who always provided all the tools necessary for me to be successful in my academic life, as also encouraged me in all the steps of my work. Their support and love were crucial for me.

Resumo

A "Internet das Coisas" Bio-Nano é um novo paradigma de rede definido como a interconexão de dispositivos nano escala. Este é um conceito revolucionário que espectavelmente permitirá uma vasta gama de aplicações. Em particular, prevê-se que os sistemas de saúde sejam transformados com a integração de redes centradas no corpo, em futuras gerações de sistemas de comunicação. Neste contexto, as comunicações moleculares (CM) emergem como a forma mais promissora de transmitir informação, devido ao facto de serem intrinsecamente biocompatíveis, eficientes em termos energéticos e robustos em condições fisiológicas.

Um dos maiores desafios é como minimizar os efeitos do ruído ambiental e reduzir a interferência intersimbólica que pode ser muito elevada num cenário de CM por difusão. A codificação de canal é um dos tipos de técnicas mais promissoras para abordar este problema. Este trabalho baseia-se na avaliação da modulação, da deteção e de novos esquemas de codificação energeticamente eficientes e de baixa complexidade aplicados em CM. Com especial foco, na implementação de códigos Tomlinson, Cercas, Hughes (TCH) como uma nova abordagem para um ambiente de CM, devido às suas particulares propriedades das palavras de código, que permitem uma deteção simplificada. Os resultados das simulações mostram que os códigos TCH são mais eficazes para estes cenários quando comparados com outras alternativas existentes, sem introduzir demasiada complexidade ou poder de processamento no sistema.

Adicionalmente, é descrita uma experiência macroscópica, que utiliza o pH como portador de informação, demonstrando que os códigos TCH propostos podem melhorar a fiabilidade para CM.

Palavras-chave: futuras redes sem fios; comunicações moleculares; difusão; codificação de canal; códigos TCH

Abstract

The "Internet of Bio-Nano Things" (IoBNT) is a new networking paradigm defined as the interconnection of nanoscale devices. IoBNT is a revolutionizing concept that will likely enable a wide range of applications, in particular, it is envisioned that healthcare systems will be transformed with the development and integration of body-centric networks into future generations of communication systems. Within this context, molecular communications (MC) emerge as the most promising way of transmitting information for in-body communications, due to being inherently biocompatible, energy-efficient, and robust in physiological conditions.

One of the biggest challenges is how to minimize the effects of environmental noise and reduce intersymbol interference (ISI) which can be very high in an MC via diffusion scenario. Analogous to traditional communications, channel coding is one of the most promising types of techniques for addressing this problem. This work is based on the study and evaluation of novel energy efficient and low complexity coding, modulation and detection schemes for MC. With a special focus on the implementation of Tomlinson, Cercas, Hughes (TCH) codes as a new attractive approach for the MC environment, due to the particular codeword properties which enable simplified detection. Simulation results show that TCH codes are more effective for these scenarios when compared to other existing alternatives, without introducing too much complexity or processing power into the system.

Furthermore, an experimental macroscale proof-of-concept is described, which uses pH as the information carrier and demonstrates that the proposed TCH codes can improve the reliability in this type of communication channel.

Keywords: future wireless networks; molecular communications; diffusion-based; channel coding; TCH codes

Contents

Acknowledgments i

Resumo iii

Abstract v

List of Figures xi

List of Tables xiii

List of Acronyms xv

List of Symbols xvii

Chapter 1. Introduction 1
1.1. Project background and motivation 1
1.2. Investigation questions and objectives 2
1.3. Contributions 3
1.4. Dissertation structure 4

Chapter 2. State of the Art and Fundamental Concepts 5
2.1. MC body-centric networks 5
2.2. Propagation techniques: Physical Principles 6
2.2.1. Diffusion 7
2.2.2. Advection 8
2.2.3. Reactions 8
2.3. System components architecture 9
2.3.1. Transmitter 9
2.3.2. Receiver 10
2.4. Physical Channel 11
2.4.1. Channel models 12
2.4.2. Addictive Noise and ISI Models 12
2.4.3. Correlation 12
2.5. Channel coding: Literature channel codes applied in MC 13
2.6. Modulation 15

2.7. Experimental works 16
2.7.1. Air-driven 17
2.7.2. Liquids-driven 17

Chapter 3. System model 19
3.1. Encoder and decoder: TCH Codes 19
3.2. Pulse generator 22
3.2.1. Types of pulses 22
3.2.2. Modulation: Concentration Shift Keying 23
3.3. Diffusion based channel 24
3.4. Detection and bit decision 26
3.4.1. Threshold-based hard decision 26
3.4.2. LLR-based soft decision 27

Chapter 4. Experimental architecture 29
4.1. Electronic components 30
4.1.1. Arduino UNO 30
4.1.2. Pump 31
4.1.3. Pump driver 32
4.1.4. pH sensor 33
4.2. Chemicals 34
4.3. Transmitter 35
4.3.1. Hardware 35
4.3.2. Software 36
4.4. Channel 40
4.5. Receiver 40
4.5.1. Hardware 40
4.5.2. Software 41

Chapter 5. Results and discussion 45
5.1. Simulation results 45
5.1.1. Simulation setup 45
5.1.2. Effects of symbol duration and type of pulses 46
5.1.3. Modulation techniques 48
5.1.4. Comparsion between different codes 49

5.1.5. Effect of code length 50
5.2. Experimental results 51
5.2.1. Experimental system validation 51
5.2.2. Overall system performance 55

Chapter 6. Conclusions 57
6.1. Future work 58

Bibliography 59

Appendix 65

List of Figures

2.1 Resume representation of a MCvD system. 5
2.2 IoBNT example in the human body, from [**11**]. 6
2.3 Propagation techniques in MC, adapted from [**16**]. 6
2.4 a) Point transmitter b) Volume Transmitter c) Ion-channel-based transmitter, adapted from [**22**]. 9
2.5 a) Passive receiver b) Absorbing receiver c) Reactive receiver, adapted from [**22**]. 11
2.6 Modulation techniques. 15

3.1 System block diagram. 19
3.2 Typical autocorrelation function of a B-TCH polynomial with n=256, [**53**]. 21
3.3 TCH decoder of maximum likelihood [**54**]. 22
3.4 Type of pulses. 23
3.5 Possible representation for the bit stream "01110010", regarding different types of CSK modulation. 24
3.6 Example illustrating the expected and observed concentrations for a diffusion-based MC channel, assuming number of molecules (N_{molec})=10000, d_0=0.3 µm, D=450 µm^2/s and a_{rx}=50 nm. 25

4.1 (a) Schematic diagram of the experimental implementation; (b) Photo of the experiment. 29
4.2 Arduino Uno. 30
4.3 12V Aquarium pump. 31
4.4 L298N motor driver. 32
4.5 L298N schematics, from [**66**]. 32
4.6 pH probe, from [**68**]. 33
4.7 pH calibration liquids. 33
4.8 (a) pH- used; (b) pH+ used. 34
4.9 Transmitter scheme. 35
4.10 Transmitter circuit. 36

4.11 How bit-0 and bit-1 are transmitted. 37
4.12 Transmitter workflow for the uncoded case. 38
4.13 Transmitter workflow for the TCH codes. 39
4.14 Receiver scheme. 40
4.15 Receiver workflow for the uncoded case. 42
4.16 Receiver workflow for the TCH codes. 43

5.1 BER comparison for all the considered channel coding methods, using BCSK modulation and with τ=2. 47
5.2 BER comparison for all the considered channel coding methods, using BCSK modulation and with τ=10. 48
5.3 BER comparison for all the considered channel coding methods, using OOK modulation. 49
5.4 TCH codewords length. 50
5.5 Calibration Arduino IDE serial monitor, code from [**68**]. 51
5.6 pH behavior in a MCvD channel. 52
5.7 Initial reception delay. 53
5.8 Example of 10 received bits with T_s=40s after initial delay removal. . . 54
5.9 Uncoded and TCH(16,10) performance for Ts=20s and Ts=40s. 56

List of Tables

1.1 Possible applications for IoBNT. 2

3.1 Fermat numbers for generating TCH codes [**53**]. 21

5.1 Simulation parameters. 45

List of Acronyms

ADC Analog to Digital Converter

B-TCH Basic TCH Polynomials
BCSK Binary Concentration Shift Keying

C-RM Cyclic Reed-Muller
CLT Central Limit Theorem
CSI Channel State Information
CSK Concentration Shift Keying

DNA Deoxyribonucleic Acid

EG-LDPC Euclidean Geometry Low Density Parity Check

FSK Frequency Shift Keying

HC Hamming Codes

IDE Integrated Development Environment
IoBNT Internet of Bio-nano Things
ISI Intersymbol Interference
ISI-mtg ISI-mitigation

LLRs Log-likelihood Ratios

MC Molecular Communications
MCvD Molecular Communications via Diffusion
MIMO Multiple-input and Multiple-output
MoSK Molecular Shift Keying

Na2Co3 Sodium Carbonate
NaHSO3 Sodium Bisulfate
NRZ Non-return to Zero
NSK Nucleotide Shift Keying

OOK On-Off Keying

PPM Pulse Position Modulation

RNA Ribonucleic Acid

RS Reed Solomon

SOCCs Self-Orthogonal Convolutional Codes

TCH Tomlinson, Cercas, Hughes

List of Symbols

η	Solvent viscosity
∇^2	Laplace operator
τ	Scaling factor
a	A scaling factor
A_0	Level for 0
A_1	Level for 1
a_{rx}	The radius of the spherical receiver
C	Concentration of the diffusing molecules
$c(d,t)$	conserved quantity described by a scalar field
D	Diffusion coefficient
d	The distance between the center of the transmitter and the center of the receiver
d_{min}	Minimum distance
d_{min}^{TCH}	TCH minimum distance
F_i	Fermat numbers
$F_{hit}(t_u, t_l)$	The probability of a molecule emitted at time t=0 to be observed at the receiver during interval [t_l, t_u]
$h[l,k]$	The number of molecules observed by the absorbing receiver during time interval k after N_{molec} are released from the transmitter at time $k-l$
H_d	Hamming distance between any two polynomials
i	Comprising codeword for each symbol
k	Information message length
k_B	Boltzmann constant
L	The assumed length of the ISI
m	Parity check bits
$m[k]N^{tx}$	The number of molecules release by the transmitter at $k-th$ symbol slot
N	Number of symbols
n	Code block length
$n[k]$	The number of noise molecules detected by the receiver at time k
N_{molec}	Number of molecules
p	A prime number

P_i	Theoretically generated set of Base Polynomials
R_p	Radius of the diffusing molecule
S	A set of CSK modulation symbols
T	Temperature
t	The number of errors that can be correct in each block
T_s	Symbol duration
t_{TCH}	TCH error-correcting capacity
$v(d,t)$	Flow velocity
$x[k]$	A concentration based binary modulation symbol emitted from the transmitter at time k
$x^i[k]$	Concentration-based binary modulation
$y[k]$	The concentration vector corrupted by noise at the k^{th} time interval

Chapter 1

Introduction

1.1. Project background and motivation

The telecommunications domain is always developing, facing all types of new demands and therefore, while fifth generation (5G) wireless networks are currently being deployed around the world, work is already underway for future evolutions such as sixth generation (6G) systems [**1**] and even beyond. One of the many new trends is nanoscale applications. Emerging as a new network paradigm is the "Internet of Bio-nano things" (IoBNT), defined as the interconnection of nanoscale devices. As in Table 1.1 this is a revolutionizing concept promising many applications [**2**–**5**], in the biomedical field, medicine, industry, environmental, agriculture and military. Many of these scenarios rely on macroscale IoBNT which is expected to evolve much faster than in-body communications and may even allow integration into the next generation of wireless networks, namely into 6G [**6**]. In the case of in-body applications, the communications in IoBNT need to be biocompatible, energy-efficient and robust in physiological conditions. This way, electromagnetic signals became a hard first choice due to issues related to biocompatibility, power and possible health hazards. An alternative to this concern is employing Molecular Communications (MC), using molecules for encoding, transmitting and receiving information. MC became the choice for these systems because:

- It is feasible, being considered easier to implement than other approaches in the near term.
- It is scalable, having an appropriate size for nanomachines.
- It is bio-compatible, allowing the integration with living systems.

Due to the growing interest in nanoscale communications and MC, IEEE has already started standardization efforts through the IEEE P1906.1 working group [**6**, **7**]. Nevertheless, MC for in-body networks is still in its infancy and it will take some time before practical systems become a reality, probably not before a seventh or higher generation (7G) of wireless networks [**6**]. In the last decade MC has been a challenge, with the investigation focusing in developing theoretical models of MC systems, by exploring the communications channels, devising new modulation and detection techniques. MC has

been captivating a lot of research interest with several novel transmitter/receiver architectures being proposed in recent years [**11**].

Applications	Examples
Medicine and healthcare	Targeted drug delivery, cooperatively release medication; Health monitoring, identify the presence of toxic substances; Regenerative medicine, rebuild damaged tissues or organs; Genetic engineering, manipulating DNA.
Industry	Quality control, identification of product defects.
Environmental	Environmental monitoring, detection of pollutants or toxins; Degradation, safe conversion of undesired materials.

TABLE 1.1. Possible applications for IoBNT.

1.2. Investigation questions and objectives

Despite MC has been developing for billions of years in nature, it is hard to control and manipulate in non-natural ways. This leads us to face a lot of challenges and concerns in order to use them in a functional communication system. Some of the identified challenges are:

- the stochasticity, where molecules have random propagation called environmental noise,
- the delay, because molecules propagation times are very long compared to electromagnetic waves,
- the range, since known techniques have very short practical ranges,
- the fragility, because biological components can be environmentally sensitive (namely to temperature, to pH and to other reagents).

These conditions pose many challenges for achieving a reliable communication. Therefore, channel coding became an important solution to help reduce these effects. This is achieved by adding redundancy to the emitted message and by making it possible to correct the errors caused by noise during transmission at the receiver, improving the reliability of the communication link.

Motivated by this, the objectives of this work are to explore ways to encode, transmit and decode data by using molecules as information carriers. For that intent, we will test

different type of pulses (non return to zero and Manchester pulses) and we will also use some known methods of encoding and decoding information borrowed from traditional wireless communications. In particular, we study the Tomlinson, Cercas, Hughes (TCH) codes, as a new concept applied to MC systems. These codes possess several properties allowing us to use them efficiently in various applications, such as error correction, synchronism, spread spectrum systems and channel estimation. The main goal of these codes is to achieve good coding gains with low implementation complexity and a smaller auto-correlation distribution, that is, higher minimum distances, while still supporting soft-decision decoding for best performance [**52**–**54**]. We will be adapting those into MC, as our main effort to improve the reliability of the system. Furthermore, an MC macro experiment will be assembled with the objective to bridge the gap between theoretical and experimental MC systems. Both, simulated and experimental systems, efficiency will be evaluated by analyzing the bit error rate (BER) and the permanence of the molecules in the channel.

1.3. Contributions

We provide a state of art in MC macroscale experiments. Also, an overview of known MC techniques regarding an end-to-end architecture. Throughout this work, we will focus on the MC via diffusion (MCvD), where the transmitted molecules propagate through passive diffusion along concentration gradients. This method of propagation is the most widely utilized for MC in the literature, because the molecule's movement do not require additional complexity or energy consumption.

We propose adequate symbol detection schemes for the MC channel which enable low complexity hard and soft decision TCH decoding. The performance of the resulting MC system is evaluated through an end-to-end simulator and also through a macroscale experimental prototype. The TCH codes BER in MC will be analyzed and compared with other existing alternatives proposed in the literature. Different pulse types will be considered in the evaluation. Regarding the macro-scale experiment, we adopt pH as an information carrier combined with Manchester pulses.

Moreover, the results obtained within the development of this project were submitted to a scientific journal. The journal was "Sensors" being a leading international, peer-reviewed, open access journal on the science and technology of sensors. The submitted article can be found on the Appendix A.

1.4. Dissertation structure

This dissertation is structured as follows: In Chapter 2, we do a state of art regarding macroscale MC experiments and a literate review of the different approaches to MC, particularly, a description of the components and techniques used in end-to-end MC systems. In Chapter 3, we describe in detail the considered system model and the used techniques. In Chapter 4, we present our developed experimental system architecture, with details on the design and the features of all the components. In Chapter 5, we simulate and experiment with the designed MC system and discuss the obtained results. Finally, in Chapter 6, we present the overalls dissertation conclusions and we also discuss possible future improvements for the developed MC system.

Chapter 2

State of the Art and Fundamental Concepts

This chapter presents the main concepts of an MC system and also the developed work so far in this area. The focus is describing all the processes represented in Figure 2.1, being them the most important for a good understanding of an MC system. This is based on the referenced work from the literature review and includes a brief theoretical explanation about the areas of study.

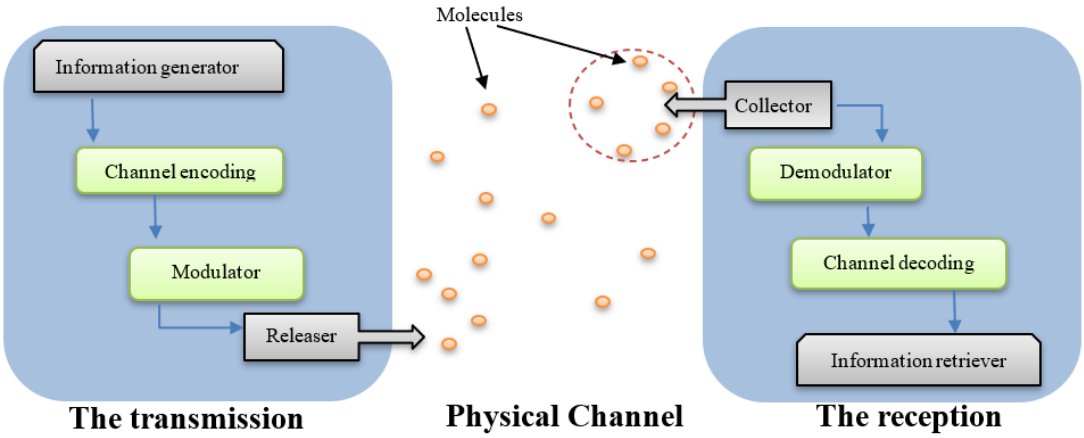


FIGURE 2.1. Resume representation of a MCvD system.

2.1. MC body-centric networks

The human body is a great source of energy, for instance, thermoelectric from body heat, vibrational from heartbeats and biochemical from perspiration. This opens many possibilities to several MC microscales/nanoscales implementations, as exemplified in Figure 2.2. Carrying information through molecules is a natural way to implement IoBNT in the human body. Utilizing nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are the most promising ways, because they are biocompatible and chemically stable [**3**,**8**–**10**]. Encoded information in DNA sequences can be known as nucleotide shift

keying (NSK). In the literature [**11**,**12**], NSK can be a potential method to boost the typically low data rates in MC up to competing with the traditional wireless communication. Carrying information through molecules in the human body can also be accomplished by varying the concentration of elemental ions, such as Na+ ions, K+ and Ca2+, which are present in many biological human systems [**13**,**14**]. Also, neurotransmitters, proteins and other molecules can be considered as potential information messengers for IoBNT [**15**].

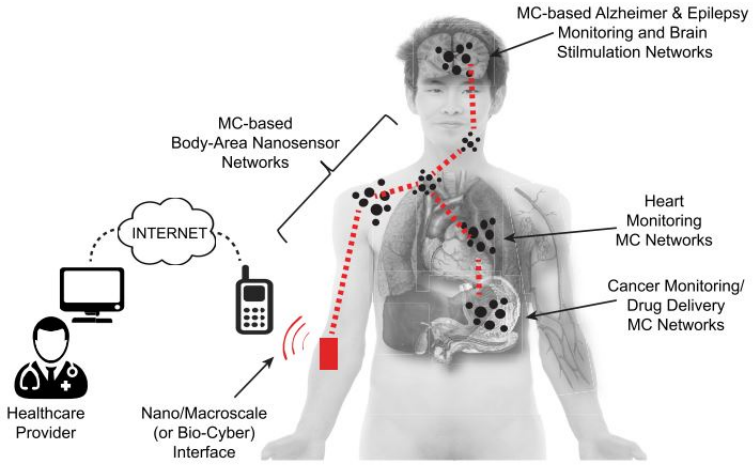


FIGURE 2.2. IoBNT example in the human body, from [**11**].

2.2. Propagation techniques: Physical Principles

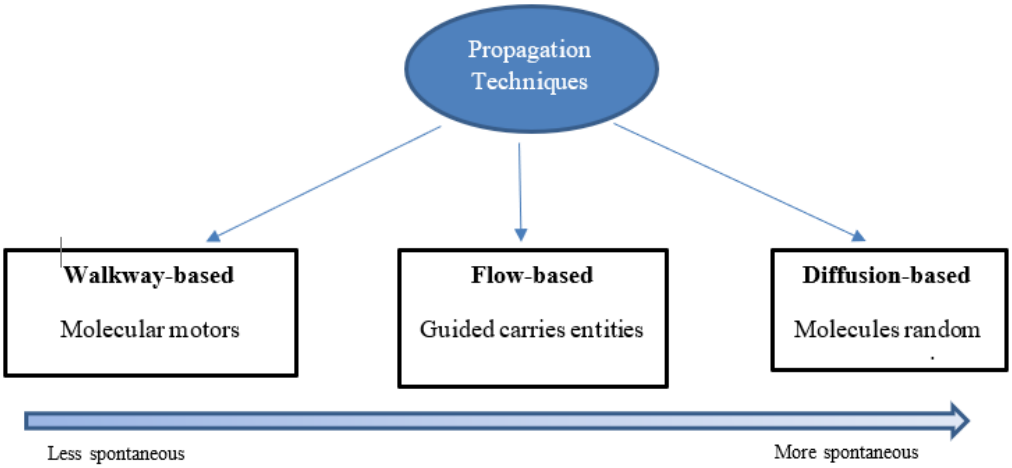


FIGURE 2.3. Propagation techniques in MC, adapted from [**16**].

One of the key challenges in molecular communications is to characterize how molecules propagate through the medium and, for a good perception of how this happens, it is

necessary to analyse some chemical and physical foundations happening at the molecular level. As Figure 2.3 shows, there are three main architectures based on the type of molecule propagation as described in [**16**, **17**]:

1) Walkway-based, where the molecules propagate in predeterminate pathways connecting the transmitter to the receiver through carrier substances;

2) Flow-based, having the molecules propagating through a guided and predictable flow;

3) Diffusion-based, where molecules propagate without any specific direction but according to a concentration gradient. This method is the most studied and used in molecular communications, this is also the propagation technique that we will implement in this work.

In the following sections, the diffusion-based propagation and other physical effects that influence it, will be clarified.

2.2.1. Diffusion

Molecular diffusion is the thermal movement of molecules in a fluid medium such as liquid or gas, the molecules flux happens from higher to lower concentration levels. The rate of this movement is a function of the environment temperature and the mass of the particles. When the concentration of these molecules becomes homogeneous in a determined fluid (dynamic equilibrium), the molecules continue to move, but since there is no concentration gradient, their movement becomes random. The resulting movement of the molecules without any specific direction is named Brownian motion: an uncontrolled movement of particles in a fluid as they constantly collide with other molecules. Diffusion is the most popular method in MC, because it is seen many times in nature, such as in synaptic communication, quorum sensing and Ca2+ signalling [**18**].

A mathematical description of diffusion is possible as described in [**19**–**21**]. Adolf Fick derived a deterministic law in 1855 based on the diffusion coefficient D (m^2/s). The random movement of molecules due to diffusion leads to the variation in C across time and space, which denotes the concentration of the diffusing molecules i.e., the average number of solute molecules per unit volume (in m^{-3}). That leads us to Fick's second law of diffusion, where ∇^2 is the Laplace operator

$$\partial C/\partial t = D\nabla^2 C. \tag{2.1}$$

The value of the diffusion coefficient depends on the environment as well as the shape and the size of the particle. For immersed spherical particles, the diffusion coefficient can be determined based on the Einstein relation, where k_B is the Boltzmann constant ($1.38\text{x}10^{-23}$ J/K), T is the temperature in Kelvin, η is solvent viscosity (Kg.m^{-1}.s^{-1}) and R_p is the radius of the diffusing molecule (in m)

$$D = \frac{(k_B T)}{6\pi\eta R_p}. \tag{2.2}$$

The particles movement can be modelled by a Wiener Process and can be accurately simulated by Monte Carlo simulation.

2.2.2. Advection

Considering that this type of communications can happen in a fluid such as blood or water, it is important to review another fundamental mechanism for particle transport in a fluid environment called advection. Transport by advection can happen by a for-induced drift or a bulk flow.

In [**22**], the flow velocity is described for advection as a vector, $v(d,t)$ that depends on position d and at time t. The change in concentration, related to time due to advective transport, is modelled by the following equation, where $c(d,t)$ is a conserved quantity described by a scalar field and $\nabla = [(\partial/(\partial x)), (\partial/(\partial y)), (\partial/(\partial z))]$

$$\partial c(d,t)/\partial t = -\nabla.(v(d,t)c(d,t)). \tag{2.3}$$

In many application scenarios, advection and diffusion are both present in the MC environment. The combined effects of both advection and diffusion are characterized by the following equation, known as the advection–diffusion effect

$$\partial c(d,t)/\partial t = D\nabla^2 c(d,t) - \nabla.(v(d,t)c(d,t)). \tag{2.4}$$

2.2.3. Reactions

Other reactions may be considered, such as molecule degradation, where, in certain moments, one specific molecule ends up disappearing. This is a natural feature of some types of molecules and its effects must be accounted for the communication design. In typical scenarios, the speed of natural degradation might be too slow compared to the desired time scale for communications. In this case, enzymes can be used to accelerate the reaction process [**22**].

