Quasi-Cyclic Short Packet (QCSP) Transmission for IoT

Kassem Saied

DOCTORAT BRETAGNE LOIRE / MATHSTIC

Université Bretagne Sud

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE BRETAGNE SUD

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : ***Télécommunication***

Par
Kassem SAIED

Quasi-Cyclic Short Packet (QCSP) Transmission for IoT

Thèse présentée et soutenue à Lorient , le 25-3-2022
Unité de recherche : Lab-STICC –UMR 6285 CNRS
Thèse N° : 621

Rapporteurs avant soutenance :

Giuseppe DURISI	Professeur, Chalmers University of Technology, Gothenburg, Sweden
Jean-Marie GORCE	Professeur, Institut National des Sciences Appliquées (INSA), Lyon

Composition du Jury :

Examinateurs :	Ghaya Rekaya-Ben OTHMAN	Professeure, Telecom Paris, Institut Polytechnique de Paris
	Jean-Baptiste DORE	Ingénieur, CEA, Laboratoire d'électronique et de technologie de l'information
	Benoit GELLER	Professeur, ENSTA, Institut Polytechnique de Paris
Dir. de thèse :	Emmanuel BOUTILLON	Professeur, Lab-STICC, Université de Bretagne Sud (UBS)
Co-dir. de thèse :	Ali Al GHOUWAYEL	Enseignant-Chercheur, École d'Ingénieurs du Numérique-EFREI Paris

Dedication

To the Almighty God...
To my Father, To my Mother...
To the unknown on earth, known in heaven...

Declaration of Originality

I, Kassem SAIED, declare that this thesis titled, "Quasi Cyclic Short Packet Transmission for IoT" and the work presented in it are my own. I confirm that:

— This work was done wholly or mainly while in candidature for a Doctoral degree at the LabSTICC, Université Bretagne Sud.

— Where I have consulted the published work of others, this is always clearly attributed.

— Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

— I have acknowledged all main sources of help.

— Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Kassem SAIED
16/1/2022

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to all persons and institutions that helped in making my Ph.D. years a valuable experience and a pleasant journey.

Firstly and foremost, I would like to express my sincere thanks, appreciation, and deepest gratitude to my supervisor, Emmanuel Boutillon. Without his support, encouragement and expertise, this work would never been achieved such completion. His insightful advice and technical guidance have shaped the way of my thinking in the most positive direction. It was a true privilege to be under his tutelage. With him, I knew what devotion, dedication, determination and discipline mean. Over the course of these few years, I am glad to have developed a strong friendship with him. In retrospect, if I had the chance of starting my PhD afresh, I would have definitely chosen the same path all over again.
Also, I would like to thank my co-supervisor, Ali Al Ghouwayel for choosing me as a candidate to this PhD. I want to thank him for his time, advice and continued encouragement throughout my years of PhD.
Emmanuel and Ali gave me the opportunity to pursue my Ph.D. degree, which had been one of my most desirable dreams until this moment. Thank you for everything.

I highly appreciate all the partners of the QCSP project within the framework of French National ANR projects (ANR-19-CE25-0013-01) which include collaboration between different academic and industrial partners. Of course, I want to thank them all for every fruitful discussion. Special thanks to Valentine Savin, Cathrine Douillard, Xavier Giraud and Louis Adrien for the unforgettable discussions and suggestions they provided to me during the QCSP quarterly progress meetings.

I would like to express my thanks to Giuseppe Durisi and Jean-Marie Gorce for their acceptance as my Ph.D. reviewers. Thank you so much for your valuable comments. I would like also to thank Ghaya-Rekaya Ben Othman, Jean-Baptiste Dore and Benoit Geller for their acceptance as my PhD committee members and for the fruitful discussion we had through my PhD defense. In addition, special thanks to the CSI members, Olivier Berder and Karine amis, for following my Doctorate situation every year. Many thanks

and appreciation goes out also to all my teachers I have had over the past years. Thanks for being there!

Moreover, I want to thank all my wonderful colleagues and friends in both the Lab-STICC and IRDL, especially Christian, Camille and Cedric for the technical discussion we had during the PhD years. I want to extend my gratitude also to all fellows and administrative staff for their help and kindness, especially Virginie Guillet for her very friendly and helpful attitude.
Furthermore, many special thanks go to my Lorient friends, my second family, Cynthia, Hassan, Asma, Joseph, Machhour, Jad, Hamza, Abbass ...
Their presence made the already beautiful Lorient even more beautiful.

Without my family I would be nothing. Foremost, and as every step in my life, I am grateful to God, Ahl Al-Bayt, my father and my mother; my parents whose unconditional love, prayers and wishes have been with me since day one. I want to thank them from the bottom of my heart for all the sacrifices they have made for me. Your influence in my life will always be remembered and cherished. Special thanks also to my lovely sister Rawan, brother Hamoudi and my sweet niece Aline. They all were my source of constant endorsement and motivation throughout achieving my goals.

To anyone I may have missed in these acknowledgments, I apologize and I thank you.

Kassem SAIED.

Table of Contents

List of Figures

List of Tables

Acronyms

QCSP	Quasi Cyclic Short Packet
CCSK	Cyclic Code Shift Keying
NB-LDPC	Non-Binary Low Density Parity Check
GF	Galois Field
DSSS	Direct-Sequence Spread Spectrum
FHSS	Frequency Hopping Spread Spectrum
OSI layers	Open System Interconnection layers
PHY layer	Physical layer
3GPP	3rd Generation Partnership Project
MAC	Media Access Control
AWGN	Additive White Gaussian Noise
SNR	Signal to Noise Ratio
FBR	Finite Block Regime
BPSK	Binary Phase Shift Keying
IoT	Internet of Things
WSN	Wireless Sensor Network
NB-FEC	Non-Binary Forward Error Correction
NB-Polar	Non-Binary Polar codes
NB-Turbo	Non-Binary Turbo Codes
LLR	Log-Likelihood Ratio
LPWAN	Low Power Wide Area Network
mMTC	massive Machine Type Communication
URLLC	Ultra-Reliable Low Latency Communication
$\boldsymbol{P}_0$	Basic Random sequence of CCSK modulation
PCM	Parity Check Matrix

VN	Variable Node
CN	Check Node
EMS	Extended Min Sum
LFSR	Linear Feedback Shift Register
FFT	Fast Fourier Transform
PDF	Probability Density Function
CDF	Cumulative Distribution Function
MC	Monte-Carlo
OM	Over Modulation
WOM	Weighted Over Modulation
SB	Syndrome Based
VNB	Variable Node Based
NoZ	Number of Zeros
DM	Direct Method
PM	Parametric Method
ML	Maximum Likelihood
GA	Genius Aided
FCI	Forward Correction Iterations
FER	Frame Error Rate
SDR	Software-Defined Radio
SDR	Universal Software Radio Peripheral
Tx	Transmitter
Rx	Receiver
ADC	Analog to Digital Converter
DAC	Digital to Analog Converter
SF	Super Frame

PARAMETERS

p	Number of bits in one symbol
q	GF order and P_0 sequence length
$\boldsymbol{H}$	NB-LDPC parity check matrix
d_c	Degree of the CN (number of VNs connected to a CN)
d_v	Degree of the VN (number of CNs connected to a VN)
K	Number of information symbols
m	Number of information bits
M	Number of redundant symbols
$N = K + M$	Code length (number of VNs)
$h_{i,j}$	Non-zero value in PCM that connects CN_i with VN_j
$M_{v_j p_i}$	Messages sent from VN_j to CN_i
$M_{p_i v_j}$	Messages sent from CN_i to VN_j
R_c	Coding Rate
R_c^*	Maximal achievable coding rate according to Polyanskiy's equation
R_m	Modulation rate
$\boldsymbol{P}_s$	CCSK shifted sequence with s positions
$\boldsymbol{F}$	QCSP Frame
$\boldsymbol{Y}_n$	Received QCSP frame of length $N \times q$ at time n
$\boldsymbol{y}_n$	Received symbol of length q at time n
S_n^f	Score function at time n and frequency f
$\boldsymbol{L}_n^f$	Correlation vector at time n and frequency f
$(\delta_n, \delta_\theta)$	Time-Frequency grid resolution
ℓ	Time bin resolution in chips; $\ell = \delta_n$ if chip period is one
p_ω	Number of frequency bins in the grid
p_Δ	Number of time bins in the grid

n	Time index at chip level
k	Time index at symbol level
f	Frequency in Hz
θ	Symbol rotation in radian, $\theta = 2\pi q f$
n_a	Time of arrival in chips, $n_a = n_c + \Delta$
n_c	Coarse time of arrival in chips
Δ	Finer time of arrival in chips
f_o	Exact frequency offset, $f_o = f_c + f_\delta$
$\theta_o = 2\pi f_o q$	Symbol rotation due to f_o
f_c, θ_c	Coarse frequency offset
f_δ, θ_δ	Finer frequency offset
f_r	Residual finer frequency offset, $f_r = f_o - \hat{f}_o$
$\hat{n}_c, \hat{n}_c$	Coarse offsets estimation
$\hat{n}_o, \hat{n}_o$	Fine offsets estimation from first step of synchronization
$\tilde{n}_o$	Better time estimation using OM
$\bar{n}_o$	Exact time estimation using VNB
$\tilde{f}_o$	Finer frequency estimation using FFT method
$\bar{f}_o$	Finest frequency estimation using either DM or PM
ϕ	Initial phase offset
$\Theta_k = 2\pi f_r q k + \phi$	Residual Phase offset
$\hat{\Theta}_k$	Phase estimation using GA method
$\tilde{\Theta}_k$	Phase estimation using DM method
$\bar{\Theta}_k$	Phase estimation using PM method
$\mathcal{P}_d$, $\mathcal{P}_{\mathrm{md}}$	Probability of correct detection, Probability of miss detection
$\mathcal{P}_s$, $\mathcal{P}_{\mathrm{ms}}$	Probability of correct synchronization, Probability of miss synchronization
$\mathcal{P}_c$, $\mathcal{P}_\epsilon$	Probability of correct decoding, Probability of miss decoding
$\mathcal{P}_{\mathrm{fa}}$	Probability of false alarm
$\mathcal{P}_\epsilon^*$	FER according to Polyanskiy's bound

Chapter 1

General Introduction

1.1 Introduction

This Ph.D. thesis is a collaborative framework between the Université de Bretagne Sud (UBS, France) and the Lebanese International University (LIU, Lebanon). It has been supervised by Prof. Emmanuel Boutillon and Dr. Ali Al Ghouwayel. The research leading to these results received funding from the French National Research Agency ANR-19-CE25-0013-01 part of the project entitled Quasi Cyclic Short Packet (QCSP) (website: https://qcsp.univ-ubs.fr/).

The Internet of things (IoT) is an increasingly growing topic of conversation, where the forecasts predict that more than 50 billion devices will be connected through IoT. Most of this traffic is reported wirelessly, which is a major challenge due to limited frequency resources. In this context, various applications are supported by the utilization of a range of technologies. The performance-oriented categories like LTE or WiFi represent the first edge of these technologies. Such categories deploy sophisticated concepts including multi-user Multi Input Multi Output (MIMO) techniques to boost throughput and spectral efficiency. However, the Low Power Wide Area (LPWA) networks form the other edge of the technologies [1], where the requirements include a large coverage area, low data rates, a small data packet size, and low energy consumption at the device side [2]. EC-GSM [3], Narrow Band-IoT [4], LTE-M [5], LoRa [6], and SigFox [7] are examples of current IoT standards under the LPWA networks.

At a system level, reducing "meta-data" throughput, (i.e. the exchange of information linked to signaling, synchronization, and identification) is the new paradigm of massive IoT networks [8] to enhance the spectral efficiency. Polyanskiy has shown in [9], that asynchronism, even with short packets, does not affect the capacity of the channel; this means that classical methods that use coordination for synchronization and collision avoiding

are far from the optimum since the energy used for coordination is simply wasted.

In an unslotted ALOHA protocol, the receiver has no information regarding the time of arrival of messages (or frames). Moreover, in the context of this Ph.D., each frame is assumed to be affected by a phase and frequency offsets. This frequency offset can be generated by the clock-Jitter of a local device or Doppler effect. It can be also generated in a purpose for different motivations which are out of the scope of this work (multi-users access for example). The issue of frame detection, synchronization, and decoding at a low Signal to Noise Ratio (SNR) has appeared again with short packet transmission. The overall joint probability of successful transmission in an asynchronous ALOHA system can be expressed as $\mathcal{P} = \mathcal{P}_d \times \mathcal{P}_s \times \mathcal{P}_c$, where $\mathcal{P}_d$ is the probability of correct detection of the frame, $\mathcal{P}_s$ is the probability of correct estimation of the synchronization parameters, and $\mathcal{P}_c$ is the probability of correction of all transmission errors by an error-correcting code. Maximizing the probability of successful transmission requires maximizing $\min(\mathcal{P}_d, \mathcal{P}_s, \mathcal{P}_c)$. In other words, the communication performance is given by the weakest probability.

In this work, we propose to use the modulation presented in [10] to transmit short packets without any additional symbol that is dedicated to detection and synchronization. It is working also at ultra-low SNR, i.e. SNR < 0 dB. This "preamble-less" frame is hereby referred to as a Quasi-Cyclic Short Packet (QCSP) frame. It is based on the use of a Cyclic Code Shift Keying (CCSK) modulation scheme [11, 12] characterized by an inherent correlation property that will help the frame detection and synchronization at the receiver side. The key idea is to consider the whole frame first as a preamble for detection and timing synchronization, then as an encoded payload for error correction decoding and information recovery. This idea is implemented owing to the cyclic property of the CCSK modulation that allows the design of efficient detection and synchronization algorithms based on the correlation of the received frame with cyclically shifted versions of a predefined pseudo-random sequence. In addition, this CCSK modulation is jointly designed with powerful Non-Binary (NB) forward error correction codes defined over a Galois Field GF(q), where $q > 2$, such as NB-Low Density Parity Check (NB-LDPC) Codes [13], NB-Turbo [14], NB Turbo Product Codes [15], and NB-Polar codes [16]. These codes benefit from excellent error-correcting performance. Moreover, due to their non-binary nature, they enable a direct mapping between codeword symbols and symbols of high order modulation. This association is usually called coded modulation. It avoids the binary marginalization required in classical Bit Interleaved Coded Modulation (BICM) [17]. In this work and without loss of generality, the NB-code considered is the NB-LDPC.

1.2 Contributions

This work is conducted in the context of wireless sensor networking, where low-cost sensors are considered. We consider the communication scenario where the sensors are sending a sporadically short message to a base station with an unslotted ALOHA protocol. Note that the QCSP frames can also be used in much more structured networks. In that case, the receiver complexity can be significantly reduced. This point is out of the scope of the thesis. The main contributions of this Ph.D. can be summed up as the following:

— Firstly, we develop a practical detection algorithm for a QCSP frame in the Additive White Gaussian Noise (AWGN) channel that does not require any prior knowledge of the time of arrival, phase, and frequency offsets. Using the tools of detection theory, we derive a theoretical model to express the probability of miss-detection and the probability of false alarm according to the QCSP structure and the channel conditions. The global detection performance is assessed based on a detailed QCSP parametric study. *This contribution is submitted in [18].*
Another contribution, not described in the Ph.D. report since it is partial. It is related to my participation in the definition of the "time sliding method". It corresponds to the implementation of correlation of the detection algorithm in the time domain instead of the frequency domain. *This contribution is published in [19].*

— Secondly, we propose a time synchronization algorithm that estimates the start of the QCSP frame with high accuracy at very low SNRs. This synchronization method mitigates the time ambiguity first at the symbol level, then at the chip level. Synchronization at symbol level is performed thanks to phase information added by an Over-Modulation sequence applied to the symbols of the frame. Synchronization at the chip level is done using side information given by the Non-Binary code structure. Each of these steps is validated by Monte Carlo simulations. *Those contributions are published in [20]. Moreover, a patent about the Over-Modulation technique has been filled.*

— Thirdly, we propose two methods for phase and frequency synchronization. The first method, called Direct Method (DM), assumes that all the CCSK symbols are demodulated correctly. With this hypothesis, the frequency and phase estimation task resumes to an estimation problem of a pure complex sinusoidal signal affected by Gaussian noise. The second method, called Parametric Method (PM), is based on the Maximum Likelihood (ML) estimation using a parametric Prob-

ability Density Function (PDF) of each phase error. The first parameter of the proposed PDF model is computed using the CCSK demodulation score ratios, where the ML method used a weighted version by privileging the symbols received with high reliability. The second parameter is from the first decoding iteration of the NB-Low Density Parity Check (NB-LDPC) decoder. The PM method is simple to process and gives a result close to the Genius Aided (GA) method, i.e. the DM method when all the transmitted symbols are considered to be known. *This contribution is submitted to the 2022 IEEE 95th Vehicular Technology Conference: VTC2022-Spring [21].*

— A practical example is given to trade-off the joint reception probabilities, where the Extended Min-Sum (EMS) is used as the decoding algorithm of the NB-LDPC in the QCSP system. We study the detection-correction trade-off and compare the obtained result to a classical Zadoff-Chu preamble-based frame [22, 23], using the LDPC (proposed in 3GPP (5G) standard) as an error-control code. After that, we add the synchronization probability to the obtained results and trade-off the detection correction and synchronization performance. The output shows that there-exist some time offsets limited to a few chips. To solve this problem, iterative NB-code operations are made for each of the chip hypotheses, and a good time synchronizing is obtained and thus QCSP frame is decoded.

Finally, a proof of concept of the proposed algorithms is applied in a real radio system. QCSP frames have been generated, transmitted through the wireless channel, then successfully detected, synchronized, and decoded at the receiver side. Although the channel is real, all the processing of the algorithms is done offline. The results obtained with this "Real-channel transmission" validate the obtained MC-simulation results.

1.3 Organization of this Manuscript

This manuscript is organized into seven chapters. The content of chapters 2 to 7 is described in the following.

Chapter 2: In this chapter, the telecommunication chain layers are first presented by defining the Open System Interconnection protocol, where the physical layer is discussed in detail. Then, a brief history of the evolution of wireless communication till the emerging of IoT technologies is provided. After that, different properties

and aspects of the IoT technologies (SigFox, LoRa, and NB-IoT) are highlighted. The challenges of short packet transmission in IoT technologies are then discussed. Finally, the related work and the thesis goal are presented.

Chapter 3: The whole communication chain of the QCSP system is briefly discussed. It is mainly based on the CCSK modulation and NB-LDPC coding combination. First, a general overview of Forward Error Correction (FEC) codes and the Galois Fields (GF), under which these codes are defined, is provided. The principle of the NB-LDPC decoding procedure is then presented. The CCSK modulation is also addressed, as well as the process for creating the best spreading sequence. Then, a novel idea of adding an over-modulation sequence to help the synchronization task is also defined. Finally, the flow of a message from the transmitter to the receiver side passing through an asynchronous AWGN channel is thoroughly addressed.

Chapter 4: The preambles-less detection method is discussed. Then, a theoretical model which permits assessing the performance of the detection method for different parameters is derived and validated. Finally, performance analysis has been performed to illustrate the effect of different QCSP parameters (length of symbols and GF order), and channel (SNR, time, and frequency offsets) on the detection performance.

Chapter 5: This chapter develops and analyzes the time, frequency, and phase synchronization methods. It is divided into two main parts. In the first part, a preamble-less time-synchronization algorithm for QCSP frames is developed. The steps of achieving blind synchronization are discussed in detail. Then, the performance results are presented.

In the second part, the phase synchronization of the residual frequency and initial phase offsets are discussed. Both the DM and PM are defined and discussed in detail. Finally, the performance results are presented.

Chapter 6: This chapter synthesizes the results of detection and synchronization with the decoding using the NB-LDPC decoder. It first recalls the definition of the estimated Shannon's limit for short packet size thanks to Polyanski's bound. Then, the problem of optimization between detection, correction, and decoding is formalized. This chapter also sets a comparison with the classical preamble frame using up-to-date code where it shows that a QCSP has higher spectral efficiency.

Finally, a verification of the theoretical algorithms is shown using an offline chain via GNU radio software and USRPs. The offline data-processing shows that a QCSP frame can be transmitted and received correctly in a real radio system, even at very low SNRs.

Chapter 7: This Chapter concludes the thesis and summarizes the main contribution. It also points out the open research problems and provides a discussion about the perspective directions for future works.

Chapter 2

State of Art

Contents

In this chapter, the Open System Interconnection (OSI) protocol for a telecommunication chain is presented at first. It consists of seven different layers, where the thesis focuses on the first layer (which is the Physical layer). In the next section, the importance of the short packet transmission for IoT technologies is shown. The necessity of the LPWA networks for wireless sensor networks is also discussed. Different properties and aspects of the IoT technologies (SigFox, LoRa, and NB-IoT) are briefly highlighted. The third section shows the challenges of the short packet transmission of IoT. Accordingly, the coordination in massive IoT networks, Shannon's theory, and metadata organization are discussed in the context of short packets. After that, the related work in this target is presented and discussed. Finally, section 5 sums up our thesis goal.

2.1 Communication chain

In telecommunication, the OSI protocol describes the way of communication between two entities or more. It defines the standards and rules of the structure of system communication, its semantics, and syntax. For organization, and facilitating successful transmission, this protocol is split into seven layers, where every single layer has its own function. Each of these layers communicates only with the layers above and below it [24]. This helps to perform troubleshooting, by determining the layer that is creating a problem to focus the effort on it.

2.1.1 Open System Interconnection layers

The OSI layers are briefly discussed from "top" to "down", from the application layer that directly interacts with the end-user, down to the physical layer that interacts with the transmission medium. Fig. 2.1 shows the different OSI layers.

Starting with the Application layer, it primarily serves as a software interface between

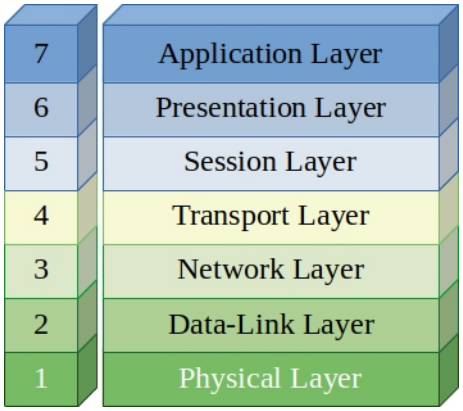


Figure 2.1 – OSI model Layers.

the user and network services. It provides protocols that allow the software to send and receive data that is presented in a meaningful and useful way to the user. Some of the application layer examples are the Hypertext Transfer Protocol (HTTP), File Transfer Protocol (FTP), Post Office Protocol (POP), Simple Mail Transfer Protocol (SMTP), and Domain Name System (DNS). The presentation layer is the name of the sixth OSI layer. It either takes data transmitted by the application layer and prepares it for transmission over the session layer or prepares data coming from the session layer and gives it to the application layer. Encoding, encryption, and compressing the data is the main job

of the presentation layer. The session layer, which is the fifth layer, establishes the communication channels between devices, known as sessions. It is responsible to open the sessions and make sure that they stay functioning while the data is transmitted. It should terminate them after the connection is finished. The session layer can also specify checkpoints during a data transfer, allowing devices to restart data transmission from the latest checkpoint if the session is stopped. On the receiving end, the transport layer receives the data sent in the session layer and divides it into "segments". It is responsible also to reassemble the segments on the receiving end and turning them back into data that is used in the session layer. The transport layer handles flow control, which involves providing data at a rate that matches the receiving device's connection speed. It also carries out an error control, which involves determining if data was received correctly or not, to request it again. The Network Layer belongs to the third layer of the OSI model. It mainly performs the transmission of data in various networks. This layer might not be so beneficial if we are transmitting the info in the same network. Routers are mainly utilized in the network layer for routing purposes. The network layer uses network addresses (typically Internet Protocol addresses) to route packets to its destination node. A connection between two physically linked nodes on a network is established and terminated via the data link layer. It divides the packets into frames and transmits them from the source to the destination. This layer is divided into two sections: Logical Link Control (LLC), which detects network protocols, conducts error checking and synchronizes frames. Also, the Media Access Control (MAC) section, which is responsible for devices connections and permissions to transmit and receive data.
The lowest layer in OSI is the Physical (PHY) layer, where our thesis algorithms are focusing on. It is the layer that interacts with the transmission medium, which may consist of wires, coaxial cable, radio link, or anything [25].

2.1.2 PHY layer

The PHY layer's purpose is to convert the raw information bits which come from the upper layers into a physical signal to be transmitted over the transmission medium. It can be simply represented as a series of 1s and 0s, in addition to taking care of bit rate control. The layer should also be able to perform the reverse operation and retrieve the data bits from the received signal. The communication between two PHY layers is modeled by a transmission system, represented in Fig. 2.2.

Usually, it is decomposed into three main parts: a transmitter (including information

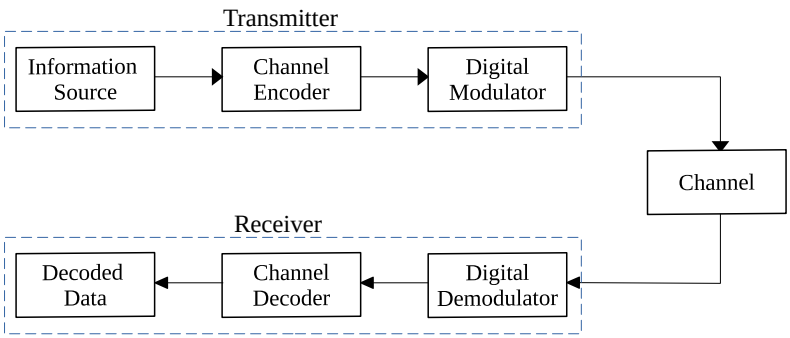


Figure 2.2 – The transmission system model.

source, channel encoder, and digital modulator), the channel of transmission, and a receiver (including digital demodulator, channel decoding, and the decoded data). In this work, a simplified base-band model is considered. The transmitter and receiver analog base-band signals are modeled by complex sampled signals. The work of this thesis is held in this layer and is mainly based on the association of channel coding and digital modulation.

2.2 IoT in wireless technologies

In this section, the evolution of different wireless technologies is briefly shown. First, the interest of IoT technologies in short packet transmissions is mentioned. Then, some Low Power Wide Area (LPWA) networks, which are dedicated to the wireless sensors in the IoT technologies, are presented.

2.2.1 Interest of IoT in short packets transmission

In wireless telephone technology, the second generation (2G) introduced digital communications between mobile phones in the 90's. Its data services enabled the various networks to provide services such as text messages, picture messages, and multimedia messages. Because of the growing connectivity and popularity of the internet, the third (3G) and fourth (4G) generations in mobile cellular networks have been developed. Several evolutions took place since the 2G era, aiming for higher rates and enhancing the user Quality of Service (QoS). With peak data speeds of 1 Gigabit per second (Gbps), 4G

(also known as Long Term Evolution (LTE)) was substantially implemented worldwide [26].