2.3. System components architecture

This section will evaluate the components of the system in Figure 2.1. The overall characterization of this system is similar to traditional radio communication systems, where we have an end-to-end chain: a transmitter, a channel and a receiver. However, the way these components are characterized in MC is significantly different from other known characterizations in electromagnetic communications. Therefore, existing models for MC transmitters, channels and receivers, will be clarified. Each of these components has unique features and roles in the system, which are outlined in the following sections.

2.3.1. Transmitter

The transmitter is responsible for processing the information to be sent. For this, data must be mapped through source coding, representing the information with the adequate number of bits and channel coding to introduce additional information bits providing redundancy and error correction capability. Then, the modulator unit encodes the information according to the physical properties of the molecules such as concentration, type, ratio, or order. Finally, controlled by a releasing mechanism, data is sent into a designated channel [**11**]. As explained in [**23**], the transmitter physical components must include a processing unit which controls the communication system and the information encoding by giving instructions to the particle generation, a particle producer or reservoir where molecules being transmitted are stored and a release mechanism which controls the release of the particles into the medium.

The following points will present some of the existing transmitters models for MC. It will be explained their geometry and release mechanisms to the medium and how they can affect the end-to-end system, as described in [**22**].

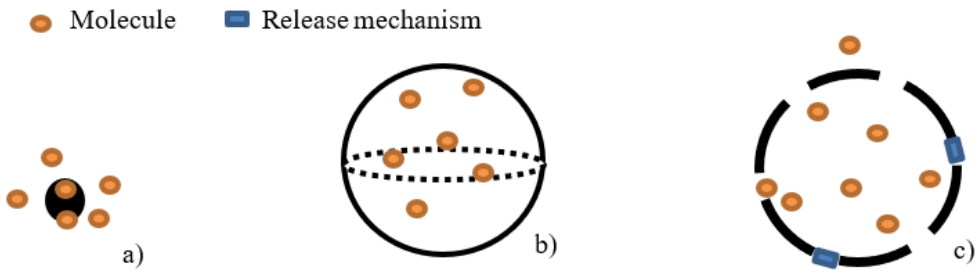


FIGURE 2.4. a) Point transmitter b) Volume Transmitter c) Ion-channel-based transmitter, adapted from [**22**].

a) Point transmitter [**24**]: The point transmitter, represented in Figure 2.4a, is the simplest and the most used in MC transmitter models. Being a point transmitter means that geometry is not considered when simulating the system or in the physical transmitter. In this model, we just assumed that molecules are produced and enter the physical channel immediately.

b) Volume transmitter [**25**]: Unlike the point transmitter, in the volume transmitter we take geometry into consideration by assuming that the molecules are uniformly distributed over the transmitter's volume, we assume a 3-D model, where molecules are physical quantities that occupy space. This model does not take the release mechanism into account and it is depicted in Figure 2.4b.

c) Ion-channel-based transmitter [**26**]: In Figure 2.4c presents the ion-channel-based transmitters. These transmitters take into account both the transmitter geometry and the release mechanism of the signalling molecules to the physical channel. These transmitters are modelled as spherical objects with ion channels embedded in their membrane and are controlled via a gating parameter. For example, taking voltage as the gating parameter, when a certain voltage is applied across the transmitter membrane, the ion channels open and a number of molecules leave the transmitter via passive diffusion.

2.3.2. Receiver

The receiver is responsible for receiving information and decoding it into the original data. For that purpose, it starts by recognizing, through external sensory units, the arrival of molecules and detects the encoded information according to a physical property of these molecules, such as concentration, type, ratio, or order. Then, a processing unit is needed to detect and decode the information, by analysing the output of the sensor measurements. For MC as explained in [**23**], the receiver physical components must include a processing unit which is responsible for controlling the communication system, processing and decoding the information. Additionally, it includes an information particle detector that detects and senses the information particles.

The following topics summarize the architectures of MC receivers, as described in [**11**, **22**].

a) Passive receiver [**27**–**29**]: The model of passive receiver is the simplest to use, since the physical sampling process is based on sampling the instantaneous number of molecules inside the reception space, at a specific sampling time. This way, research can focus only on the transport of the messages into the receiver location. This transmitter is assumed

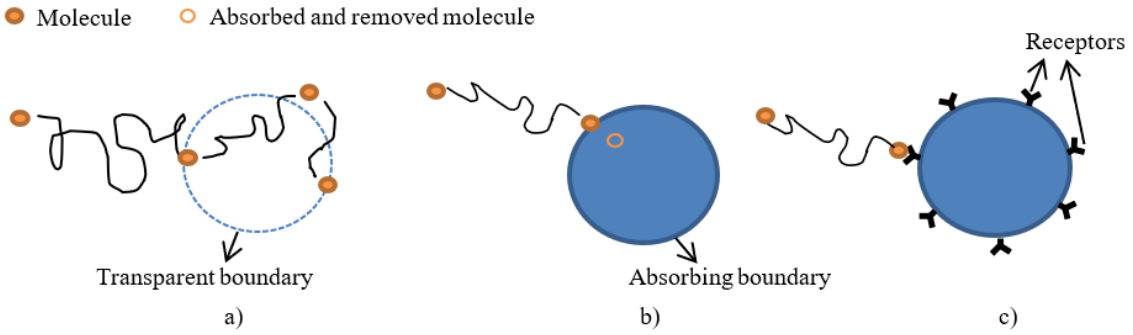


FIGURE 2.5. a) Passive receiver b) Absorbing receiver c) Reactive receiver, adapted from [**22**].

as having a transparent membrane to all kind of molecules. The problem with this model is that the same molecules can be measured multiples times because of their random movement, in and out of the measuring area. This model is represented in Figure 2.5a.

b) Absorbing receiver [**30**–**33**]: The absorbing receiver, shown in Figure 2.5b, is considered spherical, absorbing and degrading every single molecule that hits its surface. This is a more realistic scenario compared to a passive receiver because it includes a physical interaction between the receiver and the molecules. The perfect absorbing receiver would have a very high concentration of receptors with infinitely high absorption rate, in such a way that every molecule hitting the surface is consumed instantly. Unlike passive receivers which have sampling-based detectors, absorbing receivers are energy-based detectors, related to the total number of molecules absorbed by the receiver during a prespecified time period, which is usually the symbol interval.

c) Active receiver [**34**, **35**]: The reactive receiver approach, Figure 2.5c, is the most realistic in the sense that natural cells such as bacteria or neurons detect MC signals through their receptors on the cell's membrane, therefore many types of artificial biosensors work with biological receptors for higher selectivity. Here the molecules are processed via specialized receptor proteins or enzymes. Diffusion-based MC systems with reactive receivers, in most cases, can be considered as reaction–diffusion systems with finite reaction rates.

2.4. Physical Channel

Next, we will present the properties of the physical channel. We will describe deterministic and statistical models related to the received molecules. In addition, we also review the ISI and the correlation of the received signal.

2.4.1. Channel models

Channel models are described in [**22**, **23**] and can be classified as:

1) Deterministic model: it characterizes the expected number of molecules observed at the receiver.

2) Statistic model: it characterizes the actual number of observed molecules at the receiver. The following statistical models are the most commonly adopted for the number of molecules observed at the receiver for time-invariant channels:

a) Binomial model, based on the independent molecule behaviour assumption and, since any given molecule released by the transmitter is either observed by the receiver or not, a binary state model applies and the number of observed molecules following the binomial distribution;

b) Gaussian model: If the expected number of molecules observed at the receiver, is sufficiently large, then we can apply the central limit theorem (CLT). The Gaussian distribution is much more amenable to analyse than the binomial distribution;

c) Poisson model: For cases when the number of trials is large and the mean of the binomial is small, the binomial distribution can be well approximated by a Poisson distribution with the same mean.

2.4.2. Addictive Noise and ISI Models

In terms of addictive noise as in [**36**], for MC, noise can arise, for example from inconsistent pulse emissions, physical disturbances, medium influence and receiver and transmitter design.

Since diffusion is the way molecules propagate, which consists in random movements, it becomes impossible to predict the accurate arrival time of information particles at the receiver. These particles remain in the medium until they reach the receiver or until totally degradation of themselves. Hence, some particles can reach the receiver in future time slots, therefore causing interference to the corresponding symbol. This may cause an ISI increment and it can significantly degrade the channel performance. The ISI can be modelled as a sum of random variables, each following a gaussian distribution or as an additive noise modelled using a Poisson-Binomial probability mass function [**37**].

2.4.3. Correlation

Correlation is considered in [**22**] according to two types:

1) Sample correlation: For a good correlation between emitted and received information we must assume that the number of molecules, counted at different time instants, is independent of each other within one symbol interval or in different symbol intervals. The sampling interval is chosen to be large enough, in such a way that the independence of consecutive samples is ensured.

2) Mean correlation: Likewise, in the sample correlation, we can define a correlation factor between the signals at time instants t_1 and t_2. If the mean signal varies over time, then it is because the system parameters changed.

2.5. Channel coding: Literature channel codes applied in MC

Channel coding is important to provide immunity to the transmitted signal noise, introduced by the channel [**38**]. This is achieved by adding redundancy to the emitted message and by making it possible to correct the errors caused by noise during transmission, at the receiver. Channel coding improves the communication link's reliability, but it also implies a reduction in the overall transmission rate, due to implicit code rate k/n, where n is the block length and k is the message length. The codes mentioned in [**37**] applied in MC sytems, for channel coding, will be described.

1) Hamming Codes (HC) – The Hamming codes as in [**39**], are simple linear block codes. The simplest way to construct these codes is through the multiplication of the polynomial information with the polynomial generator. These can be represented by the notation (n, k), where it uses m $(m >= 2)$ parity check bits:

$$n = 2^m - 1 \tag{2.5}$$

$$k = m - n \tag{2.6}$$

The number of errors that can be corrected in each block are given by t:

$$t = \lfloor (d_{min} - 1)/2 \rfloor \tag{2.7}$$

where d_{min} is the minimum distance in which, for HC, equals to 3. Therefore, they can correct one error in each block. In [**40**] the authors demonstrated that HC improve the performance for MC systems and that an optimal length exists according to power and another one exists according to gain. It is also shown that there is a Hamming order at which they become self-defeating, making it important to study other codes.

2) Euclidean Geometry Low Density Parity Check (EG-LDPC) codes – Considering [**40**], EG-LDPC codes can be constructed based on the lines and points of Euclidean geometry and are represented as (n, k). Where n is defined as (2.5) and k as

$$k = 2^{2m} + 3^m. \tag{2.8}$$

This kind of code has the minimum distance given by

$$d_{min} = 2^m + 1 \tag{2.9}$$

LDPC codes have generally much higher gains than HC.

3) Cyclic Reed-Muller (C-RM) – In [**41**], RM codes are a class of binary codes with multiple error correction capabilities constructed as cyclic codes. They can be represented as C-RM(n, k), with block length as in (2.5). The minimum distance, d_{min}, is given by

$$d_{min} = 2^{k-n} - 1, k \geq 2, 0 \leq n < k - 1. \tag{2.10}$$

These codes can be easily encoded using the shift-register and decoded using majority logic decoding schemes. The main advantage of these codes for MC is the simplicity of the encoder.

4) Self-Orthogonal Convolutional Codes (SOCCs) – As evaluated in [**41**], SOCCs are a subset of convolutions codes. They can be represented by (n, k, m), where m is the length of input memory. The maximum error-correcting capability is t, given by (2.7). SOCCs, as demonstrated in [**40**], have higher gain than Hamming Codes.

5) Reed Solomon (RS) codes - As described in [**41**], RS are non-binary block codes that utilize symbol-based arithmetic and are highly effective against burst and random errors in real channels. They can be expressed as RS(n, k) and defined over a Galois Field $(GF(p^m))$, where p is a prime and m is a positive integer. These codes have the capacity to correct up to $(n - k)/2$ number of errors per codeword. As shown in [**42**], the RS codes are effective in MCvD because they improve substantially the bit error probability compared to uncoded MC systems and also because they can achieve higher coding gains, although implementing them could be more complex to the channel than others mention codes.

6) ISI-mitigation (ISI-mtg) codes - The ISI-mtg codes are proposed in [**43**] and are denoted as ISI-mtg(n,k). These codes were designed considering the peculiar characteristics of the molecular communication channel. They can be applicable for all the different

diffusive molecular communication channels without requiring knowledge of the channel state information (CSI). This code family does not have any consecutive bit-1s and uses an adaptive threshold for each codeword. The codewords have important properties design special considering an MC environment, which are:

- it cannot have consecutive bit-1s
- it starts with bit-0 to avoid consecutive bit-1s between two neighbors codewords
- it has to have at least one bit-1.

According to the authors in (ISI-mtg), the results using this channel code showed a significant BER improvement compared to other tested channel codes, such as the HC and the SOOCs.

2.6. Modulation

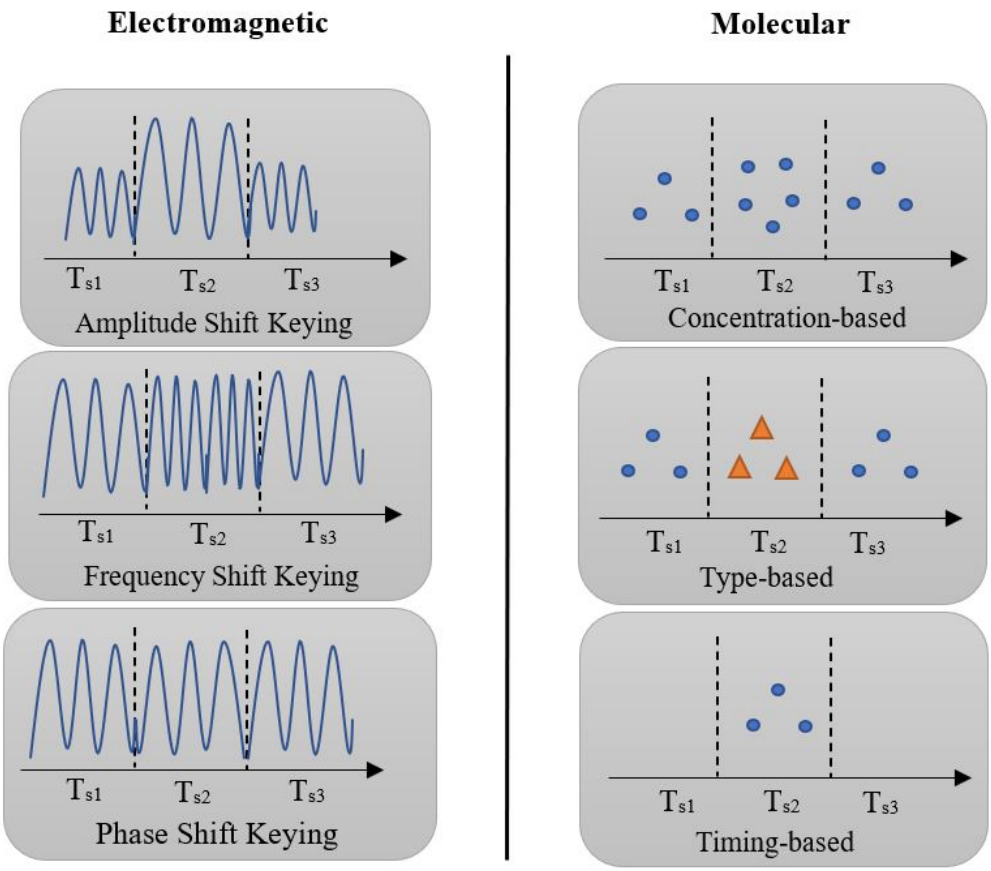


FIGURE 2.6. Modulation techniques.

In MC, to convey distinct messages, each possible message is associated with a distinct molecule signal. For this purpose, as defined in [**44**], the information can be modulated based on the properties of the molecules:

• Concentration-based is a representation technique of the information that relies on varying the concentration of the molecules.

• Type-based modulation focuses on using distinct molecules for transmitting the information.

• The Timing-based modulation encodes information on the time of release molecules.

• Spatial techniques modulation utilizes spatial diversity to convey information.

• The hybrid form of modulation combines two or more modulations previously explained.

In Figure 2.6 there is a representation of electromagnetic modulations compared with MC modulations. Next, we will explain some MC modulation techniques [**45**].

1) Molecular Shift Keying - An example of a type-based modulation is the molecular shifting keying (MoSK) [**46**, **47**], where different information means different type of released molecules. MoSK is analogous to frequency shift keying (FSK) in electromagnetic communications. Assuming that are N types of molecules available for signaling which are denoted by types $A_1, A_2, \ldots, A_N$, we can have $m[k] \in \{1, 2, \ldots, N\}$. In this case it is released $a_n[k]N^{tx}$ type-A_n molecules in the $k-th$ symbol where $a_n[k] = 1$ holds if the information symbol $m[k] = n$, otherwise, $a_n[k] = 0$ holds. The simplest scenario is when $N = 2$, in which case we have two types of molecules, one representing the bit-1 and the other the bit-0.

2) Pulse Position Modulation - The pulse Position Modulation (PPM) [**48**] is a timing-based modulation, where information is encoded in the release time of the pulses. In this scheme, the symbol interval is divided into N sub-slots and where we have $m[k] \in \{1, 2, \ldots, N\}$ as information symbols. The transmitter releases N^{tx} molecules at the beginning of the $n-th$ sub-slot only if $S[k] = n$.

2.7. Experimental works

MC systems based on diffusion have been tested on an experimental basis over the last few years. The purpose is to bridge the gap between theory and practice. Currently the experiments to study the whole architecture of MC in end-to-end systems are still carried out at the macro level instead of nano, to which MC really applies. In order to develop

a macroscopic experience that represents an MC system, we will describe some previous experiments as well as their achievements and limitations.

2.7.1. Air-driven

Several experiments have been made using the air as a communication channel between the transmitter and the receiver. The following are examples of experiences using air as the communication channel.

An experimental MC system was presented in [**43**], based on spraying and detecting alcohol in open space. Having at the transmitter part, a microcontroller controlling two spray nozzles separately and, at the receiver part, two alcohol sensors driven by two separated microcontrollers. Counting the number of released molecules was not possible on the assembled testbed, so they used the spray duration to adjust the power. One of the issues present in this experiment was the fluctuations in humidity and temperature which change the channel features, which in the molecular context may have an impact on the results obtained.

Demonstrated in [**49**] is an end-to-end MC system, where both transmitter and receiver are controlled by microcontrollers, having as main goal the transmission of short text messages using chemical signals. To modulate the channel symbols into chemical signals, they used an electronic spray called DuroBlast which has a battery-operated electrical pump that can spray a wide variety of liquid chemicals. The main challenges, found after its implementation, was a nonlinearity of the platform for being unable of achieving high transmission rates because of the simplicity of the implemented system.

2.7.2. Liquids-driven

The use of liquid transmission channels is also widely used in experiments for MC. The most popular is the use of water as a transmission medium.

For example, a water-based system is presented in [**50**] were small tubes are used as the communication channel. They implement a time-slotted binary communication system where information is carried via the pH of transmitted signals, using acids and bases to modulate 1-bits and 0-bits, respectively. The authors refer that the system was shown to achieve data rates that were better than any previous chemical communication platform.

In [**51**] a biological microscale modulator based on E. coli bacteria was used, that expresses the light-driven proton pump. The transmitted data from the induced chemical

signal, was measured by a pH sensor, which served as a receiver and it was shown that the testbed could successfully convert an optical signal representing a sequence of binary symbols into a chemical signal. A high bit rate in this testbed compared to existing biological testbeds, was also obtained.

Chapter 3

System model

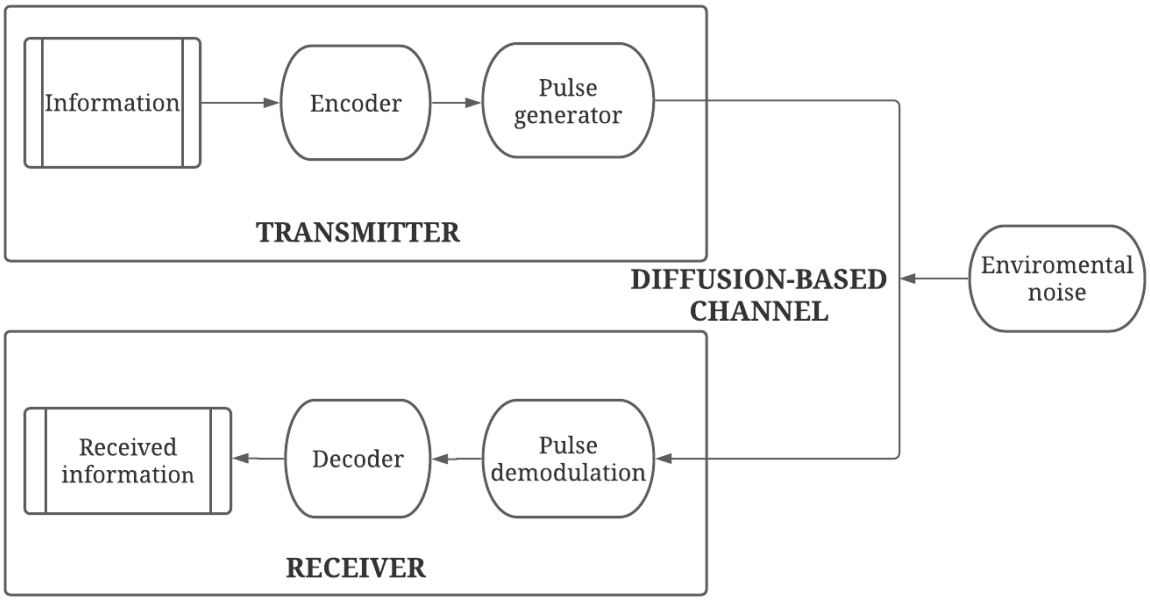


FIGURE 3.1. System block diagram.

The assumed model for this work is an end-to-end MCvD system. Figure 3.1 presents the system model using a block diagram. At the transmitter, the information bits enter the encoder block which generates, as output, coded bit sequences. These codewords then enter the pulse generator, which modulates the bits according to different concentrations of molecules and according to the selected format (NRZ, Manchester). Then, these molecules are released to the channel where they difuse until they reach the receiver where they can be affected by environmental noise. In the receiver, the readings first enter in the pulse demodulation, which demodulates them according to bit decision techniques, generating estimates of the coded bits. Finally, in the decoder, the demodulated codewords are decoded into information bits.

In the next sections, will describe in detail the functionalities of each block.

3.1. Encoder and decoder: TCH Codes

In this work we adopted TCH codes as the encoder/decoder scheme.

TCH codes are a class of binary, non-linear, non-systematic and cyclic block codes of length $n = 2^m$, where m is any positive integer. For efficiency reasons, the all-zero codeword was excluded from the code set and the inverse (binary negation) of any codeword is always another valid codeword [**52**]. TCH block codes are identified as TCH(n, k, t). With P_i being a theoretically generated set of Base Polynomials, which give the whole structure for any of these codes, we can write [**53**]

$$n = 2^m, m \in \mathbb{N} \tag{3.1}$$

$$k = m + \log_2(h) + 1 \tag{3.2}$$

$$d_{min}^{TCH} \geq 2t_{TCH} + 1 \tag{3.3}$$

$$d_{min}^{TCH} \leq H_d[P_i(x), \{P_j(x^r)\} \text{mod } n_{TCH}] \leq n_{TCH} - d_{min}^{TCH} \tag{3.4}$$

$$P_i(x) = P_j(x^r) \text{mod } n_{TCH}, i \neq j \forall t_{TCH} \in \mathbb{N} \tag{3.5}$$

where h is the number of polynomials. As usual, their error-correcting capacity, t_{TCH}, depends on the minimum distance, d_{min}^{TCH}, between the polynomials, where H_d stands for the Hamming distance between any two polynomials. TCH codes are also balanced codes, that is, the number of zeros equals the number of ones in each codeword, which is an important feature for MC systems. We can derive TCH codes of nearly any length, however, the most important class of TCH codes which originated all of them, are based on the so-called Basic TCH Polynomials (B-TCH). B-TCH polynomials with degree n are obtained by the following equation

$$P(x) = \sum_{i=0}^{\frac{p-1}{2}-1} a_i x^{K_i}, a^i \in GF(q), q = p^k, k \in \mathbb{N} \tag{3.6}$$

where the exponents K_i satisfy the equation

$$a^{K_i} = 1 + a^{2i+1}, i = 0, 1, \ldots, \frac{p-1}{2} - 1. \tag{3.7}$$

In these expressions p is a prime number obeying the following condition

$$p = n_{TCH} + 1 = 2^m + 1. \tag{3.8}$$

i	p	n_{TCH}
0	3	2
1	5	4
2	17	16
3	257	256
4	65537	65536

TABLE 3.1. Fermat numbers for generating TCH codes [**53**].

Prime numbers that verify this condition can be written as

$$F_i = 2^{2^i} + 1, i \in \mathbb{N} \tag{3.9}$$

and so they are in fact Fermat numbers (F_i). Only five code lengths obey these rules, which means that we can only generate pure TCH polynomials for codes with lengths of 2, 4, 16, 256 and 65536, as shown in Table 3.1.