5G is the fifth generation technology standard for broadband cellular networks, which cellular phone companies began deploying worldwide in 2019. The 5G not only boosts the data rate, but also improves existing applications like Machine-to-Machine (M2M) communications, and provides various new use cases [27] like the IoT [28]. This IoT which has the potential to change lifestyle and work has the potential to assist us in tackling the key global issues of population growth, energy scarcity, resource depletion, environmental pollution, etc. To accomplish this vision, objects and things must be able to detect their surroundings and exchange this knowledge with one another as well as with people to make intelligent decisions. Because of this potential, where it has positive impacts on our whole ecosystem, the interest in IoT is growing [29]. Apart from supporting increased bandwidth services, 5G brings two new application scenarios: massive Machine Type Communication (mMTC) and Ultra-Reliable Low Latency Communication (URLLC). Accordingly, it deals with a variety of data transmission demands, driven by prospective IoT applications. The last aspects of 5G communication were underway, prompting interest in the sixth Generation (6G) in the research community [30]. The above-mentioned two scenarios will be further explored since 6G is projected to accommodate a wider range of application scenarios.

Short packet communications are one of the main challenges in providing services for mMTC and URLLC applications. Consequently, future systems are considered to be different from the existing ones which are based on longer blocklength for high bandwidth [31]. Providing reliable data transmission is a necessity for many potential IoT applications in 5G and beyond. However, communication with small packets results in a significant degradation in channel coding gain, making communication reliability harder to be achieved [32].

Wireless Sensor Networks (WSNs) are an important subset class of IoT that is increasing and more demanded. It is a wireless network containing distributed independent sensor devices which are meant to monitor physical or environmental conditions. A WSN is made up of a network of interconnected small sensor nodes that interact and share information and data. These nodes collect environmental data such as temperature, pressure, humidity, and pollution levels and communicate it to a base station. Depending on the type and amount of data monitored, the latter delivers the information to a wired network or triggers an alert or an action [33, 34]. In the following decades, several billions

of items are projected to be linked and connected, with an expected ratio of more than 6 devices connected per person [35, 36]. WSN is using the same transmission channel, which is the air, as wireless local area networks for wireless transmission (WLANs). Standard access protocols are provided and available to allow nodes in a local area network to interact effectively. However, these protocols, cannot be applied directly to the WSNs. The main difference is that sensors have a very limited amount of energy (typically a battery) that drains out extremely quickly, unlike devices in local area networks. As a result, new energy-aware protocols at the MAC or PHY levels are required. There is a clear difference between a standard WLAN and a WSN, as the latter has restricted resources.

To sum up, while the majority of connections are under the cellular or legacy networks like Bluetooth or WiFi, a gap persists and should be filled between local wireless networks and cellular networks.

The aforementioned problems are addressed by Low Power Wide Area (LPWA) networks are addressing [37]. More than 10% of the anticipated 25 billion IoT connections are projected to be LPWA connections [38], illustrating the potential market for this new network. LPWA networks are rapidly entering the communities of industry and research due to their low power, long-range, and low-cost communication characteristics. It is supplying long-range communication up to 10–40 km in rural zones and 1–5 km in urban zones [39], as well as providing highly efficient and inexpensive energy (i.e. 10+ years of battery lifetime [40]). Because of its promising properties, LPWA has inspired recent experimental studies on its performance in both outdoor and indoor situations [41, 42].

2.2.2 Low Power Wide Area (LPWA) networks

LPWA networks are a revolutionary communication paradigm that is enhancing traditional cellular and short-range wireless technologies in meeting the various requirements of IoT applications. LPWA technologies provide unique capabilities such as wide-area connection for low power and low data rate devices where the classical wireless technologies were not delivering. In Fig. 2.3, the tradeoffs between LPWA and the traditional technologies are highlighted. Traditional technologies of IoT landscapes as short-range wireless networks (e.g. ZigBee, Bluetooth, Z-Wave), wireless local area networks (WLANs) (e.g. Wi-Fi), and the cellular networks (e.g. Global System for Mobile Communications (GSM), LTE, etc). Non-cellular wireless methods aren't suited for connecting low-power devices that are distributed over wide geographic areas. These technologies have a maximum range of a few hundred meters. The devices cannot be deployed or moved arbitrar-

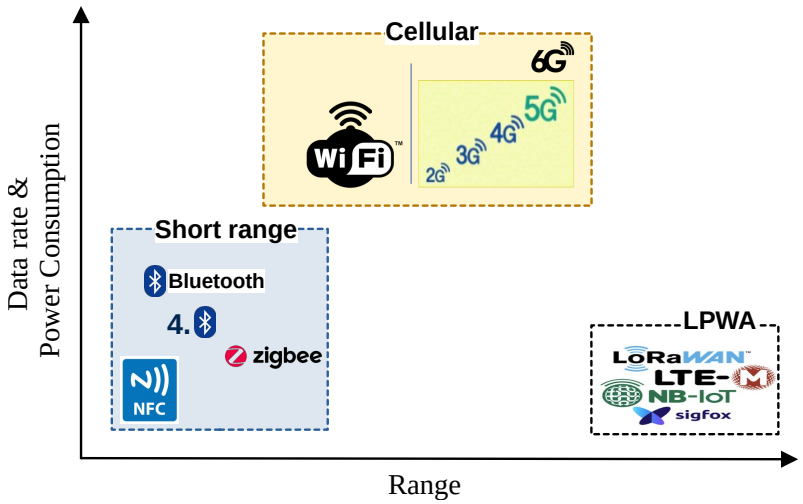


Figure 2.3 – Required data rate and power consumption vs. range capacity of radio communication technologies [43].

ily. Dense deployment of devices and gateways connected by multi-hop mesh networking expands the range of these technologies. Consequently, large-scale installations are prohibitively costly. When it comes to mMTC, legacy WLANs have lower coverage areas and higher power consumption. In conclusion, as shown in Fig. 2.3, LPWAN is highly convenient for IoT applications that just require transferring a low amount of data in long-range.

As a result, achieving long-range transmission and a battery life of 10 years or more is critical for LPWA connectivity. Low degrees of sensitivity is required by the receiver, which is depending on the data rate. Consequently, long-range communication can only be achieved when paired with a low data rate. The majority of existing LPWA technology is based on one of two ways to reduce the data rate. The first technique is to deal with narrow-band signaling to limit the amount of collected noise. The second option is to reduce the spectral efficiency of the technology being employed, for example, by using a Spreading Factor (SF) or channel coding. Therefore, it is evident that LPWA solutions are not intended to handle every IoT use case and instead focus on a certain section of the IoT landscape. This is specially investigated for use cases that do not require high data rates, are delay tolerant, have cheap cost, and require low power consumption.

Table 2.1 gives a summarized overview of LPWAN technologies SigFox, LoRa and NB-IoT for IoT. For more details about the proper techniques and some aspects of the

different technologies, the reader is referred to Appendix 8.1.

2.3 Challenges of short packets transmission in IoT

Short packet transmission is a modern field that is necessary for wireless communications. It is also widely existing in multiple IoT applications. Because of that, many challenges are rising.

2.3.1 Coordination in massive IoT network

Massive IoT connection is one of the major challenges that the 5G wireless networks and beyond are facing [44]. A simple scenario can be viewed for this massive connectivity when a large number of devices (usually thousands), linked to a base station, where each of them is sporadically active and transmits short packets. In this scenario, requiring random access protocol is mandatory, since each of the instants of sending data from the devices is unknown.
In the classical ALOHA model for random access, [45], analysis in massive access scenarios is performed usually for infinite population, i.e. number of users N tends to infinity. The protocols for the coordination and organization of the transmission take a large number of resources, which is very limited in short packet transmission in IoT. Consequently, one needs to think about the structure of the packet to examine the fundamental performance bounds. For the transmission of short packets in an unslotted aloha protocol, two information-theoretic approaches have been introduced [46]. The first approach is that the number of users in many access channels [47] is presented as a function of codeword length. This allows the identifying capabilities to be preserved even when both go to infinity. The second approach is to assume that the packet does not contain the address of the sender as usual protocols for long packets. This leads to an unsourced access scheme, and consequently the case where the same code-book is shared for all the users.
The coming proposed work follows these assumptions to undergo the short packet transmission and reduce the used resources for organization [46].

2.3.2 Shannon's theory

The channel capacity (defined also as Shannon's limit) has been introduced by Claude Shannon in 1948 [48]. It remains in information theory one of the fundamental results

Table 2.1 – Overview of LPWAN technologies: SigFox, LoRa, and NB-IoT.

	SigFox	LoRa	NB-IoT
Modulation	BPSK	CSS	QPSK
Frequency	Unlicensed ISM bands (868 MHz in Europe, 915 MHz in North America, and 433 MHz in Asia)	Unlicensed ISM bands (868 MHz in Europe, 915 MHz in North America, and 433 MHz in Asia)	Licensed LTE frequency bands
Bandwidth	100 Hz	250 kHz and 125 kHz	200 kHz
Max data rate	100 bps	50 kbps	200 kbps
Bidirectional	Limited / Half-duplex	Yes / Half-duplex	Yes / Half-duplex
Max message delay	140 Upper Link (UL), 4 Down Link (DL)	Unlimited	Unlimited
Max Payload length	12 bytes (UL), 8 bytes (DL)	243 bytes	1600 bytes
Range	10 km (urban), 40 km (rural)	5 km (urban), 20 km (rural)	1 km (urban), 10 km rural
Interference Immunity	Very high	Very high	Low
Authentication and encryption	Not supported	Yes (AES 128b)	(LTE encryption)
Adaptive data rate	No	Yes	No
FEC schemes	No	Simple Hamming codes	Conventional codes
Standardization	Collaborating with ETSI	LoRa-Alliance	3GPP

until nowadays. The channel capacity is defined as the maximal achievable rate R in the asymptotic of infinite codeword length, for a given bandwidth B, level of SNR, and an arbitrarily small amount of error. In an AWGN channel, R (or capacity in bits.s^{-1}) is classically expressed as

$$R = B \log_2(1 + \mathrm{SNR}). \tag{2.1}$$

The capacity bounds the rate of any reliable transmission in telecommunication, consequently, extensive research efforts have been focused on bringing the rate close to the capacity. Once they are very close for a given bandwidth and SNR, this indicates that the resource is being efficient and used in its optimal ability.

In the non-asymptotic regime, there are no exact formulas for the maximal achievable rate as a function of the code length N. However, in [49], Polyanskiy reformulated the problem for the Finite Block Regime (FBR). It was shown that the back-off from channel capacity can be accurately and succinctly characterized by a parameter known as channel dispersion. Specifically, the maximum achievable coding rate in FBR, denoted by R_c^* , can be tightly (for $N > 100$) approximated by

$$R_c^* \approx R - \sqrt{\frac{V}{N}} Q^{-1}(\mathcal{P}_\epsilon), \tag{2.2}$$

where R is the channel capacity (maximum rate achievable in the asymptotic regime), V is the channel dispersion, Q^{-1} is the inverse of Q function (i.e. Q-function is the tail distribution function of the standard normal distribution), and $\mathcal{P}_\epsilon$ is the error probability. We use this approximation (known as the normal approximation) in the results section to find the optimal results of $\mathcal{P}_\epsilon$ that can be found in an AWGN channel.

2.3.3 Meta data between short and long data packets

Transmitting long data packets is demonstrated via information-theoretic principles which are the basis of most of the modern developments in the design of high-speed, reliable, and efficient wireless systems. The updated versions of the wireless systems require assisting novel traffic types that deploy short packets. For instance, sensors in WSN and other devices involved in M2M communications are based on short packet transmission. Moreover, it is expected that small packets are carrying critical information that should be received, through the applications, with ultrahigh reliability and low latency. The asymptotic capacity-achieving principles are not relevant to the transmission of short

packets because of the very small amount of transferred data.

Most of the development of the existing systems depends on the fact that the size of the metadata (control information) is very small (sometimes negligible) compared to the real information payload as shown in Fig. 2.4. Therefore, the performance of the

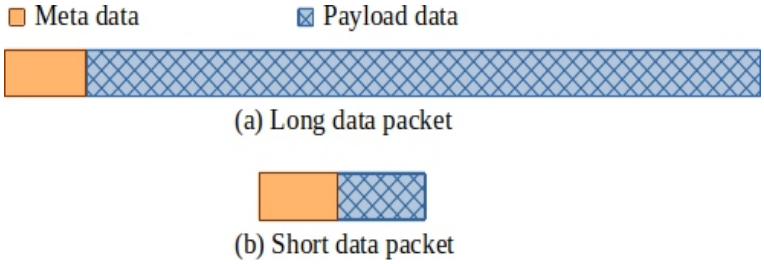


Figure 2.4 – Packet structure of data and meta-data in both long and short packets.

overall system is not affected by the heuristic methods to transfer metadata. However, when the packet is short, both the payload and metadata may have similar sizes with highly sub-optimal conventional methods to transfer the packets. At a system level, reducing "meta-data" throughput, (i.e. the exchange of information linked to signaling, synchronization, and identification) is the new paradigm that requires new guidelines for massive IoT networks [8]. Additionally, to approach Shannon capacity, the most current PHY layer design massively depends on long codes and needs to be adopted for finite codes. As a result, short packets, which are known as Finite Block Length (FBL), require new and advanced design rules.

2.4 Related works

In an unslotted ALOHA protocol, the base station has no information regarding the time of arrival of the messages (or frames). Moreover, each frame is affected by a frequency offset due to the low-cost sensor transceivers, which imposes uncertainty on the modulation frequency. The problem of frame transmission/reception at low SNR is examined in the literature.

The information theory community made significant progress in addressing the challenge of sending short packets. Information theorists have obtained some grasp of the theoretical principles underlying short-packet transmission and have metrics that allow them to measure their performance, particularly in point-to-point settings. The NB-codes are attractive for their tremendous error-correcting capabilities in low SNR environments.

Information theorists, on the other hand, have always regarded metadata design as a field outside of their expertise. So, they work on the decoding blocks taking the assumption that a frame is detected and received perfectly synchronized. Consequently, heuristic approaches are responsible for the transmission of metadata.

Polyanskiy has shown in [9], that asynchronism, even with short packets, does not affect the capacity of the channel; this means that classical methods that use coordination for synchronization and collision avoiding are far from the optimum since the energy used for coordination is simply wasted. In [9, 50, 51], the maximum channel coding rate bounds for the finite block regime, which are peers of asymptotic capacity metrics, were recently obtained. To reach the claimed throughput, effective frame detection and synchronization are also necessary before the decoding process. Blind estimators should be capable of delivering frame detection and synchronization. However, for reasons of complexity, the usage of synchronization words, which are sequences of known symbols, is favored. A frame header [52] can be used to connect preamble symbols to the information symbols. When the frame length is fixed, adding a header for frame detection and synchronization lowers both spectral efficiency and coding length. Reducing this length, on the other hand, will diminish both the detection and synchronization performance. As a result, a trade-off in the recent research fields is running to maximize the chances of getting a frame free of mistakes with optimal spectral efficiency.

Many papers propose detection, frequency, and time synchronization algorithms based on the transmission of sync words, or preambles for each frame [53–58]. The correlation between the predefined sync word and the received signal is calculated to determine the correct frame starting point. The same or similar patterns of the sync word may be present in the payload data. Hence the performance of detectors and synchronizers using sync words is constricted by the random data limit [59]. Besides, sync words consume signal energy. Thus insertion of sync words is not the best solution for codes working at very low SNRs. These classical preamble-based methods allow simplifying the receiver complexity significantly thanks to the known received information. However, the use of a preamble alleviates a significant part of bandwidth resource when the message payload is small as shown in Fig. 6.4 in the results section.

Frame synchronization is a well-studied subject. According to two separate transmission mechanisms, these techniques may be loosely split into two groups.
The first is burst transmissions, for which binary hypothesis testing is used to synchronize the data [60, 61]. The Neyman-Pearson lemma is commonly used to calculate a thresh-

old for binary hypothesis testing. In [62–64], the authors investigate optimal metrics for binary modulation over AWGN channels in terms of decreasing the synchronization error probability. In [65], the authors provide an approximation of the best metric for M-PSK modulation with phase offset error across AWGN channels. A study of the synchronization for burst transmission in the finite blocklength domain was recently suggested in [66].

Continuous transmission constitutes the second group of frame synchronization techniques, in which frames are sent one after the other. Maximum Likelihood (ML) synchronization is enabled by frame length knowledge: a metric to be maximized is computed for all feasible synchronization data locations. The measure to be utilized is determined by channel models and data distributions. For binary symmetric channels (BSC), for example, the correlation metric reduces the synchronization error probability [67], but the ideal metric for AWGN channels with binary modulation includes an extra energy correction term [68]. The authors of [59] propose ML metrics for generic M-ary phase-coherent and phase-non coherent AWGN channels as an extension of [68]. The best non-coherent detection metrics for Rayleigh fading models are given in [69]. So far, little work has been offered in terms of analytic performance (the chance of erroneous frame synchronization). We are aware of some related works that give both measurements and analytic performances, and [70] is one of them.

The engineering literature has already examined and explored some various blind (preamble-less) methods for short packets detection and synchronization [71–75], but all the proposed algorithms have proved their efficiency on positive decibel SNR values (i.e. SNR $>$ 0 dB). In [76] also, a phase synchronization method has been proposed based on the use of turbo decoder information to improve the carrier phase and offset synchronization for short packets. This method showed good performance at an SNR greater than -0.5 dB. Consequently, conventional frame synchronizers, which ignore the structure of the code, usually fail at low SNR. To improve frame acquisition performance, frame synchronization should be considered jointly with decoding [77, 78]. This is based on superimposing the preambles and headers to data symbols in the context of low latency communications and/or huge connectivity, thus the frame length can be lowered while maintaining a maximum frame synchronization length (i.e. the length of the whole frame) [79].

One more interesting topic in the literature is the trade-off between detection, synchronization, and decoding. It has been demonstrated that optimizing detection and decoding independently is sub-optimal [80], and that the probability of false alarm, misdetection,

and decoding error should be considered jointly. In [81], the authors deal with joint detection and synchronization for a Direct Sequence Spread Spectrum (DSSS) system utilizing differential encoding, and the probabilities of false alarm and misdetection for the AWGN channel are determined. It was investigated in [82] the interaction between error control coding and channel estimation for a short packet scenario in an AWGN channel and the unknown, constant gain over a block. It is demonstrated that for a single-antenna receiver, there is an optimal training length for which the required SNR can be minimized.

2.5 Thesis goal

In this Ph.D., we aim to contribute to the evolution of the IoT communication paradigm through the design of a spectrally efficient and low-power short packet-based transmission chain. This is achieved by removing the preamble from the frames, which leads to saving the associated power and bandwidth resources. Compared to the existing related works mentioned in the previous section, and to the best of our knowledge, the preamble-less detection and synchronization techniques proposed in this work are the first techniques allowing reliable frame detection and synchronization at negative SNR going down to -10.5 dB.

We consider a low-cost sensor that sporadically transmits/receives, through the AWGN channel, small messages in an unslotted asynchronous ALOHA protocol, i.e. without prior knowledge of the time of arrival and the potential carrier frequency offset of the signal. In this Ph.D. work, we propose to use the CCSK-NB-LDPC coded-modulation scheme to compensate for the use of the preamble. This frame, named QCSP, is seen as a preamble for detection and synchronization from one side and viewed as a codeword for decoding from the other side. The main focus is to develop blind detection and self-synchronization algorithms and to evaluate their performance using both theoretical mathematical tools and experimental MC simulations. After that, a proof of the theoretical concept is done through the implementation of the proposed detection and synchronization methods on a real-radio system based on SDR transceivers. This is achieved by offline processing of the received data using USRP platforms supported by GNU radio software.

Chapter 3

CCSK and NB-LDPC based Communication chain

Contents

The main objective of this chapter is to discuss the blocks of the overall communication chain in the QCSP system. It is mainly based on the association of the CCSK modulation and NB-LDPC code. First, it a gives general overview regarding the Forward Error Correction (FEC) codes and the Galois Fields (GF) under which these codes are defined. Then, it presents the principle of the decoding process of NB-LDPC code. The CCSK modulation, as well as the methodology of generating the optimal spreading sequence, are also discussed. It presents a novel idea of adding an over-modulation sequence to help

the synchronization task. Finally, it describes in detail the flow of a message $\boldsymbol{M}$ through the considered QCSP system.

3.1 Forward Error Correction codes (NB-Codes)

Nobody is tolerant of receiving incorrect data (voice, picture, video, and message or mixed) which gets contaminated due to the noise during its journey. The noise comes from different places, from the beginning of the transmitter, then passing by the channel, until the last part of the receiver. Channel coding is one of the most important techniques against the noise effect. The redundancy added at the emitter side by the encoder allows to correct a given amount of transmission error at the receiver side thanks to the FEC. There are powerful Non Binary (NB)-FEC codes that exist and defined over GF(q), where $q > 2$, such as Non-Binary Low Density Parity Check (NB-LDPC) codes [13], Non-Binary Turbo Codes (NB-Turbo) codes [14][83], NB-Turbo Product Codes [15], and Non-Binary Polar codes (NB-Polar) codes [16]. These NB codes offer a capability of error correction, thereby enabling a coding gain that allows the transmission at low power. This family of error correction codes received the attention of a considerable number of researchers in the digital communication community because of its good performance with short packet size and/or the high order modulation compatibility [84]. These codes benefit from better error-correcting performance than their binary counterpart since their non-binary nature codes are directly mapped on high order modulation avoiding binary marginalization [17]. If the cardinality q of the symbol set is equal to the cardinality q of the modulation space, then each symbol of GF(q) can be directly mapped to a symbol of the modulation.
This Coded-Modulation scheme is more efficient than the Bit Interleaved Coded Modulation (BICM) classically used binary encoded is used. This section presents the background of NB-LDPC codes and the principles of their decoding process. First, it introduces the notion of GF necessary for defining the NB-LDPC. Then, it discusses the structure of the code itself.

3.1.1 Algebraic definition of Galois Field(q)

Classical algebra is known to study the most commonly used sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$ built with arithmetic operations such as addition and multiplication. However, modern algebra is characterized by a higher level of abstraction; the concept of operation is defined as an

application that returns a symbol from two or more symbols combination. This allows the scientists to extend the definition of error correction codes to sets other than classical ensembles. This work has a special interest in the NB codes defined on GF. To present a clear definition of the Galois field, the basic algebraic structures and their internal composition laws are presented. The content of this section is mainly extracted from [85], [86], and [87].

Groups

Let G be a set of elements. A binary operation "$*$" on G is a function that assigns to a couple of elements a and b a unique element $c = a * b$ in G. A binary operation "$*$" on G is associative if, for any a , b , and $\boldsymbol{C}$ in G, $a * (b * c) = (a * b) * c$. A set G on which a binary operation "$*$" is defined, is called a group if the following conditions are satisfied:

1. The binary operation "$*$" is associative.
2. G contains an identity element ϱ of G, with $\forall a \in G$, $a * \varrho = \varrho * a = a$.
3. For any element $a \in G$, there exist another element $a' \in G$ such that $a * a' = a' * a = \varrho$; a and a' are inverse to each other.

A group G is called commutative if its operation "$*$" also satisfies the following condition: for any a and b in G, $a * b = b * a$. Finally, a group G satisfies the following properties:

— The identity element in a group G is unique. Proof: If ϱ and ϱ' are identity elements $\in G$. Then $\varrho' = \varrho' * \varrho = \varrho$.

— Every element has a unique inverse. Proof: If a' and a'' are inverse to a, then $a' = a' * \varrho = a' * a * a'' = \varrho * a'' = a''$.

Fields

Let F be a set of elements defined with two binary operations, addition "+", and multiplication "$*$". The set F together with the two binary operations "+" and "$*$" is a field if the following conditions are satisfied:

1. $(F, +)$ is a commutative group. The identity element 0 of the addition operation is called the zero elements of F.
2. $(F - \{0\}, *)$ is a commutative group. The identity element of the multiplication operation 1 is called the unit element of F.
3. Multiplication is distributive over addition; $\forall a, b$ and $c \in F, a * (b + c) = a * b + a * c$.

Table 3.1 – Primitive Polynomials.

Polynomial degree	Primitive Polynomials $\mathbb{P}_p$
1	$1 + x$
2	$1 + x + x^2$
3	$1 + x + x^3$
4	$1 + x + x^4$, $1 + x^3 + x^4$
6	$1 + x + x^6$
8	$1 + x^2 + x^3 + x^4 + x^8$

Galois Fields (GF)

A Galois field has a finite order, which is either a prime number or the power of a prime number. A field of order $q = n^p$ GF(n^p) or GF(q) contains q-elements which are denoted as $\{0, 1, \alpha, \alpha^2, \ldots, \alpha^{q-2}\}$, where α is called the primitive symbol of the field, the powers of which construct all the other elements of the field. A specific type called characteristic-2 fields represent the fields when $n = 2$. All the elements of a characteristic-2 field can be represented in a polynomial format [88].

The primitive polynomial $\mathbb{P}_p$ of the field of order 2^p is an irreducible polynomial of degree p that generates all the other polynomials. The set of polynomials defined over $\mathrm{GF}_{(2)}[x]$ modulo $\mathbb{P}_p$ defines the Galois Field GF(q), with $\mathrm{GF}_{(2)}[x]$ is the set of polynomials with coefficient in the set $\{0, 1\}$. Note that for each field GF(q), one or more primitive polynomial $\mathbb{P}_p$ of degree p over GF(q) can be found.

Table 3.1 lists some examples of the the primitive polynomials for $p \in \{1, 2, 3, 4, 6, 8\}$. For $p = 1$, the field is a binary field and for $p \geq 2$, it represents a non-binary field. Binary LDPC codes are defined over a GF(2), with 0 and 1 being the field elements. Hence NB-LDPC codes are a generalization of binary LDPC codes. Each element of a binary representation of the Galois field is represented by a polynomial with binary coefficients.

Table 3.2 shows the example for $p = 3$ while considering the primitive polynomial $1 + x + x^3$. The field consists of 8 elements and each one has a binary representation composed of the binary coefficients of the associated polynomial. With this representation, finite field addition and multiplication become polynomial addition and multiplication, where the addition is modulo-2 ($\alpha + \alpha = 0$). The result of a multiplication is realized by applying a polynomial multiplication followed by a modulo reduction using $\mathbb{P}_p$. Only the remainder of the Euclidean division is kept. In Table 3.2, all the α^j, $j \in \{0 \ldots 6\}$ have their binary representation by applying an Euclidean division by the primitive polynomial.

Table 3.2 – Binary Representation of Symbols.

Element	Binary representation	Polynomial Sum
0	000	0
α^0	100	1
α^1	010	α
α^2	001	α^2
α^3	110	$1+\alpha$
α^4	011	$\alpha + \alpha^2$
α^5	111	$1 + \alpha + \alpha^2$
α^6	101	$1 + \alpha^2$

Equation (3.1) presents the example of division of α^4 by the primitive polynomial.

$$\alpha^4 = \alpha * (\alpha^3 + \alpha + 1) + \alpha^2 + \alpha = \alpha^2 + \alpha \bmod[\mathbb{P}_p]. \tag{3.1}$$

$GF_{(2^p)}[x]$ modulo $\mathbb{P}_p$ can be considered as the polynomial representation of $GF(2^p)$. Then $GF_{(2^p)}[x] = GF_{(2)}[x]/\mathbb{P}_p[x]$, where $GF_{(2)}[x]$ is the polynomial set with coefficients in $\{0, 1\}$ and $\mathbb{P}_p[x]$ is the representation of $\mathbb{P}_p$ in $GF_{(2)}[x]$.