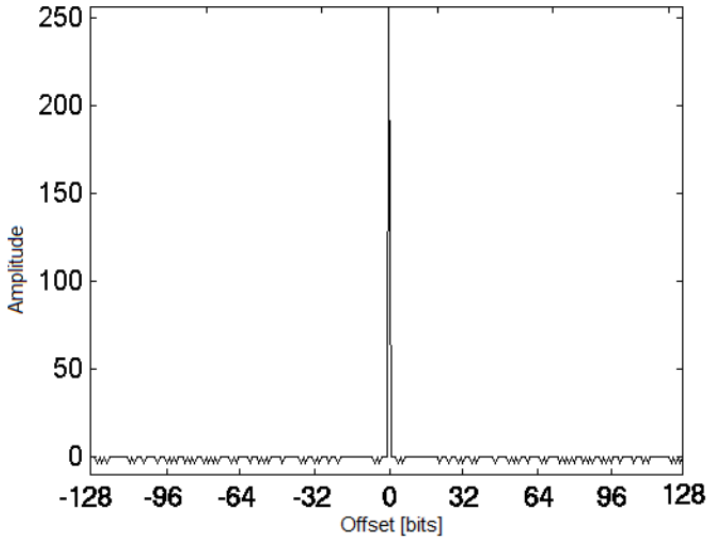


FIGURE 3.2. Typical autocorrelation function of a B-TCH polynomial with n=256, [**53**].

TCH codes originated by B-TCH polynomials have both good cross and auto-correlation. Their autocorrelation is always three-valued, with values n, 0 and –4 and its distribution is perfectly characterized for any n [**54**]. This translates into a great advantage for higher sized TCH codes, such as TCH codes length $n \geq 256$ since the autocorrelation of these sequences tend to get closer to a Dirac impulse, as shown in Figure 3.2.

Conceptually, the TCH receiver is based on a group of correlators to compare the received word. As usual and by comparing the outputs of the correlators it becomes possible to choose the sequence that corresponds to the highest correlation or the most likely word sent. In the TCH receiver, depicted in Figure 3.3, correlation is done in the frequency domain using two FFT's, a complex multiplication and an IFFT, which, together with some optimizations, accelerates the decoding process. For example and since FFT's of length n process n codewords simultanesously, the efficiency increases with the code length n. Furthermore since both, real and imaginary parts are also processed simultaneously, that further decreases the processing time to half, again doubling the receiver efficiency.

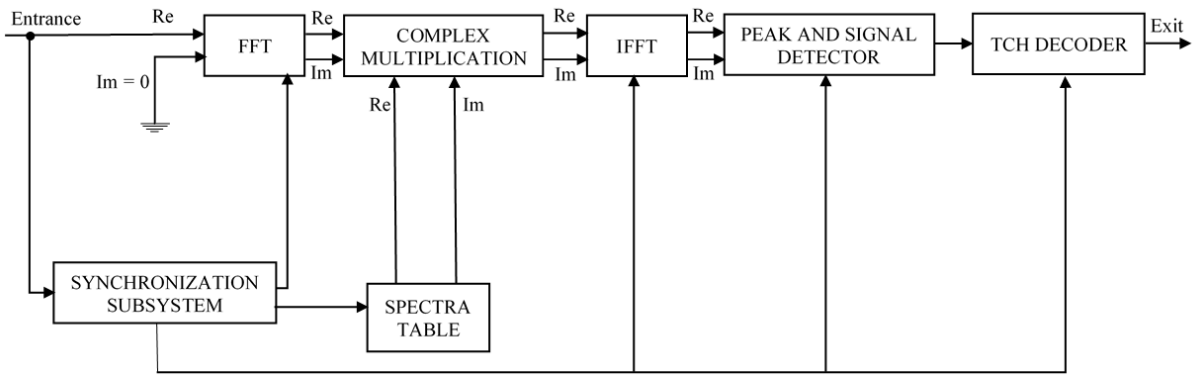


FIGURE 3.3. TCH decoder of maximum likelihood [**54**].

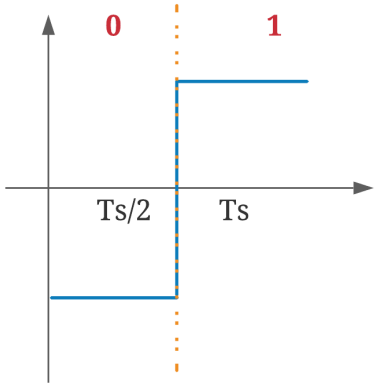
3.2. Pulse generator

3.2.1. Types of pulses

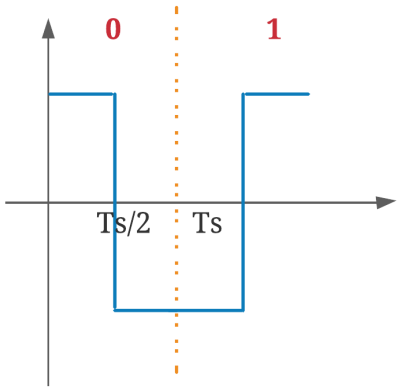
For the proposed MC system, non-return to zero (NRZ) pulses and Manchester pulses were tested. NRZ is a line code technique, characterized by the pulses not returning to zero while mapping binary data (1's and 0's). They are basic line codes as they determinate the binary data in a very simple way, attributing to the bit-1 one significant condition and to the bit-0 another significant condition, without no resting state [**55**]. Some disadvantages of using these pulses are:

- No error correction capability.
- A long string of ones and zeros may lead to loss of synchronization.
- In a system that can saturate, information could be easily lost.

In order to minimize the described challenges, other pulses were also considered, namely the Manchester pulses as in Figure 3.4b. They have their wave boundaries always between 1 and -1, therefore the decisions are usually taken in the middle of each bit, [**56**,**57**]. These features will provide more robustness to the MC communication system ensuring that the system does not overflow.



(a) NRZ



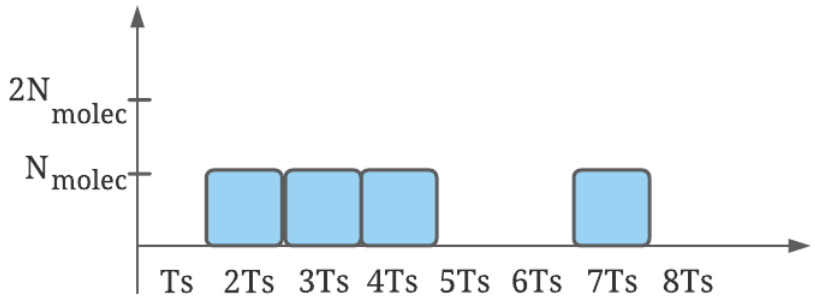
(b) Manchester

FIGURE 3.4. Type of pulses.

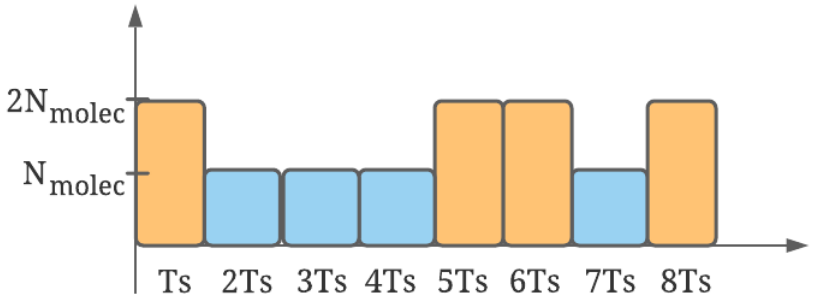
3.2.2. Modulation: Concentration Shift Keying

Concentration Shift Keying (CSK) [**58**, **59**], is a concentration-based modulation technique, analogous to amplitude shift keying (ASK) used in telecommunications. This modulation works by releasing molecules carrying information over discrete period time slots, the symbol slots (T_s). Let $m[k]N_{molec}$ denote the number of molecules released by the transmitter at $k-th$ symbol slot, where N_{molec} is the maximum number of molecules

released at a time and $m[k] \in [0, 1]$ is the information symbol. Considering that S denotes the set of CSK modulation symbols, we have $S = \{0, \frac{1}{N-1}, \ldots, \frac{N-2}{N-1}, 1\}$, where N is the number of symbols. One variation of this technique is the binary concentration shift keying (BCSK) where a symbol represents one bit of information. For instance, as depicted in Figure 3.5b, in each symbol slot, sending a certain concentration of a molecule represents bit-1 and a different concentration represents bit-0. As a special case, when $N = 2$, we have the simplest form of this technique, the on-off keying (OOK), where each symbol represents a one-bit value. Bit-1 corresponds to the transmitter releasing N_{molec} and bit-0 corresponds to the transmitter not releasing anything within the symbol slot, Figure 3.5a.



(a) OOK



(b) BCSK

FIGURE 3.5. Possible representation for the bit stream "01110010", regarding different types of CSK modulation.

3.3. Diffusion based channel

Suitable channel models are essential for the analysis and design of an MC system. Let us consider an MCvD system operating in a three-dimensional space. The model assumed

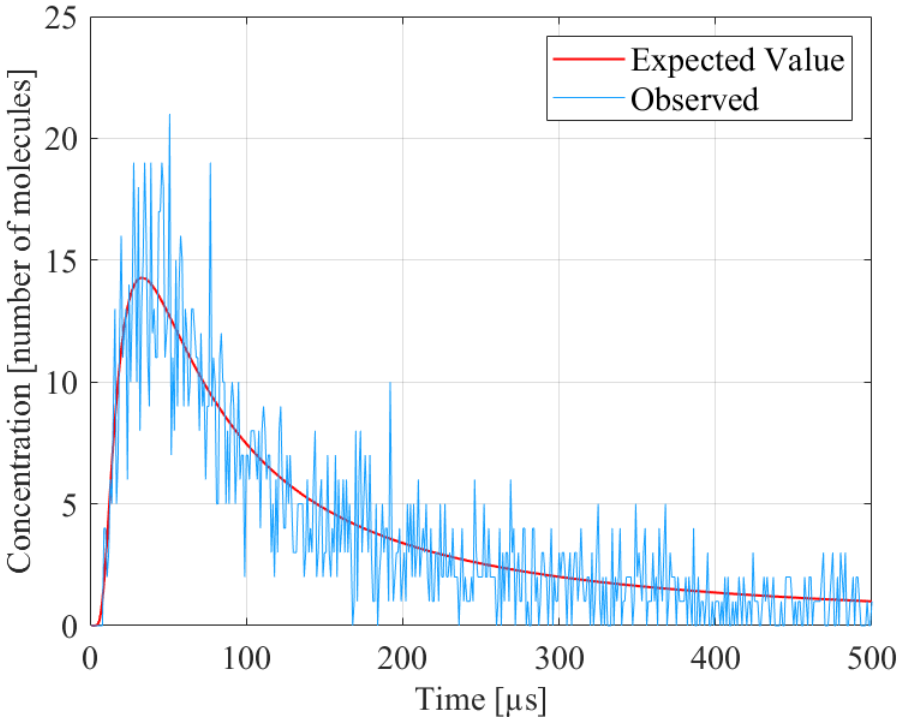


FIGURE 3.6. Example illustrating the expected and observed concentrations for a diffusion-based MC channel, assuming number of molecules (N_{molec})=10000, d_0=0.3 µm, D=450 µm^2/s and a_{rx}=50 nm.

for the average number of molecules counted at the receiver for this work is the point transmitter to a spherical absorbing receiver. An absorbing receiver is used because it is considered to be a more realistic model compared to a passive receiver, as it captures the molecules. This type of receiver counts the number of absorbed molecules during an observation window $[t_l, t_u]$. The probability, $F_{hit}(t_u, t_l)$, of a molecule emitted at time t=0 to be observed at the receiver during interval $[t_l, t_u]$, can be defined as [**22**]

$$F_{hit}(t_u, t_l) = \frac{a_{rx}}{d_0}\left[\text{erfc}\left(\frac{d_0 - a_{rx}}{\sqrt{4Dt_u}}\right) - \text{erfc}\left(\frac{d_0 - a_{rx}}{\sqrt{4Dt_l}}\right)\right] \tag{3.10}$$

where erfc is the complementary error function, a_{rx} is the radius of the spherical receiver and d_0 is the distance between the center of the transmitter and the center of the receiver. D is the diffusion coefficient as in (2.2).

The approach taken in this work assumes a diffusion-based MC system whose information bits are encoded in order to reduce ISI. The concentration vector corrupted by noise at the k^{th} time interval can be written as

$$y[k] = \sum_{l=0}^{L-1} h[l,k]x[k-l] + n[k] \tag{3.11}$$

where L is the assumed length of the ISI, $x[k] \in \{A_0, A_1\}$ (with A_0 and A_1 representing the levels for a '0' and a '1') denotes a concentration based binary modulation symbol emitted from the transmitter at time k, $h[l,k]$ is the number of molecules observed by the absorbing receiver during time interval k after N_{molec} are released from the transmitter at time $k-l$ and $n[k]$ is the number of noise molecules detected by the receiver at time k. Both $h[l,k]$ and $n[k]$ are modeled as Poisson random variables: $h[l,k] \sim Poisson(\overline{h}[l,k])$ and $n[k] \sim Poisson(\overline{n})$. Taking into account the probability model defined in (3.10), we have

$$\overline{h}[l,k] = N_{molec}.F_{hit}((l+1)T_s, l.T_s). \tag{3.12}$$

It can be shown that the received concentration at the receiver can be expressed as

$$y[k] = \underbrace{\overline{h}[0,k]x[k]}_{\text{desired signal}} + \underbrace{v[k]}_{\text{diffusion noise}} + \underbrace{I[k]}_{\text{ISI}} + \underbrace{n[k]}_{\text{environmental noise}} \tag{3.13}$$

where the diffusion noise follows the distribution $v[k] \sim Poiss_0(\overline{h}[0,k]x[k])$, i.e., it is equivalent to a Poisson random variable whose mean has been subtracted. This noise is signal dependent and accounts for the deviation of the concentration from its expected value according to the diffusion model, as illustrated in Figure 3.6.

3.4. Detection and bit decision

In this work we assume that no CSI exists at the transmitter nor the receiver. Therefore, the detection is accomplished solely based on the received observations over each encoded block duration. Two simple detection approaches are considered. The next sections, will describe them.

3.4.1. Threshold-based hard decision

The first considerated detection technique is a threshold based detection similar to [**43**], which generates hard decision symbols/bits. We consider a concentration-based binary modulation where each symbol comprising codeword i is represented as $x^i[k] \in \{A_0, A_1\}$, with A_0 and A_1 denoting the levels for a '0' and a '1', respectively. In this case an adaptive threshold is computed for the i-th codeword as

$$t^i = a y^i_{min} + (1-a) y^i_{max} \tag{3.14}$$

where $y^i_{min} = min(y^i)$ and y^i_{max}=max(y^i), with $y^i = \left[y^i[1], \ldots, y^i[N]\right]^T$ denoting the vector with the corresponding N received samples. Parameter a is a scaling factor which, considering the adoption of a balanced code with equal number of '0's and '1's, we set as a=0.5. Using this threshold, a hard decision estimate is simply obtained as

$$\hat{x}^i[k] = \begin{cases} A_0, & y^i[k] < t^i \\ A_1, & y^i[k] < t^i \end{cases} \tag{3.15}$$

which can be directly demodulated into bits.

3.4.2. LLR-based soft decision

The second detection technique relies on an ad-hoc computation of log-likelihood ratios (LLRs) for each bit. Since we assume that no prior knowledge about the statistics of the channel exists, we compute estimates for the probabilities $p_{1,k} = P(x^i[k] = A_1 \big| y^i)$ and $p_{0,k} = P(x^i[k] = A_0 \big| y^i)$ considering a uniform distribution whose bounds are defined by y^i_{min} and y^i_{max}. In this case these probabilities are obtained as

$$p_{1,k} = \frac{y^i[k] - y^i_{min}}{y^i_{max} - y^i_{min}} \tag{3.16}$$

and

$$p_{0,k} = \frac{y^i_{max} - y^i[k]}{y^i_{max} - y^i_{min}}. \tag{3.17}$$

Soft values can then be computed as the conditional expected value of the symbols which can be written as

$$\tilde{x}^i[k] = E(x^i[k] \big| y^i) = (A_1 - A_0) p_{1,k} + A_0. \tag{3.18}$$

Taking into account that the LLRs for the individual bits can be expressed as

$$\lambda^i[k] = log\left(\frac{p_{1,k}}{p_{0,k}}\right) \tag{3.19}$$

we can write

$$p_{1,k} = \frac{1}{1 + e^{-\lambda^i[k]}} \tag{3.20}$$

which allows us to compute the soft-values as

$$\hat{x}^i[k] = \frac{(A_1 - A_0)}{1 + e^{-\lambda^i[k]}} + A_0. \tag{3.21}$$

Note that in the case of OOK, with $x^i[k] \in \{0, A\}$, this expression can be rewritten as

$$\hat{x}^i[k] = \frac{A}{2}\left(\tanh\left(\frac{\lambda^i[k]}{2}\right) + 1\right). \tag{3.22}$$

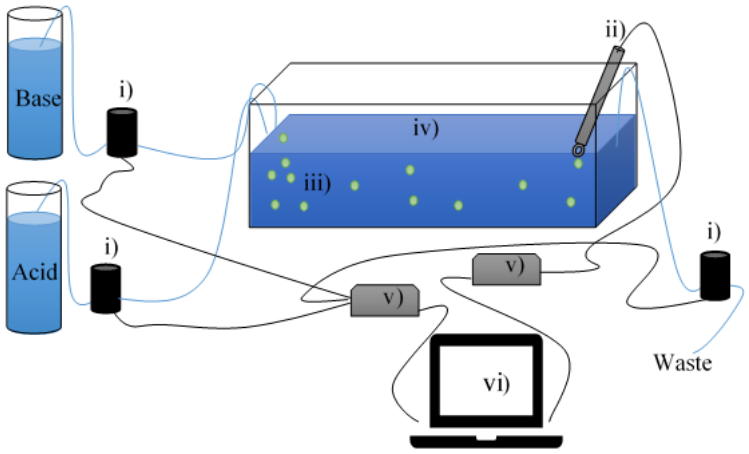
In the case of a polar modulation where $x^i[k] \in \{-A, A\}$, (3.21) reduces to

$$\hat{x}^i[k] = A\left(\tanh\left(\frac{\lambda^i[k]}{2}\right)\right). \tag{3.23}$$

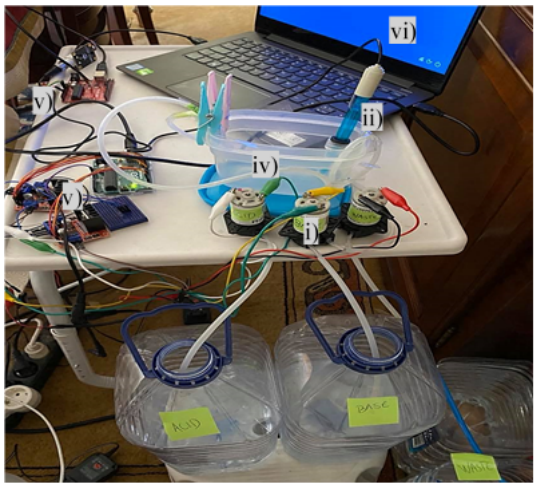
These soft estimates then allow simple soft-decision decoding based on squared Euclidean distance minimization.

Chapter 4

Experimental architecture



(a)



(b)

FIGURE 4.1. (a) Schematic diagram of the experimental implementation; (b) Photo of the experiment.

To help validate the performance of the proposed TCH codes in MC, we implemented a macroscale experiment. In our experiment, we used pH levels to test the molecules as information carriers. The different pH measurements are indicators of how acidic or basic the water is: pH's above 7 indicate a basic solution and below 7 an acidic solution,

on a scale between 0 and 14. The pH measures the amount of free hydrogen (H+) and hydroxyl (OH-) in the water. This means that water with more free hydrogen ions is acidic and water with more free hydroxyl ions is basic. The scale on which pH is measured is logarithmic, which means that increasing or decreasing an integer value changes the concentration by a factor of ten [**60**–**62**]. Thus, we use an acid and a base to send information. The used base solution has a pH level of 9 and the used acid solution has a pH of 3. Utilizing pH could constitute a problem because of the limited scale, from 0 to 14. For example, in the case of bit-1, this is represented by the base and bit-0 by the acid. So, if many consecutive bit-1's are transmitted, information could be lost. To avoid this, we decided to use Manchester codes in this experiment to transmit the bits. Thus, bit-1 is represented by sending an acid and then a base and bit-0 is represented by sending a base and then an acid. This prevents the information from being lost as it ensures the same amount of acid and base solution is released in the channel independently of the information sequence.

Next, we will explain the elements of the system, namely, all the used components, the chemicals, the transmitter, the channel and the receiver. The main diagram and a photo of the experiment are presented in Figure 4.1.

4.1. Electronic components

4.1.1. Arduino UNO

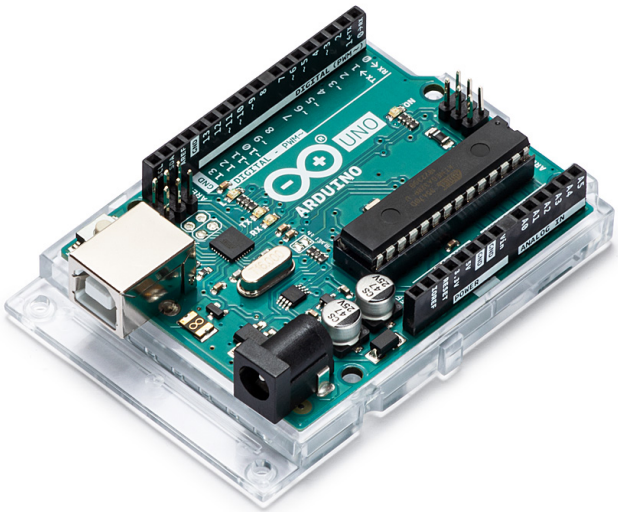


FIGURE 4.2. Arduino Uno.

Arduino is an open-source electronic prototyping board, as shown in Figure 4.2, and is composed not only by the hardware part but also by the software. The software part is developed using the C/C++ language. The Arduino board is capable of interfacing different sensors and wireless communication devices through input/output pins [**63**]. As in [**64**] the main advantages of utilizing this platform are:

• Low cost - Arduino boards compared to other micontrollers are relatively inexpensive.

• Easy to program - The Arduino integrated development environment (IDE) is for all types of users, simple for beginners but complex enough to advanced users.

• Cross-plataform - IDE runs on Windows, Macintosh OSX and Linux operating systems.

• Large number of tutorials, articles and ready-made projects on the internet - The Arduino software is an open-source tool.

4.1.2. Pump

FIGURE 4.3. 12V Aquarium pump.

The peristaltic pumping principle [**65**], called peristalsis, is based on alternating compression and relaxation of the tube, generating a flow. A rotating roller runs along the length of the tube creating a temporary seal between the suction and the discharge sides of the pump. As the pump roller rotates this sealing pressure moves along the tube forcing the product to discharge. When the pressure is released, the tube creates a vacuum, which draws the product from the suction side of the pump. The used pumps were 12V aquarium peristaltic pumps working at 5000 RPM, Figure 4.3.

4.1.3. Pump driver

Pump drivers are circuits used to run a pump or a motor. A driver takes low-current control signals and turns them into higher-current signals that can drive the pumps. The used driver model was the L298N, Figure 4.4. The L298N is a dual H-Bridge driver which allows speed and direction control of two DC pumps at the same time. The module can drive DC motors that have voltages between 5 and 35V [**66**]. The schematics of how the L298N driver works are detailed in Figure 4.5.

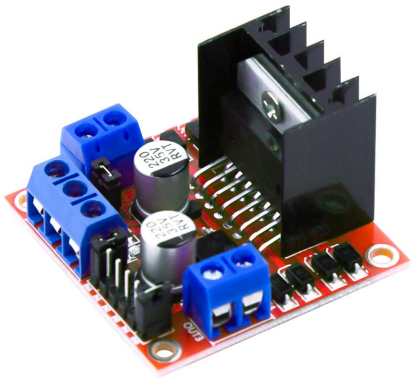


FIGURE 4.4. L298N motor driver.

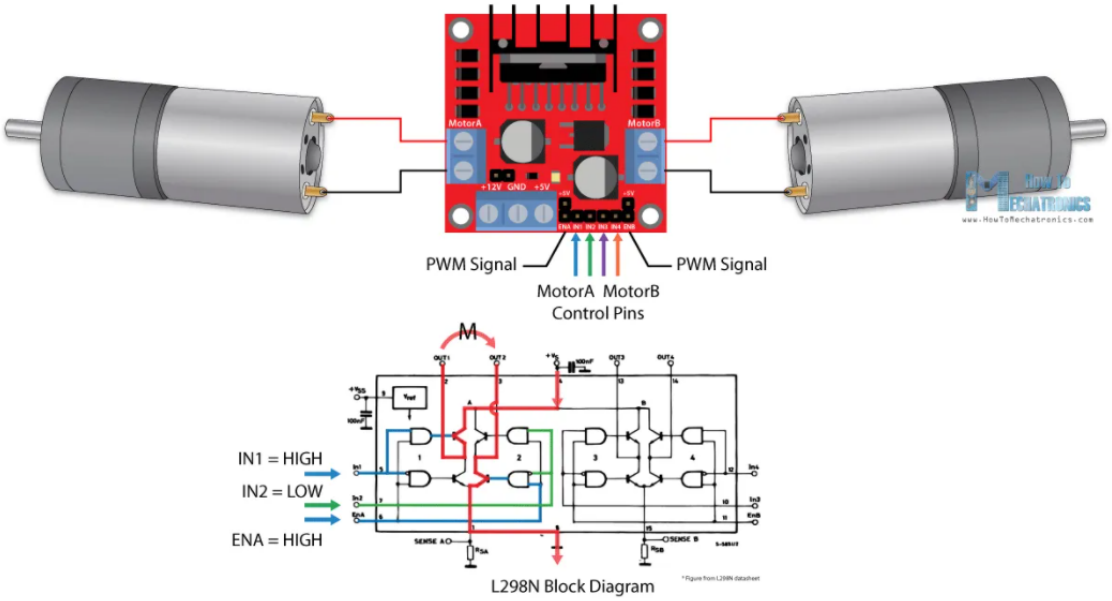


FIGURE 4.5. L298N schematics, from [**66**].