3.1.2 Non-Binary Low Density Parity Check Codes

LDPC codes with symbols that belong to the binary Galois field ($p = 1$) are said binary, while if $p \geq 1$, that is called NB-LDPC codes and the matrix products of the parity equations are made using the internal composition laws of the Galois field. The NB-LDPC is the extension of the LDPC codes. Binary LDPC codes have asymptotic performances approaching the Shannon limit [89]. However, for small or medium size codewords, the performance of the binary LDPC codes degrades considerably. It is shown in [90] that this loss can be compensated by using NB-LDPC codes of high cardinality. In addition, the high cardinality of the codes ensures better resistance to packet errors [91], [48]. This improvement of the performance can be intuitively explained by the fact that several bits are grouped into a single non-binary symbol. As a result, the erroneous bits are confined to fewer non-binary symbols, and subsequently, the parity constraints are affected by fewer errors. Nevertheless, the improvement of the performances by increasing the order of the Galois field is accompanied by an exorbitant increase of the decoding complexity. This constitutes a brake on the practical implementation of the NB-LDPC codes.

An LDPC code is a part of linear block codes family [92] with the particularity of being defined by a sparse Parity Check Matrix (PCM), which means a matrix containing only a small number of non-zero elements. From that, a generator matrix $\boldsymbol{G}$ is generated to be used at the transmitter side for the encoding of $K-$length information message $\boldsymbol{U}$ to obtain a codeword $\boldsymbol{C}$. The transmitted message $\boldsymbol{C}$ is obtained as the following

$$\boldsymbol{C} = \boldsymbol{U}\boldsymbol{G}. \tag{3.2}$$

The PCM generally denoted by $\boldsymbol{H}$ with dimensions $(M \times N)$. Its number of lines denoted by M corresponds to the number of parity check constraints of the code, and its number of columns denoted by N corresponds to the length of the codewords. A codeword consists of K information symbols and $M = N - K$ redundant symbols added by the encoder to protect and correct data. The parity check constraints of the matrix $\boldsymbol{H}$ must be respected by the codewords in the construction. Thus, a message $\boldsymbol{C}$ of length N is a codeword if and only if $\boldsymbol{C} \times \boldsymbol{H}^T = 0$, where $\boldsymbol{H}^T$ is the transposed matrix of $\boldsymbol{H}$. All the entries of the matrix $\boldsymbol{H}$ belong to finite Galois field GF($q = 2^p$) discussed in the previous section.

As an example, consider the following PCM $\boldsymbol{H}$ of size 4×6

$$\begin{pmatrix} h_{0,0} & 0 & 0 & h_{0,3} & 0 & h_{0,5} \\ h_{1,0} & h_{1,1} & h_{1,2} & 0 & 0 & 0 \\ 0 & h_{2,1} & 0 & h_{2,3} & h_{2,4} & 0 \\ 0 & 0 & h_{2,3} & 0 & h_{3,4} & h_{3,5} \end{pmatrix} \tag{3.3}$$

Hence a codeword $\boldsymbol{C} = [c_0; c_1; c_2; c_3; c_4; c_5]$ satisfies the following four equations

$$\begin{aligned} h_{0,0}c_0 + h_{0,3}c_3 + h_{0,5}c_5 &= 0 \\ h_{1,0}c_0 + h_{1,1}c_1 + h_{1,2}c_2 &= 0 \\ h_{2,1}c_1 + h_{2,3}c_3 + h_{2,4}c_4 &= 0 \\ h_{3,2}c_2 + h_{3,4}c_4 + h_{3,5}c_5 &= 0 \end{aligned} \tag{3.4}$$

In addition, the PCM of NB-LDPC code can be represented by a bipartite graph known as a Tanner graph [93]. The bipartite graph describes the code structure and also helps to perform the decoding algorithms, especially the iterative ones. A bipartite graph is composed of two sets of nodes. Each node is connected only to other nodes of the other set. For the NB-LDPC codes, the two sets of nodes are the Check Nodes (CNs) and variable nodes (VNs). The CN referrers to one row of the PCM and VN referrers to one

column or equivalently a symbol of the codeword. Thus, the bipartite graph associated with an LDPC code is represented by a $M \times N$ dimension $\boldsymbol{H}$ parity-check matrix. It represents the relation between M CNs and N VNs. If a check node CN_j is connected to a VN_i, the element of the j^{th} row and i^{th} column of the PCM will be non-null. The matrix of the given example in (3.3) can be represented by the bipartite graph in Fig. 3.1.

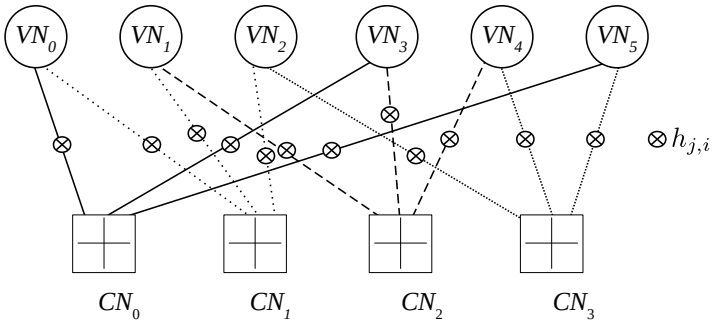


Figure 3.1 – Tanner Graph of a NB-LDPC code.

The number of non-null symbols in the j^{th} column of the PCM is denoted by $d_v(j)$, $j \in \{0 \dots N-1\}$, and the number of non-null symbols in the i^{th} row is denoted $d_c(i)$, $i \in \{0 \dots M-1\}$. A NB-LDPC code is regular if d_v and d_c are constant (i.e. $\forall j, i;\ d_v(j) = d_v$ and $d_c(i) = d_c$) respectively, for all the columns and all rows of the matrix, otherwise, the code is said to be irregular. The regularity of a PCM can be defined using its associated bipartite graph. The code is regular if d_v and d_c that define the number of variable node connections and check node connections, respectively, are constant. In the case of a regular NB-LDPC code, the code rate R_c of the code can be expressed as a function of d_v and d_c as follows

$$R_c = \frac{K}{N} = \frac{N-M}{N} \geq 1 - \frac{d_v}{d_c}. \tag{3.5}$$

3.1.3 Iterative decoding algorithms for NB-LDPC codes

The Believe Propagation (BP) decoding algorithms are based on the bipartite graph defined by the NB-LDPC code [94]. They are also called message-passing algorithms because, at each iteration, messages are transmitted from CNs to VNs and vice versa. We distinguish two types of messages:

— Intrinsic or *a priori* messages are computed from the channel observations. They are called intrinsic because the information they contain only comes from the channel. At the initialization stage, these messages are directly sent to all the CNs.
— Extrinsic messages are computed from messages coming from other branches of the graph. Outgoing extrinsic messages from a VN are computed from an intrinsic message and extrinsic messages from the connected CNs. Outgoing extrinsic messages from a CN are computed from incoming extrinsic messages (from the connected VNs) and with the local parity constraint.

The decoder should be able to converge on a valid codeword after a finite number of iterations. In practice, the decoding algorithm can be stopped according to two criteria. The simplest is to set the number of iterations independently of the convergence of the decoder. The second criterion, which permits to reduce the latency of the decoder, consists in stopping the decoding as soon as it converges to a valid codeword (an estimated codeword $\hat{C}$ is valid if it satisfies the syndrome $\hat{\boldsymbol{C}}.\boldsymbol{H}^T = 0$). However, to avoid an infinite execution in case the decoder fails to converge to a valid codeword, a maximum number of iterations is fixed.

In the BP algorithm, the exchanged messages are *a posteriori* probabilities calculated on the symbols of the codeword. However, the BP algorithm [95] suffers from a prohibitive computational complexity, dominated by $O(q^2)$, which mainly comes from the calculations carried out during the update of the parity constraints.

Barnault *et al.* proposed in [96] the FFT-BP algorithm in which the updates of the parity constraints are made in the frequency domain. This transforms the convolution products into simple multiplications. Thus, additional operations of Fourier transform, direct and inverse, are added between the VNs and the CNs to ensure the transition from the probability domain to the frequency domain, and vice versa. Although the complexity of the FFT-BP algorithm is considerably reduced to the order of $O(q\log(q))$, a large number of multiplications remains necessary to perform the update of the nodes in the graph.

The log-BP algorithm [97] performs the four decoding steps in the logarithmic domain to allow a hardware layout less sensitive to quantization errors, and therefore better suited to fixed-point arithmetic. However, the update of the CNs always requires a large amount

of calculation and the complexity of the decoder remains dominated by $O(q^2)$. A direct combination of the FFT-BP and log-BP algorithms is not advantageous because the calculation of the Fourier transform is very complex in the logarithmic domain.

To simultaneously benefit from the advantages of the FFT-BP and log-BP algorithms, Song et al. proposed in [98] the log-BP-FFT algorithm. In this algorithm, the VNs are processed in the logarithmic domain. The extrinsic messages of the VNs undergo a double transformation to pass from the logarithmic domain to the probabilities domain and from the probabilities domain towards the frequency domain in which the CNs are processed. The extrinsic messages of the CNs in turn undergo a double transformation to return to the logarithmic domain of the VNs. However, the log-BP-FFT algorithm requires look-up tables to ensure the conversion between the probabilities domain and the logarithmic domain. These tables have the disadvantage of consuming a lot of memory resources, a consumption that increases with the degree of parallelism of the decoder.

The BP, FFT-BP, log-BP and log-BP-FFT algorithms are optimal decoding algorithms because they do not use any mathematical approximation to reduce the complexity of the decoding. The BP algorithm and its variants guarantee optimal decoding performance but they are not of great interest for hardware implementation. Therefore, other algorithms based on approximations of the BP algorithm are proposed to ensure a reasonable performance/complexity trade-off. We cite mainly the algorithm Min-Sum [97] and its variant EMS (Extended Min-Sum) [99, 100]. A detailed comparison of the optimal and suboptimal algorithms cited above can be found in [101]. Besides the EMS algorithm, there is also the Min-Max algorithm [102] which can be considered as an approximation of the Min-Sum algorithm, and which consequently provides poorer performance.
In the sequel, decoding performance is given for the EMS algorithm with floating-point representation

3.2 CCSK modulation

CCSK is a Direct Sequence Spread Spectrum (DSSS) modulation technique. A spread spectrum modulation is a technique that uses a wider transmission spectrum than the spectrum of the original signal. The use of these techniques has its roots in military applications that seek to hide the transmitted signal to reduce the probability of intercep-

tion. The spread spectrum modulation is also widespread in civil wireless applications. This is due to their immunity to interference and the possibility of multiple access communications links using different spread sequences [11][12]. The spread spectrum modulation can be classified into two categories:

— The Frequency Hopping Spread Spectrum modulations (FHSS): In this case, the signal is switched by repetition of several sub-carriers generated by a frequency synthesizer controlled by a generator of pseudo-random sequences. A receiver not knowing the frequencies of the sequences cannot intercept communication.

— DSSS: In this case, each data bit is transmitted as a pseudo-random sequence of n chips such as $T_b = nT_c$, where T_b is the duration of one bit and T_c is the duration of a chip. Therefore, the transmission band is wider than the data band since $T_c < T_b$.

CCSK modulation [11] is an alternative to the orthogonal modulation which allows simplifying treatments using 2^p sequences. It is mainly constructed from a single pseudo-random sequence called the fundamental sequence. Each sequence is obtained with one or more cyclic shifts of the root sequence.

Throughout this section, denote by $\mathbb{Z}_q \stackrel{\Delta}{=} \{0, 1, \ldots, q-1\}$ the set of integers comprised between 0 and $q-1$, where $q = 2^p$ is a power of 2. Also, identify $\mathbb{Z}_q \cong \mathbb{Z}_2^p \stackrel{\Delta}{=} \{0,1\}^p$, by identifying an integer to its binary representation, $k \in \mathbb{Z}_q \cong (k(0), \ldots, k(p-1)) \in \mathbb{Z}_2^p$.

The CCSK modulation uses a PN sequence

$$\boldsymbol{P}_0 = (P_0(0), P_0(1), \ldots, P_0(q-1)), \tag{3.6}$$

which is written in short as $\boldsymbol{P}_0 = (P_0(i))_{i=0,\ldots,q-1}$, where $P_0(i) \in \{-1, +1\}, \forall i = 0, \ldots q-1$.

An example of a mapping is considered in Table 3.3 over $\mathbb{Z}_8$. The CCSK modulation is constructed from a basic sequence of length 8 with $\boldsymbol{P}_0 = \{+1+1+1-1+1-1-1-1\}$. Then CCSK modulation is applied on each of the c_k encoded symbols, such that $\boldsymbol{P}_{c_k}$ is the circularly right shifted sequence of $\boldsymbol{P}_0$ by c_k positions which corresponds to the GF(q) symbol c_k, i.e. the element of $\mathbb{Z}_q$.

To formalize, the CCSK modulation maps an element $s \in \mathbb{Z}_q$ to the sequence $\boldsymbol{P}_s$, defined as the circular right shift of $\boldsymbol{P}_0$ by s positions, that

$$\boldsymbol{P}_s = (\boldsymbol{P}_0(i - s \bmod q))_{i=0,1,\ldots,q-1}. \tag{3.7}$$

It should be noted here that a symmetrical mapping can be defined using circular left shift. In this case "−" will be replaced in (3.7) by "+". Throughout this report, right

Table 3.3 – CCSK codes of GF(8).

$c_k \in \mathbb{Z}_8$	CCSK sequence $\boldsymbol{P}_{c_k}$
0	$\mathbf{+1}+1+1-1+1-1-1-1$
1	$-1+\mathbf{1}+1+1-1+1-1-1$
2	$-1-1+\mathbf{1}+1+1-1+1-1$
3	$-1-1-1+\mathbf{1}+1+1-1+1$
4	$+1-1-1-1+\mathbf{1}+1+1-1$
5	$-1+1-1-1-1+\mathbf{1}+1+1$
6	$+1-1+1-1-1-1+\mathbf{1}+1$
7	$+1+1-1+1-1-1-1+\mathbf{1}$

circular shift are used, and integers p and q is referred to the dimension and length of the CCSK modulation respectively.

3.3 Generation of initial CCSK sequences

The task of this section is linked to the definition of a spreading sequence that maximizes the probability of good detection and synchronization. The impact of different types of spreading sequences has a direct effect on the overall performance system. In the following, two methods are presented to generate the basic sequences. The first method is the Linear Feedback Shift Register (LFSR) and the second method is the genetic algorithm [103]. Based on the output simulation tests, the genetic algorithm has better max-to-side-lobe correlation performance than the LFSR. Consequently, The genetic algorithm is used to generate the basic sequences for the CCSK modulation in the sequel.

3.3.1 Linear Feedback Shift Register

A finite-length pseudo-random sequence $\boldsymbol{P}_0$ can be generated by using a LFSR [104], as illustrated in Fig. 3.2, where

— $g_1, g_2, \ldots, g_{p-1}$ are the coefficients of the *feedback polynomial*

$$g(x) = \sum_{i=0}^{p-1} g_i x^i, \quad \text{with } g_0 = g_{p-1} = 1. \tag{3.8}$$

— $s_{p-1}, \ldots, s_1, s_0$ is the *state* of the shift register, initialized as $(0, \ldots, 0, 1)$.

— $\bigoplus$ operations represent the exclusive or (XOR) gates.

— BPSK denotes the binary phase shift keying modulation of the binary output sequence ($0 \longrightarrow +1, 1 \longrightarrow -1$)

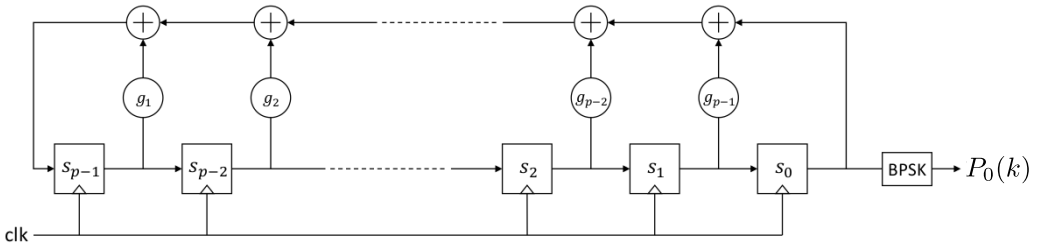


Figure 3.2 – LFSR sequence (Fibonacci representation).

The LFSR is maximum-length, i.e. its state cycles through all possible $q-1$ states (except the all-zero state), if and only if the feedback polynomial is primitive. In this case, the generated sequence is referred to as an *m-sequence*. An *m-sequence* is periodic, of period $q-1$, that is, $P(i) = P(i+q-1), \forall i \geq 0$. In particular, $P(0) = P(q-1)$. Moreover, the cyclic cross-correlation function takes on only two values, as follows

$$\theta(k) \stackrel{\Delta}{=} \sum_{i=0}^{q-2} P(i)P(k+i) = \begin{cases} q-1, & \text{if} k = 0 \\ -1, & \text{if} k = 1, \ldots, q-2 \end{cases} \tag{3.9}$$

Going back to the CCSK modulation, we take the PN sequence $\boldsymbol{P}_0$ form (3.6) to be defined as

$$\boldsymbol{P}_0 = (P(0), P(1), \ldots, P(q-2), P(q-1)), \tag{3.10}$$

where $\boldsymbol{P}$ is an *m-sequence*. Since the length of $\boldsymbol{P}_0$ is q, which is q greater than the period of $\boldsymbol{P}$, the cross-correlation function of $\boldsymbol{P}_0$ is different from the cross correlation function of $\boldsymbol{P}$. Thus, defining

$$\theta_0(k) \stackrel{\Delta}{=} \sum_{i=0}^{q-1} \boldsymbol{P}_0(i)\boldsymbol{P}_0(k+i \bmod q), \quad \forall k = 0, \ldots, q-1, \tag{3.11}$$

we have $\theta_0(0) = q$, but for $k \neq 0$, the absolute value $|\theta_0(k)|$ may exceed 1 (although its value is still small with respect to q).

The feedback polynomials used throughout this report for various p values, together with the minimum, maximum and average $|\theta_0(k)|$ value, for $k \neq 0$, are given in Table 3.4.

Table 3.4 – Feedback polynomials for various p values.

p	Feedback polynomial $g(x)$	min $\|\theta_0(k)\|$	max $\|\theta_0(k)\|$	ave $\|\theta_0(k)\|$
$p = 6$	$g(x) = x^6 + x^5 + x^4 + x + 1$	0	12	2.48
$p = 7$	$g(x) = x^7 + x^3 + 1$	0	20	4.94
$p = 8$	$g(x) = x^8 + x^4 + x^3 + x^2 + 1$	0	25	6.95
$p = 9$	$g(x) = x^9 + x^4 + 1$	0	32	10.07
$p = 10$	$g(x) = x^{10} + x^3 + 1$	0	56	13.32
$p = 11$	$g(x) = x^{11} + x^2 + 1$	0	76	19.85
$p = 12$	$g(x) = x^{12} + x^9 + x^3 + x^2 + 1$	0	132	28.16

3.3.2 Genetic algorithm

A genetic algorithm is a type of metaheuristic inspired by an evolution that is known to generate a good solution for a complex optimization problem [105]. In the frame of the QCSP project, a Matlab implementation of the GA algorithm has been developed to generate the sequences $\boldsymbol{P}_0$ of size $q = 64$ up to $q = 4048$. The explanation of the genetic algorithm is explicit in the Matlab code and its comments which is given in [103]. Note that many other parameters can be used and that several attempts can be done to improve the result. The result found for the $\boldsymbol{P}_0$ sequence are given for $q = 64$ up to $q = 2048$ are also given in Appendix 8.2. Three types of norms can be used to measure the correlation distance: $\boldsymbol{L}_1$, $\boldsymbol{L}_2$, and $\boldsymbol{L}_\infty$ respectively defined as

$$\begin{aligned} L_1(\theta) &= \frac{1}{q-1}\sum_{k=1}^{q-1}|\theta(k)| \\ L_2(\theta) &= \frac{1}{q-1}\sum_{k=1}^{q-1}\theta(k)^2 \\ L_\infty(\theta) &= max\{|\theta(k)|, k = 1, \ldots q-1\}. \end{aligned} \tag{3.12}$$

Table 3.5 gives the metric of 3 types of sequences: the one obtained by LFSR: (PN_L), the one obtained by genetic algorithm with the balanced number of -1 and +1 (PN_B), and the one obtained with GA algorithm without the balanced constraint (PN_U). Based on the output results, the proposition is to use the sequence PN_B for all CCSK simulations.

Table 3.5 – Correlation distance output for basic sequences of different size q.

	L_1			L_2			L_∞		
q	PN_L	PN_B	PN_U	PN_L	PN_B	PN_U	PN_L	PN_B	PN_U
64	2.44	1.25	1.50	16.75	6.00	6.00	12	8	4
128	4.91	3.00	2.62	44.12	15.50	15.75	20	8	8
256	6.92	4.81	4.33	82.93	38.87	32.19	24	20	12
512	10.05	6.84	6.37	181.16	78.56	67.50	32	28	20
1024	13.31	9.37	9.73	312.58	142.66	148.97	56	48	32
2048	19.84	13.60	13.43	701.26	295.45	287.56	76	56	68

3.4 System model

This section describes in detail the overall communication principle of the QCSP system model. The principle of a CCSK modulation is first presented in the context of its association with NB-codes. Then, the Over-Modulation (OM) scheme that is added at the transmitter side is described. Also, the demodulation of a QCSP frame and the procedure of LLR generation is illustrated. Finally, the effect of the channel at the receiver side is presented when neither time nor frequency information is available. Fig. 3.3 shows a simplified description of the overall communication link. The different blocks of this chain at both the transmitter and receiver sides are discussed in the following sub-sections.

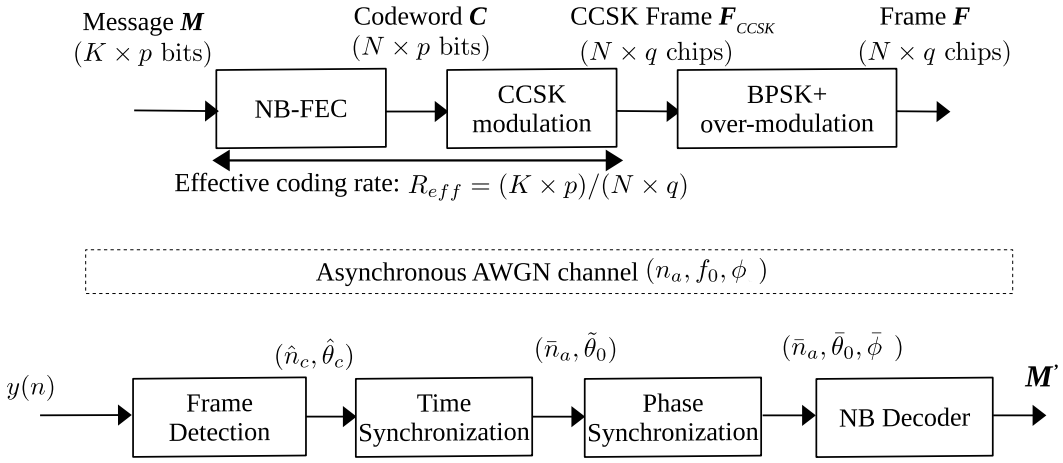


Figure 3.3 – Overall communication principle, with $\theta_i = 2\pi f_i q$.

3.4.1 Transmitter side

Consider a NB code defined over GF(q) of q elements, $q = 2^p$, with K symbols of information and a total length N symbols. The code rate is thus $R_c = K/N$ and the input of the NB-code is a binary message $\boldsymbol{M}$ of size $m = K \times p$ information bits, equivalently K GF(q) symbols. The encoder (NB-LDPC code is considered) generates a codeword $\boldsymbol{C}$ of N GF(q) symbols

$$\boldsymbol{C} = [c_0, c_1, \ldots, c_{N-1}], \text{ with } c_k \in \mathrm{GF}(q), \ k = 0, 1, \ldots, N-1. \tag{3.13}$$

Since $\boldsymbol{H}$ is the PCM associated to the NB-LDPC FEC code, then, as defined in section 3.1.2, the codeword $\boldsymbol{C}$ verifies $\boldsymbol{C} \times \boldsymbol{H}^T = 0$.

For the goal of direct sequence spread spectrum technique, the CCSK modulation uses a pseudo-random binary sequence $\boldsymbol{P}_0 = \{P_0(i)\}_{i=0,1,\ldots,q-1}$ as defined in section 3.3 of length q, where $P_0(i) \in \{-1, 1\}$, with good auto-correlation properties. The CCSK modulation maps an element c_k of the encoded GF(q) symbols to the sequence $\boldsymbol{P}_{c_k}$ defined as the circular right shift of $\boldsymbol{P}_0$ by c_k positions

$$\boldsymbol{P}_{c_k} = \{P_0(i - c_k \bmod q)\}_{i=0,1,\ldots,q-1}. \tag{3.14}$$

With this convention, the CCSK consists of mapping each c_k to a circular c_k-right-shift of $\boldsymbol{P}_0$ such that

$$\forall c_k \in \boldsymbol{C} : \mathrm{CCSK}(c_k) = \boldsymbol{P}_{c_k}. \tag{3.15}$$

The link between the element c_k of GF(q) and the integer shift value c_k is done by considering the element of GF(q) as polynomial over $\mathrm{GF}_{(2)}[X]/\mathbb{P}_p[X]$ as in (3.1). The binary representation of this polynomial gives the binary representation of the integer value c_k. The CCSK modulation rate can be defined as $R_m = p/q$, and the overall effective coding rate R_{eff} is given by $R_{eff} = R_c \times R_m = \frac{K}{N} \times \frac{p}{q}$. Since BPSK modulation is used, the effective spectral efficiency S_e is R_{eff} bits per channel use.

The CCSK frame $\boldsymbol{F}_{CCSK}$ is thus defined as the concatenation of N CCSK symbols, i.e. $\boldsymbol{F}_{CCSK} = [\boldsymbol{P}_{c_0}, \boldsymbol{P}_{c_1}, \ldots, \boldsymbol{P}_{c_{N-1}}] = \coprod_{k=0}^{N-1} \boldsymbol{P}_{c_k}$, where $\coprod$ represents the concatenation operation.

Since a noisy environment is targeted, a symbol OM as discussed in the following section is added to the CCSK symbols in-order to help the synchronization process. Before transmission, the generated frame $\boldsymbol{F}$ is composed of $N \times q$ chips and obtained after the Binary Phase Shift Keying (BPSK). Finally, it is shaped by a half-raised cosine filter with

a roll-off factor equal to 0.35.

3.4.2 Over-Modulation (OM)

Over modulation is proposed in order to enhance the synchronization process in a very noisy environment. In fact, at very low SNRs, determining when the sequence starts at the exact symbol level is not a trivial task. To help this task, an additional modulation (called Over-Modulation) is used at symbol level to generate a known pattern of phase shift. This known pattern of phase shift is helping the time synchronization task at the receiver side. Thus, instead of transmitting $\boldsymbol{F}_{CCSK}$ as defined before, we send the over-modulated QCSP frame $\boldsymbol{F}$, defined as (see Fig. 3.4)

$$\begin{aligned} \boldsymbol{F} &= [(-1)^{b_0}\boldsymbol{P}_{c_0}, (-1)^{b_1}\boldsymbol{P}_{c_1}, \ldots, (-1)^{b_{N-1}}\boldsymbol{P}_{c_{N-1}}] \\ &= \coprod_{k=0}^{N-1} (-1)^{b_k}\boldsymbol{P}_{c_k}, \end{aligned} \tag{3.16}$$

where $\boldsymbol{B} = [b_0, b_1, ..., b_{N-1}]$ with $b_k \in \{0, 1\}$ is a sequence with good auto-correlation properties. The OM keeps the phase of the i^{th} symbol unaffected when $b_i = 0$, and shifts it by π radian when $b_i = 1$. The over-modulation sequences for size $N = 60$ and $N = 120$ are given in Appendix 8.3.

Although not studied in the Ph.D., it is also possible (and maybe interesting) to use complex over-modulation sequences. For example, a CAZAC sequence ("Constant Amplitude Zero Auto-correlation") is like a Zadoff-Chu sequence or a Barker sequence.

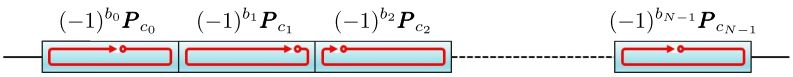


Figure 3.4 – Symbolic representation of a QCSP Frame.

3.4.3 Channel model: time, frequency, phase offsets and AWGN noise

In this work, we assume a low-cost sensor that sporadically transmits/receives small messages in an unslotted asynchronous ALOHA protocol, i.e. without prior knowledge of the time of arrival and the potential carrier frequency offset of the signal.

Let T_c and $T = q \times T_c$ (in seconds) be the duration of a chip and a CCSK symbol respectively. Half raised cosine filter is applied at the receiver side. The frequency offset is assumed to be small enough to guaranty no interference between chips. The receiver over-samples the incoming signal with O samples per chip. In other words, the clock frequency F_e of the receiver Analog to Digital Converter (ADC) is equal to $F_e = O/T_c$, with O the over-sampling factor (typically between 4 and 8). Indexing the time by duration T_c of a chip (i.e. O clock cycles), it is possible to determine the time of arrival t_a as a real $x_a = t_a/T_c$ and by decomposing x_a as

$$x_a = n_a + r_a/O + \epsilon_a, \tag{3.17}$$

where $n_a = \lfloor x_a \rfloor$, the integer part of x_a representing the time in number of chips, r_a the closest index of the clock cycle within a chip ($r_a \in \{0, 1, \ldots, O-1\}$) and ϵ_a is the residual timing synchronization error (with $\epsilon_a \in [-\frac{1}{2O}, \frac{1}{2O}]$).

In the sequel, we consider that the oversampling factor is high enough so ϵ_a is negligible and can be considered equal to 0. Moreover, we also assume that by testing in parallel all the O hypotheses of the r_a value and by keeping the best one, we can always manage to set r_a equal to 0. This is tested through MC simulations, and validated by the real-data radio reception in section 6.3.4.

Carrier frequency errors are also considered, leading to a frequency offset F_o affecting the received frame. In T_c seconds, the frequency offset generates a rotation $\frac{T_c F_o}{2\pi}$ radians between two consecutive chips. In the sequel, a normalized frequency offset $f_o = F_o T_c$ is used. The impact of f_o is to generate a rotation $\theta_o = 2\pi f_o q$ radians between two chips separated by a symbol duration. Finally, the initial phase offset ϕ is unknown too where $\phi \in [0,\ 2\pi]$. It is also assumed that there is enough time between each message to ensure no interference. In summary, the frame is received at chip index n_a can be defined as

$$\begin{aligned} y(n) &= e^{j(n\frac{\theta_o}{q}+\phi)}\boldsymbol{F}(n-n_a) + z(n), \text{ if } n \in [\![n_a, n_a + Nq - 1]\!] \\ &= z(n), \text{ otherwise.} \end{aligned} \tag{3.18}$$

Without any prior information, θ and ϕ are supposed to be uniformly distributed in their respective interval ranges. The $z(n)$ are realisations of a complex AWGN $\mathcal{N}(0, \sigma)$, with zero mean and standard deviation $\sigma = \sqrt{10^{-\mathrm{SNR}/10}}$. The whole received frame is denoted $\boldsymbol{Y}$.

3.4.4 Demodulation of a QCSP frame

The AWGN channel adds to the signal a complex white noise of power σ^2. The SNR of the channel is defined at the chip level as $\frac{1}{\sigma^2}$. In a full synchronous channel (i.e. assumed to be synchronized by the correct estimation of the offset parameters), the received frame $\boldsymbol{Y}$ can be represented as the sum of the frame $\boldsymbol{F}$ and a vector of AWGN $\boldsymbol{Z}$.

Each received block message (CCSK symbol) $\boldsymbol{y}_n$ is of length q and starts at time n, can be defined as $\boldsymbol{y}_n = (y(n), y(n+1), \ldots, y(n+q-1))$. This is demodulated by correlating $\boldsymbol{y}_n$ with each of the q possible shifted sequences $\boldsymbol{P}_s$, $s = 0, 1, \ldots, q-1$ to generate the vector $\boldsymbol{L}_n = (L_n(0), L_n(1), \ldots, L_n(q-1))$ where the s components are defined as

$$L_n(s) \cong \left(\sum_{i=0}^{q-1} y_n(i) P_s^*(i) \right) = \left(\sum_{i=0}^{q-1} y_n(i) P_0^*(i - s \bmod q) \right), \tag{3.19}$$

for $s = 0, \ldots, q-1$, where $P_s^*(i)$ is the conjugate of $P_s(i)$, i.e. if $\boldsymbol{X}^*$ represents the conjugate vector of $\boldsymbol{X}$ then $\boldsymbol{X}^* = \mathrm{Re}(\boldsymbol{X}) - j\mathrm{Im}(\boldsymbol{X})$ [1]. Since $P_s(i) = P_0(i - s \mod q)$ by construction, the log-likelihood vector $\boldsymbol{L}_n$ is the circular correlation between the received block message $\boldsymbol{y}_n$ of length q and the spreading sequence $\boldsymbol{P}_0$, i.e.

$$\boldsymbol{L}_n = \boldsymbol{y}_n \star \boldsymbol{P}_0^*, \tag{3.20}$$

where $\star$ denotes the circular correlation. This computation can be efficiently computed in the frequency domain [89],

$$\boldsymbol{L}_n = \mathrm{IFFT}(\mathrm{FFT}(\boldsymbol{y}_n) \odot \mathrm{FFT}^*(\boldsymbol{P}_0)), \tag{3.21}$$

where operator $\odot$ denotes the element-wise (or Hadamard) product of two vectors.

From $\boldsymbol{L}_n$, the Log-Likelihood Ratio (LLR) related to $\boldsymbol{y}_n$ is computed as $\mathrm{Re}(\frac{2\boldsymbol{L}_n}{\sigma^2})$ [89] and the full frame LLR vector is sent to the NB decoder which retrieve the message $\boldsymbol{M'}$.

The channel model defined in section 3.4.3 is asynchronous, with time, frequency, and phase offsets. Thus, before performing the aforementioned process, it is required to detect the presence of a new frame at the received side (frame detection block in Fig. 3.3). Once a frame is detected, it is also required to estimate precisely its time of arrival, its frequency offset and its initial phase to compensate them before performing demodulation (time and

1. In case of a BPSK modulation, there is no difference, but a CCSK modulation can also be considered using a CAZAC sequence where complex values exist.

phase synchronization blocks in Fig. 3.3). Finally, once LLR is generated, the NB-LDPC decoder is used to correct the potential residual error (NB-decoder block in Fig. 3.3).

3.5 Conclusion

In this chapter, the definition of FEC techniques is recalled at the beginning. The theory of finite fields, and the LDPC codes in their binary and non-binary versions are mentioned. The main algorithms used in the decoding of NB-LDPC codes are also discussed. Then, the DSSS and CCSK techniques have been described. Afterward, the system model of the NB-LDPC and CCSK chain has been shown, where the format of the CCSK frame has been defined. Finally, the channel model, as well as the time of arrival formulation, have been described, along with a quick overview of the QCSP frame demodulation.
It is worth mentioning here that the CCSK and NB-LDPC are currently used in space. The recent Quasi-Zenith Satellite System (Japaneses GPS enhancement system) is using a CCSK modulation [106, 107]. Moreover, DeiDou (Chinese GPS-like system) has recently incorporate a NB-LDPC code to protect the data [108]. The use of these two techniques in recent space application show the potential of their association for IoT system on earth like the proposed QCSP frame.

The overall processes (detection, synchronization and decoding) are addressed in detail in the next chapters. This is discussed while showing the effect of each of the QCSP parameters and channel effect on each process separately, and jointly on the overall performance.

Chapter 4

Detection

Contents

This chapter discusses in detail the detection method that is processed to acquire the QCSP frame. It first gives a general overview of the detection problem. Then, it describes the score function used to assess if a new frame is arrived or not, and presents the concept of time and frequency research grid. Then, a formal performance model of the detection algorithm is described. The theoretical model is validated through a comparative study with experimental results obtained with MC simulations over a complex AWGN channel. Finally, the effect of different parameters that affect the CCSK-based system is discussed.

4.1 Detection problem

The detection problem being considered in this study is to decide, based on the observation of $N \times q$ received samples of $\boldsymbol{Y}$ in equation (3.18), if a frame is detected or not.

Two hypotheses are achieved: either present ($\mathcal{H}1$) or not ($\mathcal{H}0$).

The task is to develop a reliable score function $S(\boldsymbol{Y})$ that takes high values when $\mathcal{H}1$ is fulfilled, and low values when $\mathcal{H}0$ is true. Then, for a given observation, it is possible to decide by comparing $S(\boldsymbol{Y})$ to a threshold U_0 to decide whether a new frame is present or not. In detection theory, the detector can give one of the following four different cases:

— Miss Detection probability: $\mathcal{P}_{\text{md}} = \mathcal{P}(S(\boldsymbol{Y}) < U_0 |\ \mathcal{H}1)$ is the probability that the system takes an erroneous decision by signaling the absence of any frame while a frame in fact exists.

— Correct detection: $\mathcal{P}(S(\boldsymbol{Y}) \geq U_0 |\ \mathcal{H}1)$ correctly detects an existing frame (the probability of correct detection is equal to $1 - \mathcal{P}_{\text{md}}$).

— False alarm: $\mathcal{P}_{\text{fa}} = \mathcal{P}(S(\boldsymbol{Y}) \geq U_0 |\ \mathcal{H}0)$ takes an erroneous decision by signaling the existence of a frame while the frame in fact does not exist.

— Correct Absence: $\mathcal{P}(S(\boldsymbol{Y}) < U_0 |\ \mathcal{H}0)$ correctly indicates the absence of a frame (the probability of correct absence is equal to $1 - \mathcal{P}_{\text{fa}}$).

Based on this definition, we obtain

$$\mathcal{P}_{\text{fa}} = \int_{U_0}^{+\infty} f_{\mathcal{H}0}(x)dx, \qquad \mathcal{P}_{\text{md}} = \int_{-\infty}^{U_0} f_{\mathcal{H}1}(x)dx, \tag{4.1}$$

where $f_{\mathcal{H}0}$ and $f_{\mathcal{H}1}$ are the probability density functions of the random variable $S(\boldsymbol{Y})$ given that $\mathcal{H}0$ is true, $\mathcal{H}1$ is true, respectively. Note that when only part of a frame is inside the detector filter, the output $S(\boldsymbol{Y})$ may become greater than U_0, triggering potentially early or late detection. Since $S(\boldsymbol{Y})$ is maximized under hypothesis $\mathcal{H}1$, it is natural to consider only this hypothesis in the detection theory study. Note that once detected, a synchronization task estimates the time, frequency, and phase offsets of the received frame (see chapter 5).

Fig. 4.1 illustrates three different threshold values that correspond to various probabilities of false alarm $\mathcal{P}_{\text{fa}} = 10^{-4}, 10^{-6}$ and 10^{-10} versus the output of the correlation filter over a Gaussian channel. It can be clearly inferred from Fig. 4.1 that the threshold value U_0 allows a trade-off between $\mathcal{P}_{\text{fa}}$ and $\mathcal{P}_{\text{md}}$. In fact, in a perfect detector, both should be equal to zero to decide perfectly the presence or not of a new frame. In practice, high value of U_0 decreases $\mathcal{P}_{\text{fa}}$ but increases $\mathcal{P}_{\text{md}}$, while low value of U_0 has the symmetrical effect. For example, at threshold value $U_0 = 1200$ that corresponds to $\mathcal{P}_{\text{fa}} = 10^{-4}$, the probability of miss detection is approximately $\mathcal{P}_{\text{md}} = 10^{-4}$. This value increases to $\mathcal{P}_{\text{md}} = 5 \times 10^{-3}$ for U_0 corresponding to $\mathcal{P}_{\text{fa}} = 10^{-10}$. Thus, the value of U_0 is selected according to the system

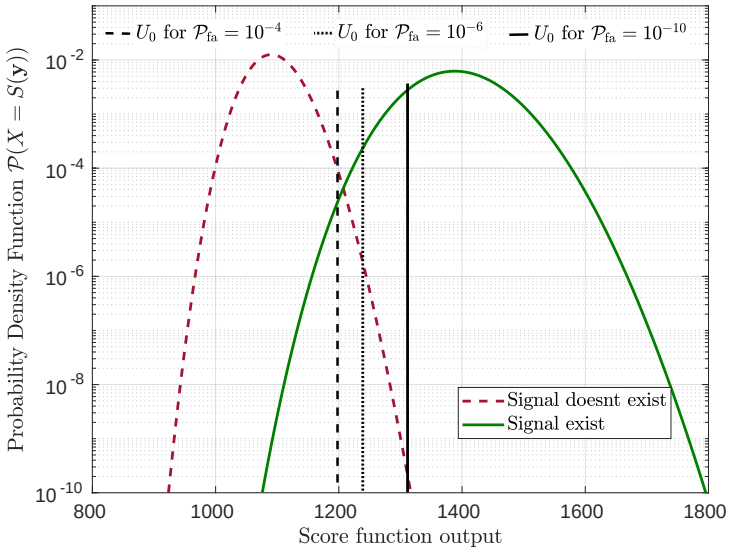


Figure 4.1 – Detection problem illustration.

requirements. In the sequel $\mathcal{P}_{\text{fa}}$ is set to 10^{-6}. We will try to minimize $\mathcal{P}_{\text{md}}$ by proposing an efficient score function, i.e. a score function that is not computationally intensive to be calculated and allows to have low values of $\mathcal{P}_{\text{md}}$. The following sections describe first the proposed score function. Then, the probability density functions $f_{\mathcal{H}0}$ and $f_{\mathcal{H}1}$ are formally derived as a function of the numerous parameters of the problem: number of frame's symbols N, CCSK sequence $\boldsymbol{P}_0$ and its length q, variance of the complex AWGN channel (σ^2), time delay Δ and the frequency offset f_o.

4.2 Score function calculation

This section discusses in detail the proposed score function $S_n(\boldsymbol{Y})$, which is the main metric used in the detection algorithm used to detect the CCSK frames. From the $y(n)$ received samples where $n \in \mathbb{N}$, it is possible to extract the vector $\boldsymbol{Y}_n$ corresponding to the arrival of a frame at time n, i.e.

$$\boldsymbol{Y}_n = \Big(y(n+l)\Big)_{l=0,1,\ldots,N\times q-1} = \coprod_{k=0}^{N-1} \boldsymbol{y}_{n+kq}, \tag{4.2}$$

where $\boldsymbol{y}_x = (y(x), y(x+1), \ldots, y(x+q-1))$.

Using FFT operations as in Eq. (3.21), cross-correlation is performed between each

received block $\boldsymbol{y}_{n+kq}$ for $k = 0, 1..., N-1$ and the reference sequence $\boldsymbol{P}_0$. Without loss of generality, assume that $\Delta \in [0, q/2]$, be the time shift (in number of chips) between the effective time of arrival of the frame and the receiver. Note that the same effect exits when $\Delta \in [-q/2, 0]$.

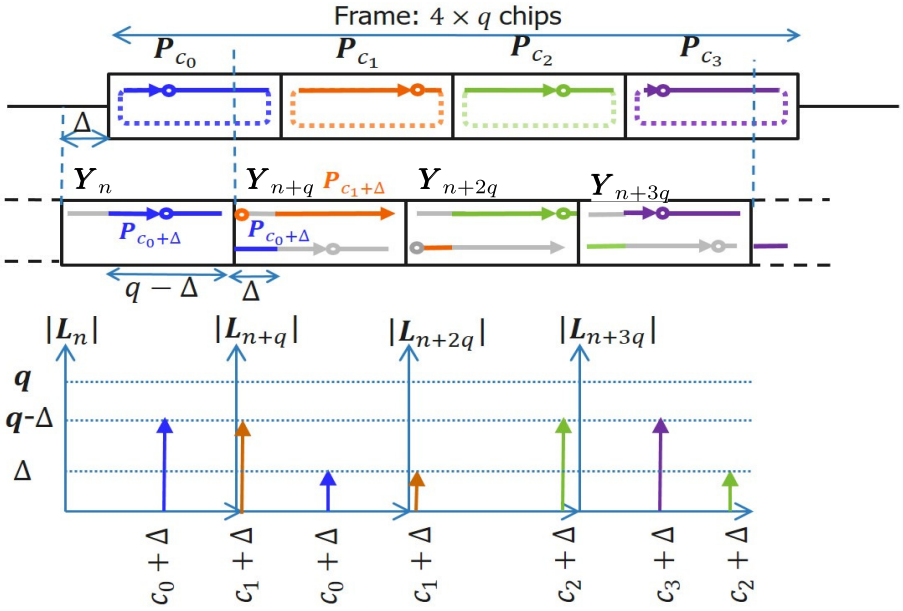


Figure 4.2 – Illustration of the frame detection principle.

The best way to explain the score function characterizing the detection method is by giving an example. Consider a frame containing $N = 4$ sequences, as in Fig. 4.2, each of length q. The symbols (c_0, c_1, c_2, c_3) are associated with the four CCSK sequences $(\boldsymbol{P}_{c_0}, \boldsymbol{P}_{c_1}, \boldsymbol{P}_{c_2}, \boldsymbol{P}_{c_3})$, and a distinct color is associated with each symbol. In vector $\boldsymbol{y}_n$, there are $q - \Delta$ chips that are aligned with the first symbol of the received message of the frame, i.e. $\boldsymbol{P}_{c_0}$ ($\boldsymbol{P}_0$ sequence circularly shifted by c_0 chips). Relatively to $\boldsymbol{y}_n$ and because of the delay Δ, the first Δ chips are null; then, the sequence starts at time $c_0 + \Delta \pmod q$ which is presented at the receiver side as the structure of another sequence $\boldsymbol{P}_{c_0+\Delta}$. Hence, $q - \Delta$ are aligned with the CCSK sequence $\boldsymbol{P}_{c_0+\Delta}$. Thus, the correlation vector $\boldsymbol{L}_{n+kq}$ related to vector $\boldsymbol{y}_{n+kq}$ gives for $k = 0$, $\boldsymbol{L}_n$ that has a spike of height $q - \Delta$ at index $c_0 + \Delta \pmod q$. Similarly, for $k = 1$, the vector $\boldsymbol{y}_{n+q}$ has Δ chips that are aligned with the first symbol $\boldsymbol{P}_{c_0}$ with an offset of $c_0+\Delta$ chips (the sequence $\boldsymbol{P}_{c_0+\Delta}$). Thus, corresponding to Δ chips of the sequence $\boldsymbol{P}_{c_0+\Delta}$, the correlation vector $\boldsymbol{L}_{n+q}$ has a spike of height Δ at index $c_0 + \Delta \pmod q$. Moreover, $\boldsymbol{y}_{n+q}$ contains $q - \Delta$ chips aligned with the second symbol of the

received message, which gives a spike of height $q - \Delta$ for $\boldsymbol{L}_{n+q}$ in position $c_1 + \Delta \pmod q$ (which is the correlation with the sequence $\boldsymbol{P}_{c_1+\Delta}$ and so on).

Hence, the received block $\boldsymbol{y}_{n+kq}$ has $q-\Delta$ chips of correlation with the CCSK sequences $\boldsymbol{P}_{c_k+\Delta}$ and Δ chips with the other sequence $\boldsymbol{P}_{c_{k-1}+\Delta}$. $\boldsymbol{y}_n$ is a special case as it only has $q-\Delta$ correlation value with the CCSK sequence $\boldsymbol{P}_{c_0+\Delta}$.

Thus, the score function can be obtained using a detection filter $S_n(\boldsymbol{Y})$ (corresponds to $\boldsymbol{Y}_n$) of length N acting as forward accumulator

$$S_n(\boldsymbol{Y}) = \sum_{k=0}^{N-1} M_{n+kq}, \tag{4.3}$$

where

$$M_{n+kq} = \max\{|L_{n+kq}(i)|, i = 0, 1, \ldots q-1\}, \tag{4.4}$$

and $\boldsymbol{L}_{n+kq}$ is the correlation vector between the received block $\boldsymbol{y}_{n+kq}$ and the q CCSK symbols, that is calculated based on (3.21).

In the absence of noise with optimized $\boldsymbol{P}_0$, i.e. the auto-correlation properties are $\langle \boldsymbol{P}_s, \boldsymbol{P}_{s'} \rangle \ll q$ for $s \neq s'$, the filter output gives $S_n(\boldsymbol{Y}) = N \times (q - \Delta)$. In order to draw benefits from the second maximum shown in Fig. 4.2, it is possible to add two consecutive correlation vectors before taking its maximum (SC method, for Sum of Correlation). The score function becomes

$$S_n^{(\mathrm{SC})}(\boldsymbol{Y}) = \sum_{k=0}^{N-2} \max(|\boldsymbol{L}_{n+kq} + \boldsymbol{L}_{n+kq+1}|). \tag{4.5}$$

This method is not examined in the report, but it is worth mentioning that, compared to the score function $S_n(\boldsymbol{Y})$, $S_n^{(\mathrm{SC})}(\boldsymbol{Y})$ gives a slight improvement of detection capacity when Δ is closed to $q/2$, and gives a few dB penalty when Δ is equal to 0. It is also more sensitive to a frequency offset since the duration of coherent integration is multiplied by 2.

Fig. 4.3 gives an example of the calculation of $\boldsymbol{L}_n$ for a data stream $\mathbf{y}$ composed first of 326 random bits (represents the noise), then a QCSP frame composed of $N = 4$ symbols each of length $q = 64$ chips, starting at chip index $n_a = 327$ and finally, an additional 320 random bits. The continuous gray curve is the output of the maximum of the correlation vectors $\max(|\boldsymbol{L}_n|)$ at each chip index n. The blue lines represent the output of correlation function $\max(|\boldsymbol{L}_{64k}|)$ for each $n = kq$, i.e. $\max(|\boldsymbol{L}_{64k}|)$, and the orange lines

for $\max(|\boldsymbol{L}_{64k+32}|)$, $k \in \mathbb{N}$. Correlation outputs are effectively computed using (3.21). The

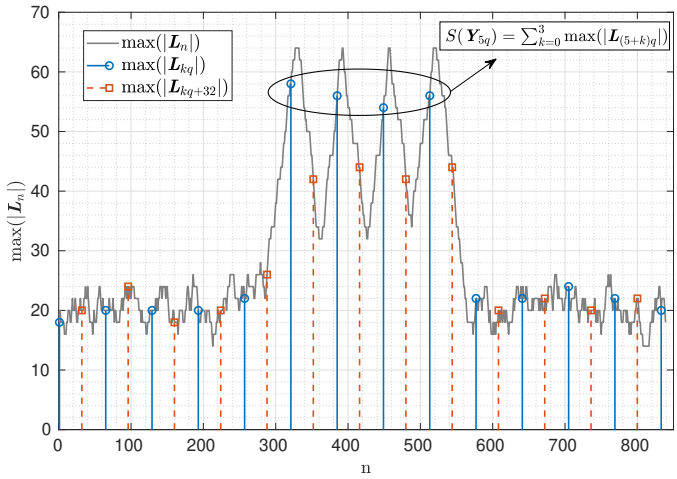


Figure 4.3 – Example of the score function calculation for a CCSK frame of size $N = 4$, $q = 64$ starting at time $n_a = 327 = 5q + \Delta$, where $\Delta = 7$. Random bits are assumed to be transmitted before and after the CCSK frame.

values used to compute the score function $S_{n=5\times64}(\boldsymbol{Y})$ which correspond to the minimum time offset error $\Delta = n_0 - n = 7$ are also shown. It is worth noticing that the time sliding windows method proposed in [19] computes the score function $S_n(\mathbf{Y})$ for every value of n, thus giving an optimal time bin size of size $\ell = 1$. This method has a complexity of the order of $O(q)$ per chip. It is thus applicable for small values of q ($q \leq 128$, typically). For lower complexity, and good performance, we have proposed time-frequency space decomposition as discussed in the next section to trade off the performance-complexity challenge.

To conclude, in the presence of AWGN noise, the detector compares $S_{n=kq}(\boldsymbol{Y})$ to a threshold U_0 to assess, or not, the arrival of a frame. The following section evaluates the detection performance of the QCSP frame in an asynchronous AWGN channel.

4.3 Time and Frequency decomposition

The blind detection algorithm splits the time and frequency domain into a regular grid composed of bins. Each bin defined by a time span $\delta_n = \ell T_c$ where ℓ is the number of chips inside the duration δ_n, and a frequency span of size δ_θ. For the sake of notation clarity and simplicity, we assume in the sequel that the chip period is normalized, $T_c = 1$.

For the aim of explaining the detection process, assume the following example. Without loss of generality, frequency offset θ is assumed to be bounded between $-\pi$ and π, i.e. giving at maximum a half clockwise or counterclockwise rotation per CCSK symbol. Then, it is possible to divide this frequency interval into p_ω sub-intervals each of size $\delta_\theta = \frac{2\pi}{p_\omega}$, and associated to a score filter $S_n^{\omega(r)}$ with $\omega(r) = \pi(-1+\frac{2r+1}{p_\omega})$, $r = 0, 1, \ldots, p_\omega - 1$. Therefore, the maximum distance between θ and the closest $\omega(r)$ value is bounded by a maximum error $e_m = \frac{\delta_\theta}{2}$ or $e_m = \frac{\pi}{p_\omega}$ radian. For example, when $p_\omega = 4$, $e_m = \frac{\pi}{4}$ which corresponds to $1/8^{th}$ of residual rotation per CCSK symbol. For the time uncertainty, a similar approach is used to limit the computational resource: the CCSK size q is divided into p_Δ sections of length $\ell = \frac{q}{p_\Delta}$. Every ℓ chips ($\ell \leq q$ typically), the last $N \times q$ received chips are extracted to form the vector $\boldsymbol{Y}_{\gamma\ell} = (y(\gamma\ell + i))_{i=0,1,\ldots,N\times q-1}$ (with γ is an index of time corresponds to ℓ). Then, at the entry of the $\omega(r)$ frequency detector, $\boldsymbol{Y}_{\gamma\ell}^{\omega(r)} = \boldsymbol{Y}_{\gamma\ell} \odot \boldsymbol{E}_{\omega(r)}$, where $\boldsymbol{E}_{\omega(r)} = (e^{-j\omega(r)i/q})_{i=0,1,\ldots,N\times q-1}$ is computed in order to compensate the frequency offset before entering the detector. In summary, for every CCSK symbol (thus q received chips), $p_\omega p_\Delta$ circular correlations (3.19) are computed and the $p_\omega p_\Delta$ score functions are updated.