4.1.4. pH sensor

FIGURE 4.6. pH probe, from [**68**].

FIGURE 4.7. pH calibration liquids.

We used a 5V pH probe with a pH measurement's errors of ± 0.1 at 25℃, Figure 4.6, featuring, as most pH probes, a glass bulb filled with strong electrolytes and a silver wire inside. This allows the pH electrode to measure the difference in potentials between the two sides of the glass electrode. This signal is then converted into an electric potential which allows reading the pH value through a connection to a signal conversion board [**67**]. Therefore, it is essential that this bulb is properly stored. Every time, if the pH probe is not reading, it should be clean and placed inside its plastic protection. Also, because this is a fragile sensor, with an estimated lifetime of six months, all the wires connecting the

pH probe should be disconnected guaranteeing that no current is passing through the pH probe.

An important aspect of most pH probes is that they need to be calibrated before being used. This calibration must be done regularly, preferably always between experiments. To do that we submerge the pH probe into standard sterile solutions, establishing in the code the corresponding voltage to the pH value that we are diving the pH probe. A neutral and an acid or basic solutions are enough for a good calibration. In this work we used a sterile acid solution with pH=4 and a neutral solution with ph=7 to do the pH probe calibration. These solutions are shown in Figure 4.7.

4.2. Chemicals

Understanding and choosing the right chemicals to carry information is essential in this type of system. In this study, two types of chemicals were used. One to lower the pH (pH-) and one to raise the pH (pH+), Figure 4.8. The pH- chosen was a granular sodium bisulfate (NaHSO3) compound and the pH+ was a granular sodium carbonate (Na2Co3) compound. Certain quantities of these chemicals were dissolved in different water recipients until solutions with the desirable pH degree were achieved, namely pH=3 and pH=9. For reaching pH=3 the pH- was used, Figure 4.8a, and the metric to lower a tenth of pH was mix 10g of this chemical for each cubic meter of water. We used the pH+ to increase the pH until reaching pH=9, Figure 4.8b, and the metric to increase a tenth of pH was dissolve 5g of this chemical for each cubic meter of water. After the right quantities of the chemicals were defined, they could be released to the channel, thus carrying the information to the receiver.

(a) (b)

FIGURE 4.8. (a) pH- used; (b) pH+ used.

4.3. Transmitter

4.3.1. Hardware

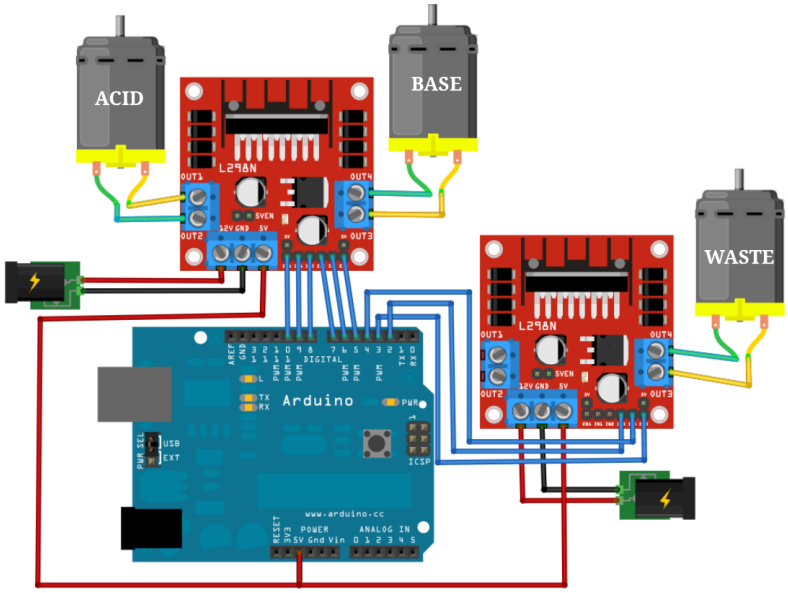


FIGURE 4.9. Transmitter scheme.

The transmitter components are:

- two containers with the acid solution, the basic solution and another container for the channel fluids waste;
- silicone tubes with ≈ 4 mm of diameter;
- jumper wires: male-to-male, male-to-female and crocodile clips;
- a small bread board;
- an Arduino Uno microcontroller;
- three 12V aquarium pumps;
- three pump drivers, the L298N.

The transmitter hardware is composed of two containers, one for the basic solution and another for the acid solution. These containers with the solutions are connected through a silicone tube to peristaltic pumps. These pumps feature a flexible tube, giving an open flow path for the liquids to pass, Figure 4.1. Because we have an acid and a basic solution, we used two peristaltic pumps to transmit the solutions individually and avoid contamination. There is also a third pump that removes the liquid, working at the same speed as the other two pumps mentioned above. Since these pumps work with a high current, a driver is used. We control when they are on or off and how fast they run

by using the microcontroller connected to a computer, as in Figure 4.9. According to the transmitted information, a certain quantity of an acid solution or a basic solution is released to the channel. The connections between the components and the ports used in the assembling of the transmitter are detailed in Figure 4.10.

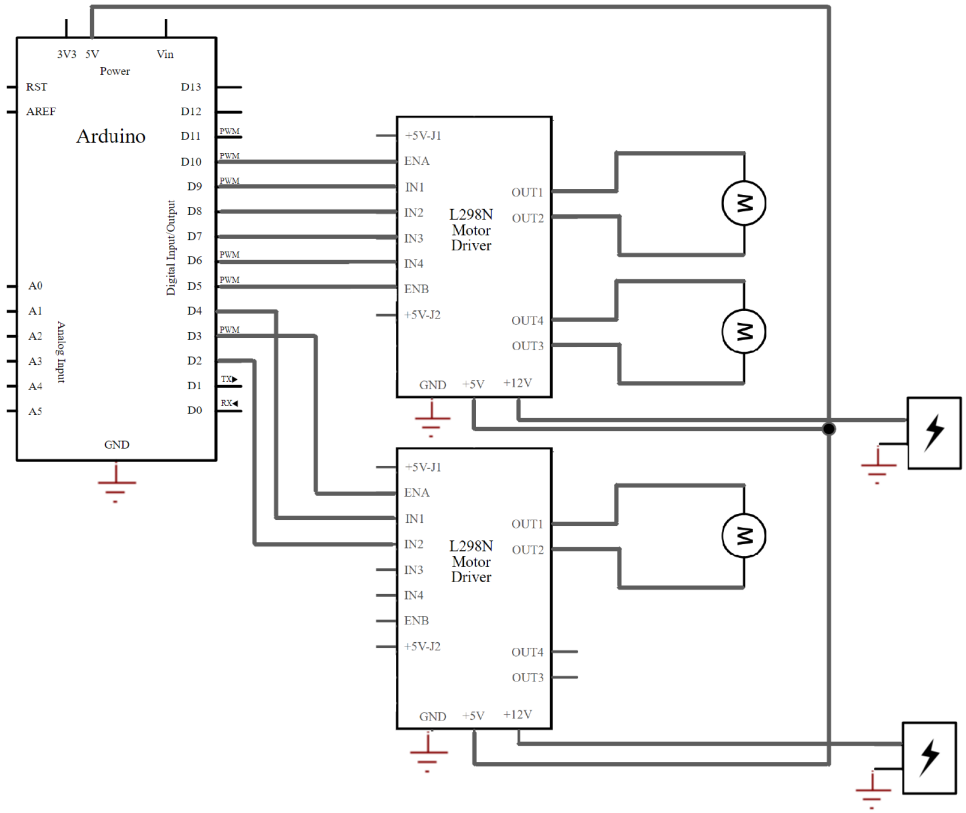


FIGURE 4.10. Transmitter circuit.

4.3.2. Software

We used the MATLAB platform to program the transmitter functions. As shown in Figure 4.11, for transmitting a bit we define two moments: a first moment that starts at the beginning of the symbol and goes to $T_s/2$ and a second moment that goes from $T_s/2$ to T_s. To transmit the bit-0, we start by sending a base solution during the beginning until halfway of the first moment, followed by a period until the end of the first moment when nothing is released to the channel. In the second moment, an acid solution is sent to the channel until halfway of the second moment followed by a period until the end of the symbol where nothing is released. To transmit bit-1, the opposite happens, in the first moment an acid is emitted followed by a period when nothing is released and in the second moment a base is emitted followed by a period when nothing is released.

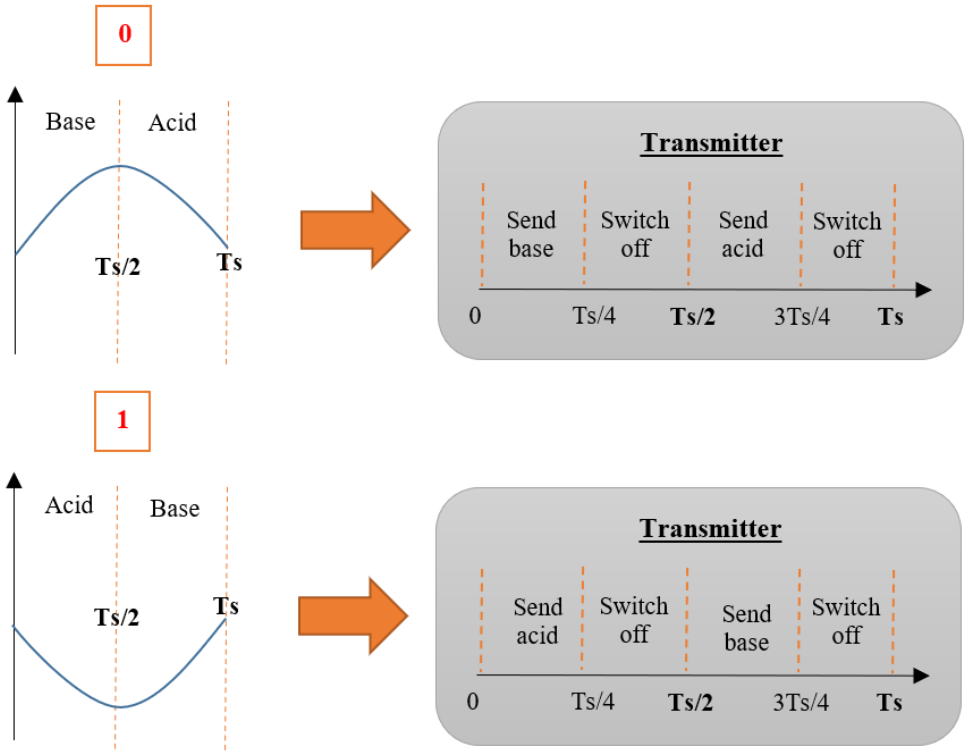


FIGURE 4.11. How bit-0 and bit-1 are transmitted.

Two different scenarios were implemented regarding channel coding: an uncoded scenario and one with TCH(16,10) codes. The transmitter workflow for the uncoded case is detailed in Figure 4.12, and can be explained in simple steps:

1) We start by generating the information bits and modifying the *Channel coding* variable to "No", which means no channel coding is being applied.

2) Next, we create a cycle that only finishes when all the bits are transmitted, we do that by starting to count the time just before the program enters in this cycle and by defining the ending as the number of bits multiplied by the bit period (T_s).

3) During this time and for each bit, we verify if it is a 0 or a 1. We do that by confirming if it is greater than 0 or not, then according to this result, bit-0 or bit-1 are transmitting complying to Figure 4.11.

4) Once all the bits are sent, all the pumps must turn off in order to ensure that nothing else is being released or taken from the channel.

The TCH(16,10) codes scenario is presented in Figure 4.13. For this case, the variable *Channel coding* must be turned on to "Yes". All the process is similar to the uncoded case except that before the bits enter the transmission cycle they are converted to codewords where redundancy is added. Therefore, we do not consider the total number of bits but the total number of the codewords to be transmitted.

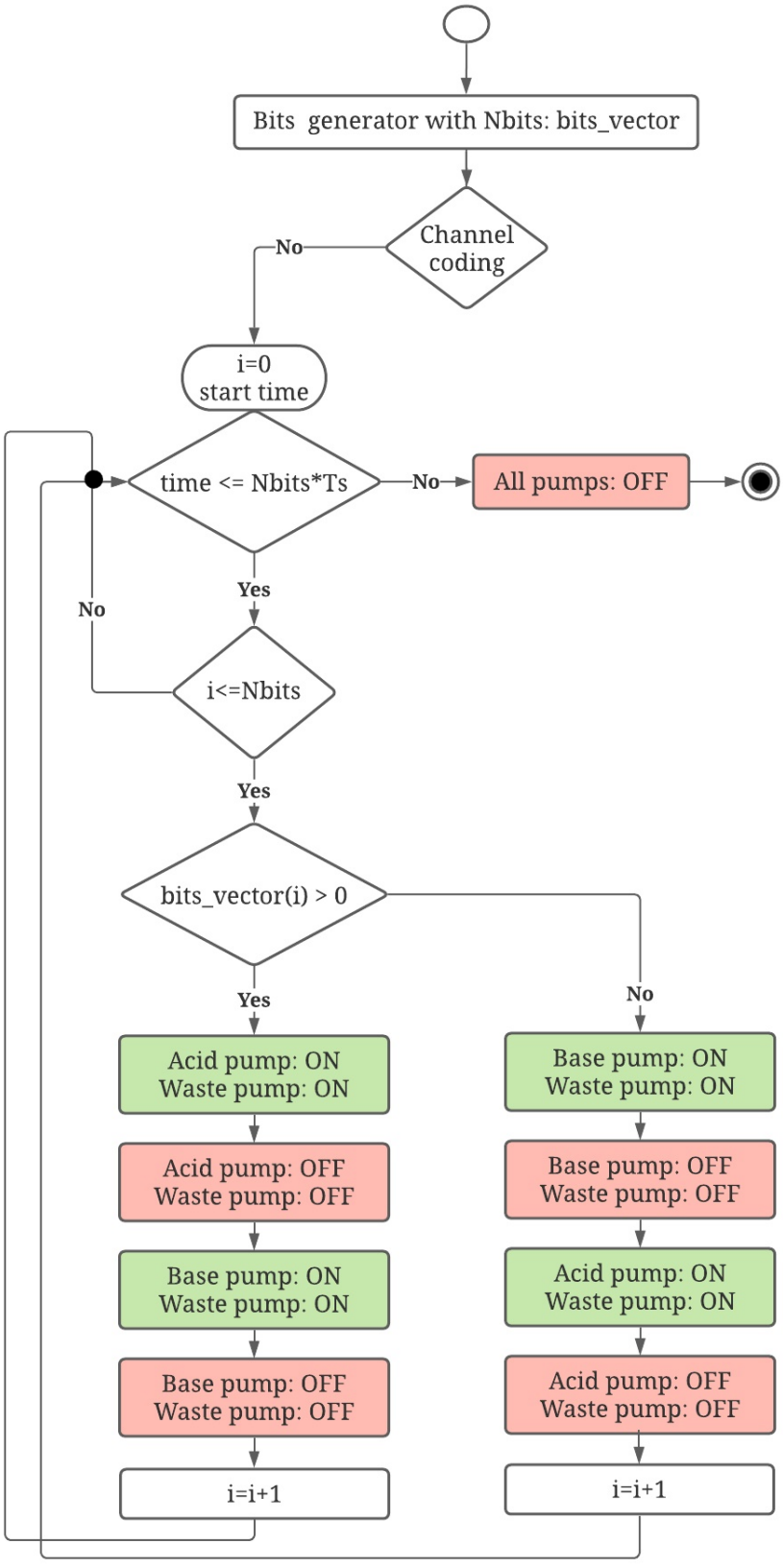


FIGURE 4.12. Transmitter workflow for the uncoded case.

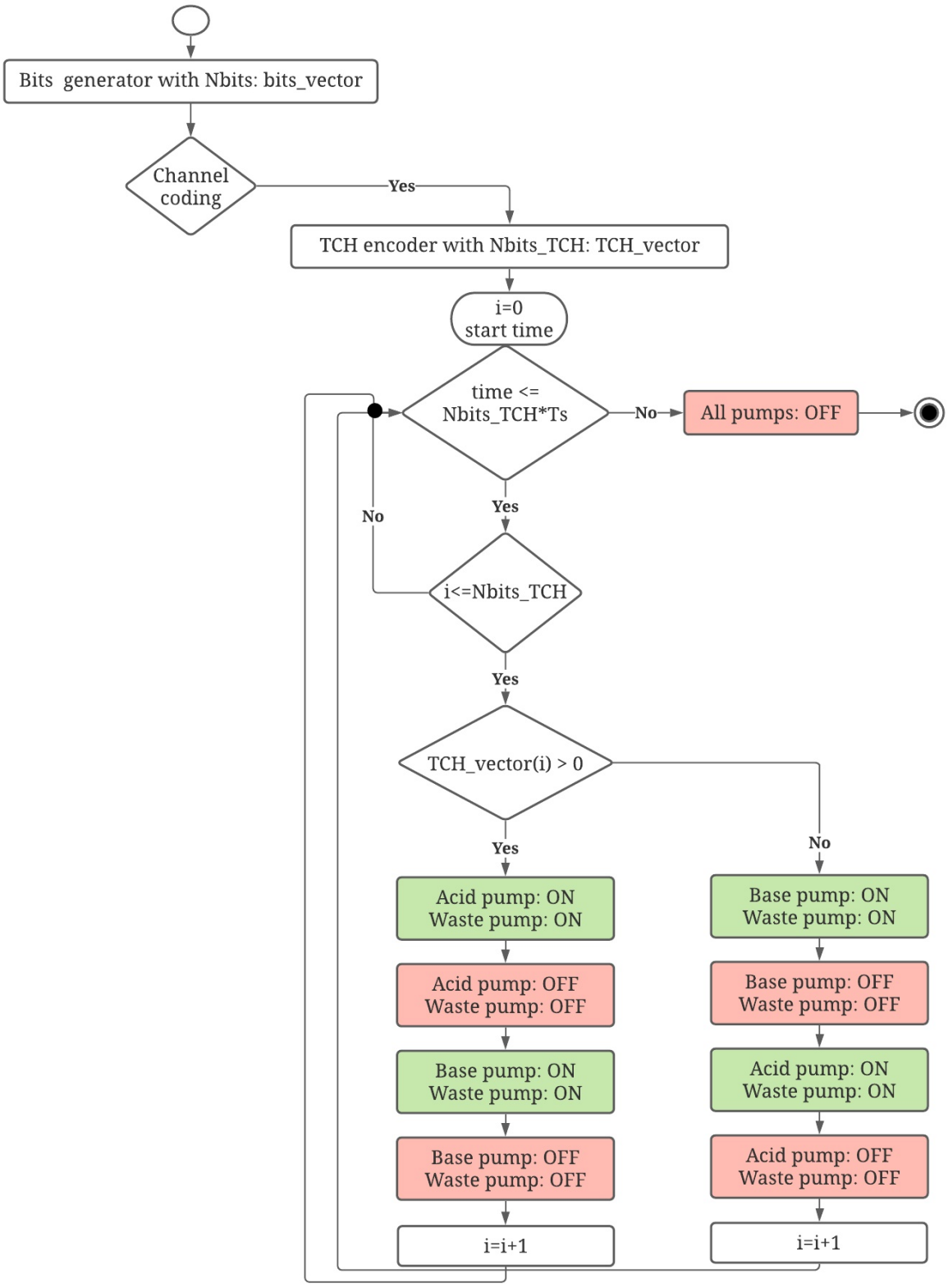


FIGURE 4.13. Transmitter workflow for the TCH codes.

4.4. Channel

As represented in Figure 4.1, we use a plastic container to simulate the channel. We decided to fill the container with only 20ml of tap water with a pH around 7. On one side of the container, we placed the tubes sending the acid or basic solution, one at a time (never simultaneously). On the other side of the container, 6cm apart, we placed the pH probe and another pump to keep the container's liquid volume always the same. Although these pumps generate some flow, the system process goes on via diffusion. One advantage of using acids and bases is that they are chemicals that cancel each other and this property is very important in closed loops because the average concentration of these chemicals would remain constant assuming the same transmission concentration of bases and acids.

4.5. Receiver

4.5.1. Hardware

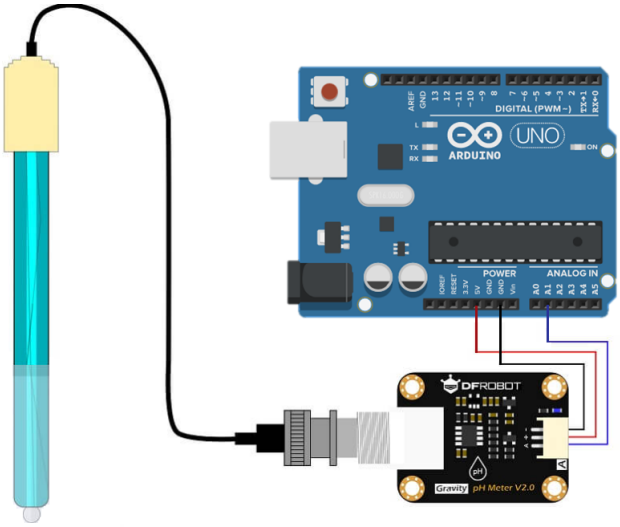


FIGURE 4.14. Receiver scheme.

The main components of the receiver are:

- an Arduino Uno;
- jumper wires: male-to-male;
- a signal conversion board;
- a ph probe.

The receiver hardware uses a different microcontroller, also an Arduino Uno, connected to the main computer. This board function is to read the pH levels in the water through the pH probe, as shown in Figure 4.14. The pH probe is connected to a signal conversion

board which is connected to the Arduino Uno microcontroller, that digitizes the pH reading through a 10-bit analog to digital converter (ADC).

4.5.2. Software

As the transmitter, in the receiver software we used the MATLAB platform to program and also two cases regarding channel coding were implemented. The first case, in Figure 4.15, is a representation of how the receiver works for the uncoded case and can be summarized in the following steps:

1) We start to change the variable *Channel coding* to "No", meaning that we are approaching the channel without any type of channel coding.

2) The next step is during the transmission period plus the maximum system delay time (which is $T_s/2$), read the pH measurements. To know the right pH measurement is read the input of the Arduino analog pin to which the pH probe is connected. Then, that value is adjusted using the voltage values of the calibration, and according to expressions given by the pH probe manufacture turned into values within the pH scale (from 0 to 14). Next, the pH values are saved in a vector with all the pH measurements within the total sample period, rx_pH. We defined the sample period as being one sample per second. To have an accurate sample vector, according to the established sample period, the computer processing delay had to be removed when adding the sample in the right position to the rx_pH vector.

3) After the rx_pH is created, the initial delay had to be removed, to fairly compare the emitted with the received signal. Since no synchronization was implemented, this removal was done by an ad-hoc technique.

4) Next, we add the difference between slopes within each $T_s/2$, to estimate the pH variation within that period. After this, we have two samples per bit. The soft decision technique uses this two samples per bit to decode. The used soft decision was described in Chapter 3.4.2.

5) Finally, the number of errors between the emitted with and received signal is counted.

Figure 4.16 details the decoding of the received information using TCH(16,10) as the implemented channel coding. When channel coding is applied, the variable *Channel coding* is changed to "Yes". The receiver procedures are similar to the uncoded case. The difference now is that, instead of considering the number of emitted bits, we consider the

total number of codewords. Also, before counting the number of errors between the emitted and received signal, the TCH decoder is applied in order to remove the redundancy inserted by the encoder, thus transforming the codewords into bits.