Thus, each bin (γ, r) corresponds to an arrival hypothesis $\mathcal{H}(\gamma\ell, \omega(r))$ of the frame, i.e. can be either $\mathcal{H}0$ or $\mathcal{H}1$, with a coarse time and frequency precision $(\gamma\ell,\ \omega(r))$. Each time, a value of $S_{\gamma\ell}^{\omega(r)}$ is calculated and compared to the threshold value U_0, to assess (hypothesis $\mathcal{H}1$) or not (hypothesis $\mathcal{H}0$) the arrival of a frame within the bin $(\gamma\ell,\ \omega(r))$. The whole detection system is shown in Fig. 4.4.

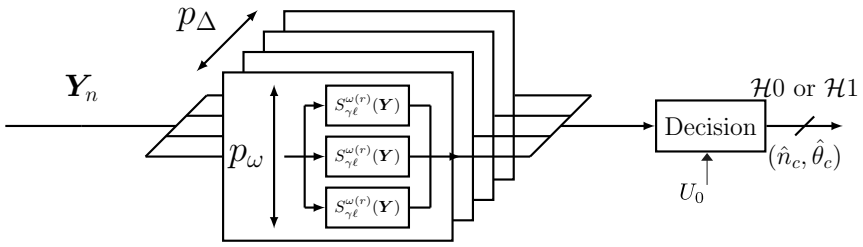


Figure 4.4 – Complete theoretical detection system.

As in Eq. (3.18) in the system model, consider a frame arriving at chip index n_a with a frequency offset θ. The phase offset ϕ is unknown too, but has no effect in this context since only correlation norms are used in detection. The arrival time index n_a can be represented as $n_a = n_c + \Delta$, with coarse time offset $n_c = \gamma_c\ell$ and finer chip offset Δ,

$-\delta_n/2 < \Delta \leq \delta_n/2$. The frequency offset θ_o is also be represented as $\theta_o = \theta_c + \theta_\delta$, with $\theta_c = \omega(r_c)$ and a finer frequency offset θ_δ bounded by $-\delta_\theta/2 < \theta_\delta \leq \delta_\theta/2$. Note that the frame is optimally detectable in the bin (γ_c, r_c) since in this bin, the time and frequency offset errors are minimized.

To lessen the notations in this chapter, the frame $\boldsymbol{Y}_{n_c}^{\omega(r_c)}$ processed at bin (γ_c, r_c) that corresponds to the correct coarse time and frequency offsets is denoted as $\boldsymbol{Y}$, and can be re-defined as the chip level as

$$y(n) = e^{j(n\frac{\theta_\delta}{q}+\phi)}\boldsymbol{F}(n-\Delta) + z(n), \tag{4.6}$$

where $\theta_\delta = 2\pi f_\delta q$ and $z(n)$ are independent realizations of a complex Gaussian noise $\mathcal{N}(0, \sigma^2)$ of zero mean and variance σ^2, $\phi \in [0, 2\pi]$, $\Delta \in \{-\ell/2, \ldots, \ell/2\}$ and $\theta_\delta \in [-\delta_\theta/2, \delta_\theta/2]$.

In the case of the reception of a frame in the optimal bin (hypothesis $\mathcal{H}1$), the baseband transmission model is thus a function of three parameters: the time offset Δ, the finer frequency offset of rotation θ_δ and the standard deviation σ of the complex AWGN. In the case of no reception (Hypothesis $\mathcal{H}0$), the baseband transmission model is simply

$$y(n) = z(n). \tag{4.7}$$

When the estimated couple $(\hat{n}_c = \hat{\gamma}_c\ell,\ \hat{\theta}_c = \omega(\hat{r}_c))$ is the same as the real coarse time of arrival and frequency offset couple (n_c, θ_c) then the hypothesis $\mathcal{H}(\hat{\gamma}_c\ell, \omega(\hat{r}_c))$ appears to be verified if the level of noise is not too high. Fig. 4.5 shows in 3D the values of $S_{\gamma\ell}^{\omega(r)}(\boldsymbol{Y})$ for very small resolution of grid size, i.e. $\delta_n = 1$ and $\delta_\theta = \pi/32$, for a frame of length $N = 60$ without noise and with a CCSK sequence of length $q = 64$, affected by $n_a = 20$ chips and $\theta_o = \frac{9\pi}{8}$ radian (cyan colored point). In the figure, we first used the detection algorithm $S_{\gamma\ell}^{\omega(r)}(\boldsymbol{Y})$ to assess or not the arrival of a new frame with limited computational resource. The time and the frequency space are split into bins of size $(\delta_n, \delta_\theta) = (q/4, \pi/2)$, as shown in the black lines decomposition of the grid. Note that when a bin is triggered (i.e. the associated score function is above a threshold), some bins of the time and frequency neighbourhood can also be triggered. The $(\hat{\gamma}_c, \hat{r}_c)$ index information related to the bin with the highest score function is used to generate the coarse time synchronisation $\hat{n}_c = \hat{\gamma}_c\ell$ and coarse frequency synchronisation $\hat{\theta}_c = \omega(\hat{r}_c)$. The couple $(\hat{n}_c, \hat{\theta}_c)$ (the bin delimited by the white rectangle) is thus sent to the synchronisation process. In this example, the remained fine offsets is $\Delta = 4$ and $\theta_\delta = \pi/8$.

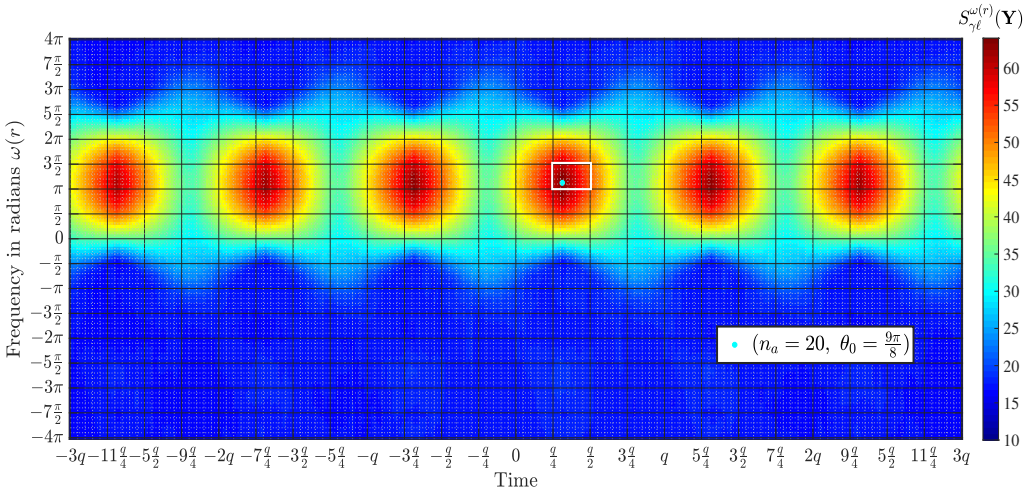


Figure 4.5 – Values of $S_{\gamma\ell}^{\omega(r)}(\boldsymbol{Y})$ for an arriving QCSP frame in a noiseless channel.

Since the work done at very low SNRs, we may have some time and frequency ambiguity (triggering a maximum score value in the wrong bin, i.e. $n_c \neq \hat{\gamma}_c\ell$ and/or $\theta_c \neq \omega(\hat{r}_c)$). Consequently, the received samples in a window of size $2N$ around the estimated time of arrive $\hat{n}_c$ is sent to the synchronisation unit. More specifically, the samples $y(\hat{n}_c - Nq/2), y(\hat{n}_c - Nq/2+1), \ldots, y(\hat{n}_c + qN + qN/2)$ with coarse synchronization parameters $(\hat{n}_c, \hat{\theta}_c)$ are sent to the synchronisation unit. This is discussed in detail in the next chapter that describes the synchronization process.

It is worth noticing here that we can replace the frequency domain computation of the correlation vector $\boldsymbol{L}$ with a "time sliding" computation [19] which has better performance. The name "time sliding" comes from the computation scheduling that uses the circular property of the CCSK modulation to reduce dramatically the computation burden.

4.4 Theoretical model

In this section, we derive the formal performance model of the frame detection algorithm discussed in the previous section. This model allows avoiding costly estimation performance through MC simulation. It gives the insight to analyze the impact of each parameter on the detection performance. In this section, variable $\boldsymbol{n}$ is omitted from expressions $\boldsymbol{L}_{n+kq}$ and $\boldsymbol{y}_{n+kq}$ to lighten the notations.

4.4.1 Correlation expressions

Let us first express the exact expression of $L_{kq}(s)$, see (3.21) for each value of s. Then, we derive the probability law of $|L_{kq}(s)|$ with and without signal.

Definitions and notations

Define vector-operators with vectors $\boldsymbol{g} = [g_0\ g_1\ \dots\ g_{L-1}]$, and $\boldsymbol{h} = [h_0\ h_1\ \dots\ h_{L-1}]$:

— Sectioning a vector from index a to b: $\boldsymbol{g}_a^b = [g_a\ g_{a+1}\ \dots\ g_b]$.

— Concatenation of two vectors $\boldsymbol{g}$ and $\boldsymbol{h}$: $\boldsymbol{g} \coprod \boldsymbol{h} = [g_0\ \dots\ g_{L-1}\ h_0\ \dots\ h_{L-1}]$.

— Linear Right and Left shifts of vector $\boldsymbol{g}$ by Δ positions,

$$\mathcal{R}^{\Delta}(g) = \mathbf{0}_0^{\Delta-1} \coprod \boldsymbol{g}_0^{L-\Delta-1},$$
$$\mathcal{L}^{\Delta}(g) = \boldsymbol{g}_{\Delta}^{L-1} \coprod \mathbf{0}_0^{\Delta-1},$$

where $\mathbf{0}_0^{\Delta-1}$ is a zero vector of length Δ.

Based on the discussion in previous sections, frame $\boldsymbol{Y}$ at the detected bin can be rewritten in vector-operational form as

$$\boldsymbol{Y} = \mathrm{e}^{j\varphi}\left(\mathcal{R}^{\Delta}(\boldsymbol{F}) \odot \boldsymbol{\Phi}\right) + \boldsymbol{Z}, \tag{4.8}$$

where φ is the initial phase offset, $\mathcal{R}^{\Delta}(\boldsymbol{F})$ the delayed CCSK frame by Δ chips, and $\boldsymbol{\Phi} = \{\mathrm{e}^{j2\pi f_\delta n}\}_{0\le n\le Nq-1}$ a vector representing the effect of frequency offset f_δ over the whole frame. $\boldsymbol{Z}$ is the complex AWGN vector: $\boldsymbol{Z} = \boldsymbol{Z}_I + j\boldsymbol{Z}_Q$, where $\boldsymbol{Z}_I$ and $\boldsymbol{Z}_Q$ follow Normal distribution $\mathcal{N}(0, 2\sigma^2)$.

Due to the specific structure of the CCSK modulation (all the sequences are cyclically shifted versions of the reference sequence $\boldsymbol{P}_0$), the delayed Frame $\mathcal{R}^{\Delta}(\boldsymbol{F})$ in (4.8) can be expressed as

$$\mathcal{R}^{\Delta}(\boldsymbol{F}) = \left(\mathbf{0}_0^{\Delta-1} \coprod (\boldsymbol{P}_{c_0})_0^{q-\Delta-1}\right) \coprod \left(\coprod_{k=1}^{N-1} \left((\boldsymbol{P}_{c_{k-1}})_{q-\Delta}^{q-1} \coprod (\boldsymbol{P}_{c_k})_0^{q-\Delta-1}\right)\right). \tag{4.9}$$

Finally, the received vector $\boldsymbol{y}_0$ can be written as

$$\boldsymbol{y}_0 = \mathrm{e}^{j\varphi}\mathcal{R}^{\Delta}\left(\boldsymbol{P}_{c_0}\right) \odot \boldsymbol{\Phi}_0^{q-1} + \boldsymbol{Z}_0^{q-1}, \tag{4.10}$$

and $\boldsymbol{y}_{kq}$, $k > 0$ as

$$\boldsymbol{y}_{kq} = \mathrm{e}^{j\varphi} \left\{ \mathcal{L}^{q-\Delta} \left(\boldsymbol{P}_{c_{k-1}} \right) + \mathcal{R}^{\Delta} \left(\boldsymbol{P}_{c_k} \right) \right\} \odot \boldsymbol{\Phi}_{kq}^{kq+q-1} + \boldsymbol{Z}_{kq}^{kq+q-1}. \tag{4.11}$$

Exact expression of $L_{kq}(s)$

Taking into consideration the expression of $\boldsymbol{y}_{kq}$ defined in (4.11) and the linearity property of the scalar product, the correlation $L_{kq}(s) = \langle \boldsymbol{y}_{kq}, \boldsymbol{P}_s \rangle$ can be expressed as

$$L_{kq}(s) = L_{kq}(s)^{-} + L_{kq}(s)^{+} + z_{kq}(s), \tag{4.12}$$

where

$$\begin{aligned} L_{kq}^{-}(s) &= e^{j\varphi} \langle \mathcal{L}^{q-\Delta} \left(\boldsymbol{P}_{c_{k-1}} \right) \odot \boldsymbol{\Phi}_{kq}^{kq+q-1}, \boldsymbol{P}_s \rangle \\ &= e^{j\psi_k} \sum_{n=0}^{\Delta-1} P(n - c_{k-1} - \Delta) P(n-s) \mathrm{e}^{j2\pi f_\delta n}, \end{aligned} \tag{4.13}$$

$$L_{kq}^{+}(s) = e^{j\psi_k} \sum_{n=\Delta}^{q-1} P(n - c_k - \Delta) P(n-s) \mathrm{e}^{j2\pi f_\delta n}, \tag{4.14}$$

and

$$z_{kq}(s) = \langle \boldsymbol{Z}_{kq}^{kq+q-1}, \boldsymbol{P}_s \rangle. \tag{4.15}$$

The phase offset $\psi_k = \varphi + kq2\pi f_\delta$ represents the sum of the initial phase shift φ and the contribution of the residual frequency offset f_δ on the k^{th} received block Y_{kq}.

Let us analyze (4.12), (4.13) and (4.14) in particular useful cases.
a) When $k = 0$, (4.12) can be reduced to $L_0(s) = L_0^{+}(s) + z_0(s)$.
b) When $s = c_{k-1} + \Delta$, (4.13) gives

$$L_{kq}^{-}(c_{k-1} + \Delta) = e^{j\psi_k} \sum_{n=0}^{\Delta-1} e^{j2\pi f_\delta n} = e^{j\psi_k^{-}} \left(\frac{\sin(\pi f_\delta \Delta)}{\sin(\pi f_\delta)} \right), \tag{4.16}$$

where $\psi_k^{-} = \psi_k + \pi f_\delta (\Delta - 1)$.
c) When $s = c_k + \Delta$, (4.14) gives

$$L_{kq}^{+}(c_k + \Delta) = e^{j\psi_k^{+}} \left(\frac{\sin(\pi f_\delta (q - \Delta))}{\sin(\pi f_\delta)} \right), \tag{4.17}$$

where $\psi_k^+ = \psi_k + \pi f_\delta(q + \Delta - 1)$.
d) In the particular case where $c_{k-1} = c_k = c$, when $s = c + \Delta$

$$L_{kq}(c + \Delta) = e^{j(\psi_k + \pi f_\delta(q-1))} \left(\frac{\sin(\pi f_\delta q)}{\sin(\pi f_\delta)} \right) + z_{kq}(s). \tag{4.18}$$

e) It is worth adding that when there is no phase and frequency offset ($\varphi = 0$ and $f_\delta = 0$), then (4.16), (4.17) and (4.18) give $L_{kq}^-(c_{k-1}+\Delta) = \Delta$, $L_{kq}^+(c_k+\Delta) = (q-\Delta)$ and $L_{kq}(c+\Delta) = q + z_{kq}(s)$, respectively, as shown in Fig. 4.2.

From the formal expression of $L_{kq}(s)$ for any value of s, it is possible to derive the exact probability law of $\max\left(|\boldsymbol{L}_{kq}|\right)$ used to compute $S(\boldsymbol{y})$ in (4.3).

Finally, according to (4.15), $z_{kq}(s)$ is the sum of q independent Gaussian Random Variables (GRV) $\mathcal{N}(0, 2\sigma^2)$ multiplied by +1 or by -1. Thus, $z_{kq}(s)$ is a realization of Gaussian distribution of law $\mathcal{N}(0, 2q\sigma^2)$.

Probability law of $L_{kq}(s)$

Under the hypothesis $\mathcal{H}0$ (no signal), the terms L_{kq}^- and L_{kq}^+ of (4.12) are null and thus, for each s, $L_{kq}(s) = z_{kq}(s)$ is a GRV of law $\mathcal{N}(0, 2q\sigma^2)$ as defined before.

Under the hypothesis $\mathcal{H}1$ (signal exists), when $k > 0$, $L_{kq}(s) = L_{kq}^-(s) + L_{kq}^+(s) + z_{kq}(s)$. The first two terms are deterministic. Their sum can be expressed in polar coordinate as $L_{kq}^-(s) + L_{kq}^+(s) = \rho_k(s)e^{j\theta_k(s)}$, and thus $L_{kq}(s)$ is a GRV of law $\mathcal{N}(\rho_k(s)e^{j\theta_k(s)}, 2q\sigma^2)$. Since we are interested in the absolute value of $L_{kq}(s)$, the phase $\theta_k(s)$ has no impact. The value of $\rho_k(s) = |L_{kq}^-(s) + L_{kq}^+(s)|$ takes particular values for $s = c_{k-1}+\Delta$ and $s = c_k+\Delta$, as shown in (4.16) and (4.17).
For the first symbol, when $k = 0$, $L_0(s) = L_0^+(s) + z_0(s)$, and thus $\rho_0(s) = |L_0^+(s)|$.

In next subsections, the distributions of of the absolute values $|L_{kq}(s)|$, $s = 0, 1, \ldots, q-1$, the absolute value of each of the GRVs are derived.

4.4.2 Probability distributions of $|\boldsymbol{L}_{kq}(s)|$ and maximum of $|\boldsymbol{L}_{kq}(s)|$

In this section we discuss the Probability Density Function (PDF) as well as the Cumulative Distribution Function (CDF) of $|L_{kq}(s)|$ the absolute value of each of the GRVs representing the elements of the correlation vector $L_{kq}(s)$, $s = 0, 1, \ldots, q-1$, defined in previous section. Then we derive the PDF of the maximum value of $|L_{kq}(s)|$ in both hypothesis $\mathcal{H}0$ and $\mathcal{H}1$.

PDF and CDF of the absolute value of $L_{kq}(s)$, $|L_{kq}(s)|$

The dependency of $|L_k(s)|$ on the index $k > 0$ depends only on the couple values (c_{k-1}, c_k). It is thus convenient to replace (c_{k-1}, c_k) by (a, b) to lighten notation. With this notation, $L_{(a,b)}(s)$ is GRV of law $\mathcal{N}(\rho_{(a,b)}(s)e^{j\theta_{(a,b)}(s)}, 2q\sigma^2)$, where $\rho_{(a,b)}(s)$ and $\theta_{(a,b)}(s)$ are the module and the phase of $L^-_{(a,b)}(s)+L^+_{(a,b)}(s)$, respectively. Thus, $|L_{(a,b)}(s)|$ is a Rician distribution with the following PDF and CDF [109]

$$\begin{aligned} f_{|L_{(a,b)}(s)|}(x) &= \frac{2x}{q\sigma^2}\mathrm{e}^{\left(-\frac{x^2+\rho_{(a,b)}(s)^2}{q\sigma^2}\right)} I_0\left(\frac{2x\rho_{(a,b)}(s)}{q\sigma^2}\right), \\ F_{|L_{(a,b)}(s)|}(x) &= 1 - Q_1\left(\frac{\rho_{(a,b)}(s)}{\sigma\sqrt{q/2}}, \frac{x}{\sigma\sqrt{q/2}}\right), \end{aligned} \tag{4.19}$$

where $x \in [0, +\infty[$, $I_0(z)$ is the modified Bessel function of the first kind with order zero and Q_1 is the Marcum Q-function. For a given couple $a = c_{k-1}$ and $b = c_k$, $F_{|L_{(a,b)}(s)|}(x)$ is plotted in Fig. 4.6 for $s = c_{k-1} + \Delta$, $s = c_k + \Delta$ and the other $q - 2$ cases when $s \neq c_{k-1} + \Delta$, $s \neq c_k + \Delta$.

PDF and CDF of the maximum value of $|L_{kq}(s)|$ for $\mathcal{H}1$

Define our first hypothesis of the proposed theoretical model. According to (4.15), for any couple (s, s'), we have the inter-correlation $\mathbb{E}[z_{kq}(s), z_{kq}(s')]$ between $z_{kq}(s)$ and $z_{kq}(s')$ equal to $\langle \boldsymbol{P}_s, \boldsymbol{P}_{s'} \rangle$. Since $z_{kq}(s)$ and $z_{kq}(s')$ are both Gaussian variables of zero mean, they are independent if, and only if, $\mathbb{E}[z_{kq}(s), z_{kq}(s')] = 0$. This hypothesis is assumed in the rest of the report since the sequence $\boldsymbol{P}_0$ is carefully selected so that $s \neq s' \Rightarrow \langle \boldsymbol{P}_s, \boldsymbol{P}_{s'} \rangle \ll q$. In others words, variables $z_{kq}(s)$ are considered to be independent to each others.

At first, consider $k > 0$ and let $M_{(a,b)}$ be defined as the maximum of the absolute values of $L_{(a,b)}(s)$, i.e. $M_{(a,b)} = \max\{|L_{(a,b)}(s)|, s \in GF(q)\}$. The independence hypothesis of the $z_{(a,b)}(s)$ variables also implies the independence of the $|z_{(a,b)}(s)|$ variables. Thus, the CDF of the $M_{(a,b)}$ denoted by $F_{M_{(a,b)}}$ is defined as the product of the elementary CDFs of each element $F_{|L_{(a,b)}(s)|}$, $s = 0, 1, ..., q - 1$

$$F_{M_{(a,b)}}(x) = \prod_{s=0}^{q-1} F_{|L_{(a,b)}(s)|}(x), \tag{4.20}$$

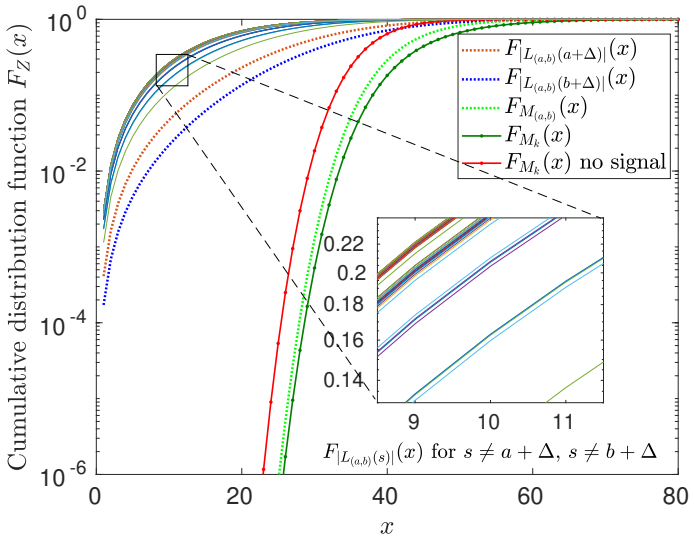


Figure 4.6 – Illustration of different CDF equations for a given GF(64) received block $\boldsymbol{y}_{kq}$ at SNR = −7 dB, $\Delta = 24$ chips and $\theta_o = \pi/4$.

for $x \in [0, +\infty[$. All the CDF functions implied in (4.20) are plotted in Fig. 4.6 for a given couple $a = c_{k-1}$ and $b = c_k$. Since all couples (a, b) are equiprobable. The average value of $F_{M_k}(x)$ is given by marginalizing $F_{M_{(a,b)}}(x)$ over all possible couples, i.e.

$$F_{M_k}(x) = \frac{1}{q^2} \sum_{(a,b)} F_{M_{(a,b)}}(x), \tag{4.21}$$

as shown in Fig. 4.6 also.

When $k = 0$, M_0 depends only on c_0 and we can replace the index 0 by the value (b) to be consistent with the previous notation, i.e. $M_0 = M_{(b)}$. Thus, $F_{M_0}(x)$ obtained as

$$F_{M_0}(x) = \frac{1}{q} \sum_{(b)} \prod_{s=0}^{q-1} F_{|L_{(b)}(s)|}(x). \tag{4.22}$$

The PDF of the maximum value of the absolute correlation vector denoted by f_{M_k} can be obtained by taking the derivative of F_{M_k}.

$$f_{M_k}(x) = \frac{dF_{M_k}(x)}{dx}. \tag{4.23}$$

The detection filter described in (4.3) takes the sum of N maximum values over a

window of N blocks Y_{kq}. Thus the score function can be expressed as

$$S = \sum_{k=0}^{N-1} M_k. \tag{4.24}$$

In the sequel, we assume that the M_k, $k = 0, 1, \ldots, N-1$, are independent and identically distributed random variables with common probability density function f_{M_k}. This is an approximation because two consecutive values $|L_{kq}(s)|$ and $|L_{k+1}(s)|$ are not necessarily uncorrelated since the same c_k value is used in both of them. Nevertheless, considering the set of couple L_{2k}, $k = 1..N/2$ are thoroughly random, as for the set $L_{2K+1}, k = 0, ..., N/2-1$. If N is not too small, the space is explored almost randomly. Thus, the PDF of the random variable S can be defined as the convolution of f_{M_k}, $k = 0, 1, \ldots, N-1$

$$\begin{aligned} f_S(x) &= f_{M_0}(x) * f_{M_1}(x) * \cdots * f_{M_{N-1}}(x) \\ &= f_{M_0}(x) * f_{M_k}^{*(N-1)}(x), \end{aligned} \tag{4.25}$$

where $f_{M_k}^{*(N-1)}(x)$ is the $(N-1)$-fold convolution power of $f_{M_k}(x)$ and $x \in [0, +\infty[$. It is worth mentioning that as the number of symbols N in a packet increases, f_S converges to normal distribution according to central limit theorem. Under the hypothesis $\mathcal{H}1$, $f_S(x)$ is denoted as $f_S^{\mathcal{H}1}(x)$.

CDF and PDF of the maximum value of $|L_{kq}(s)|$ for $\mathcal{H}0$

The distribution of $L_{kq}(s)$ when no frame has been transmitted was given as GRV $\mathcal{N}(0, 2q\sigma^2)$. In this case, the absolute value of the complex number $L_{kq}(s)$ is a random variable following the Rayleigh distribution [109], where the CDF and PDF of $|L_{kq}(s)|$ are given in (4.26) for $x \in [0, +\infty[$

$$\begin{aligned} F_{|L_{kq}(s)|}(x) &= 1 - \mathrm{e}^{\left(-\frac{x^2}{q\sigma^2}\right)}, \\ f_{|L_{kq}(s)|}(x) &= \frac{2x}{q\sigma^2}\mathrm{e}^{\left(-\frac{x^2}{q\sigma^2}\right)}. \end{aligned} \tag{4.26}$$

Note that (4.26) is just a particular case of (4.19) when $\rho = 0$. The analysis done in section 4.4.2 can be applied again. The PDF of the maximum value of $|L_{kq}(s)|$ can be

obtained by calculating first its CDF,

$$F_{M_k}(x) = \prod_{s=0}^{q-1} F_{|L_{kq}(s)|}(x) = \left[1 - \mathrm{e}^{\left(-\frac{x^2}{q\sigma^2}\right)}\right]^q, \tag{4.27}$$

for $x \in [0, +\infty[$, that is also illustrated in Fig. 4.6, and then finding its derivative $f_{M_k}(x)$ such that

$$f_{M_k}(x) = \frac{2x}{\sigma^2}\mathrm{e}^{\left(-\frac{x^2}{q\sigma^2}\right)}\left[1 - \mathrm{e}^{\left(-\frac{x^2}{q\sigma^2}\right)}\right]^{q-1}. \tag{4.28}$$

Finally, under hypothesis $\mathcal{H}0$ the PDF of the random variable S, sum of M_k, can be defined as the convolution of f_{M_k}, $k = 0, 1, \ldots, N-1$

$$f_S^{\mathcal{H}0}(x) = f_{M_k}^{*N}(x), \tag{4.29}$$

which is the N-fold convolution power of $f_{M_k}(x)$.

4.4.3 Confirmation of the theoretical model by Monte Carlo simulation

In the previous section the PDFs $f_S^{\mathcal{H}1}(x) \backsim \mathcal{P}(X = S(y)|\ \mathcal{H}1)$ in (4.25) and $f_S^{\mathcal{H}0}(x) \backsim \mathcal{P}(X = S(y)|\ \mathcal{H}0)$ is derived in (4.29) over AWGN channel when the CCSK frame exists or is absent, respectively. To check the validity of the hypothesis taken to build the theoretical model, it is compared with the MC simulation, when 10^6 CCSK frames are transmitted, in case of a frame length $N = 20$ GF(64) symbols over AWGN channel at SNR = -10 dB. Two different scenarios are tested, the first one (see Fig. 4.7a) assesses perfect synchronization conditions ($\Delta = 0, \theta_\delta = 0$), and the second case (see Fig. 4.7b) is considered for $\Delta = q/4$ and $\theta_\delta = \pi/2$. As it can be seen in both cases, the probability distribution functions in the theoretical model fit exactly the MC simulation. It is worth noting that in the theoretical model we can go through very low values of probabilities (here 10^{-10}) without the need to run 10^{10} iterations of a MC simulation for transmitting 10^{10} CCSK frames for example. Thus, the detection performance can be found through the derived theoretical model without the need to conduct extensive MC simulations.

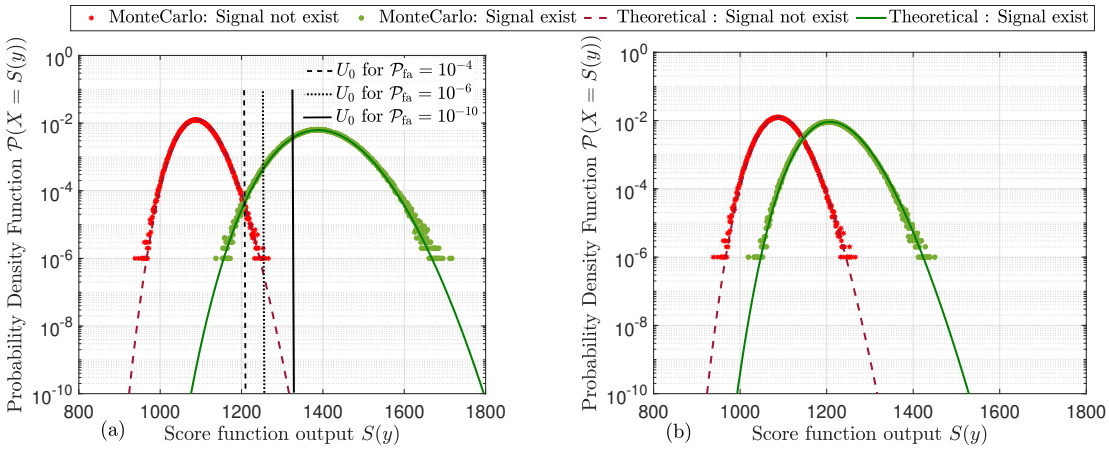


Figure 4.7 – MC and Theoretical PDFs in both hypothesis $\mathcal{H}0$ and $\mathcal{H}1$, for a CCSK frame of N=20 symbols in GF(64) for the first scenario **(a)** SNR = -10 dB, $\Delta = 0$, no frequency offset and for the second scenario **(b)** SNR = -10 dB, $\Delta = 16$, $\theta_\delta = \pi/2$.

4.5 Performance analysis

In this section, we assess the detection performance of the system, at very low SNR, while considering the impact of the QCSP parameters: Galois field order q and the number of CCSK symbols in a frame N. After that, we examine the effect of the time and frequency offsets on the system performance in an asynchronous channel. All the upcoming results are obtained thanks to MC simulation that stopped after obtaining 100 frames of errors. All the results are also confirmed by the theoretical performance model.

4.5.1 Effect of GF(q) order, $q = 2^p$

In this section, we study the effect of the sequence length q, or in other terms the Galois Field order $q = 2^p$. For that, we fix the following set of parameters needed for generating a QCSP frame and vary the value of q to illustrate its effect on the detection performance:

— Number of CCSK symbols N: $N = 60$ and $N = 120$.

— Threshold value U_0: determined to get a $\mathcal{P}_{fa}$ of 10^{-6}.

— Perfect time and frequency synchronization: $\Delta = 0$, $\theta_\delta = 0$.

Fig. 4.8 shows the simulations results of $\mathcal{P}_{md}$ vs. SNR for $q = 2^p$ ranging from $p = 6$ up to $p = 12$, for two different frame lengths $N = 60$ and $N = 120$. For $N = 60, q = 64$, $\mathcal{P}_{md}$ is plotted for three different values of $\mathcal{P}_{fa}$: 10^{-4}, 10^{-6} and 10^{-10}. As expected, $\mathcal{P}_{md}$

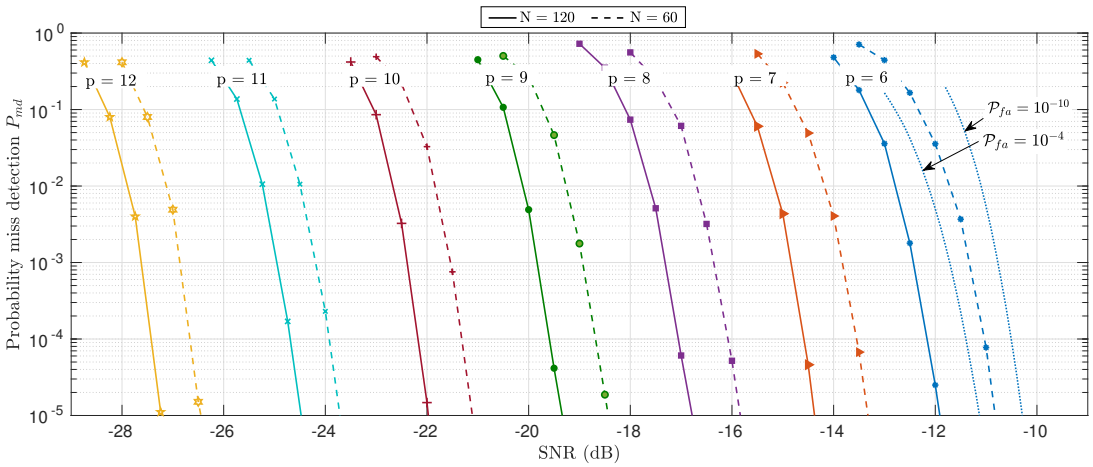
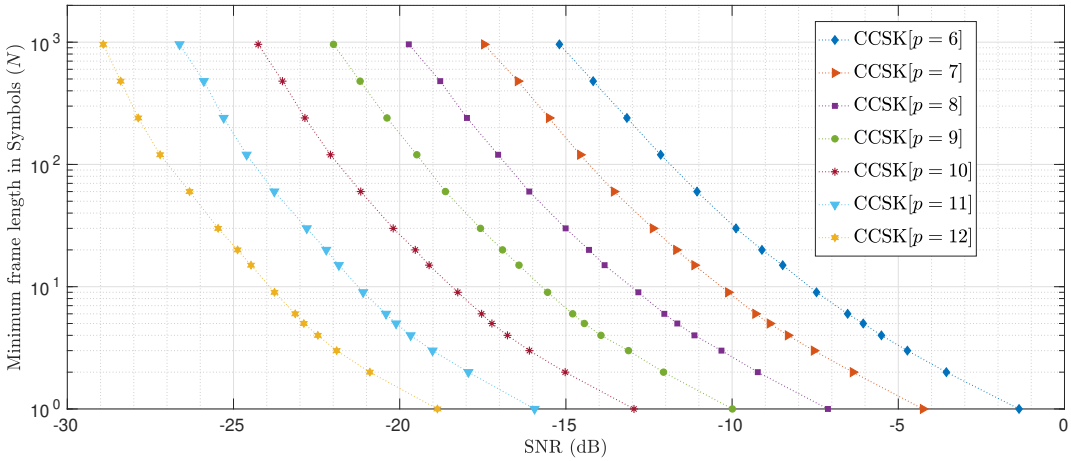


Figure 4.8 – $\mathcal{P}_{\text{md}}$ and $\mathcal{P}_{\text{fa}}$ as function of SNR for a CCSK frame of $N = 60, 120$ symbols for different $q = 2^p$ orders, in an ideally synchronized channel.

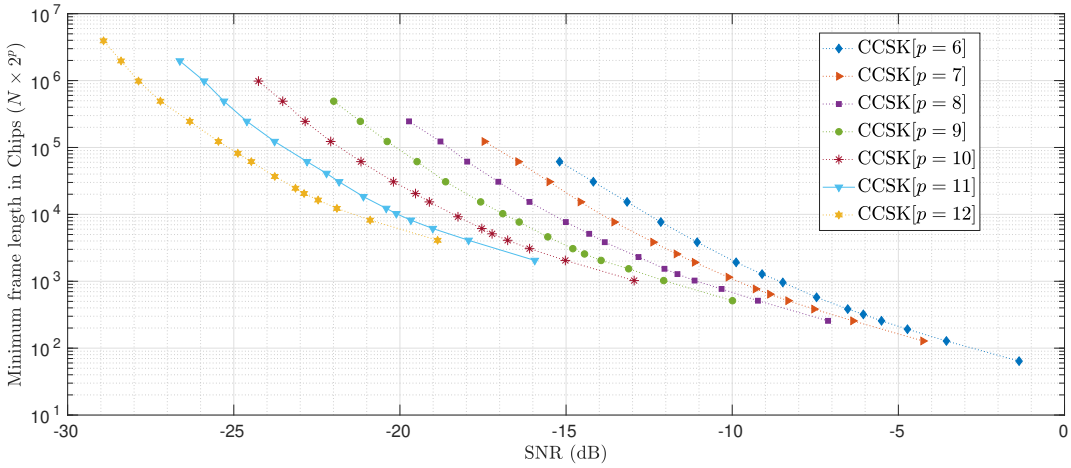
increases when $\mathcal{P}_{\text{fa}}$ decreases, i.e. when the threshold U_0 value increases. So the value of U_0 in the system is selected based on the desired trade-off $\mathcal{P}_{\text{fa}}$ vs $\mathcal{P}_{\text{md}}$ according to the application requirements. This observation is valid for $q > 64$, but the corresponding curves of $\mathcal{P}_{\text{md}}$ are omitted for the sake of figure simplicity. As shown for $N = 60$ curves, the SNR required to obtain an acceptable $\mathcal{P}_{\text{md}}$ of the order of 10^{-4} is -11.05 dB when $q = 64$, and decreases as q increases to go down to -26.75 dB when $q = 4096$. We can see from the figure also, that the performance is shifted 1 dB in average for $N = 120$ symbols over $N = 60$. This is an important result that shows the higher impact of the length of the spreading sequence on the detection performance and that the proposed detector can operate reliably at a very low SNRs. Therefore, the pseudo random sequence length q of the CCSK modulation can be chosen depending on the target application that corresponds to the desired $\mathcal{P}_{\text{md}}$ and $\mathcal{P}_{\text{fa}}$.

4.5.2 Effect of the CCSK frame length

This section examines the minimum number of symbols N (or in chips $N \times q$) in the QCSP frame that minimizes the energy for reliable transmission of a frame of finite length. In order to interpret this challenge, we need first to find the minimum length of a QCSP frame (in symbols and chips) for given probabilities of detection ($\mathcal{P}_{\text{md}}$ and $\mathcal{P}_{\text{fa}}$).



(a) CCSK frame length in Symbols N.



(b) CCSK frame length in chips $N \times 2^p$.

Figure 4.9 – Minimum frame length of a QCSP frame, needed to guarantee $\mathcal{P}_{\text{md}} \leq 10^{-4}$ and $\mathcal{P}_{\text{fa}} \leq 10^{-6}$ at different SNR, for different CCSK order p in an ideally synchronized channel.

Fig. 4.9 shows the minimum frame length needed to guarantee $\mathcal{P}_{\text{md}} = 10^{-4}$ and $\mathcal{P}_{\text{fa}} = 10^{-6}$, in an ideally synchronized channel (no frequency and no time offset), as function of SNR, for $p = 6$ (right-most curve) to $p = 12$ (left-most curve). Fig. 4.9a represents the results in symbols and 4.9b in chips. Each point in 4.9b is calculated in chips as $N_q = N \times q$. The flat region explains that at high SNRs, the corresponding minimal value is needed to guarantee both ($\mathcal{P}_{\text{md}} \leq 10^{-4}$ and $\mathcal{P}_{\text{fa}} \leq 10^{-6}$). This value is one GF($2^p$) symbols, and $q = 2^p$ for each GF order p. It is worth noticing from Fig. 4.9 that the length q of the sequence $\boldsymbol{P}_0$ has higher impact on the detection performance than the number of CCSK symbols.

4.5.3 Effect of time and frequency offset

The effect of both time and frequency shifts (see (4.6)) on the detector performance is discussed in this section. We consider the frame of length $N = 120$ symbols and of order $p = 6$. Fig. 4.10 plots the minimum SNR needed, for predefined probabilities ($\mathcal{P}_{\text{fa}} = 10^{-6}$ and $\mathcal{P}_{\text{md}} = 10^{-4}$), as a function of temporal offsets Δ for different frequency offsets θ_δ. The figure is divided into two mirrored sides. The left hand side (worst-case scenario) presents the result of the worst case scenario when the couple time and frequency offsets equal $|\Delta|$ and $|\theta_\delta|$ respectively, while the right hand side (average-case scenario) presents the result of the average case when the couple time and frequency offsets are uniformly seen in $[-\Delta, \Delta]$ and $[-\theta_\delta, \theta_\delta]$ respectively, where $\theta_\delta = 0, \pi/4, \pi/2$ and π. The latter is a more accurate scenario since Δ and θ_δ are uniformly distributed in their respective intervals.

We observe that the rotation of a CCSK frame during q chips by $\theta_\delta = \pi/4$ radian degrades the minimum required SNR by less than 0.1 dB, whereas a half rotation when $\theta_\delta = \pi$ degrades by more than 2.5 dB. For that, the frequency bin size of the time-frequency grid decomposition discussed in section 4.3 is chosen to be $\delta_\theta = \pi/2$. In other words, several filters N_F need to be performed in parallel, one for each frequency hypothesis δ_θ to decrease the impact of ω_0 to be in the range of $[-\pi/4, \pi/4]$. [1]

An interesting trade-off between detection performance and QCSP system complexity exists. Based on the application requirements, we can adjust the system complexity to the performance requirements. Consequently, either we can work on lower complexity, but with lower SNR, or vise-versa. A maximal temporal offset equals to $|\Delta| = 32$ corresponds to a temporal bin length $\ell = 64$. Similarly, a maximal frequency offset equals to $|\omega_0| = \pi$

1. To reduce the overall complexity, we propose to use a similar method to the one proposed by Akopian in [110] for the detection of a GPS signal.

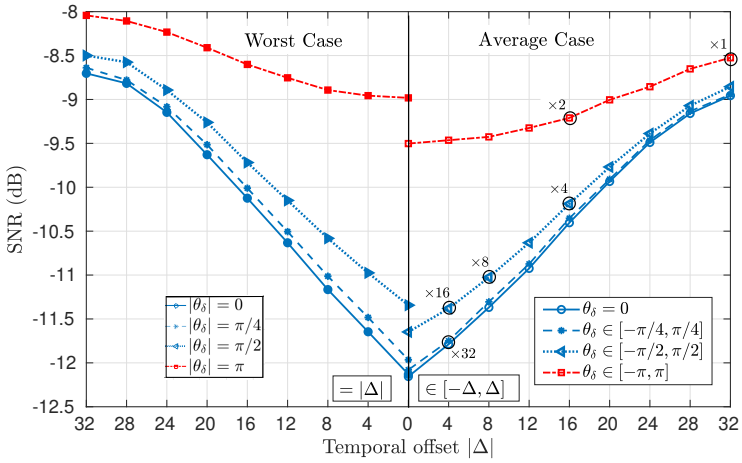


Figure 4.10 – Minimum SNR required as function of different Δ and θ_δ values, for defined probabilities ($\mathcal{P}_{\text{fa}} = 10^{-6}$ and $\mathcal{P}_{\text{md}} = 10^{-4}$), in a CCSK frame of $N = 120$ and order $p = 6$.

corresponds to a frequency bin of size $\delta_\theta = 2\pi$. For this grid size, the required SNR is equal to -8.5 dB. The associated complexity of the decoder for a bin of size $(\delta_\theta, \ell) = (2\pi, 64)$ can be denoted by 1 (see circle noted ×1 in Fig. 4.10). There are two possibilities to halve the bin size in order to get better detection performance. This can be achieved by halving the frequency dimension, i.e. using bins of size $(\delta_\theta, \ell) = (\pi, 64)$. In this case, the required SNR for detection is -8.8 dB. However, it is more efficient to halve the time dimension, i.e. using bins of size $(\delta_\theta, \ell) = (2\pi, 32)$. In the latter case, the required SNR for detection is reduced down to -9.25 dB. This solution is indicated by the circle ×2 (to reflect that the number of bins is doubled) in Fig. 4.10. By continuing recursively this process, the optimal solutions of complexities ×4 , ×8, ×16 and ×32 are indicated in Fig. 4.10. The associated SNR are -10.2 dB, -11 dB, -11.4 dB and -11.8 dB, respectively. In the sequel, the ×32 solution with bin size $(\delta_\theta, \ell) = (\pi/2, 8)$ is considered.

4.6 Conclusion

In this chapter, the proposed preamble-less detection method is developed. It is based on the score function calculated through the correlation of the incoming sequence of samples with the pre-defined CCSK sequences. A theoretical model has been derived, simulated, and validated through a comparison with the experimental MC simulations. This theoretical model permits quick assessment of the performance of the detection method for different parameters values without the need to conduct extensive MC simulations.

Performance analysis has been also performed to illustrate the effect of the different parameters, i.e. length of the CCSK-based frame and the length of the CCSK sequence (Galois filed order q). For a particular case (i.e. $N = 120$ symbols and $q = 64$ chips), the simulations results showed that a reliable detection can be obtained at -7.5 dB where a phase shift uniformly distributed between $-\pi, +\pi$ and a chip delay up to $q/2$ are considered. This performance can be enhanced to reach -11.78 dB by using parallel filters, but this will be on the price of the complexity where it will be $\times 32$.
The next chapter, includes the completion of the system design by discussing the time and frequency synchronization aspects.

Chapter 5

Synchronization

Contents

In this chapter, we propose to use the QCSP structure to transmit short packets without any additional symbol dedicated to help the synchronization process. Time, frequency, and phase synchronization are considered. The key idea is to consider the whole frame as a preamble for the synchronization aspects. The proposed synchronization algorithms process the received samples and compute first the exact start of the frame, then the frequency and phase offsets. These synchronization parameters are calculated with high accuracy at very low SNRs, thanks to the internal structure of the QCSP frame: the CCSK modulation, the Over-Modulation (OM) at symbol level, and finally the information comes from the NB-LDPC codes.
This chapter is divided mainly into two parts. First, we illustrate the time and frequency synchronization process. Then we discuss the phase synchronization of the remained residual frequency and initial phase offsets. In both synchronizations, each of the steps is addressed and discussed by details from the theoretical perspective. The related performance results are found by MC simulation. Then, it is discussed and analyzed. Finally, a conclusion summarizes the content of the chapter.

5.1 Time-frequency synchronization

This section presents the first step of the synchronization process which is time synchronization. It also gives the first finer estimation of the frequency offset, which can decrease its effect on the coherent demodulation. In the following, we first present the problem statement. Then, we illustrate the successive steps of the blind synchronization process in detail. The main focus of this section is on the two levels that mitigate the time synchronization ambiguity: at symbol level thanks to the OM of the CCSK symbols, and at chip-level thanks to the NB-LDPC structure. Finally, we show some results and interpret them.

5.1.1 Problem statement

In a noisy channel, with very low SNR, the maximum value of the score function at the optimal bin (n_c, θ_c) gives a first estimation of the coarse time and frequency offsets, i.e. $(\hat{n}_c = \hat{\gamma}_c \ell,\ \hat{\theta}_c = \omega(\hat{r}_c))$. This does not simply matches the hypothesis that represents the coarse synchronization parameters all the times, (i.e. $\mathcal{H}(\hat{n}_c \neq n_c, \hat{\theta}_c \neq \theta_c)$). Moreover, even if the resolution of bin size in the time-frequency grid is decreased, the hypothesis of

the corresponding finer time and frequency offsets may not be matched $\mathcal{H}(\hat{n}_a \neq n_a, \hat{\theta}_o \neq \theta_o)$. Fig. 5.1 shows in 3D the values of $S_n^\theta(\boldsymbol{Y})$ for the same previous frame (in section 4.3) of length $N = 60$ with a CCSK sequence of length $q = 64$ but at SNR = -10 dB. Also the frame is received at time index $n_a = 20$ chips and affected by a frequency offset $\theta_o = \frac{9\pi}{8}$ radian. Recall here that the values of $S_n^\theta(\boldsymbol{Y})$ for very small resolution of grid size, i.e. $\delta_n = 1$ and $\delta_\theta = \pi/32$. The axes of the figure are limited as the following: n is limited to the interval $[n_a - \frac{N}{2}q, n_a + \frac{N}{2}q]$ and θ to the interval $[\theta_o - 2\pi, \theta_o + 2\pi]$. The figure is shown from different angles of view to be more illustrative. Fig. 5.1a shows S_n^θ as function of time and frequency change. Fig.5.1b shows S_n^θ only as function of time and Fig. 5.1c as function of frequency. Finally, Fig. 5.1d shows the grid image in 2-D as in Fig. 4.5.
The pink dot represents the maximum score value, and the real-time and frequency offsets are represented in the cyan dot. It is clear in this example that the maximum hypothesis, even with very small bin resolution in the grid, doesn't fit the real-time and frequency shifts. To be noted also, this maximum value could be one of the different values of the peaks shown in the figure (the dark red peaks). To alleviate that ambiguity, the synchronization block is fed by a sequence window of size $2N$ symbols around the estimated time of arrival.

For more illustration, we show the same outcome score values but with different scales to better visualize the different offsets. Fig. 5.2 shows the results of $S_n^\theta(\boldsymbol{Y})$ with a zoomed figure, where $n \in [n_a - 3q, n_a + 3q]$ and $\theta \in [\theta_o - \frac{\pi}{4}, \theta_o + \frac{\pi}{4}]$. Finally, Fig. 5.3 presents the output of $S_n^\theta(\boldsymbol{Y})$ for $n \in [n_a - q/2, n_a + q/2]$ and $\theta \in [\theta_o - \frac{\pi}{4}, \theta_o + \frac{\pi}{4}]$.

To conclude, time and frequency estimation by only using the maximum of the score function alone do not give satisfactory results at low SNR. To get a more accurate synchronization process, additional side information should be used to suppress both symbol and chip time synchronization errors.

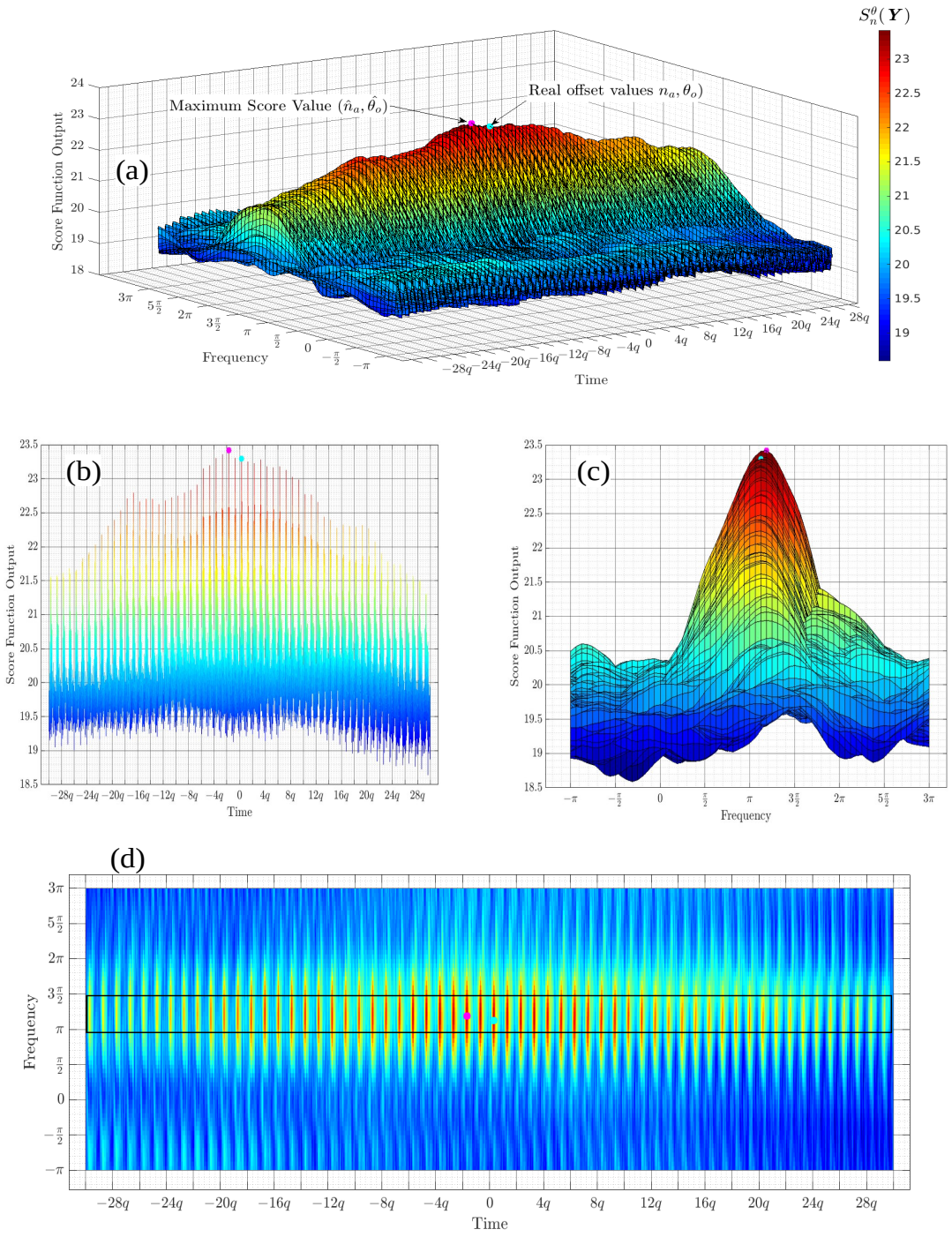


Figure 5.1 – Values of $S_n^\theta(\boldsymbol{y})$ for an arriving QCSP frame in a very Low SNR channel.