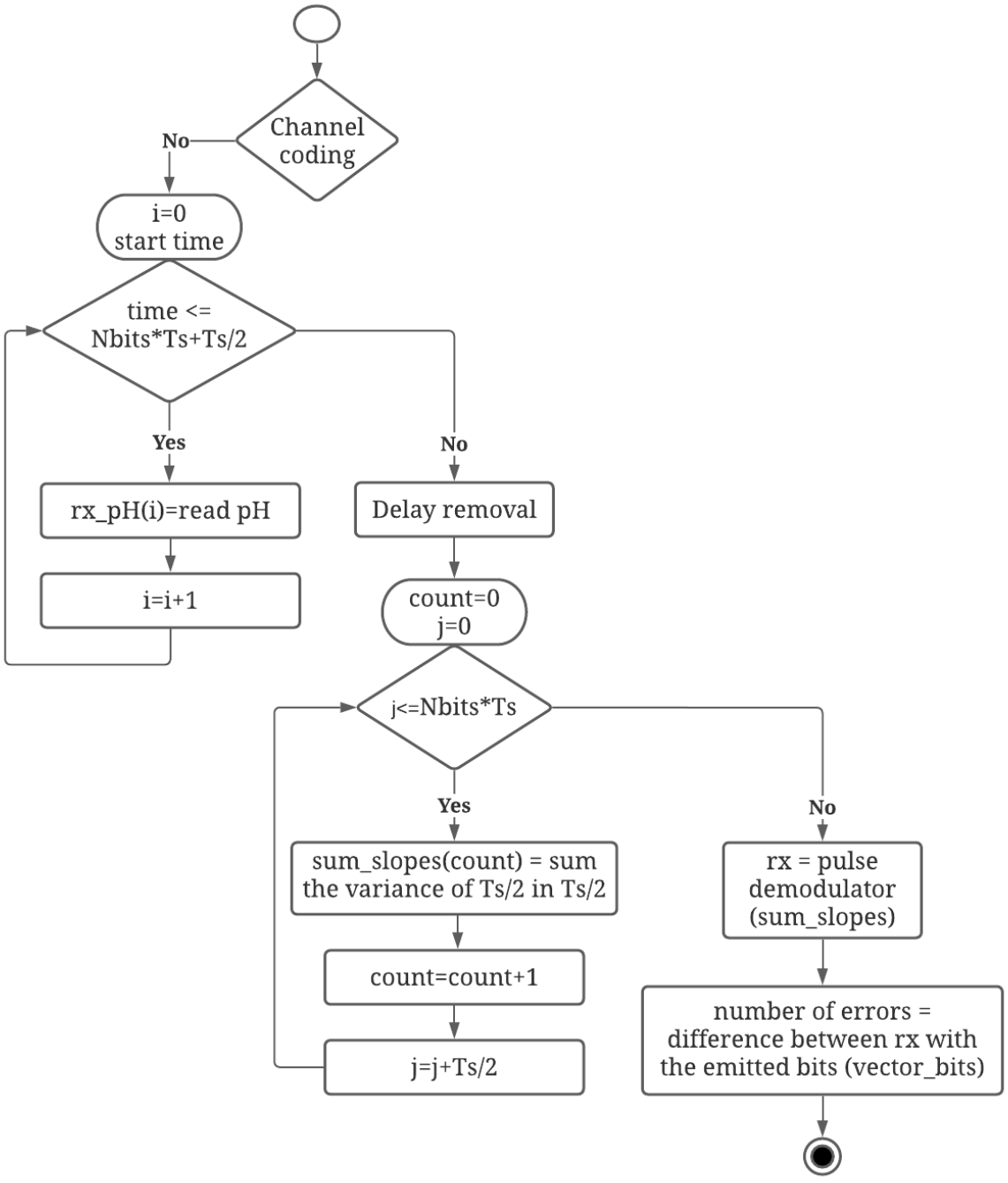


FIGURE 4.15. Receiver workflow for the uncoded case.

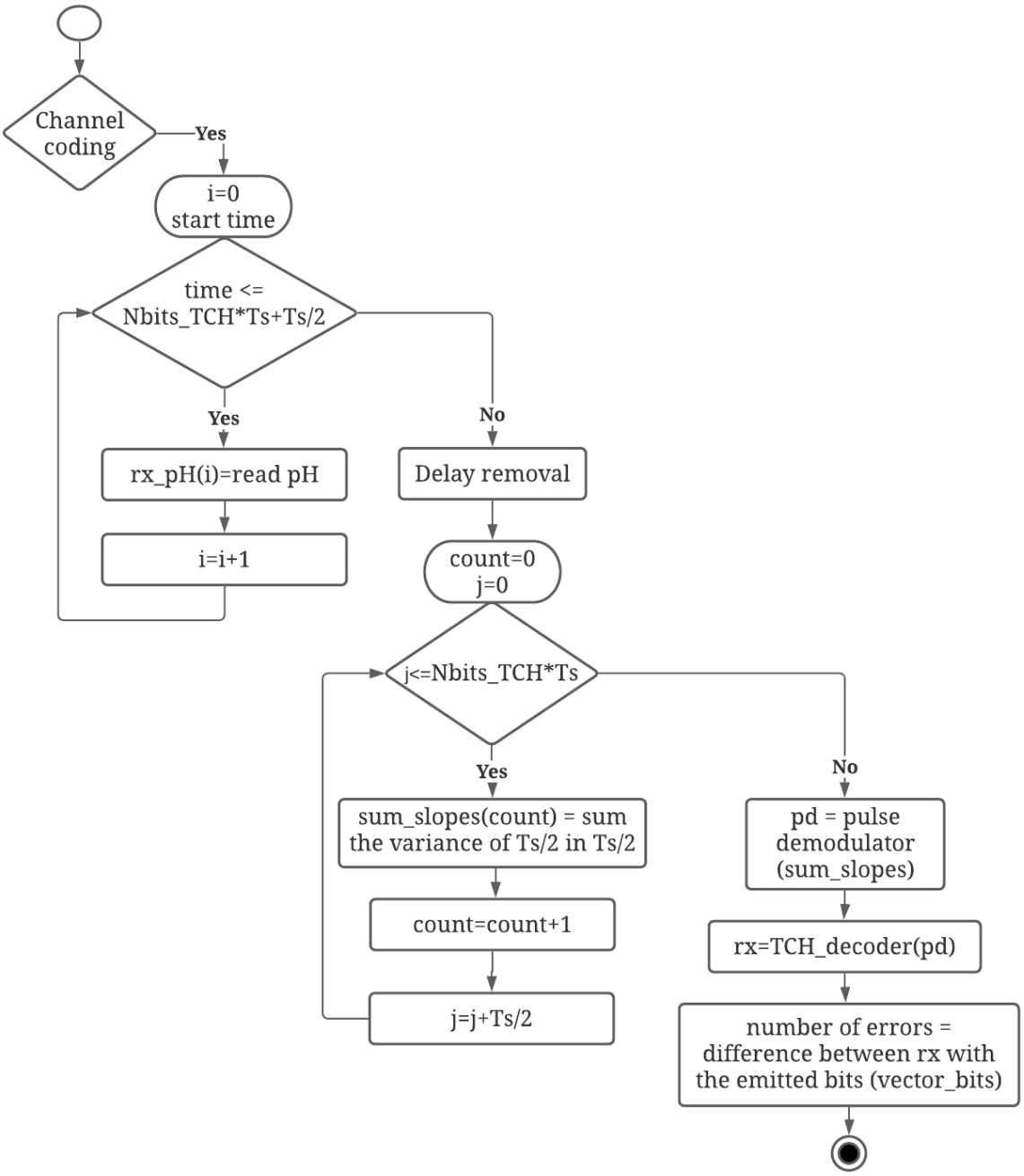


FIGURE 4.16. Receiver workflow for the TCH codes.

Chapter 5

Results and discussion

To evaluate the proposed MC scheme based on TCH codes, several simulations with different configurations were tested, followed by an experimental evaluation. In the next sections, are detailed both the simulation and experimental results and discussion.

5.1. Simulation results

5.1.1. Simulation setup

Parameter	Value
Deterministic model	Point transmitter to a spherical absorbing receiver
Statistical model	Gaussian approximating Poisson
Paired $tx - rx$ link distance (d_0)	10 (micrometers)
Diffusion Coefficient (D)	79.4 (micrometers^2/seconds)
Transmitter radius and receiver radius (a_{tx}, a_{rx})	5 (micrometers)
Symbol duration (T_s)	$\tau \frac{(d_0)^2}{6D}$ (seconds)

TABLE 5.1. Simulation parameters.

The purpose of these simulations was to test different modulations and bit encoding solutions combined with TCH channel codes. We compared the performance of some

previously studied codes for MC against the TCH codes which we propose here as a potential alternative for MC scenarios. For the channel correcting codes, we chose to test the Hamming Codes (HC) as benchmark because they are basic and easy to implement and the ISI-mtg codes from [**43**] because they were designed specifically for an MC environment. Finally, we test the TCH codes for different codes lengths. For the uncoded scenario and for the HC we used threshold hard decision. For the TCH codes and for the ISI-mtg codes, LLR soft decisions were employed. Both these bit decisions are detailed in Chapter 3.4.

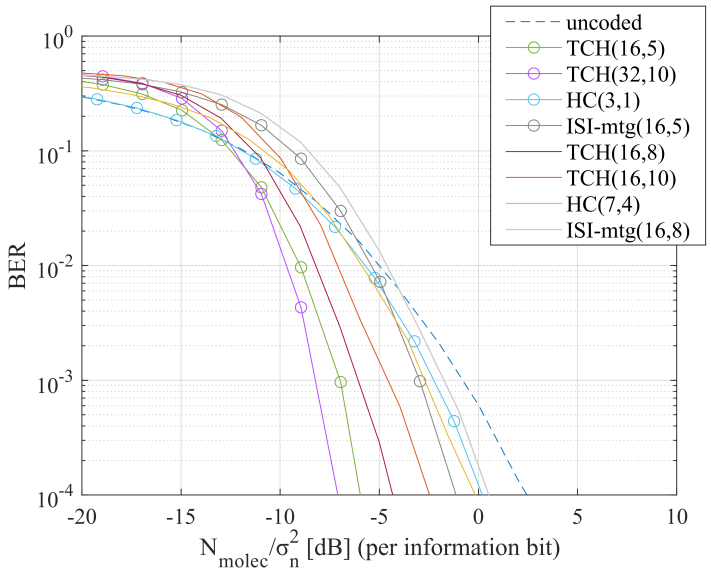
The main parameters used in the simulations are described in Table 5.1. The symbol duration (T_s) can have a critical impact on the behaviour of the system. To evaluate the impact of this parameter we considered different values for the scaling factor τ which is used to define the symbol duration, as shown in Table 5.1. The simulations were performed using NRZ and Manchester pulses, at different symbol durations, with two distinct types of modulation: OOK and BCSK. In the case of OOK modulation we have $x[k] \in \{0, 1\}$, whereas in the BCSK we use $x[k] \in \{-1, 1\}$. For the BCSK we assume the existence of destructor molecules that can reduce the concentration of the carrier molecule.

It is important to emphasize that while the channel model may not be the most adequate for this particular scenario, the proposed scheme does not depend on a specific channel and does not require channel knowledge. Therefore, using this channel model still enables an evaluation of the system behaviour, which is then additionally validated through experimental tests. The results will be presented as the relation between the BER and N_{molec}/σ_n^2 (per information bit) in dB, where σ_n^2 denotes the variance of the environmental noise.

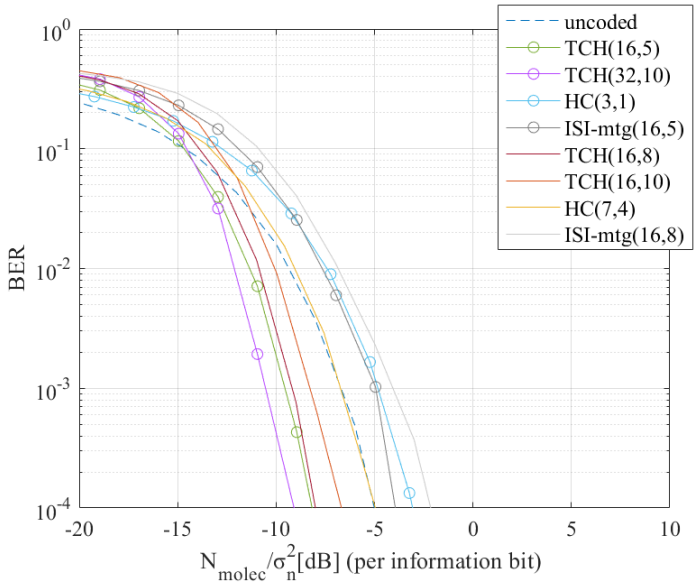
5.1.2. Effects of symbol duration and type of pulses

Figure 5.1a and Figure 5.2a present the results for NRZ pulses, whereas in Figure 5.1b and Figure 5.2b present the results for the tested Manchester pulses. Both Figure 5.1 and Figure 5.2 were simulated using BCSK modulation. When Manchester codes were used, the tested channel coding methods had an overall better performance. The big difference between these figures relies on the τ value, which in Figure 5.1 is 2, while in Figure 5.2 is increased to 10. Analyzing the results for the different τ values, it is noticeable a BER improvement in Figure 5.2 which has a higher τ value and thus a higher T_s. Increasing the T_s means that the molecules will have more time to diffuse and reach the receiver, reducing

the ISI. It is important to note however that a higher T_s also means a transmission rate decrease.

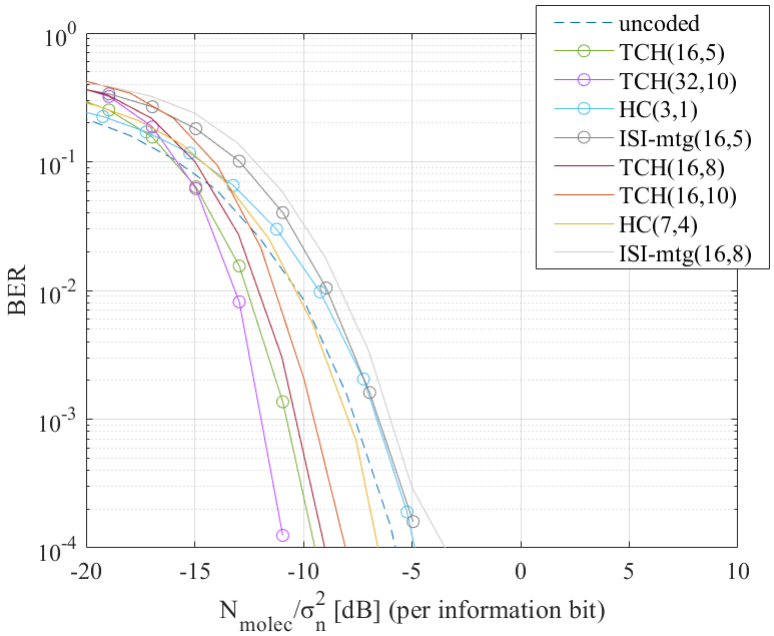


(a) NRZ

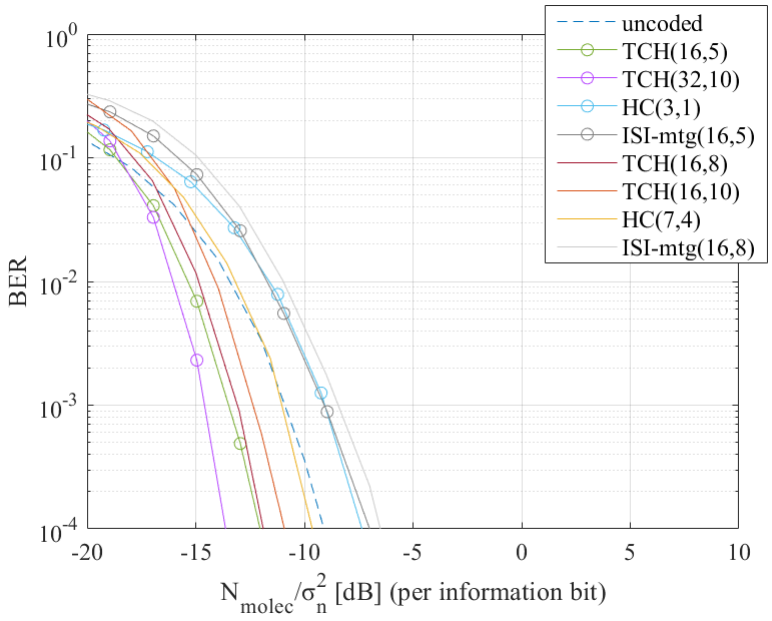


(b) Manchester

FIGURE 5.1. BER comparison for all the considered channel coding methods, using BCSK modulation and with τ=2.



(a) NRZ



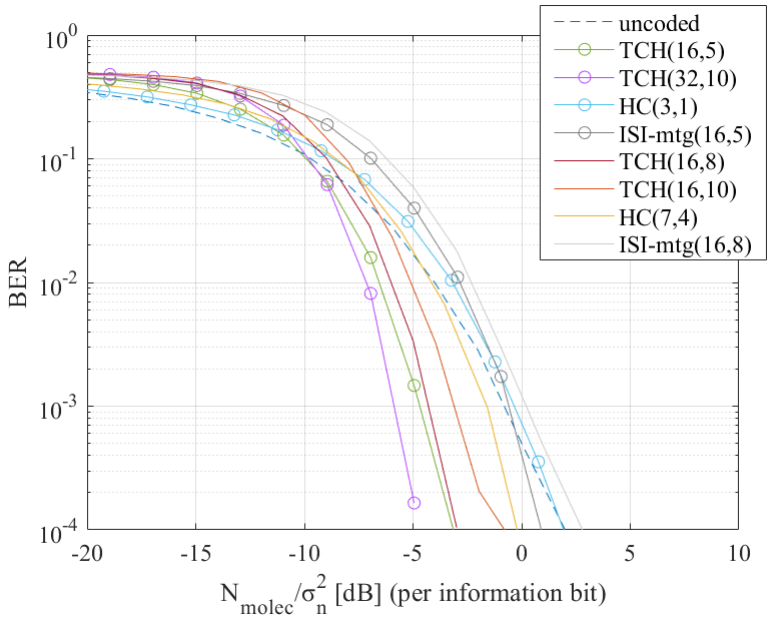
(b) Manchester

FIGURE 5.2. BER comparison for all the considered channel coding methods, using BCSK modulation and with τ=10.

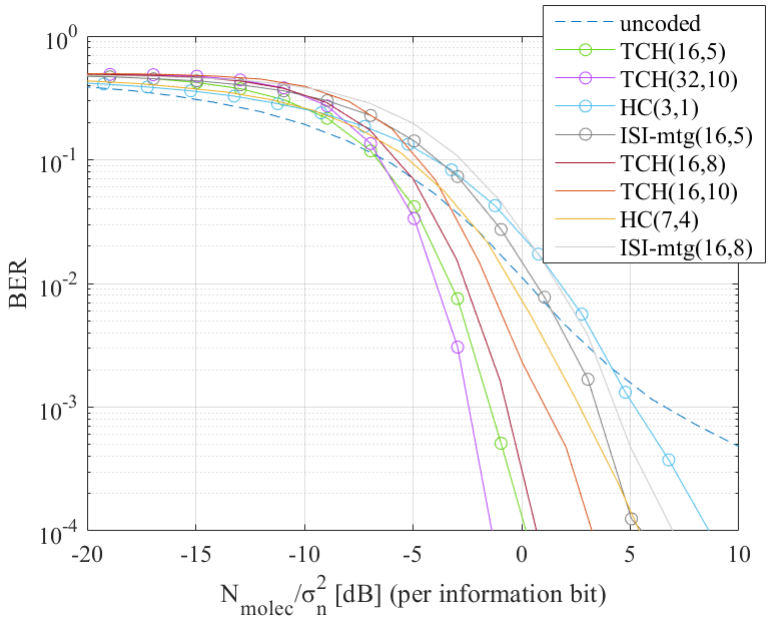
5.1.3. Modulation techniques

We tested the system with the OOK modulation, whose results are shown in Figure 5.3. In Figure 5.3a are the results for NRZ pulses with τ=10 and in Figure 5.3b are the results for Manchester pulses at τ=2. Comparing Figure 5.2a with Figure 5.3a and Figure 5.1b

with Figure 5.3b, we can see that the tested channel codes had worse BER results in Figure 5.3, when OOK modulation was used.



(a) NRZ, τ=10



(b) Manchester, τ=2

FIGURE 5.3. BER comparison for all the considered channel coding methods, using OOK modulation.

5.1.4. Comparsion between different codes

To ensure a fair comparison between the tested codes, similar sets of code rates were used comprising both lower codes rates and higher codes rates. We remind that lower coding

rates may obtain better performances but decrease the information bit rate. Higher coding rates can improve the bit rate sacrificing some BER performance.

In Figures 5.1, 5.2 and 5.3 we can observe that TCH codes clearly had a lower BER than codes with similar rates:

• TCH(16,5) and TCH(32,10) codes performed better than HC(3,1) and ISI-mtg(16,5). For instance, in Figure 5.1b for a BER reference of 10^{-2} it can be observed a gain between 4 dB to 5 dB in terms of N_{molec}/σ_n^2;

• TCH(16,8) and TCH(16,10) codes performed better than HC(7,4) and ISI-mtg(16,8). This is illustrated, for example, in Figure 5.2a for a BER reference of 10^{-2}, they had a gain between 1 dB to 2 dB compared with the HC(7,4) and a gain between 3 dB to 4 dB compared with the ISI-mtg(16,8).

Furthermore, TCH codes with higher codes rates had a better BER performance than other codes with lower rates:

• TCH(16,8) and TCH(16,10) were able to perform better than the HC(3,1) and the ISI-mtg(16,5). For example, in Figure 5.3a and considering a BER reference of 10^{-2}, it can be observed a gain between 2 dB to 3 dB in terms of N_{molec}/σ_n^2.

5.1.5. Effect of code length

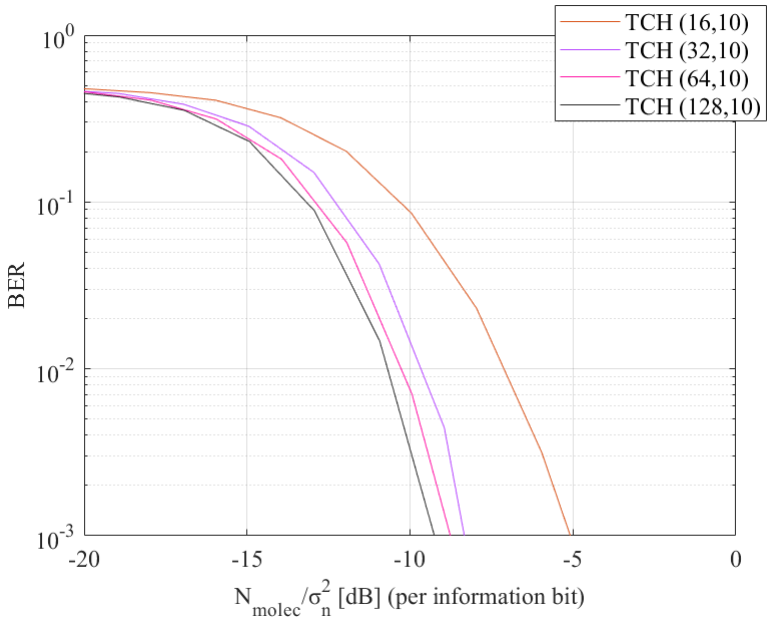


FIGURE 5.4. TCH codewords length.

In Figure 5.4 are shown the results for different TCH codewords lengths. As expected, the TCH codes with longer codewords, namely the TCH(128,10) when compared to the

TCH(16,10), achieve better BER performances. It is important to notice that longer codewords cause an additional delay. The results of our simulations proved that the TCH codes achieved a better BER performance against the other tested channel correcting codes, independently of the symbol duration, type of pulse and modulation.

5.2. Experimental results

To add value to our simulation results we developed a macro scale proof-of-concept MCvD system using TCH codes to encode information. Because this experiment was performed on a macro scale testbed some parameters previously adopted for the simulation had to be adjusted. Next, we present the results of the assembled experiment using the uncoded scenario against using TCH codes. Also detailed, are the experimental system validations needed to have a viable system.

5.2.1. Experimental system validation

At the beginning of this experiment it was important to do several tests to ensure the reliability of the system:

1) pH probe calibration
2) Channel response to pH changes
3) Bits reception

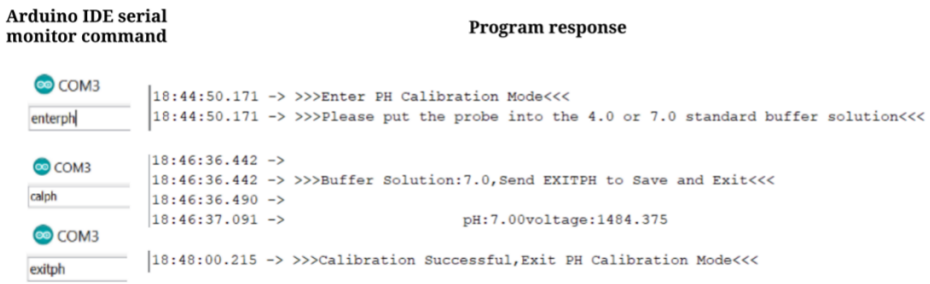


FIGURE 5.5. Calibration Arduino IDE serial monitor, code from [**68**].

Regarding aspect number 1), a calibration program was used in the Arduino IDE to calibrate the pH probe. This program was developed by the pH probe manufacturers [**68**] and made available to the users. The main goal of this program is to establish a voltage value for the pH calibrations solutions. First, we submerged the probe into one standard calibration solution (ph=4 or pH=7). Next as in Figure 5.5, we run the program writing in the serial monitor of the Arduino IDE, by order, the following keywords:

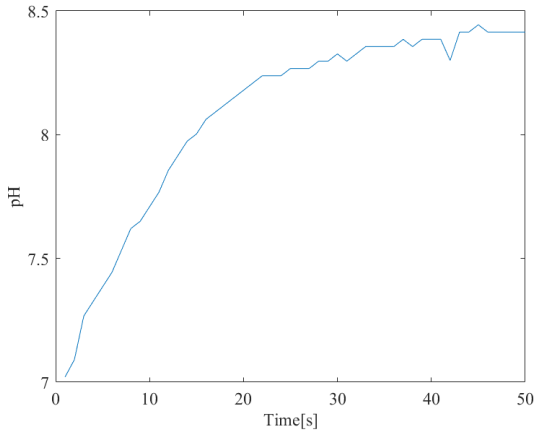
• "enterph" - with this keyword the program will enter the calibration mode, asking for the user to submerge the pH probe into one of the calibrations solutions;

• "calph" - this command will establish which voltage belongs to the submerged calibration solution;

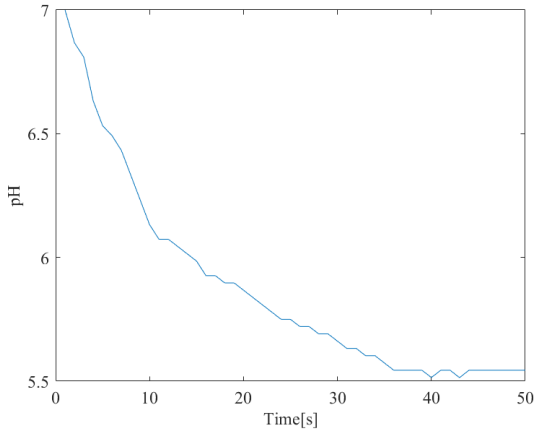
• "exitph" - writing this will end the calibration mode and save the voltage value.