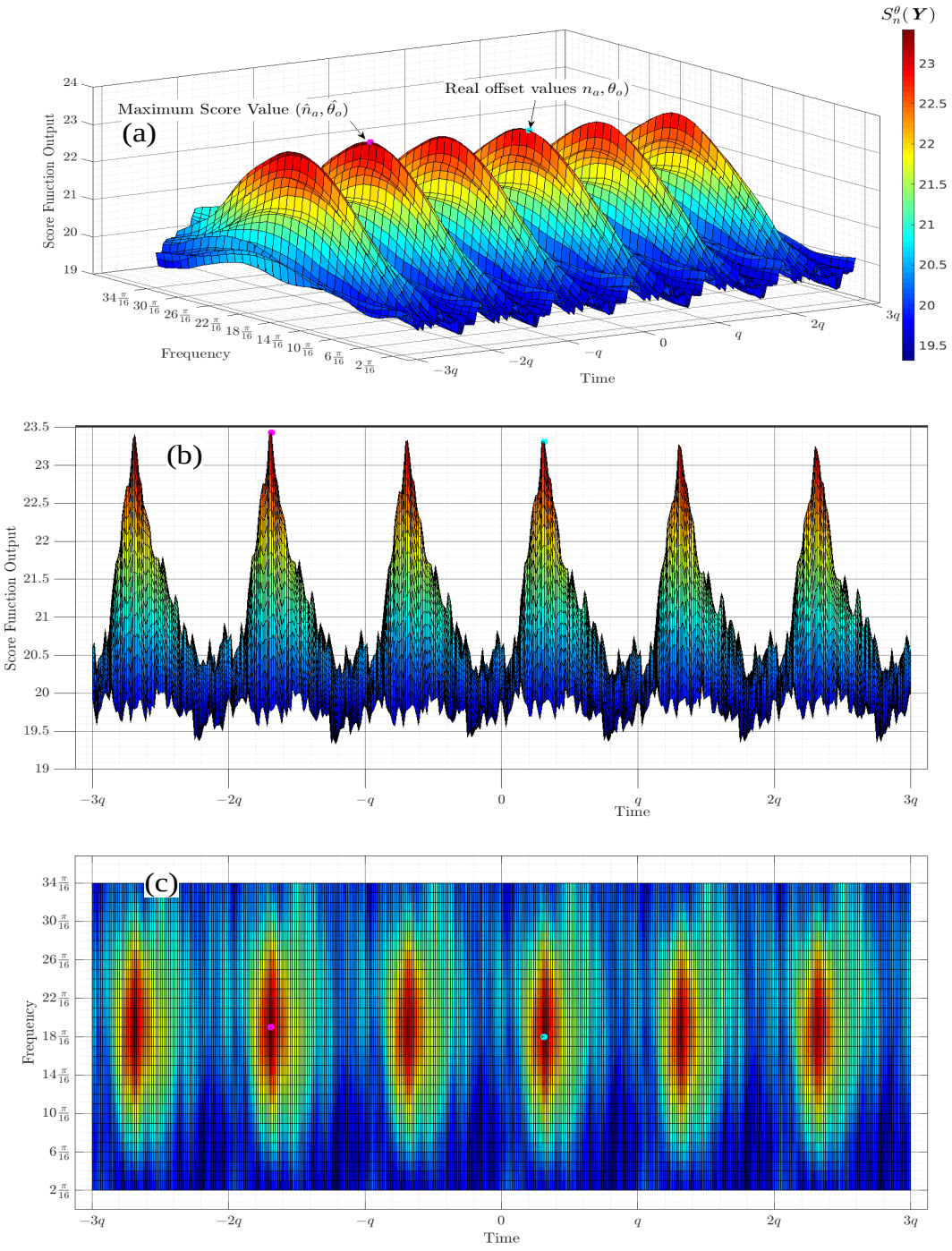


Figure 5.2 – Values of $S_n^\theta(\boldsymbol{y})$ for an arriving QCSP frame in a very Low SNR channel, with $n \in [n_a - 3q, n_a + 3q]$ and $\theta \in [\theta_o - \frac{\pi}{4}, \theta_o - \frac{\pi}{4}]$.

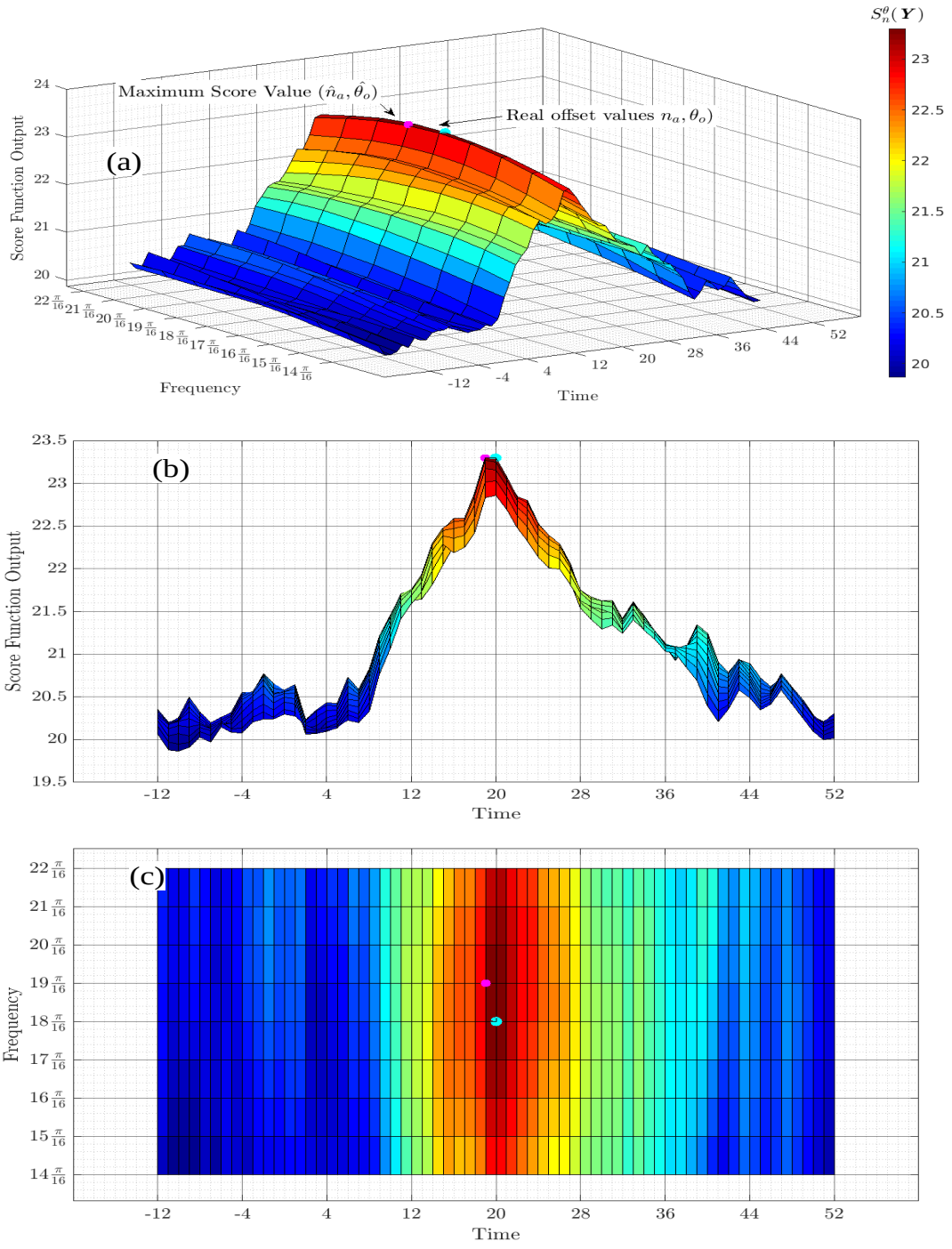


Figure 5.3 – Values of $S_n^{\theta}(\boldsymbol{y})$ for an arriving QCSP frame in a very Low SNR channel, with $n \in [n_a - q/2, n_a + q/2]$ and $\theta \in [\theta_o - \frac{\pi}{4}, \theta_o - \frac{\pi}{4}]$.

5.1.2 Algorithm specifications

This section presents the successive steps of the blind time and frequency synchronization algorithm, which comes directly after the detection process. The input of the algorithm is a set of $2N$ successive received samples that are assumed to contain a frame. Also, the algorithm is fed by the coarse estimations, time offset $\hat{n}_c$ and the coarse frequency offset $\hat{f}_c = \frac{\hat{\theta}_c}{2\pi q}$, which are the outcomes from the detection block. The output of the algorithm is the updated estimation of the correct time of arrival $\bar{n}_a$, the frequency offset $\tilde{f}_o$ and an information of the initial phase $\tilde{\phi}$.
The synchronization process is divided into several steps (or tasks) performed sequentially. Each step adds some additional knowledge that allows coherent demodulation of the received frame. Those steps are:

1. Detection on a locally finer grid resolution to improve both the time and frequency estimation. Since the output of the score function is convex as shown in Fig. 5.1, 5.2, 5.3, it is sufficient to make the time then frequency detection successively instead of parallel searches. This is proved by simulation, and it consequently reduces the complexity from $p_\Delta \times p_\omega$ to $p_\Delta + p_\omega$, where p_Δ and p_ω are number of the finer time and frequency hypothesis. The finer estimated time offset goes from $\hat{n}_c$ to $\hat{n}_a$, and frequency offset from $\hat{f}_c$ to $\hat{f}_o$.
2. Symbol synchronization (from $\hat{n}_a$ to $\tilde{n}_a$) to reduce the error on-chip synchronization to a few units thanks to the over-modulation sequence.
3. Fine frequency synchronization by suppressing as much as possible from the remaining frequency offset error ($\hat{f}_o$ to $\tilde{f}_o$), thanks to the FFT method. This step includes also the first estimation of the initial phase offset (from ϕ to $\tilde{\phi}$).
4. Exact chip synchronization (from $\tilde{n}_a$ to $\bar{n}_a$) using the NB-code properties.

Fig. 5.4 shows the different steps of the successful synchronization process, from the initial coarse time/frequency estimation $(\hat{n}_c, \hat{f}_c)$ to the final correct time/frequency estimation $(\bar{n}_a, \tilde{f}_o) \approx (n_a, f_o)$. The objective of this work is to illustrate and study a blind effective synchronization algorithm that can be optimized in the future. So far, it is worth noticing that there is no guaranty that the method is optimal specifically in terms of complexity. However, this can open the path for the blind synchronization approach for short packet transmission.

To illustrate the different phases of the synchronization process, performance results are given at each step of the synchronization, after transmitting 10^4 frames of length

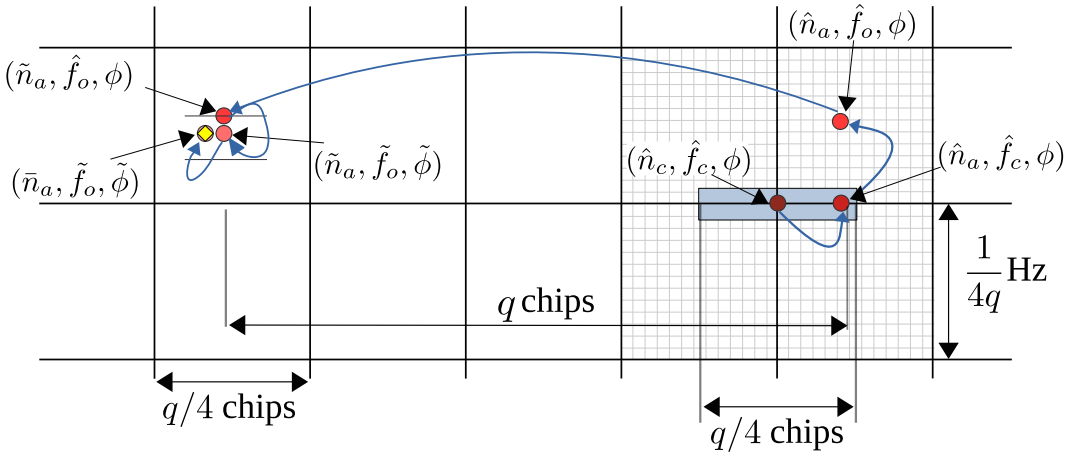


Figure 5.4 – Example of a successful synchronization process from step 1 to step 5.

$K = 20$ symbols of information, an NB-Code of rate 1/3 over GF(64) (i.e. 60 CCSK symbols). The frames are transmitted in an AWGN channel with an SNR of -10 dB, a delay n_a randomly selected between $-q/2$ and $q/2$, a frequency offset f_o and initial phase offset ϕ with uniform distribution.

5.1.2.1 Detection on the local bin with finer time resolution

The detection algorithm performs the time detection with a granularity of ℓ chips, i.e. $\ell = q/4$ chips in our example. The coarse chip synchronization perform the same operation but with a granularity of one chip in a time window $[\hat{n}_c - \ell, \hat{n}_c + \ell]$ around the chip of index $\hat{n}_c$ of the detected frame. Therefore, the estimation of time offset $\hat{n}_a$ is

$$\hat{n}_a = \underset{n \in [\hat{n}_c - \ell, \hat{n}_c + \ell]}{\operatorname{argmax}} \{S_n^{\hat{f}_c}(\boldsymbol{y})\}, \tag{5.1}$$

with $\hat{f}_c$, the coarse frequency estimation. In the rest of this section, we use f instead of θ (i.e. $f = \frac{\theta}{2\pi q}$).

Figure 5.5 shows the histogram of the chip synchronization error $\hat{e}_0 = n_a - \hat{n}_a$ obtained after transmitting 10^4 of the predefined QCSP frame. We can figure out from the first step of synchronization that there are more than 4,000 frames out of 10^4 that are not correctly synchronized at the chip level. By analyzing the probability distribution in the figure, we can distinguish that there exist two different types of errors. First, we have the chip errors that are roughly multiple of q chips (i.e. the size of a CCSK symbol in the QCSP frame), which is defined as the symbols synchronization error s. For this type of synchronization

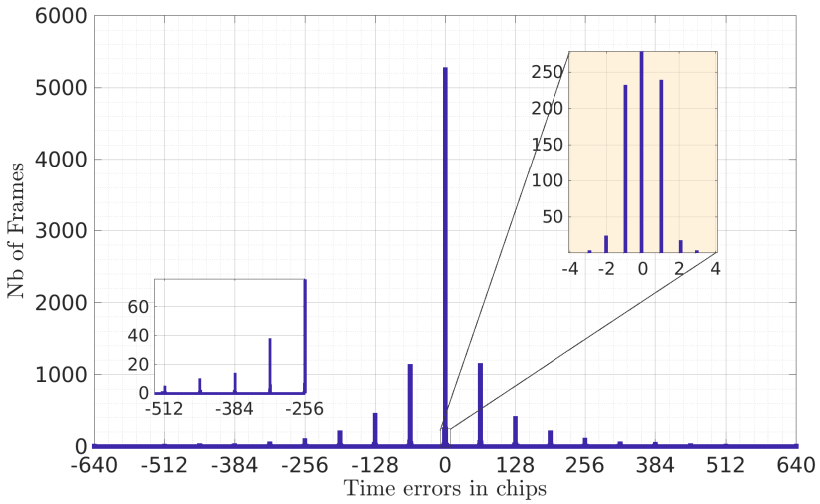


Figure 5.5 – probability distribution of chip synchronization errors obtained with transmission of 10^4 frames.

errors, a delay of length up to 17 CCSK symbols is observed in this simulation. To generalize, we suppose $-\frac{N}{2} \leq s \leq \frac{N}{2}$. The second type of chip synchronization error is small chip offset (between -4 and 4 chips, typically) around each q peak, which is occurring at indices multiple of q (i.e. kq with k is an integer $\in [-N/2, N/2]$). This type of error is defined as the residual chip offsets r, where $-4 \leq r \leq 4$. In summary, the time synchronization error in chip is equal to

$$e = sq + r. \tag{5.2}$$

For solving these synchronization errors, s is estimated by taking profit of the over modulation process, and the residual chip offsets r are found thanks to the non-binary code structure.

5.1.2.2 Detection on the local bin with finer frequency resolution

The detection algorithm performs the frequency detection with a granularity of $\delta_f = \frac{1}{4q}$ Hz, which corresponds to $\delta_\theta = \frac{\pi}{2}$ radian as in section 4.3. It is time to refine the coarse frequency estimation $\hat{f}_c$ given by the detection. The first step of this method is to perform the search of the optimal frequency with a grid of size $\delta_f = \frac{1}{32q}$ instead of $\frac{1}{4q}$ (It is worth noticing here that $\frac{1}{32q}$ can be decreased or increased based on the optimization between

complexity and performance). So, $\hat{f}_o$ is estimated as

$$\hat{f}_o = \arg\max\{S^f_{\hat{n}_a}(\mathbf{y}), f \in \{\hat{f}_c + \frac{i}{32q}\}_{i=-8,\ldots,8}\}. \tag{5.3}$$

Figure 5.6 shows the distribution of the frequency error estimation $f_o - \hat{f}_o$ over 10^4 frames same as the predefined QCSP frame.

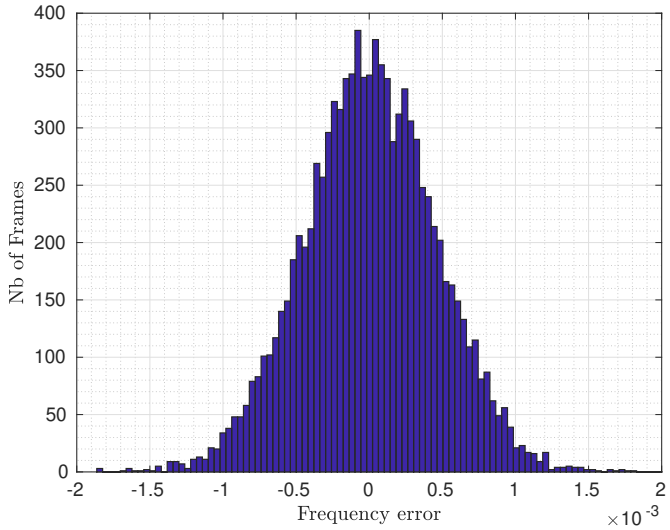


Figure 5.6 – Distribution of frequency error estimation $f_o - \hat{f}_o$ over 10^4 transmitted frames.

We can figure out from Fig. 5.6 that, in the majority of the cases, the frequency estimation error lies between $\pm 10^{-3}$. It corresponds roughly to a full rotation every 8 CCSK symbols ($1/(8q) = 0.97 \times 10^{-4}$). This residual frequency offset is too high for coherent demodulation along with the whole frame. However, it could be sufficient to have good performance in the time synchronization at the symbol level.

5.1.2.3 Symbol synchronization

In order to enhance the synchronization process at the symbol level, a symbol OM, defined in (3.16), has been performed at the transmitter side. The auto-correlation of the over-modulation sequence helps the receiver to estimate correctly the time of arrival of the frame at symbol level (estimation of the parameter s in (5.2)). The OM generates a pre-defined phase pattern (a known sequence of ± 1: 1 no phase change, and -1 (π rotation)) within the sequence of the symbols being transmitted. This phase pattern is

expected to be recovered even when the residual frequency offset causes a rotational effect on the symbols being decoded.

Going back to (4.4), it is possible to add also the phase information to the maximum correlation value $M_n^{\hat{f}_o}$. By defining $d_n^{\hat{f}_o}$ as

$$d_n^{\hat{f}_o} = \arg\max\{|L_n^{\hat{f}_o}(i)|, i = 0, 1, \dots q-1\}, \tag{5.4}$$

we can define $\gamma_n^{\hat{f}_o}$ as

$$\gamma_n^{\hat{f}_o} = (-1)^{b_k} L_n^{\hat{f}_o}(d_n^{\hat{f}_o}), \tag{5.5}$$

and one should note that, by construction, $M_n^f = |\gamma_n^f|$.

Let us determine first the exact value of γ_{n_a+kq} given in (5.5) with the hypothesis that the hard decision d_{n_a+kq} given in (5.4) is correct, i.e. $d_{n_a+kq} = c_k$. In that case, according to (3.19),

$$\gamma_{n_a+kq} = (-1)^{b_k} \sum_{i=0}^{q-1} y(n_a + kq + i) P_{c_k}(i). \tag{5.6}$$

By replacing, $y(n_a+kq+i)$ by its value given in (3.18) taking into consideration the updated frequency synchronization $\hat{f}_o$, we have $\boldsymbol{y}_{n_a+kq}^{\hat{f}_o} = ((-1)^{b_k} P_{c_k}(i) e^{j(2\pi(f_o-\hat{f}_o)(n_a+kq+i)+\phi)} + z(n_a + kq + i))_{i=0,1,\dots,q-1}$. Let the residual frequency offset be $f_r = f_o - \hat{f}_o$. Thus, (5.6) will be

$$\begin{aligned}
\gamma_{n_a+kq}^{\hat{f}_o} &= (-1)^{b_k} e^{j(2\pi f_r kq+\phi)} \sum_{i=0}^{q-1} e^{j2\pi f_r i} P_{c_k}(i)^2 + \sum_{i=0}^{q-1} z(n_a + kq + i) P_{c_k}(i) \\
&= (-1)^{b_k} e^{j(2\pi f_r kq+\phi)} \frac{1 - e^{j2\pi f_r q}}{1 - e^{j2\pi f_r}} + Z_{n_a+kq}, \\
&= (-1)^{b_k} e^{(\omega k+\varphi)} \frac{\sin(\pi f_r q)}{\sin(\pi f_r)} + Z_{n_a+kq},
\end{aligned} \tag{5.7}$$

with $\omega = 2\pi f_r q$ and $\varphi = \phi + \pi f_r(q-1)$. One can note that if $|f_r q| \ll 1$, then $\frac{\sin(\pi f_r q)}{\sin(\pi f_r)} \approx q$ (between 63.56 and 64 for $q = 64$ and $|f_r| \leq 10^{-3}$ for example). Finally, Z_{n_a+kq} represents a realization of an AWGN of zero mean and $\sqrt{q}\sigma$ as standard deviation. In case of wrong decision (i.e. $d_{n_a+kq} \neq c_k$), we have

$$\gamma_{n_a+kq}^{\hat{f}_o} = (-1)^{b_k} e^{j(2\pi f_r kq+\phi)} \sum_{i=0}^{q-1} e^{j2\pi f_r i} P_{c_k}(i) P_{d_k}(i) + Z_{n_a+kq}, \tag{5.8}$$

where $< P_{d_k}, P_{c_k} >$ is close to zero by construction.

For the sake of notation simplicity, we remove the superscript $\hat{f}$ corresponds to frequency in the rest of this subsection. Let us define the $\mathbf{\Gamma}_{n_a}$ as $\mathbf{\Gamma}_{n_a} = (\gamma_{n_a}, \gamma_{n_a+q}, \ldots, \gamma_{n_a+q(N-1)})$. In the absence of wrong decision, the term by term vector multiplication of $\mathbf{\Gamma}_{n_a}$ and $\boldsymbol{B}$ gives,

$$\mathbf{\Gamma}_{n_a} \odot \boldsymbol{B} \approx q e^{j\varphi}(1, e^{j\omega}, e^{j2\omega}, \ldots, e^{j(N-1)\omega}) + \boldsymbol{Z}_{n_a}, \tag{5.9}$$

with $\boldsymbol{Z}_{n_a}$ a vector of N CAWGN samples of zero mean and $\sqrt{q}\sigma$ standard deviation. Fig. 5.7 shows the output of 5.9 at very high SNRs in the absence of wrong decisions. In

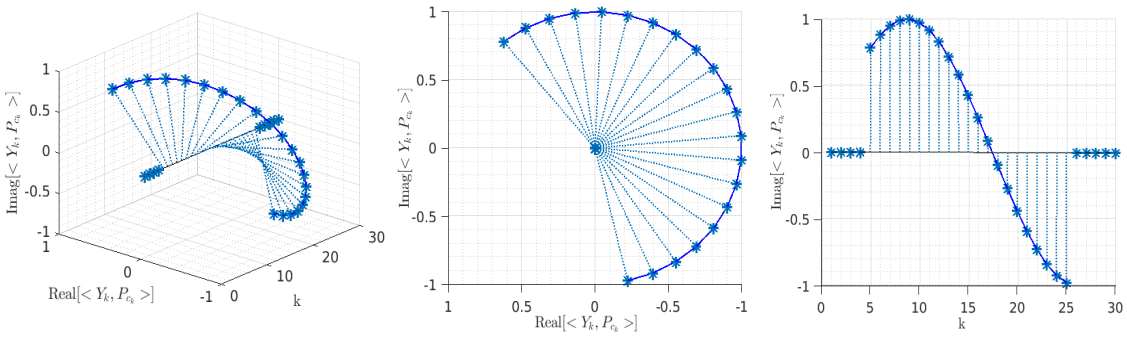


Figure 5.7 – Pattern of the point-by-point multiplication of maximum values of the correlator and $\boldsymbol{B}$ for correct decisions.

summary, in the absence of wrong decision, $\mathbf{\Gamma}_{n_a} \odot \boldsymbol{B}$ is a pure sinusoidal vector of length N and frequency ω affected by AWGN. This property is used to suppress time ambiguity at the symbol level. In fact, the initial time estimation of n_a is affected by s symbol errors and r chip errors. By neglecting the chip errors in this first step of the algorithm, we get $\tilde{n}_a = n_a + sq$. If $s \neq 0$, the vector $\mathbf{\Gamma}_{n_a+sq} \odot \boldsymbol{B}$ does not generate a pure sinusoidal of length N as shown in Fig. 5.8, but it includes a sequence of $N - s$ successive components of $(e^{jk\omega}(-1)^{b_k}(-1)^{b_{k-s}})_{k\in[\max(0,s),\min(N,N+s)]}$ that contains no regular pattern thanks to the choice of $\boldsymbol{B}$. It is thus possible to estimate the value of s by selecting the value $\tilde{s}$ that makes the hypothesis "$\mathbf{\Gamma}_{n_a+sq}$ is a pure sinusoidal affected by noise" the more likely. This estimated value $\tilde{s}$ gives an updated version of the estimation of the arrival time n_a: $\tilde{n}_a = n_a + \tilde{s}q$.