We repeat the process for the other calibration solution, in order to have a good pH probe calibration. These voltage values was then passed to our main control program in MATLAB.

It is important to underline that a good calibration is essential to guarrantee the reliability of the results. Also, it is important that frequent calibrations are made since this sensor has a high sensitivity to current changes, such as the computer power connection. In this experience, we follow this good use principle and calibrated the pH probe each time that it was used.



(a) pH=9 emitted for 30s



(b) pH=3 emitted for 30s

FIGURE 5.6. pH behavior in a MCvD channel.

Concerning point 2), analyzing the channel response is an extremely important aspect, since this work assumes an MCvD system. This implies to study how the used molecules, pH molecules, behave when released into the system. One of the found obstacles was when using a base of pH=10 and an acid of pH=4, the system responded faster when the base was transmitted than when the acid was transmitted. This problem may be related to the chemicals used when combined with tap water. This could have led to the formation of non-desirable compounds affecting the pH readings. To solve this we decided to weaken the base and to strengthen the acid. Thus, we decrease both pH solutions, to guarantee that they still could cancel each other. The changed values were for the base solution pH=9 and for the acid solution pH=3. In Figure 5.6a is the channel response when a base is transmitted, considering pH=9. Figure 5.6b is the channel behavior when an acid is emitted to the channel, considering pH=3. Both these experiments were conducted under the same circumstances and the solution (acid or basic) was emitted during 30 seconds. Figure 5.6 show the molecule's reception after the emitting period when they reach the receiver. It is visible one of the diffusion effects which makes some of the molecules to arrive late due to their random movement, making the reading to take some time to establish a value. This effect is one of the main causes for ISI and can result in bit errors at the receiver.

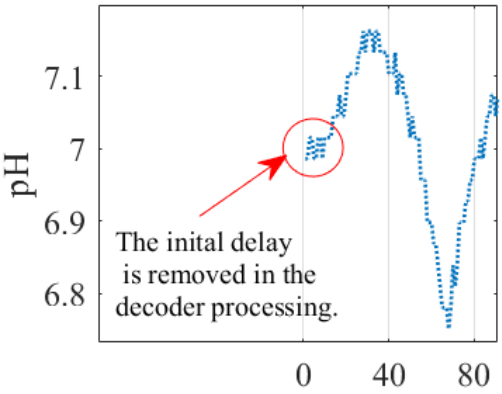


FIGURE 5.7. Initial reception delay.

Finally, analyzing 3), the way the reception of the bits happens is a crucial part of how the decoder will work. The released molecules take time to propagate and therefore reaching the receiver (initial delay), Figure 5.7. One important step comparing the emitted with the received signal is to do a delay removal. Since no synchronization between the transmitter and the receiver was implemented, an ad-hoc method was utilized in order to guarantee an accurate comparison. This was based on checking when the received signal

started to change its initial stable values. Then, the initial stable values are retrieved from the signal. The decoder starts to decode from the changing values until the end of the sampling period.

In Figure 5.8, we show an example of a received signal in Manchester format with the corresponding sent bitstream, after the initial delay removal. As this Figure shows, at an experimental level the received molecules waveform is not regular. The waveform for the same amount of the molecules transmitted could have different sizes. This could happen because of the phenomenons explained before: the environmental noise and the random movement of the molecules which could cause a high ISI, generating the irregularities in the received waveform. An example of a potential error in the received molecules when decoding into bits is, for instance, the third bit which is a 0 and could be decoded as an error since the corresponding wave is not centered within the T_s.

In this Figure, is also visible the effect of the Manchester format in the bit transmissions. When a change happens from different bits (bit-1 to bit-0 or bit-0 to bit-1) the size of the wave increases. Therefore, the pH values have a bigger fluctuation. It is important to underline that although we see a bigger pH difference when these bits transactions occur, they will never compromise the system by going besides pH scale buffer. This means, with the characteristics of the adopted system, no information will be lost because of these fluctuations.

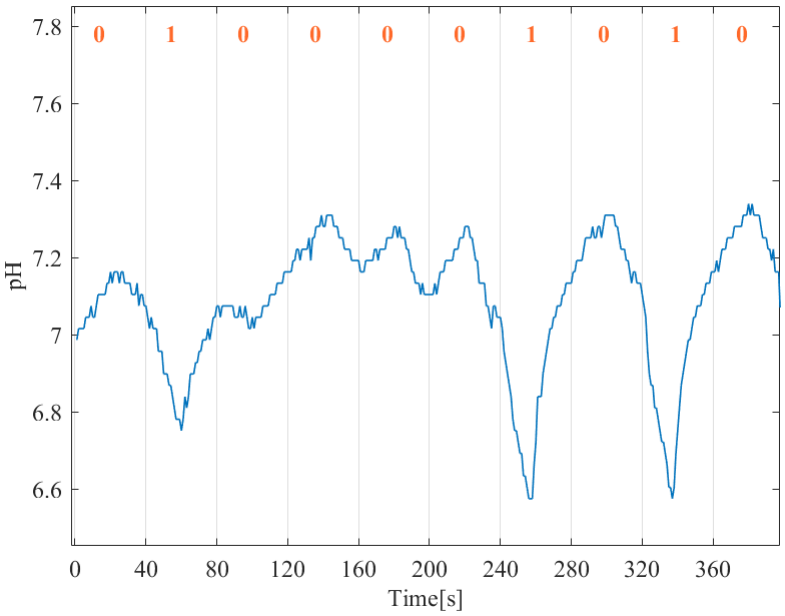


FIGURE 5.8. Example of 10 received bits with T_s=40s after initial delay removal.

5.2.2. Overall system performance

The experimental evaluation was carried out by testing the communication without channel correcting codes and by testing it with TCH(16,10) codes. We chose to work with the TCH(16,10) codes because they have a relatively high coding rate while still maintaining good error capabilities. In this experiment, we used Manchester pulses, for better channel robustness, combined with BCSK modulation. The bit decision applied was the LLR-based soft decision as in Chapter 3.4.2. In our experiment, we collected data from 100 transmissions, where each transmission comprised a block with 100 randomly generated bits. To test how the symbol interval (T_s) could affect the received bitstream, we performed the experience for T_s=20s and then for T_s=40s. Is it important to notice that due to the adopted experimental setup, there could be other existing compounds formed by the mixing of the chemicals with the tap water used, which can interfere with the channel characteristics and affect the pH probe readings. Nevertheless, the implemented communication scheme does not require channel state knowledge to accomplish the detection.

In Figure 5.9, it can be observed that when a larger T_s is adopted the bit error rate decreases both in the uncoded and coded cases. This was also observed in the simulation results, in Chapter 5.1.2. and was expected since, although affecting the transmission rate, a higher symbol interval gives the system more time to receive most of the late molecules, thus decreasing the ISI. According to the obtained results, TCH codes corrected $\approx 30\%$ errors when $T_s = 40s$ and $\approx 45\%$ errors when $T_s = 20s$ comparing to the uncoded case. Analyzing these results, it is clear that fewer errors were obtained when TCH codes were applied, confirming the TCH codes efficiency in correcting errors in an MC environment.

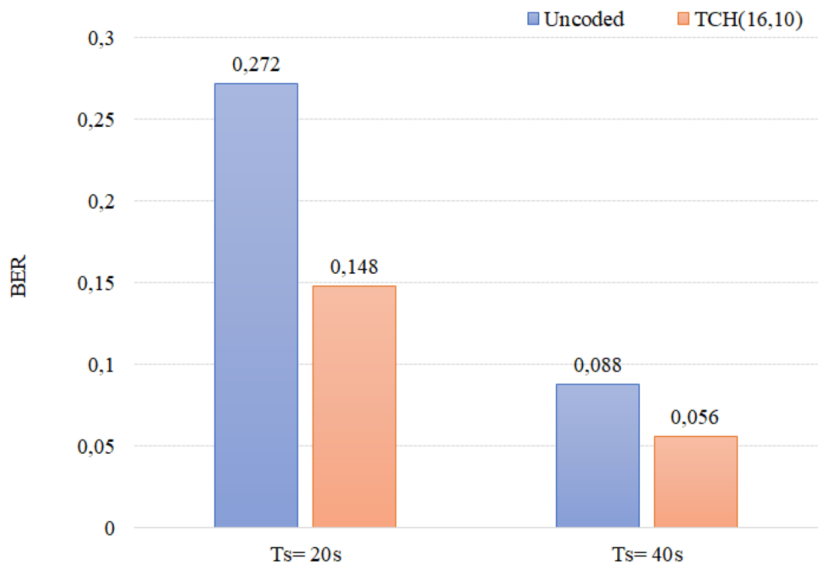


FIGURE 5.9. Uncoded and TCH(16,10) performance for Ts=20s and Ts=40s.

Chapter 6

Conclusions

In this dissertation, an overview of existing MC systems and their constituent components was undertaken. Also, a state of the art regarding macroscale MC experiments was presented. Molecular communication is described as a strong branch of communication engineering, where the applications and the challenges greatly differ, having great possibilities of being introduced into next future generations of wireless communications.

As the main objetcive of this dissertation, we studied the adoption of TCH codes as a low complexity promising solution for enabling reliable MC. Adequate detection and decoding methods which do not require CSI acquisition were presented and evaluated through end-to-end simulations. Furthermore, to help fulfill major gaps identified in the literature between theoretical and experimental MC systems and to validate the reliability achieved with the proposed approach, we implemented a macroscale MC platform. The adopted channel was an aqueous solution. To generate chemical pulses based on the digital information and transmit them we used the pH value of the medium to differentiate the different information bits, as this is a good approach to a biological environment for in-body communications, such as the human blood. At the transmitter, different encoded pulses were considered namely, NRZ and Manchester, combined with either OOK or BCSK modulations. Other channel codes were evaluated as benchmarks for the TCH codes, namely, Hamming codes and the ISI-mtg codes that were designed specifically for this context.

There are big challenges when implementing MC systems, in particular, because these systems are hard to predict/model. Therefore, it is difficult to avoid errors when receiving the bits. The results obtained, both from simulation and physical implementation, showed that the newly introduced MC scheme based on TCH codes was able to correct errors without introducing more complexity. In fact, TCH codes with higher rates were able to perform better than other studied codes with lower rates. In addition, Manchester pulses for coded and uncoded situations proved to be essential when operating in a channel that can saturate, as in the case of the pH-based experimental tests. These pulses also proved to work better in the simulation than the NRZ pulses, giving more robustness

to the communication in an MC channel with an absorbing receiver. The results are a promising possible scenario to be implemented in the many applications of future IoBNT.

6.1. Future work

There is a long way of research in this area until would be possible to have fully functional IoBNT networks. Concerning our work, for the future other aspects could be tested. One important feature of the TCH codes is the synchronism capability. In the simulation, we assumed a perfect synchronism and in the experimental system, this was adjusted using an ad-hoc technique. For that reason, synchronization between the transmitter and the receiver should be added. Also, other types of modulation could be tested regarding the type of molecules and the timing, aiming to increase the rate. A method that could also help to increase the system rate is the implementation of multiple-input and multiple-output (MIMO) techniques.

At the experimental level, regarding our setup, more bits should be tested and better BER rates achieved. The MC field in the macroscale is still in its infancy in regarding to understanding it and exploiting it. Therefore, the opportunities for making experiences in this field are many. For instance, in this work, we tested a liquid-driven channel but also an air-driven channel using other types of molecules carriers and sensors could be tried and analyzed. Also, an experimental MIMO scheme should be very interesting to implement to increase the data rate, since in MC and especially in macro scale, it is very low.

Bibliography

[1] M. Z. Chowdhury, M. Shahjalal, S. Ahmed and Y. M. Jang, "6G Wireless Communication Systems: Applications, Requirements, Technologies, Challenges and Research Directions," in IEEE Open Journal of the Communications Society, vol. 1, pp. 957-975, 2020, doi: 10.1109/OJCOMS.2020.3010270.

[2] Y. Lu, R. Ni and Q. Zhu, "Wireless Communication in Nanonetworks: Current Status, Prospect and Challenges," in IEEE Transactions on Molecular, Biological and Multi-Scale Communications, vol. 6, no. 2, pp. 71-80, Nov. 2020, doi: 10.1109/TMBMC.2020.3004304.

[3] F. Akyildiz, M. Pierobon and S. Balasubramaniam, "Moving forward with molecular communication: from theory to human health applications [point of view]," in Proceedings of the IEEE, vol. 107, no. 5, pp. 858-865, May 2019, doi: 10.1109/JPROC.2019.2913890.

[4] J. C. Anderson, E. J. Clarke, A. P. Arkin and C. A. Voigt, "Environmentally controlled invasion of cancer cells by engineered bacteria," J. Mol. Biol., vol. 355, no. 4, pp. 619–627, 2006.

[5] C. M. J. Pieterse and M. Dicke, "Plant interactions with microbes and insects: From molecular mechanisms to ecology," Trends Plant Sci., vol. 12, no. 12, pp. 564–569, 2007.

[6] S. Canovas-Carrasco, A. Garcia-Sanchez and J. Garcia-Haro, "The IEEE 1906.1 Standard: Some Guidelines for Strengthening Future Normalization in Electromagnetic Nanocommunications," in IEEE Communications Standards Magazine, vol. 2, no. 4, pp. 26-32, December 2018, doi: 10.1109/MCOMSTD.2018.1700082.

[7] K. Yang, D. Bi, Y. Deng, R. Zhang, M. M. U. Rahman, N. A. Ali, M. A. Imran, J. M. Jornet, Q. H. Abbasi and A. Alomainy, "A comprehensive survey on hybrid communication in context of molecular communication and terahertz communication for body-centric nanonetworks," IEEE Trans. Mol. Biol. Multi-Scale Commun., vol. 6, no. 2, pp. 107– 133, 2020.

[8] S. Hiyama and Y. Moritani, "Molecular communication: Harnessing biochemical materials to engineer biomimetic communication systems," Nano Commun. Netw., vol. 1, no. 1, pp. 20–30, May 2010.

[9] C. J. Donnelly, M. Fainzilber and J. L. Twiss, "Subcellular communication through RNA transport and localized protein synthesis," Traffic, vol. 11, no. 12, pp. 1498–1505, 2010.

[10] M. Mittelbrunn and F. Sánchez-Madrid, "Intercellular communication: Diverse structures for exchange of genetic information," Nature Rev. Mol. Cell Biol., vol. 13, no. 5, pp. 328–335, 2012.

[11] M. Kuscu, E. Dinc, B. A. Bilgin, H. Ramezani and O. B. Akan, "Transmitter and Receiver Architectures for Molecular Communications: A Survey on Physical Design With Modulation, Coding and Detection Techniques," in Proceedings of the IEEE, vol. 107, no. 7, pp. 1302-1341, July 2019, doi: 10.1109/JPROC.2019.2916081.

[12] L. Organick et al., "Random access in large-scale DNA data storage," Nature Biotechnol., vol. 36, no. 3, pp. 242–248, 2018.

[13] D. E. Clapham, "Calcium signaling," Cell, vol. 131, no. 6, pp. 1047–1058, Dec. 2007.

[14] T. Nakano, T. Suda, M. Moore, R. Egashira, A. Enomoto and K. Arima, “Molecular communication for nanomachines using intercellular calcium signaling,” in Proc. 5th IEEE Conf. Nanotechnol., Jul. 2005, pp. 478–481

[15] A. Jonsson, T. A. Sjöström, K. Tybrandt, M. Berggren and D. T. Simon, “Chemical delivery array with millisecond neurotransmitter release,” Sci. Adv., vol. 2, no. 11, 2016, Art. no. 1601340.

[16] S. M. Abd El-atty, A. El-taweel and S. El-Rabaie, ”Aspects of nanoscale information transmission in nanonetworks-based molecular communication,” 2016 Fourth International Japan-Egypt Conference on Electronics, Communications and Computers (JEC-ECC), 2016, pp. 131-134, doi: 10.1109/JEC-ECC.2016.7518985.

[17] M. Pierobon and I. F. Akyildiz, ”A physical end-to-end model for molecular communication in nanonetworks,” in IEEE Journal on Selected Areas in Communications, vol. 28, no. 4, pp. 602-611, May 2010, doi: 10.1109/JSAC.2010.100509.

[18] D. Bi, A. Almpanis, A. Noel, Y. Deng and R. Schober, ”A Survey of Molecular Communication in Cell Biology: Establishing a New Hierarchy for Interdisciplinary Applications,” in IEEE Communications Surveys and Tutorials, doi: 10.1109/COMST.2021.3066117. interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[19] S. Kadloor, R. S. Adve and A. W. Eckford, ”Molecular Communication Using Brownian Motion With Drift,” in IEEE Transactions on NanoBioscience, vol. 11, no. 2, pp. 89-99, June 2012, doi: 10.1109/TNB.2012.2190546.

[20] C. Zoppou and J. H. Knight, “Analytical solution of a spatially variable coefficient advection–diffusion equation in up to three dimensions,” Appl. Math. Model., vol. 23, no. 9, pp. 667–685, 1999.

[21] A. A. Kornyshev, A. M. Kuznetsov, E. Spohr and J. Ulstrup, “Kinetics of proton transport in water,” J. Phys. Chem. B, vol. 107, no. 15, pp. 3351–3366, 2003.

[22] V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel and R. Schober, ”Channel Modeling for Diffusive Molecular Communication—A Tutorial Review,” in Proceedings of the IEEE, vol. 107, no. 7, pp. 1256-1301, July 2019, doi: 10.1109/JPROC.2019.2919455.

[23] S. Qiu, “Molecular Communication Systems: Design, Modelling and Experimentation,” PhD thesis, University of Warwick, 2017

[24] N. Farsad, H. B. Yilmaz, A. Eckford, C.-B. Chae, and W. Guo, “A comprehensive survey of recent advancements in molecular communication,” IEEE Commun. Surveys Tuts., vol. 18, no. 3, pp. 1887–1919, 3rd Quart., 2016.

[25] A. Noel, D. Makrakis and A. Hafid, “Channel impulse responses in diffusive molecular communication with spherical transmitters,” in Proc. Biennial Symp. Commun., 2016. [Online]. Available: http://arxiv.org/abs/1604.04684.

[26] H. Arjmandi, A. Ahmadzadeh, R. Schober and M. N. Kenari, “Ion channel based bio-synthetic modulator for diffusive molecular communication,” IEEE Trans. Nanobiosci., vol. 15, no. 5, pp. 418–432, Jul. 2016

[27] L. S. Meng, P. C. Yeh, K. C. Chen and I. F. Akyildiz,"On receiver design for diffusion-based molecular communication," IEEE Trans. Signal Process., vol. 62, no. 22, pp. 6032–6044, Nov. 2014.

[28] W. Guo et al., "Molecular communications: Channel model and physical layer techniques," IEEE Wireless Commun. Mag., vol. 23, no. 4, pp. 120–127, Aug. 2016.

[29] A. Noel, K. C. Cheung and R. Schober, "Diffusive molecular communication with disruptive flows," in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2014, pp. 3600–3606.

[30] D. Arifler and D. Arifler, "Monte Carlo analysis of molecule absorption probabilities in diffusion-based nanoscale communication systems with multiple receivers," IEEE Trans. Nanobiosci., vol. 16, no. 3, pp. 157–165, Apr. 2017.

[31] Y. Deng, A. Noel, W. Guo, A. Nallanathan and M. Elkashlan, "Analyzing large-scale multiuser molecular communication via 3-D stochastic geometry," IEEE Trans. Mol. Biol. Multi-Scale Commun., vol. 3, no. 2, pp. 118–133, Jun. 2017.

[32] A. Akkaya, H. B. Yilmaz, C. Chae and T. Tugcu, "Effect of Receptor Density an on Signal Reception in Molecular Communication via Diffusion With an Absorbing Receiver," in IEEE Communications Letters, vol. 19, no. 2, pp. 155-158, Feb. 2015, doi: 10.1109/LCOMM.2014.2375214.

[33] Y. Deng, A. Noel, W. Guo, A. Nallanathan and M. Elkashlan, "Analyzing Large-Scale Multiuser Molecular Communication via 3-D Stochastic Geometry," in IEEE Transactions on Molecular, Biological and Multi-Scale Communications, vol. 3, no. 2, pp. 118-133, June 2017, doi: 10.1109/TMBMC.2017.2750145.

[34] A. Ahmadzadeh, H. Arjmandi, A. Burkovski and R. Schober, "Comprehensive reactive receiver modeling for diffusive molecular communication systems: Reversible binding, molecule degradation and finite number of receptors," IEEE Trans. Nanobiosci., vol. 15, no. 7, pp. 713–727, Oct. 2016.

[35] C. T. Chou, "Impact of receiver reaction mechanisms on the performance of molecular communication networks," IEEE Trans. Nanotechnol., vol. 14, no. 2, pp. 304–317, Mar. 2015.

[36] McGuiness, Daniel (2020) Macro-Scale Molecular Communications. PhD thesis, University of Liverpool.

[37] A. O. Kislal, B. C. Akdeniz, C. Lee, A. E. Pusane, T. Tugcu and C. B. Chae, "ISI-Mitigating Channel Codes for Molecular Communication Via Diffusion," IEEE Access, vol. 8, pp. 24588–24599, 2020, doi: 10.1109/ACCESS.2020.2970108.

[38] M. S. Leeson and M. D. Higgins, "Forward Error Correction for Molecular Communications," Nano Communication Networks, vol. 3, no. 3, pp. 161–167, 2012. doi:10.1016/j.nancom.2012.09.001.

[39] Y. Lu, M. D. Higgins and M. S. Leeson, "Diffusion based molecular communications system enhancement using high order hamming codes," 2014 9th International Symposium on Communication Systems, Networks and Digital Sign (CSNDSP), 2014, pp. 438-442, doi: 10.1109/CSNDSP.2014.6923869.

[40] Y. Lu, M. D. Higgins and M. S. Leeson, "Comparison of Channel Coding Schemes for Molecular Communications Systems," in IEEE Transactions on Communications, vol. 63, no. 11, pp. 3991-4001, Nov. 2015, doi: 10.1109/TCOMM.2015.2480752.

[41] Y. Lu, M. D. Higgins and M. S. Leeson, “Self-orthogonal convolutional codes (SOCCs) for diffusion-based molecular communication systems,” in IEEE International Conference on Communications, Sep. 2015, vol. 2015-September, pp. 1049–1053, doi: 10.1109/ICC.2015.7248461.

[42] M. B. Dissanayake, Y. Deng, A. Nallanathan, E.M. N. Ekanayake and M. Elkashlan,“Reed Solomon Codes for Molecular Communication with a Full Absorption Receiver,” IEEE Communications Letters, vol. 21, no. 6, pp. 1245–1248, Jun. 2017, doi: 10.1109/LCOMM.2017.2671858.

[43] A. O. Kislal, B. C. Akdeniz, C. Lee, A. E. Pusane, T. Tugcu and C. -B. Chae, ”ISI-Mitigating Channel Codes for Molecular Communication Via Diffusion,” in IEEE Access, vol. 8, pp. 24588-24599, 2020, doi: 10.1109/ACCESS.2020.2970108.

[44] M. Ş. Kuran, H. B. Yilmaz, I. Demirkol, N. Farsad and A. Goldsmith, ”A Survey on Modulation Techniques in Molecular Communication via Diffusion,” in IEEE Communications Surveys and Tutorials, vol. 23, no. 1, pp. 7-28, Firstquarter 2021

[45] Vahid Jamali,”Design and Analysis of Molecular Communication Systems,” PhD thesis, Friedrich-Alexander University Erlangen-Nuremberg, 2019

[46] G. Aminian, M. Mirmohseni, M. N. Kenari and F. Fekri, “On the capacity of level and type modulations in molecular communication with ligand receptors,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Hong Kong, China, Jun. 2015, pp. 1951–1955.

[47] S. Galmés and B. Atakan, “Performance analysis of diffusion-based molecular communications with memory,” IEEE Trans. Commun., vol. 64, no. 9, pp. 3786–3793, Sep. 2016.

[48] N. Garralda, I. Llatser, A. Cabellos-Aparicio, E. Alarcón and M. Pierobon, “Diffusion-based physical channel identification in molecular nanonetworks,” Nano Commun. Netw., vol. 2, no. 4, pp. 196–204, 2011.

[49] N. Farsad, W. Guo and A. W. Eckford, “Tabletop molecular communication: Text messages through chemical signals,” PLoS ONE, vol. 8, no. 12, Dec. 2013, doi: 10.1371/journal.pone.0082935.

[50] N. Farsad, D. Pan and A. Goldsmith, ”A Novel Experimental Platform for In-Vessel Multi-Chemical Molecular Communications,” GLOBECOM 2017 - 2017 IEEE Global Communications Conference, 2017, pp. 1-6, doi: 10.1109/GLOCOM.2017.8255058.

[51] L. Grebenstein et al., “Biological Optical-to-Chemical Signal Conversion Interface: A Small-Scale Modulator for Molecular Communications,” IEEE Transactions on Nanobioscience, vol. 18, no. 1, pp. 31–42, Jan. 2019, doi: 10.1109/TNB.2018.2870910

[52] F. A. B. Cercas, M. Tomlinson and A. A. Albuquerque, ”TCH: A New Family of Cyclic Codes Length 2/sup m/,” Proceedings. IEEE International Symposium on Information Theory, 1993, pp. 198-198, doi: 10.1109/ISIT.1993.748512.

[53] F. Cercas, J. C. Silva, N. Souto and R. Dinis, ”Optimum bit-mapping of TCH codes,” 2009 International Workshop on Satellite and Space Communications, 2009, pp. 92-96, doi: 10.1109/IWSSC.2009.5286417.

[54] F. Cercas, ”A new family of codes for simple receiver implementation”, Ph.D, Instituto Superior Técnico, Lisbon, 1996.

[55] RF Wireless World, "WRZ vs NRZ — Comparison between RZ and NRZ Line Coding", [Online] Available: https://www.rfwireless-world.com/Terminology/Comparison-between-RZ-and-NRZ-Pulse-Shapes.html (visited 17/10/2021).

[56] T. A. Vu and P. Van Thanh, "A simple diagram for data transmission using Manchester code," 2016 International Conference on Advanced Technologies for Communications (ATC), 2016, pp. 453-456, doi: 10.1109/ATC.2016.7764825.

[57] B. R. Saltzberg, "An improved Manchester code receiver," IEEE International Conference on Communications, Including Supercomm Technical Sessions, 1990, pp. 698-702 vol.2, doi: 10.1109/ICC.1990.117167.

[58] M. ¸S. Kuran, H. B. Yilmaz, T. Tugcu and B. Özerman, "Energy model for communication via diffusion in nanonetworks," Nano Commun. Netw., vol. 1, no. 2, pp. 86–95, 2010.

[59] M. U. Mahfuz, D. Makrakis and H. T. Mouftah, "On the characteristics of concentration-encoded multi-level amplitude modulated unicast molecular communication," in Proc. 24th Can. Conf. Elect. Comput. Eng. (CCECE), Niagara Falls, ON, Canada, May 2011, pp. 312–316.

[60] B. Arifin, "GChem: Assisting students learning in general chemistry," 2012 IEEE Symposium on Humanities, Science and Engineering Research, 2012, pp. 323-327, doi: 10.1109/SHUSER.2012.6268865.

[61] J. Y. Choi, H. G. Pandit and R. R. Rhinehart, "A Process Simulator for pH Control Studies," 1993 American Control Conference, 1993, pp. 2594-2595, doi: 10.23919/ACC.1993.4793362.

[62] N. Farsad and A. Goldsmith, "A molecular communication system using acids, bases and hydrogen ions," 2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2016, pp. 1-6, doi: 10.1109/SPAWC.2016.7536834.

[63] A. Nayyar and V. Puri, "A review of Arduino board's, Lilypad's and Arduino shields," 3rd International Conference on Computing for Sustainable Global Development (INDIACom), 2016.

[64] Arduino, "Arduino - Introduction," 2018, [Online] Available: https://www. arduino.cc/en/Guide/Introduction, (visited 2/11/2021).

[65] Verder, "How do peristaltic pumps work?", [Online] Available: https://www.verderliquids.com/ch/fr/how-do-peristaltic-pumps-work/, (visited 2/11/2021).

[66] How to Mechatronics, "L298N Motor Driver – Arduino Interface, How It Works, Codes, Schematics", 2017, [Online] Available: https://howtomechatronics.com/tutorials/arduino/arduino-dc-motor-control-tutorial-l298n-pwm-h-bridge/, (visited 2/11/2021).

[67] Pharmaceutical guidelines, "Working Principle of pH Meter", 2014, [Online] Available: https://www.pharmaguideline.com/2015/08/principle-and-working-of-pH-probes.html, (visited 10/11/2021).

[68] DFRobotics, "Gravity: Analog pH Sensor / Meter Kit For Arduino", 2017, [Online] Available: https://www.dfrobot.com/product-1025.html, (visited 10/11/2021).

Appendix

MDPI

Article

Low Complexity Channel Codes for Reliable Molecular Communication via Diffusion

Sofia Figueiredo [1,‡], Nuno Souto [1,2,‡] and Francisco Cercas [1,2,‡*]

[1] Telecommunications and Computer Engineering, ISCTE-University Institute of Lisbon, 1649-026 Lisboa, Portugal; sadfo@iscte-iul.pt (S.F.); francisco.cercas@iscte-iul.pt (F.C.)
[2] Instituto de Telecomunicações, 1649-026 Lisboa, Portugal; nuno.souto@lx.it.pt
* Correspondence: francisco.cercas@iscte-iul.pt
‡ These authors contributed equally to this work.

Abstract: It is envisioned that healthcare systems of the future will be revolutionized with the development and integration of body-centric networks into future generations of communication systems, giving rise to the so-called "Internet of Bio-nano things". Molecular communications (MC) emerge as the most promising way of transmitting information for in-body communications. One of the biggest challenges is how to minimize the effects of environmental noise and reduce the inter-symbol interference (ISI) which in an MC via diffusion scenario can be very high. To address this problem, channel coding is one of the most promising techniques. In this paper, we study the effects of different channel codes integrated into MC systems. We provide a study of Tomlinson, Cercas, Hughes (TCH) codes as a new attractive approach for the MC environment due to the codeword properties which enable simplified detection. Simulation results show that TCH codes are more effective for these scenarios when compared to other existing alternatives, without introducing too much complexity or processing power into the system. Furthermore, an experimental proof-of-concept macroscale testbed is described, which uses pH as the information carrier, and which demonstrates that the proposed TCH codes can improve the reliability in this type of communication channel.

Keywords: future wireless networks; 6G; molecular communications; diffusion-based; channel coding; TCH codes

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Sensors* **2021**, *1*, 0. https://doi.org/

Received:
Accepted:
Published:

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The telecommunications domain is always developing, facing all types of new demands, and therefore, while fifth generation (5G) wireless networks are currently being deployed around the world, work is already underway for future evolutions such as sixth-generation (6G) systems [1] and even beyond. Emerging as a new network paradigm is the "Internet of Bio-Nano Things" (IoBNT), defined as the interconnection of nanoscale devices. This is a revolutionizing concept promising many applications in the biomedical field, medicine, industry, environment, agriculture, and military [3]. Communications in IoBNT need to be biocompatible, energy-efficient, and robust in physiological conditions, making electromagnetic signals a difficult choice due to issues related to biocompatibility, power, and possible health hazards. An alternative to this problem is Molecular Communications (MC), using molecules for encoding, transmitting, and receiving information. MC are feasible, being considered easier to implement than other approaches in the near term. They are scalable having an appropriate size for nanomachines, and they are bio-compatible allowing the integration with living systems [2]. MC have been captivating a lot of research interest with several novel transmitter/receiver architectures being proposed in recent years. Due to the growing interest in nanoscale communications and MC, IEEE has already started standardization efforts through the IEEE P1906.1 working group [4,5]. Nevertheless, MC for in-body

networks is still in its infancy, and it will take some time before practical systems become a reality, probably not before a seventh or higher generation (7G) of wireless networks [5]. However, MC are also relevant for several industry and environment applications such as smart infrastructure monitoring, quality control, hazardous environment monitoring, plants and animals monitoring [6]. Many of these scenarios rely on macroscale MC which is expected to evolve much faster than in-body MC, and may even allow it to be integrated into the next generation of wireless networks, namely into 6G [5].

Regarding this type of communications, diffusion-based MC is one of the most studied channels because of its flexibility and because it is suitable for a wide number of scenarios. The diffusion-based architecture is characterized by the random movement of molecules (Brownian motion) from higher concentrations to less [7,8]. The system structure adopted for MC is similar to traditional radio communication systems, characterized by an end-to-end chain, [9,10]: a transmitter, a channel, and a receiver. A few MC systems based on diffusion have been experimentally tested over the last few years. The purpose is to bridge the gap between theory and practice. Due to current technology limitations, most experiments are still carried out at the macro level instead of nano. For instance, an air-driven experiment was described in [15], which was based on spraying and detecting alcohol in open space. Some difficulties identified in this experiment were the fluctuations in humidity and temperature that change the channel features, which in the molecular context may have an impact on the results obtained. In a liquid-driven experiment presented in [16], the authors used small tubes as the communication channel where information is carried out via the pH. The authors refer that the system was shown to achieve better data rates than any previous chemical communication platform tested.

MC have some limitations, such as: the stochasticity where molecules have random propagation called environmental noise; the large delays due to the much longer propagation times of molecules when compared to the speed of light; the range since known MC techniques have very short practical ranges; and the fragility because biological components can be environmentally sensitive, [2,17]. These conditions pose many challenges for achieving a reliable communication. Therefore, channel coding became an important solution to help reduce these effects. This is achieved by adding redundancy to the emitted message, and by making it possible to correct the errors caused by noise during transmission at the receiver, improving the reliability of the communication link. Some studies have been made, for instance, in [18,19] where authors demonstrated that Hamming codes are able to improve the performance of MC systems. However it was also shown that there is a Hamming order at which they become self-defeating, making it important to study other codes. Another proposed solution relies on the "ISI-mitigating channel codes" (ISI-mtg) [15], which is a code that has properties specially adapted to the MC channel environment, namely, the codewords do not contain consecutive '1's, they always start with '0', and they contain at least one bit '1'. This code proved to present a significant improvement in terms of bit error rate (BER).

Motivated by the work above, in this paper, we study Tomlinson, Cercas, Hughes (TCH), as a new concept applied to MC systems. These codes possess several properties allowing us to use them efficiently in various applications, such as error correction, synchronism, spread spectrum systems, and channel estimation. The main goal of these codes is to achieve good coding gains with low implementation complexity, and a smaller auto-correlation distribution, that is, higher minimum distances, while still supporting soft-decision decoding for best performance [20–22]. Motivated by this, we propose adequate symbol detection schemes for the MC channel which enable low complexity hard and soft decision TCH decoding. The performance of the resulting MC systems is evaluated through an end-to-end simulator and also through a macro scale experimental prototype. The efficiency of the encoder and decoder will be analyzed and compared with other existing alternatives proposed in the literature. Different pulse types and modulations will be considered in the evaluation. Regarding the macro-scale

experiment, we adopt pH as an information carrier combined with Manchester pulses. This paper is organized as follows: Section II describes the system model. Section III presents the characteristics of the TCH transmitter and receiver. Section IV presents the results of the simulated and experimental system. Finally, Section V summarizes the main conclusions of the paper.

2. Materials and Methods

2.1. Diffusion based channel

Let us consider a diffusion based MC system operating in a three dimensional space. The model assumed for the average number of molecules counted at the receiver for this work is the point transmitter to a spherical absorbing receiver, [11–14]. This type of receiver counts the number of absorbed molecules during an observation window $[t_l, t_u]$. The probability, $F_{hit}(t_u, t_l)$, of a molecule emitted at time t=0 to be observed at the receiver during interval [t_l, t_u], can be defined as [11]

$$F_{hit}(t_u, t_l) = \frac{a_{rx}}{d_0}\left[erfc\left(\frac{d_0 - a_{rx}}{\sqrt{4Dt_u}}\right) - erfc\left(\frac{d_0 - a_{rx}}{\sqrt{4Dt_l}}\right)\right] \tag{1}$$

where erfc is the complementary error function, a_{rx} is the radius of the spherical receiver and d_0 is the distance between the center of the transmitter and the center of the receiver. D is the diffusion coefficient of the i^{th} molecule, in $m^2 s^{-1}$, which can be calculated using

$$D = \frac{K_B T}{6\pi\eta R_p} \tag{2}$$

where $K_B = 1.38 \times 10^{-23}$ is the constant of Boltzmann, T is the temperature in Kelvin, η is the fluid viscosity, and R_p is the radius of the particle.

The approach taken in this work assumes a diffusion-based MC system whose information bits are encoded in order to reduce ISI. The concentration vector corrupted by noise at the k^{th} time interval can be written as

$$y[k] = \sum_{l=0}^{L-1} h[l,k]x[k-l] + n[k] \tag{3}$$

where L is the assumed length of the ISI, $x[k] \in \{A_0, A_1\}$ (with A_0 and A_1 representing the levels for a '0' and a '1') denotes a concentration based binary modulation symbol emitted from the transmitter at time k, $h[l,k]$ is the number of molecules observed by the absorbing receiver during time interval k after N_{molec} are released from the transmitter at time $k-l$ and $n[k]$ is the number of noise molecules detected by the receiver at time k. Both $h[l,k]$ and $n[k]$ are modeled as Poisson random variables: $h[l,k] \sim Poisson(\overline{h}[l,k])$ and $n[k] \sim Poisson(\overline{n})$. Taking into account the probability model defined in (1), we have

$$\overline{h}[l,k] = N_{molec}.F_{hit}((l+1)T_s, l.T_s). \tag{4}$$

It can be shown that the received concentration at the receiver can be expressed as

$$y[k] = \underbrace{\overline{h}[0,k]x[k]}_{\text{desired signal}} + \underbrace{v[k]}_{\text{diffusion noise}} + \underbrace{I[k]}_{\text{ISI}} + \underbrace{n[k]}_{\text{environmental noise}} \tag{5}$$

where the diffusion noise follows the distribution $v[k] \sim Poiss_0(\overline{h}[0,k]x[k])$, i.e., it is equivalent to a Poisson random variable whose mean has been subtracted. This noise is signal dependent and accounts for the deviation of the concentration from its expected value according to the diffusion model, as illustrated in Figure 1.

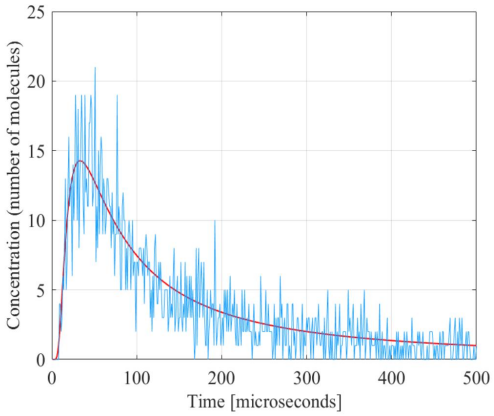
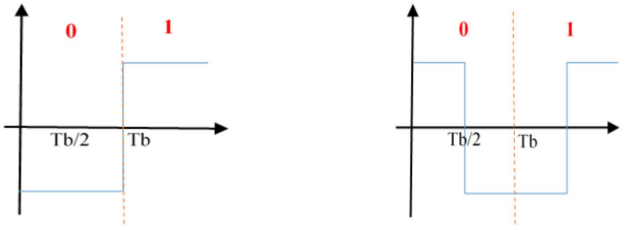


Figure 1. Example illustrating the expected and observed concentrations for a diffusion-based MC channel, assuming number of molecules (N_{molec})=10000, d=0.3 µm, D=450 and a_{rx}=50 nm.

2.2. Types of pulses

For the proposed MC system, in addition to the usual non-return to zero (NRZ) pulses exemplified in Figure 2a, Manchester pulses are also considered. These types of pulses, as in Figure 2b, have their wave boundaries always between 1 and -1, therefore the decisions are usually taken in the middle of each bit, [23,24]. These features will provide more robustness to the MC communication system in the experimental implementation.



(a) NRZ **(b)** Manchester

Figure 2. Type of pulses.

2.3. Modulation

In MC, to convey distinct messages, each possible message is associated with a distinct molecule signal. To accomplished this, as defined in [25], the information can be modulated based on the properties of the molecules which allows different types of MC schemes, namely: Concentration-based, Type-based, Timing-based, Spatial, or a combination of these (hybrid).

The concentration-based techniques are a representation of the information that relies on varying the concentration of the molecules. This is achieved by making the amount of released molecules carry information over discrete period time slots, the symbol slots. The simplest form of this approach, analogous to traditional communications, is the on-off keying (OOK), where each symbol represents a one-bit value. Bit-1 corresponds to the transmitter releasing a fixed number of molecules and bit-0 corresponds to the transmitter not releasing anything within the symbol slot. Another possible concentration-based technique is the Binary Concentration Shift Keying (BCSK) where in each symbol slot, sending a certain concentration of a molecule represents bit-1

and a different concentration represents bit-0. In this work we will consider both OOK and BCSK.

2.4. Detection and bit decision

In this paper we assume that no channel state information (CSI) exists at the transmitter nor the receiver. Therefore, the detection is accomplished solely based on the received observations over each encoded block duration. Two simple detection approaches are considered. The first one is a threshold based detection similar to [15], which generates hard decision symbols/bits. We consider a concentration-based binary modulation where each symbol comprising codeword i is represented as $x^i[k] \in \{A_0, A_1\}$, with A_0 and A_1 denoting the levels for a '0' and a '1', respectively. In this case an adaptive threshold is computed for the i-th codeword as

$$t^i = a y^i_{min} + (1-a) y^i_{max} \tag{6}$$

where and y^i_{min}=min(y^i) , with y^i_{max}=max(y^i), $y^i = \left[y^i[1], \ldots, y^i[N]\right]^T$ denoting the vector with the corresponding N received samples. Parameter a is a scaling factor which, considering the adoption of a balanced code with equal number of '0's and '1's, we set as a=0.5. Using this threshold, a hard decision estimate is simply obtained as

$$\hat{x}^i[k] = \begin{cases} A_0, & y^i[k] < t^i \\ A_1, & y^i[k] < t^i \end{cases} \tag{7}$$

which can be directly demodulated into bits. The second detection technique relies on an ad-hoc computation of log-likelihood ratios (LLRs) for each bit. Since we assume that no prior knowledge about the statistics of the channel exists, we compute estimates for the probabilities $p_{1,k} = P(x^i[k] = A_1 | y^i)$ and $p_{0,k} = P(x^i[k] = A_0 | y^i)$ considering a uniform distribution whose bounds are defined by y^i_{min} and y^i_{max}. In this case these probabilities are obtained as

$$p_{1,k} = \frac{y^i[k] - y^i_{min}}{y^i_{max} - y^i_{min}} \tag{8}$$

and

$$p_{0,k} = \frac{y^i_{max} - y^i[k]}{y^i_{max} - y^i_{min}}. \tag{9}$$

Soft values can then be computed as the conditional expected value of the symbols which can be written as

$$\tilde{x}^i[k] = E(x^i[k] | y^i) = (A_1 - A_0) p_{1,k} + A_0. \tag{10}$$

Taking into account that the LLRs for the individual bits can be expressed as

$$\lambda^i[k] = log\left(\frac{p_{1,k}}{p_{0,k}}\right) \tag{11}$$

we can write

$$p_{1,k} = \frac{1}{1 + e^{-\lambda^i[k]}} \tag{12}$$

which allows us to compute the soft-values as

$$\hat{x}^i[k] = \frac{(A_1 - A_0)}{1 + e^{-\lambda^i[k]}} + A_0. \tag{13}$$

Note that in the case of OOK, with $x^i[k] \in \{0, A\}$ this expression can be rewritten as

$$\hat{x}^i[k] = \frac{A}{2}\left(\tanh\left(\frac{\lambda^i[k]}{2}\right) + 1\right). \tag{14}$$

In the case of a polar modulation where $x^i[k] \in \{-A, A\}$, (12) reduces to

$$\hat{x}^i[k] = A\left(\tanh\left(\frac{\lambda^i[k]}{2}\right)\right). \tag{15}$$

These soft estimates then allow simple soft-decision decoding based on squared Euclidean distance minimization.

2.5. TCH codes

Channel coding is based on correcting up to several errors, by providing immunity to the transmitted signal. Channel coding improves the communication link's reliability, but it also implies a reduction in the overall transmission rate, due to implicit code rate k/n. TCH codes are a class of binary, non-linear, non-systematic, and cyclic block codes of length $n = 2^m$, where m is any positive integer. For efficiency reasons, the all-zero codeword was excluded from the code set and the inverse (binary negation) of any codeword is always another valid codeword [20]. TCH block codes are identified as TCH(n, k, t) where n represents the code length, k identifies the number of information bits in a code word and t is its error-correcting capacity. With P_i being a theoretically generated set of Base Polynomials, which give the whole structure for any of these codes, we can write [21]

$$n = 2^m, m \in \mathbb{N} \tag{16}$$

$$k = m + \log_2(h) + 1 \tag{17}$$

$$d_{min}^{TCH} \geq 2t_{TCH} + 1 \tag{18}$$

$$d_{min}^{TCH} \leq H_d[P_i(x), \{P_j(x^r)\} \bmod n_{TCH}] \leq n_{TCH} - d_{min}^{TCH} \tag{19}$$

$$P_i(x) = P_j(x^r) \bmod n_{TCH}, i \neq j \forall t_{TCH} \in \mathbb{N}. \tag{20}$$

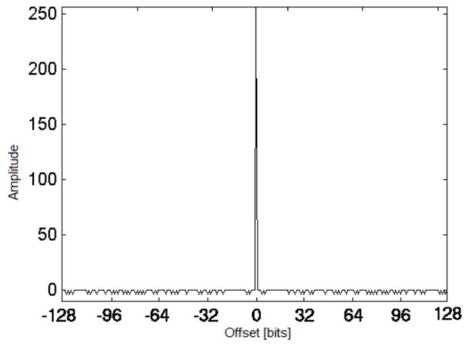


Figure 3. Typical autocorrelation function of a B-TCH polynomial with n=256, [21].

As usual, their error-correcting capacity, t_{TCH}, depends on the minimum distance, d_{min}^{TCH}, between the polynomials, where H_d stands for the Hamming distance between any two polynomials. TCH codes are also balanced codes, that is, the number of zeros equals the number of ones in each codeword, which is an important feature for MC systems. We can derive TCH codes of nearly any length, however, the most important class of TCH codes which originated all of them, are based on the so-called Basic TCH

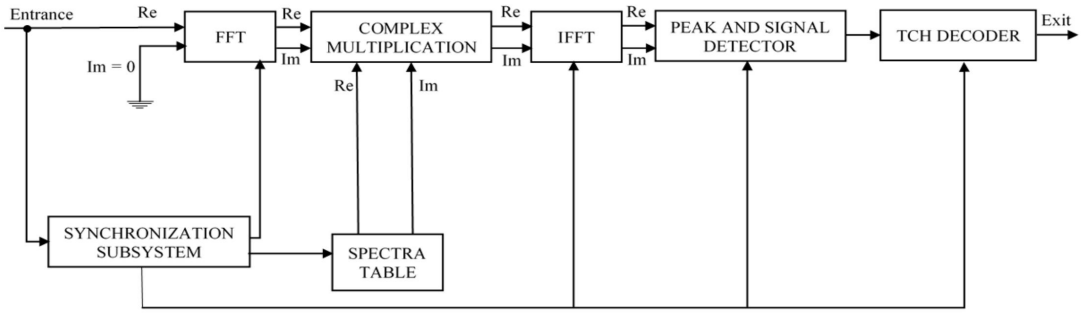


Figure 4. TCH decoder [22].

Polynomials (B-TCH). B-TCH polynomials with degree n are obtained by the following equation

$$P(x) = \sum_{i=0}^{\frac{p-1}{2}-1} a_i x^{K_i}, a^i \in GF(q), q = p^k, k \in \mathbb{N} \tag{21}$$

where the exponents K_i satisfy the equation

$$a^{K_i} = 1 + a^{2i+1}, i = 0, 1, \ldots, \frac{p-1}{2} - 1. \tag{22}$$

In these expressions p is a prime number obeying the following condition

$$p = n_{TCH} + 1 = 2^m + 1. \tag{23}$$

Prime number that verify this condition can be written as

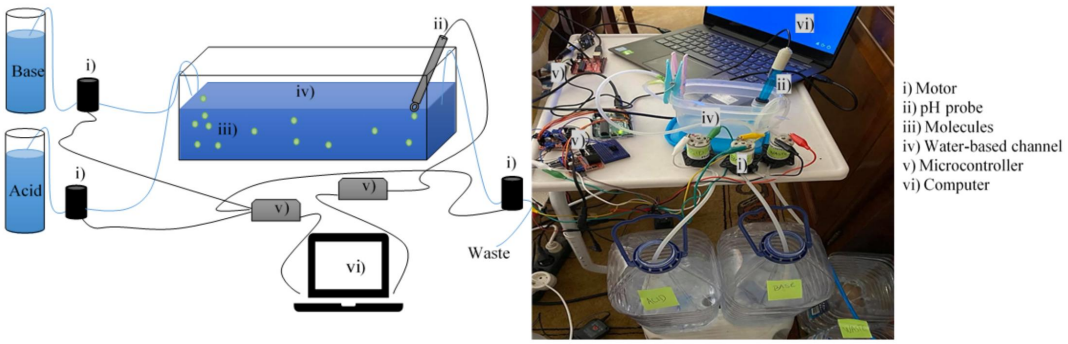
$$F_i = 2^{2^i} + 1, i \in \mathbb{N} \tag{24}$$

and so they are in fact Fermat numbers (F_i). Only five code lengths obey these rules, which means that we can only generate pure TCH polynomials for codes with lengths of 2, 4, 16, 256, and 65536. TCH codes originated by B-TCH polynomials have both good cross and auto-correlation. Their autocorrelation is always three-valued, with values n, 0, and –4, and its distribution is perfectly characterized for any n [22]. This translates into a great advantage for higher sized TCH codes, such as TCH codes length $n \geq 256$ since these sequences tend to get closer to a Dirac impulse, as shown in Figure 3.

Conceptually, the TCH receiver is based on a group of correlators to compare the received word. As usual, and by comparing the outputs of the correlators it becomes possible to choose the sequence that corresponds to the highest correlation or the most likely word sent. In the TCH receiver, depicted in Figure 4, correlation is done in the frequency domain using two FFT's, a complex multiplication and an IFFT, which, together with some optimizations, accelerates the decoding process. For example, and since FFT's of length n process n codewords simultanesously, the efficiency increases with the code length n. Furthermore since both, real and imaginary parts are also processed simultaneously, that further decreases the processing time to half, again doubling the receiver efficiency.

2.6. Experimental setup

To help validate the performance of the proposed TCH-based MC communications, we implemented a macroscale testbed. In our experiment, we used pH levels to test the molecules as information carriers. The different pH measurements are indicators



(a) **(b)**

Figure 5. (a) Schematic diagram of the experimental implementation; (b) Photo of the experiment.

of how acidic or basic the water is: pH's above 7 indicate a basic solution and below 7 an acidic solution, on a scale between 0 and 14. The pH measures the amount of free hydrogen (H+) and hydroxyl (OH-) in the water. This means that water with more free hydrogen ions is acidic and water with more free hydroxyl ions is basic. The scale on which pH is measured is logarithmic, which means that increasing or decreasing an integer value changes the concentration by a factor of ten [26–28]. Thus, we use an acid and a base to send information. The used base solution has a pH level of 9 and the used acid solution has a pH of 3. Utilizing pH could constitute a problem because of the limited scale, from 0 to 14. For example, in the case of bit 1 being represented by the base and the bit-0 by the acid, if many consecutive bit-1's were transmitted, information could be lost. To avoid this, we decided to use Manchester codes in this experiment to transmit the bits. Thus, bit-1 is represented by sending an acid and then a base, and bit-0 is represented by sending a base and then an acid. This prevents the information from being lost as it ensures the same amount of acid and base solution is released in the channel independently of the information sequence.

Next, we will explain the constituents of the system namely, the transmitter, the channel, and the receiver. The main diagram of the experiment is presented in Figure 5a.

2.6.1. Transmitter

The transmitter, in Figure 5, is composed of two containers, one for the basic solution and another for the acid solution. These containers with the solutions are connected through a silicone tube to peristaltic pumps. These pumps feature a flexible tube, giving an open flow path for the liquids to pass. The pumping principle, called peristalsis, is based on alternating compression and relaxation of the tube generating a flow. The pumps we chose were 12V aquarium pumps working at 5000 RPM. Because we have an acid and a basic solution, we used two peristaltic pumps to transmit the solutions individually and avoid contamination. We control when they are on or off and how fast they run by using a microcontroller (a simple Arduino UNO microcontroller was adopted) connected to a computer. According to the information being transmitted, an acid solution or a basic solution is released to the channel.

2.6.2. Channel

As represented in Figure 5, we use a container to simulate the channel. We decided to fill the container with only 20ml of water with a pH around 7. On one side of the container we placed the tubes sending the acid or basic solution, one at a time (never simultaneously). On the other side of the container, 6cm apart, we placed the pH probe and another pump to keep the container's liquid volume always the same. Although these pumps generate some flow, the system will still mainly work via diffusion. One

advantage of using acids and bases is that they are chemicals that cancel each other, and that property is very important in closed loops because the average concentration of these chemicals would remain constant assuming the same transmission concentration of bases and acids.

2.6.3. Receiver

The receiver uses a different microcontroller, also an Arduino Uno, connected to the main computer. This board has two functions: one is removing liquid with a pump, working at the same speed as the transmitter pump and the other function is to read the pH levels in the water through the pH probe, as shown in Figure 5. We used a 5V pH probe, featuring, as most pH probes, a glass bulb filled with strong electrolytes and a silver wire inside. This allows the pH electrode to measure the difference in potentials between the two sides of the glass electrode. This signal is then converted into an electric potential which allows reading the pH value. The pH probe is connected to a pH meter, which is connected to the Arduino Uno microcontroller, that digitizes the pH reading through a 10-bit analog to digital converter (ADC). This probe can introduce a pH measurements error of ± 0.1 at 25°C.

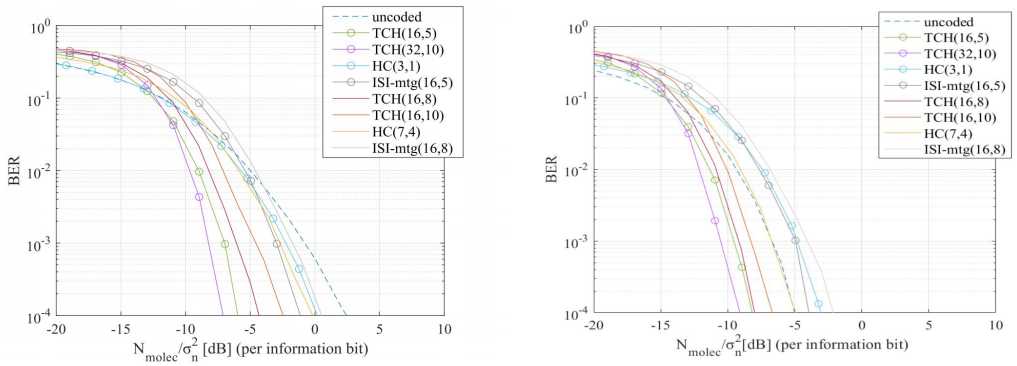
3. Results and Discussion

3.1. Simulation results

Table 1. Simulation parameters

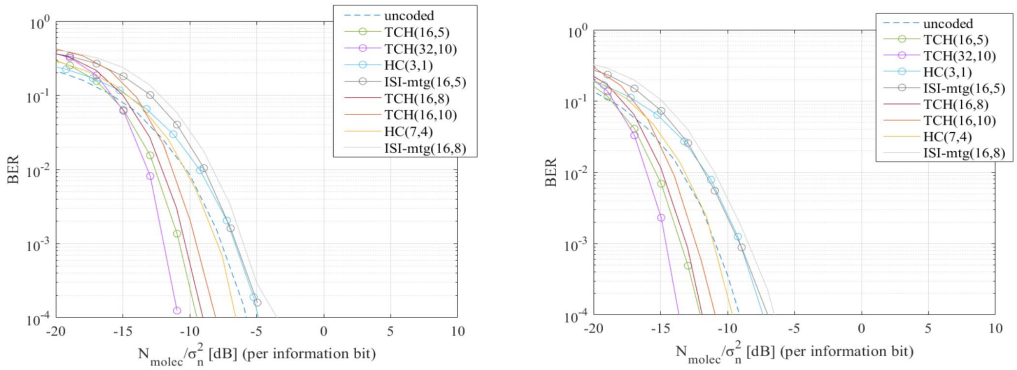
Parameter	Value
Deterministic model	Point transmitter to a spherical absorbing receiver
Statistical model	Gaussian approximating Poisson
Paired $tx - rx$ link distance (d_0)	10 (micrometers)
Diffusion Coefficient (D)	79.4 (micrometers^2/seconds)
Transmitter radius and receiver radius (a_{tx}, a_{rx})	5 (micrometers)
Symbol duration (T_s)	$\tau \frac{(d_0)^2}{6D}$ (seconds)

The purpose of this simulation is to test different modulations and bit encoding solutions combined with TCH channel codes. We will be comparing the performance of some previously studied codes for MC against the TCH codes which we propose here as a potential alternative for MC scenarios. For the channel correcting codes, we choose to test the Hamming Codes (HC) because they are basic and easy to implement, and the ISI-mtg codes from [15] because they were designed specifically for an MC environment. Finally, we test the TCH codes for different codes lengths. To guarantee a fair comparison between these codes, similar sets of code rates were used comprising both lower codes rates, and higher codes rates. We remind that lower coding rates may obtain better performances but decrease the information bit rate. Higher coding rates can improve the bit rate sacrificing some BER performance. For the uncoded scenario and for the HC we used threshold hard decision. For the TCH codes and for the ISI-mtg



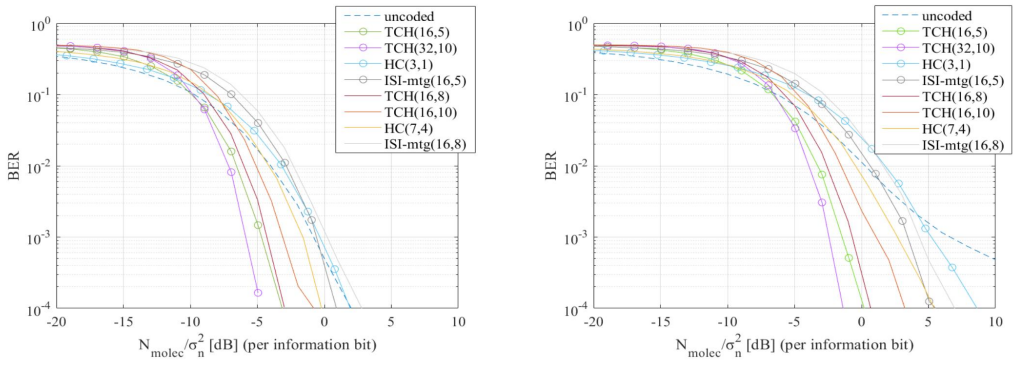
(a) NRZ **(b)** Manchester

Figure 6. BER comparison for all the considered channel coding methods, using BCSK modulation and with τ=2.



(a) NRZ **(b)** Manchester

Figure 7. BER comparison for all the considered channel coding methods, using BCSK modulation and with τ=10.



(a) NRZ, τ=10 **(b)** Manchester, τ=2

Figure 8. BER comparison for all the considered channel coding methods, using OOK modulation.

codes, LLR soft decisions were employed. Both these bit decisions are detailed in Section II. D. The main parameters used in the simulations are described in Table 1. The symbol duration can have a critical impact on the behaviour of the system. To evaluate the impact of this parameter we will consider different values for the scaling factor τ which is used for defining the symbol duration, as shown in Table 1. Tests will be performed for a small symbol duration because it is when there are more errors, becoming very noticeable how channel codes help to correct them. The simulations were performed using NRZ and Manchester pulses, at different symbol durations, with two distinct types of modulation: OOK and BCSK. In the case of OOK modulation we have $x[k] \in \{0, 1\}$, whereas in the BPSK we use $x[k] \in \{-1, 1\}$. For the BPSK we assume the existence of a destructor molecules that can reduce the concentration of the carrier molecule. It is important to emphasize that while the channel model may not be the most adequate for this particular scenario, the proposed scheme does not depend on a specific channel and does not require channel knowledge. Therefore, using this channel model still enables an evaluation of the system behaviour, which is then additionally validated through experimental tests. The results will be presented as the relation between the BER and N_{molec}/σ_n^2 (per information bit) in dB, where σ_n^2 denotes the variance of the environmental noise.

In Figure 6a and Figure 7a present the results for when NRZ pulses were tested, whereas in Figure 6b and Figure 7b present the results for when Manchester pulses were tested. Both Figure 6 and Figure 7 were simulated using BCSK modulation. When Manchester codes were used, the tested channel coding methods had an overall better performance. The big difference between these figures relies on the τ value, which in Figure 6 is 2, while in Figure 7 is increased to 10. Analyzing the results for the different τ values, it is noticeable a BER improvement in Figure 7 which has a higher τ value, and thus a higher T_s. Increasing the T_s means that the molecules will have more time to diffuse and reach the receiver, reducing the ISI. It is important to note however that a higher T_s also means a transmission rate decrease. We tested the system with the OOK modulation, Figure 8. In Figure 8a are the results for NRZ pulses with τ=10 and in Figure 8b are the results for Manchester pulses at τ=2. Comparing Figure 7a with Figure 8a and Figure 6b with Figure 8b, the results shown that the tested channel codes methods had worse BER results in Figure 8, when OOK modulation was used.

In the figures we can observe that TCH(16,5) and TCH(32,10) codes, clearly perform better than codes with similar rates, HC(3,1) and ISI-mtg(16,5). Furthermore, TCH codes with higher codes rates, TCH(16,8) and TCH(16,10) are able to perform better than other codes with lower rates namely, the HC(3,1) and the ISI-mtg(16,5). As expected, the TCH codes with the same code rate but longer codewords, namely the TCH(32,10) when compared against the TCH(16,5), achieve better BER performances. It is important to note that longer codewords incur in an additional delay. The results of our simulations proved that the TCH codes achieved a better BER performance against the other tested channel correcting codes, independently of the symbol duration, type of pulse and modulation.

3.2. Experimental results

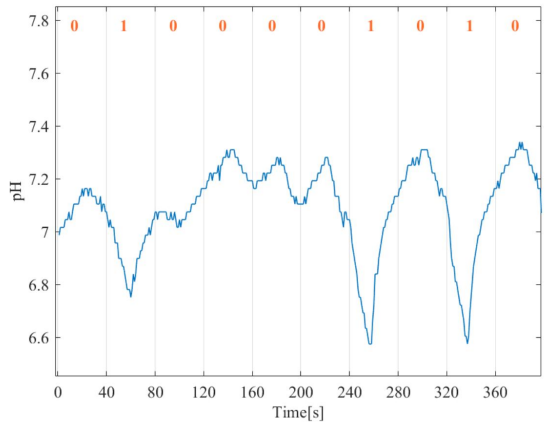


Figure 9. Example of 10 received bits with T_s=40s after initial delay removal.

To add validation to our simulation results we performed a macro scale MC experiment using TCH codes. Because this experiment was performed on a macro scale testbed some parameters previously adopted for the simulation had to be adjusted. This evaluation, was carried out by testing the channel without channel correcting codes and by testing the channel with a TCH(16,10) code. We chose to work with the TCH(16,10) code because it has a relative high coding rate while still maintaining good error capabilities. In this experiment we used Manchester pulses, for better channel robustness, combined with BCSK modulation. For transmitting a bit we define two moments: a first moment that starts at the beginning of the symbol and goes to $T_s/2$, and a second moment that goes from $T_s/2$ to T_s. To transmit the bit-0, we start by sending a base solution during the beginning until halfway of the first moment, followed by a period until the end of the first moment when nothing is released to the channel. In the second moment, an acid solution is sent to the channel until halfway of the second moment followed by a period until the end of the symbol where nothing is released. To transmit bit-1, the opposite happens, in the first moment an acid followed by a period when nothing is released, and in the second moment a base followed by a period when nothing is released. In our experiment, we collected data from 100 transmissions, where each transmission comprised a block with 100 random generated bits. To test how the symbol interval (T_s) could affect the received bit stream, we performed the experience for T_s=20s and then for T_s=40s. Is it important to underline due to the adopted experimental setup, there may exist other components contained in the chemicals used which can interfere with the channel characteristics and affect the pH probe reading. Still, the implemented communication scheme does not require channel state knowledge to accomplish the detection.

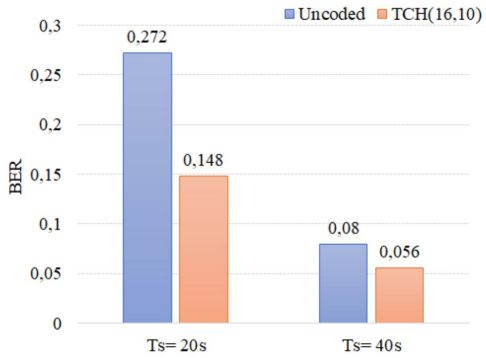


Figure 10. Uncoded and TCH(16,10) performance for *Ts*=20s and *Ts*=40.

The first released molecules take time to reach the receiver (initial delay). In Figure 9 we show an example of the received bit stream after the initial delay removal. In Figure 10, it can be observed that when a larger T_s is adopted the error rate decreases both in the uncoded and coded cases. This was also observed in the simulation results, Section V. A., and was expected since, although affecting the transmission rate, a lower symbol time gives the system more time to receive most of the late molecules, thus decreasing the ISI. Also, analyzing the results in Figure 10, it is clear that fewer errors were obtained when TCH codes were applied, confirming the TCH codes efficiency on correcting errors in an MC environment.

4. Conclusions

In this paper, we studied the adoption of TCH codes as a low complexity promising solution for enabling reliable MC. Adequate detection and decoding methods which do not require CSI acquisition were presented and evaluated through end-to-end simulations. Furthermore, to validate the reliability achieved with the proposed approach, we implemented a macroscale MC platform where the channel is an aqueous solution and the information to be transmitted is differentiated by the pH value of the medium, as this is a good approach to a biological environment for in-body communications. At the transmitter, different encoded pulses were considered namely, NRZ and Manchester, combined with either OOK and BCSK modulations. Other channel codes were evaluated as benchmarks for the TCH codes, namely, Hamming codes and the ISI-mtg codes that were designed specifically for this context.

The results obtained, both from simulation and physical implementation, showed that the new introduced coding scheme based on TCH codes, performed better without introducing more complexity to the system. In fact, TCH codes with higher rates were able to perform better than other codes with lower rates. In addition, Manchester pulses for coded and uncoded situations proved to be essential when operating in a channel that can saturate, as in the case of the pH based experimental tests. These pulses also proved to work better in the simulation than the NRZ pulses, giving more robustness to the communication in an MC channel with an absorbing receiver.

Author Contributions: Conceptualization, N.S., F.C. and S.F.; Methodology, N.S. and F.C.; Software, S.F. and N.S.; Validation, F.C. and N.S.; Formal Analysis, F.C. and N.S.; Investigation, N.S. and S.F.; Resources, S.F. and N.S.; Data Curation, S.F.; Writing – Original Draft Preparation, S.F.; Writing – Review and Editing, N.S., F.C. and S.F.; Visualization, S.F.; Supervision, N.S. and F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: This work is funded by FCT/MCTES through national funds and when applicable co-funded by EU funds under the project UIDB/50008/2020.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. M. Z. Chowdhury, M. Shahjalal, S. Ahmed and Y. M. Jang. 6G Wireless Communication Systems: Applications, Requirements, Technologies, Challenges, and Research Directions. *IEEE Open Journal of the Communications Society*, **2020**, *1*, 957-975, doi: 10.1109/OJCOMS.2020.3010270.
2. M. Kuscu, E. Dinc, B. A. Bilgin, H. Ramezani and O. B. Akan. Transmitter and Receiver Architectures for Molecular Communications: A Survey on Physical Design With Modulation, Coding, and Detection Techniques. *Proceedings of the IEEE*, **2019**, *107, no. 7*, 1302-1341, doi: 10.1109/JPROC.2019.2916081.
3. F. Akyildiz, M. Pierobon and S. Balasubramaniam. Moving forward with molecular communication: from theory to human health applications [point of view]. *Proceedings of the IEEE*, **2019**, *107, no. 5*, 858-865, doi: 10.1109/JPROC.2019.2913890.
4. K. Yang, D. Bi, Y. Deng, R. Zhang, M. M. U. Rahman, N. A. Ali, M. A. Imran, J. M. Jornet, Q. H. Abbasi, and A. Alomainy. A comprehensive survey on hybrid communication in context of molecular communication and terahertz communication for body-centric nanonetworks. *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, **2020**, *6, no. 2*, 107– 133.
5. S. Canovas-Carrasco, A. Garcia-Sanchez and J. Garcia-Haro. The IEEE 1906.1 Standard: Some Guidelines for Strengthening Future Normalization in Electromagnetic Nanocommunications. *IEEE Communications Standards Magazine*, **2018**, *2, no. 4*, 26-32, doi: 10.1109/MCOMSTD.2018.1700082.
6. Y. Lu, R. Ni and Q. Zhu. Wireless Communication in Nanonetworks: Current Status, Prospect and Challenges. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, **2020**, *6, no. 2*, 71-80, doi: 10.1109/TMBMC.2020.3004304.
7. S. M. Abd El-atty, A. El-taweel and S. El-Rabaie. Aspects of nanoscale information transmission in nanonetworks-based molecular communication. *2016 Fourth International Japan-Egypt Conference on Electronics, Communications and Computers (JEC-ECC)*, **2016**, 131-134, doi: 10.1109/JEC-ECC.2016.7518985.
8. S. Kadloor, R. S. Adve and A. W. Eckford. Molecular Communication Using Brownian Motion With Drift. *IEEE Transactions on NanoBioscience*, **2012**, *11, no. 2*, 89-99, doi: 10.1109/TNB.2012.2190546.
9. M. Pierobon and I. F. Akyildiz. A physical end-to-end model for molecular communication in nanonetworks. *IEEE Journal on Selected Areas in Communications*, **2010**, *28, no. 4*, 602-611, doi: 10.1109/JSAC.2010.100509.
10. D. Bi, A. Almpanis, A. Noel, Y. Deng and R. Schober. A Survey of Molecular Communication in Cell Biology: Establishing a New Hierarchy for Interdisciplinary Applications. *IEEE Transl. J. Magn. Japan*, **1987**, *2*, 740–741, doi: 10.1109/COMST.2021.3066117.
11. V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel and R. Schober. Channel Modeling for Diffusive Molecular Communication—A Tutorial Review. *Proceedings of the IEEE*, , *107, no. 7*, 1256-1301, doi: 10.1109/JPROC.2019.2919455.
12. A. Akkaya, H. B. Yilmaz, C. Chae and T. Tugcu. Effect of Receptor Density and Size on Signal Reception in Molecular Communication via Diffusion With an Absorbing Receiver. *IEEE Communications Letters*, **2015**, *19, no. 2*, 155-158, doi: 10.1109/LCOMM.2014.2375214.
13. Y. Deng, A. Noel, W. Guo, A. Nallanathan and M. Elkashlan. Analyzing Large-Scale Multiuser Molecular Communication via 3-D Stochastic Geometry. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*,*3, no. 2*, 118-133, **2017**, doi: 10.1109/TMBMC.2017.2750145.
14. H. B. Yilmaz, G. Suk and C. Chae. Chemical Propagation Pattern for Molecular Communications. *IEEE Wireless Communications Letters*, **2017**, *6, no. 2*, 226-229, doi: 10.1109/LWC.2017.2662689.
15. A. O. Kislal, B. C. Akdeniz, C. Lee, A. E. Pusane, T. Tugcu and C. -B. Chae. ISI-Mitigating Channel Codes for Molecular Communication Via Diffusion. *IEEE Access*, **2020**, *8*, 24588-24599, doi: 10.1109/ACCESS.2020.2970108.
16. N. Farsad, D. Pan and A. Goldsmith. A Novel Experimental Platform for In-Vessel Multi-Chemical Molecular Communications. *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, **2017**, 1-6, doi: 10.1109/GLOCOM.2017.8255058.
17. B. Tepekule, A. E. Pusane, H. B. Yilmaz, C. Chae and T. Tugcu. ISI Mitigation Techniques in Molecular Communication. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, **2015**, *1, no. 2*, 202-216, doi: 10.1109/TMBMC.2015.2501745.
18. Y. Lu, M. D. Higgins and M. S. Leeson. Diffusion based molecular communications system enhancement using high order hamming codes. *2014 9th International Symposium on Communication Systems, Networks and Digital Sign (CSNDSP)*, **2014**, 438-442, doi: 10.1109/CSNDSP.2014.6923869.
19. Y. Lu, M. D. Higgins and M. S. Leeson. Comparison of Channel Coding Schemes for Molecular Communications Systems. *IEEE Transactions on Communications*, **2015**, *63, no. 11*, 3991-4001, doi: 10.1109/TCOMM.2015.2480752.
20. F. A. B. Cercas, M. Tomlinson and A. A. Albuquerque. TCH: A New Family of Cyclic Codes Length 2/sup m/. *Proceedings. IEEE International Symposium on Information Theory*, **1993**, 198-198, doi: 10.1109/ISIT.1993.748512.
21. F. Cercas, J. C. Silva, N. Souto and R. Dinis. Optimum bit-mapping of TCH codes. *2009 International Workshop on Satellite and Space Communications*, **2009**, 92-96, doi: 10.1109/IWSSC.2009.5286417.
22. F. Cercas. A new family of codes for simple receiver implementation. Ph.D, Instituto Superior Técnico, Lisbon, 1996.

23. T. A. Vu and P. Van Thanh. A simple diagram for data transmission using Manchester code. *2016 International Conference on Advanced Technologies for Communications (ATC)*, **2016**, 453-456, doi: 10.1109/ATC.2016.7764825.
24. B. R. Saltzberg. An improved Manchester code receiver. *IEEE International Conference on Communications, Including Supercomm Technical Sessions*, **1990**, *2*, 698-702, doi: 10.1109/ICC.1990.117167.
25. M. Ş. Kuran, H. B. Yilmaz, I. Demirkol, N. Farsad and A. Goldsmith. A Survey on Modulation Techniques in Molecular Communication via Diffusion. *IEEE Communications Surveys and Tutorials*, **2021**, *23, no. 1*, 7-28, doi: 10.1109/COMST.2020.3048099.
26. B. Arifin. GChem: Assisting students learning in general chemistry. *2012 IEEE Symposium on Humanities, Science and Engineering Research*, **2012**, 323-327, doi: 10.1109/SHUSER.2012.6268865.
27. J. Y. Choi, H. G. Pandit and R. R. Rhinehart. A Process Simulator for pH Control Studies. *1993 American Control Conference*, **1993**, 2594-2595, doi: 10.23919/ACC.1993.4793362.
28. N. Farsad and A. Goldsmith. A molecular communication system using acids, bases and hydrogen ions. *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, **2016**, 1-6, doi: 10.1109/SPAWC.2016.7536834.