Our first attempt is then to select the value s that maximizes the maximum module of the Fast Fourier Transform of the $\mathbf{\Gamma}_{\hat{n}_a+sq}(s) \odot \boldsymbol{B}$ vector, knowing that $\hat{n}_a$ is the assumption

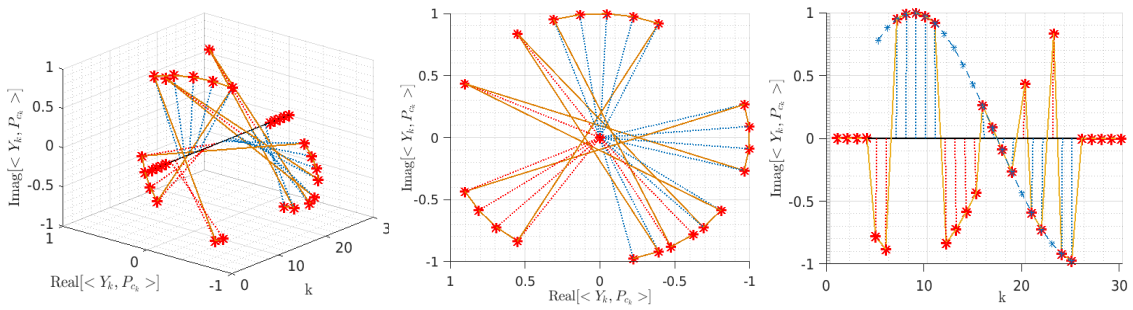


Figure 5.8 – Pattern of the point-by-point multiplication of maximum values of the correlator and $\boldsymbol{B}$ for wrong decisions.

of n_a thanks to step 1. In other words, the ambiguity on symbol position is solved by taking $\tilde{s}$ as

$$\tilde{s} = \underset{s\in[-N/2,N/2]}{\text{argmax}} \{\max\{|\text{FFT}(\boldsymbol{\Gamma}_{\hat{n}_a+sq} \odot \boldsymbol{B})|\}\}. \tag{5.10}$$

This method is simple and suppresses the symbol ambiguity in the large majority of the cases. Nevertheless, when the decisions on several symbols are wrong, the phase of the associated symbols is also wrong, making the proposed method prone to some synchronization errors. To mitigate this problem, we propose to weigh the values of γ_s by a coefficient that indicates the reliability of the decision. Let $d_{n,2}$ be the index of the second-highest decision in (3.19), and $\epsilon_n = L_n(d_{n,2})$ i.e.

$$d_{n,2} = \underset{i=0,1,\ldots,q-1, i\neq d_n}{\text{argmax}} \{|L_n^f(i)|\}. \tag{5.11}$$

The relative ratio R_n between $|\gamma_n|$ and $|\epsilon_n|$ defined as $R_n = \frac{|\gamma_n|-|\epsilon_n|}{|\gamma_n|}$ is a good indicator of the reliability of the decision as shown in Fig. 5.9. For example, $R_n \approx 0$ means that the decisions d_n and $d_{n,2}$ have very close values, thus d_n is not a reliable decision. On the contrary, R_n close to one indicates a very reliable decision, since a higher peak occurred at index d_n due the correlation match. Let $\boldsymbol{A}_n$ be the vector $\boldsymbol{A}_n = (R_n, R_{n+q}, \ldots, R_{n+(N-1)q})$, the Weighted OM (WOM) algorithm is thus given as

$$\tilde{s} = \underset{s\in[-N/2,N/2]}{\text{argmax}} \{\max\{|\text{FFT}(\boldsymbol{A}_{\hat{n}_a+sq} \odot \boldsymbol{\Gamma}_{\hat{n}_a+sq} \odot \boldsymbol{B})|\}\}. \tag{5.12}$$

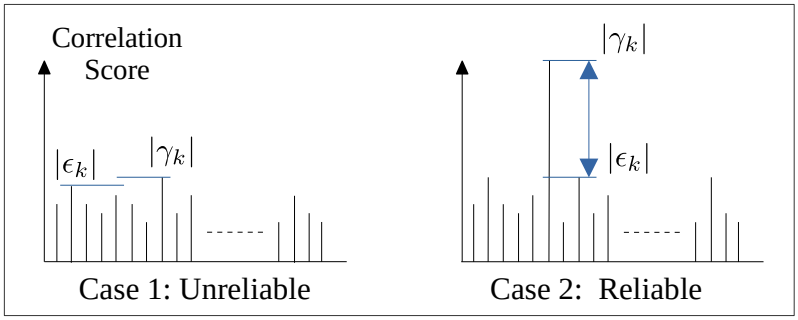


Figure 5.9 – Illustration of the maximum magnitude of correlation values for both wrong (left) and correct (right) CCSK detected frame.

Fig. 5.10 shows the value of $\max\{|\text{FFT}(\boldsymbol{A}_{\hat{n}_a+sq} \odot \boldsymbol{\Gamma}_{\hat{n}_a+sq} \odot \boldsymbol{B})|$, where four received frames are considered with $N = 64$, $q = 64$ at SNR of -10 dB. The initial coarse time estimations $\hat{n}_a$ given by the detection algorithm for the four frames are affected by a synchronization time error $s \times q$ with $s = -9, 0, 2$ and 7 respectively. In each case, (5.12) allows a correct estimation of s.

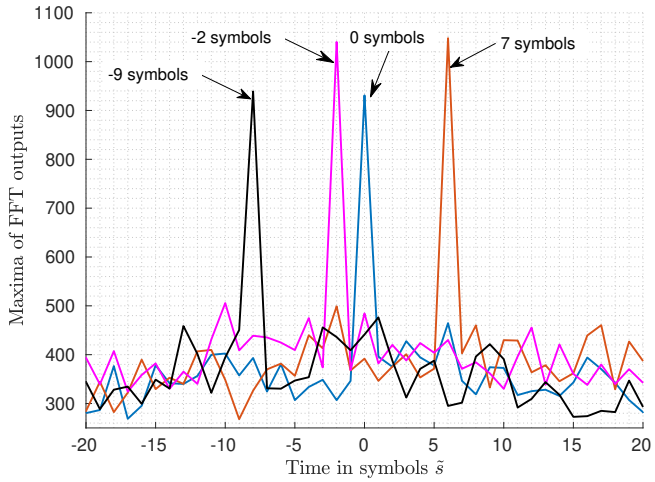


Figure 5.10 – Illustration of equation (5.12) over 4 different received frames, $N = 60, q = 64$, at -10 dB.

Fig. 5.11 shows the probability distribution of the chip synchronization error $\tilde{e} = n_a - \tilde{n}_a$ over 10^4 frames. Compared to Fig. 5.5, the first observation is that all the frames are successfully synchronized at the symbol level thanks to the first processing step. At this end of the synchronization process, more than 90% of the frames are correctly

synchronized in time. The remaining 10% are due to errors of few chips, with $\tilde{e} \in [-4, 4]$. As mentioned before, these chip offsets are solved with the help of the parity checks of the NB-LDPC decoder. This step is presented and discussed in detail in the next step of the synchronization process.

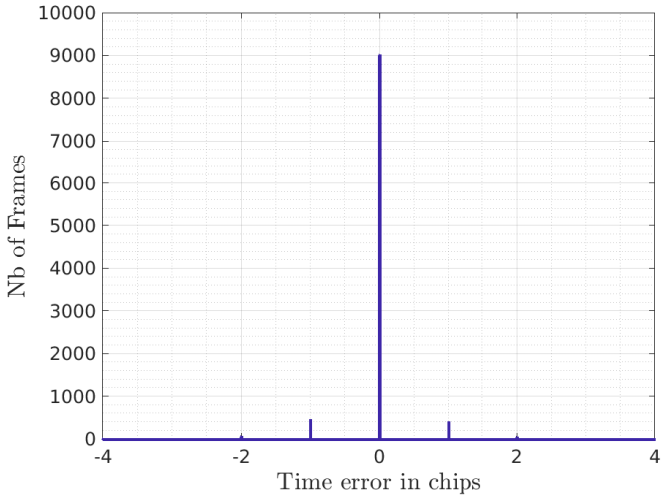


Figure 5.11 – Probability distribution of chip synchronization error $\hat{r}$ obtain with the transmission of 10^4 frames after symbol synchronization process.

To conclude, two hypotheses have been assumed to justify the mathematical model used to determine the best estimation $\tilde{s}$ of the parameter s. The first hypothesis is the fact that $n_a = n_c + sq$, while in fact $n_a = n_c + sq + r$, with r having the distribution given in Fig. 5.11. The second hypothesis is not explicitly formulated but, due to the channel noise, not all the values of d_{n_a+kq}, are correctly estimated, leading to values of γ_{n_a+kq} different than the expression given in (5.7). Nevertheless, MC simulations show that, even if the hypothesis used to justify the mathematical model is not fully exact, the method remains efficient in practice.

5.1.2.4 Finer frequency synchronization using FFT method

After guarantee of good symbol time synchronization, we use a direct method to estimate the remained frequency offset f_r. According to (5.9), the vector $\mathbf{\Gamma}$ is a pure sinusoidal vector of length N, with unknown frequency $\omega/(2\pi)$ and an unknown initial phase φ that is corrupted by an AWGN. The estimation of ω and φ is a classical problem of signal processing. First, ω is estimated as $\bar{\omega}$ thanks to the near maximum likelihood

estimator using Fast Fourier Transform given in [111]. Different near Maximum Likelihood frequency estimators for a sinusoidal wave are also given and compared in the appendix of the deliverable in [112]. This helps us also to have information about the phase offset $\tilde{\varphi}$, consequently the initial phase offset $\tilde{\phi} = \tilde{\varphi} - \pi \tilde{f}_r(q-1)$ as shown in (5.7). It is discussed in detail in the next part of this chapter which corresponds to the Phase synchronization.

Fig. 5.12 gives the histogram of the frequency error estimation $\tilde{f}_o - f_o$ using the Near-ML FFT method after the transmission of 10^4 predefined QCSP frames.

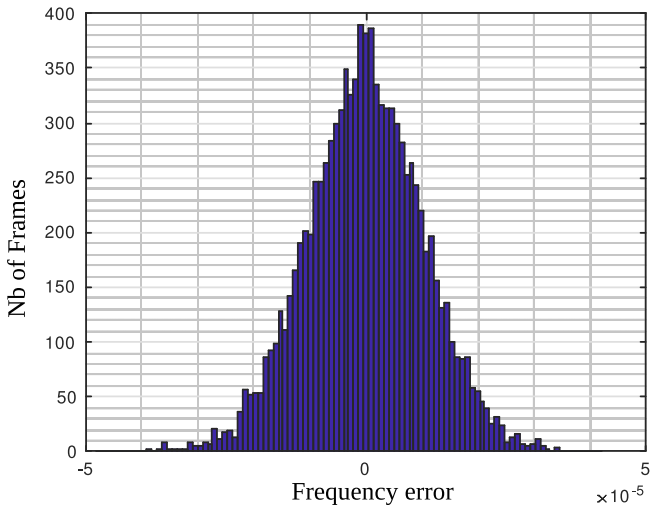


Figure 5.12 – Histogram of the final frequency estimation error after 10^4 transmissions.

.

5.1.2.5 Coded aided fine chip synchronization

The objective of the chip synchronization is to estimate the remaining chip errors $r = n_a - \tilde{n}_a$. To do so, we first adopt the Syndrome-Based (SB) time synchronization presented in [79] [70] to the proposed QCSP frame. Then, we improve this method by replacing the hard metric, which is the check node syndrome calculation, with a soft Variable Node-Based (VNB) metric.

At first, it is important to describe some features of the NB-LDPC code defined by the parity check matrix $\boldsymbol{H}$ of $N-K$ rows and N columns. The code is assumed, as an example for illustration, to be regular with weights $d_v = 2$ and $d_c = 3$ (code rate 1 - 2/3 = 1/3). Let $\mathcal{M}(j)$ be the set of $d_c = 3$ non-null positions of the j^{th} row of $\boldsymbol{H}$. In other words, $\mathcal{M}(j)$ represents the set of variable nodes (a, b, c) connected to check node j as in Fig. 5.13.

The j^{th} parity check equation for input decision vector $\boldsymbol{D}_n = (d_n, d_{n+q}, \ldots, d_{n+(N-1)q})$ is defined as

$$\sum_{i \in \mathcal{M}(j)} h(j,i) D_n(i) = 0. \tag{5.13}$$

In the absence of decoding error, $\boldsymbol{D}_{n_a}$ is a codeword and thus, all the parity checks are fulfilled ($\boldsymbol{D}_{n_a}$ verifies $\boldsymbol{D}_{n_a} \times \boldsymbol{H}^T = 0$). Also, if the number of detection errors is low enough, only few parity checks will not be fulfilled. On the other side, if $n \neq n_a$, $\boldsymbol{D}_n$ can be seen as a random vector, and thus, the number of non-verified parity checks will be in average close to $N - K$ since each check node has 1 chance over q to be fulfilled. Let us define $\mathrm{NoZ}_{\boldsymbol{H}}(\boldsymbol{D}_n)$ (Number of Zero) the function that counts the number of satisfied parity checks (or number of zeros in the syndrome) when $\boldsymbol{D}_n$ is the intrinsic decisions and $\boldsymbol{H}$ is the PCM. Consequently, the syndrome-aided chip synchronization, over multi hypotheses $r \in [-\frac{q}{8}, \frac{q}{8}]$), is given as

$$\hat{r} = \underset{r \in [-\frac{q}{8}, \frac{q}{8}]}{\operatorname{argmax}} \{\mathrm{NoZ}_{\boldsymbol{H}}(\boldsymbol{D}_{\tilde{n}_a + r})\}. \tag{5.14}$$

This method is efficient as long as the number of decoding errors is low enough. Nevertheless, it is not necessarily always the case and thus, this method fails sometimes. To mitigate this problem, we propose to replace the SB method seen at the check node level with a VNB method treating a "soft syndrome" seen at the variable node level. The idea is to perform one decoding iteration of the code with the hard decision vector $\boldsymbol{D}_n$. This decoding iteration generates $d_v = 2$ check to variable messages for each variable. The message $M_{j \to i}$ sent by check node j to variable node i message is defined as in [113]

$$M_{j \to i}(\boldsymbol{D}_n) = h(j,i)^{-1} \sum_{i' \in \mathcal{M}(j), i' \neq i} h(j,i') D(i'), \tag{5.15}$$

and shown in Fig. 5.13.

Let $\boldsymbol{X}, \boldsymbol{X}_2$ and $\boldsymbol{X}_3$ be three $\mathrm{GF}(q)$ vectors of length N. Let us define the score function $G(\boldsymbol{X}, \boldsymbol{X}_2, \boldsymbol{X}_3)$ as

$$G(\boldsymbol{X}, \boldsymbol{X}_2, \boldsymbol{X}_3) = \sum_{j=0}^{N-K} \sum_{i \in \mathcal{M}(j)} f(M_{j \to i}(\boldsymbol{X}), X(i), X_2(i), X_3(i)), \tag{5.16}$$

where $f(m, x, x_2, x_3)$ is a function of $\mathrm{GF(q)}^4$ to the real number that associate a value 1 if

$m = x$, 0.9 if $m = x_2$, 0.8 if $m = x_3$, and 0 otherwise. The new proposed method is thus:

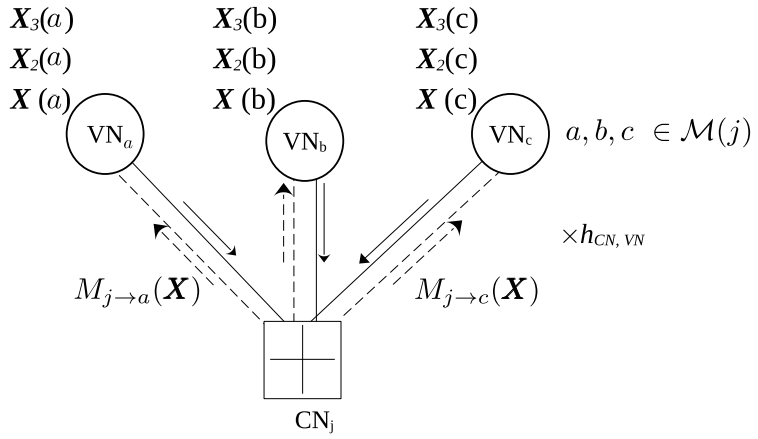


Figure 5.13 – Illustration of (5.16) for one CN example.

$$\hat{r} = \underset{r\in[-\frac{q}{8},\frac{q}{8}]}{\operatorname{argmax}}\{G(\boldsymbol{D}_{\tilde{n}_a+r},\ \boldsymbol{D}_{\tilde{n}_a+r,2},\ \boldsymbol{D}_{\tilde{n}_a+r,3})\}, \tag{5.17}$$

where $\boldsymbol{D}_{n,2} = (d_{n,2},\ d_{n+q,2}, \ldots,\ d_{n+(N-1)q,2})$, and $\boldsymbol{D}_{n,3} = (d_{n,3},\ d_{n+q,3}, \ldots,\ d_{n+(N-1)q,3})$ is the set of the third most reliable decisions, i.e.

$$d_{n,3} = \underset{i=0,1,\ldots,q-1, i\notin\{d_n,d_{n,2}\}}{\operatorname{argmax}} \{|L_n(i)|\}. \tag{5.18}$$

The addition of the values 1, 0.9, 0.8 and 0 corresponds to the following analysis. In case of correct synchronization (i.e. correct hypothesis r) without any decision error in $\boldsymbol{D}_{\tilde{n}_a+r}$, the output of the check nodes $\boldsymbol{M}_{j\to i}(\boldsymbol{D}_{\tilde{n}_a+r})$ equal $\boldsymbol{D}_{\tilde{n}_a+r}$. If $\boldsymbol{D}_{\tilde{n}_a+r}$ contains some errors, most probably the outputs of parity checks may have some elements in $\boldsymbol{D}_{\tilde{n}_a+r,2}$, and less probable in $\boldsymbol{D}_{\tilde{n}_a+r,3}$. In worst case scenarios, the output is neither element of the three different vectors. It is worth noticing here that in the case of good synchronization and with no errors on the first decisions $\boldsymbol{D}_{\tilde{n}_a+r}$, the output of the score function G is: $G(\boldsymbol{D}_{n_a},\ \boldsymbol{D}_{n_a,2},\ \boldsymbol{D}_{n_a,3}) = (N-K)\times 3$. Note that this method is simple and efficient, but more elaborated methods can also be proposed as future work.
Fig. 5.14 illustrates the use of equation (5.17) by applying it over 4 independent received QCSP frames affected by the chips errors -3, 0, 0 and 3, respectively. Then, in each of the cases, (5.17) solve the problem.

Algorithm 1 summarizes the Code-aided chip synchronization method.

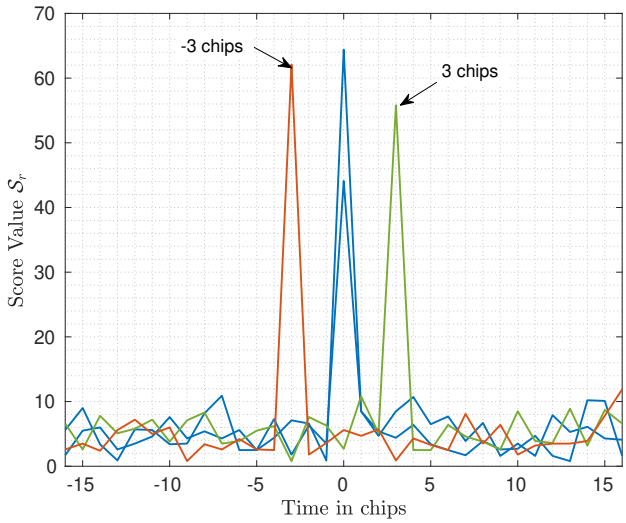


Figure 5.14 – Illustration of (5.17) over 4 different cases.

After this chip synchronization process, we found out that, at SNR = -10.25, all the 10^4 received QCSP frames have been perfectly synchronized.

Algorithm 1: Chip Synchronization method.

Input: $\boldsymbol{H}$, Intrinsic information $\boldsymbol{D}_{\tilde{n}_a+r}$, $\boldsymbol{D}_{\tilde{n}_a+r,2}$ and $\boldsymbol{D}_{\tilde{n}_a+r,3}$
Initialization;
for *hypothesis* $r \in [-\frac{q}{8}, \frac{q}{8}]$ **do**
    $G(r) = 0$; Initialization of Score function:
    foreach $PC\,\mathrm{CN}_j$ *where* $j \in [0, N-K-1]$ **do**
        Check node CN_j Processing:
        foreach $i \in \mathcal{M}(j)$ **do**
            $M_{j\to i}(\boldsymbol{D}_{\tilde{n}_a+r}) = h(j,i)^{-1} \sum_{i' \in \mathcal{M}(j), i' \neq i} h(j,i')\; d_{\tilde{n}_a+r}(i')$;
            Calculation of Score function
            if $M_{j\to i}(\boldsymbol{D}_{\tilde{n}_a+r}) = d_{\tilde{n}_a+r}(i)$ **then**
                $G(r) = G(r) + 1$;
            else if $M_{j\to i}(X) = d_{\tilde{n}_a+r,2}(i)$ **then**
                $G(r) = G(r) + 0.9$;
            else if $M_{j\to i}(X) = d_{\tilde{n}_a+r,3}(i)$ **then**
                $G(r) = G(r) + 0.8$;
        end
    end
end
Output: $\hat{r} = \mathrm{argmax}_{r\in[-\frac{q}{8},\frac{q}{8}]}\{G(r)\}$; Best hypothesis.

5.1.3 Results curve

This section presents the simulation results of the QCSP receiver implemented using the different combinations of the symbol synchronization methods (OM, WOM) and the chip synchronization methods (SB, VNB). The MC simulations are run over an AWGN channel with stopping criteria of 100 frames of error, NB-LDPC encoder with coding rate $R_c = 1/3$, $q = 64$, and time and frequency shifts considered to be uniform randomly-distributed. The blue curve in Fig. 5.15 shows the miss-detection probability $\mathcal{P}_{\text{md}}$ of the blind detection method as presented in chapter 4, where the time and frequency bins being considered are $\ell = q/4$ and $\delta_\theta = \pi/2$, with a probability of false alarm $P_{fa} = 10^{-6}$. The second step is to feed the synchronization block with the frames being successfully detected and then assess the synchronization performance. Note that the frames being processed contain a residual frequency error bounded by $\pm 10^{-5}$. The solid red curve shows the

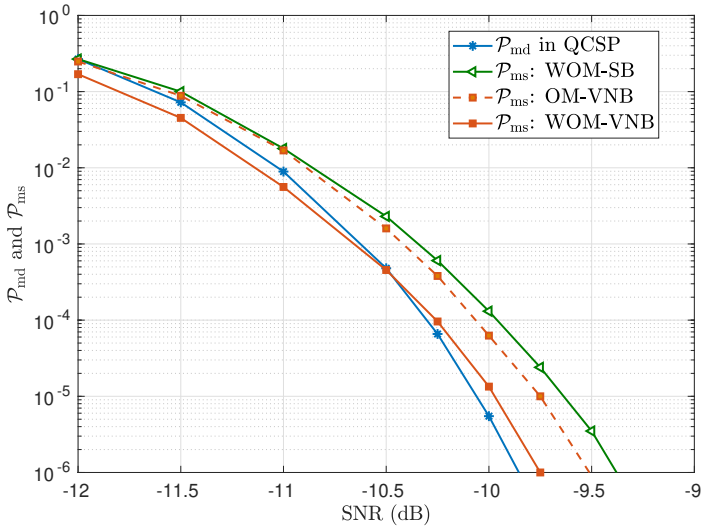


Figure 5.15 – $\mathcal{P}_{\text{md}}$ and $\mathcal{P}_{\text{ms}}$ vs SNR for the QCSP receiver using several combinations of the symbol and chip synchronization methods.

probability of miss-synchronization $\mathcal{P}_{\text{ms}}$ obtained using the WOM-VNB method, where a gain of 0.25 dB is obtained with respect to the dashed red curve representing the OM-VNB case. This gain shows the impact of the weighting technique in the OM method. When comparing the WOM-VNB (red curve) to WOM-SB (green curve), a gain of 0.5 dB is noticed which shows the efficiency of the VNB technique as compared to the SB one. To the best of our knowledge, this is the first blind-synchronization algorithm that reaches a $\mathcal{P}_{\text{ms}}$ of the order of 10^{-5} at very low SNR (-10 dB).

Fig. 5.16 shows the type of the time synchronization errors in both WOM-SB and WOM-VNB. We can see that the VNB method decreases the errors at the chip level. Consequently, the types of error at symbol and chip-level are balanced, where the overall time miss synchronization probability is enhanced.

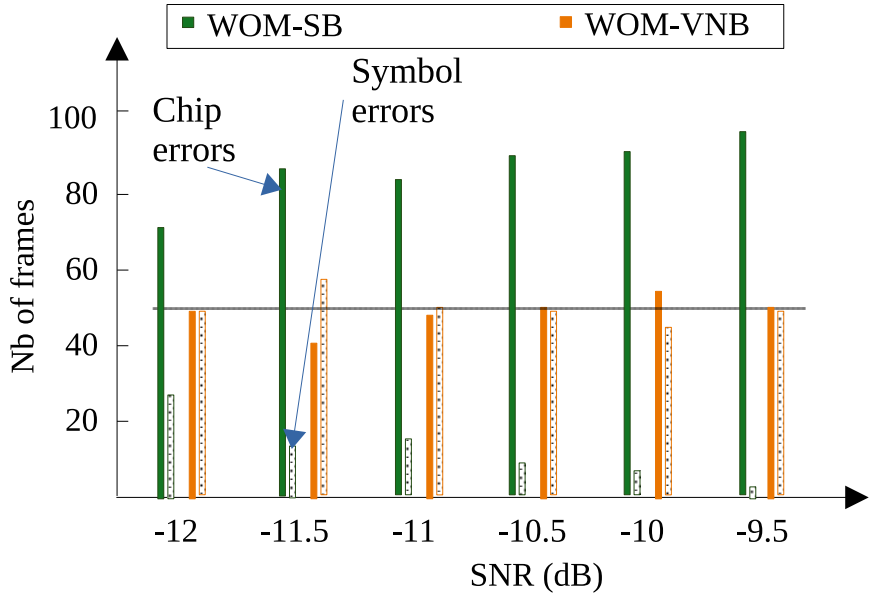


Figure 5.16 – Type of the time synchronization errors: at chip level or symbol level.

5.2 Phase synchronization

In this section, we propose a phase and frequency synchronization technique in the context of CCSK modulation and NB-LDPC codes. Two methods are proposed: the first method, called Direct Method (DM), assumes that all the CCSK symbols are demodulated correctly. In that case, the frequency and phase estimation task resumes to an estimation problem of a pure complex sinusoidal signal affected by Gaussian noise. It is a well-known problem and a near ML estimator that can be obtained for low cost using the FFT [111]. Nevertheless, at low SNR (-10 dB typically), detection errors appear. Those errors alleviate the quality of the DM and generate up to 0.5 dB of degradation on the overall performance. The second method is called Parametric Method (PM). It is based on the ML estimation using a parametric PDF of each phase error. The two parameters of the proposed PDF model are computed using information coming from the non-coherent CCSK demodulation process and information coming from one decoding iteration of the NB-LDPC decoder. The PM is simple to process and gives a result close to the Genius Aided (GA) method, i.e. the DM when all the transmitted symbols are considered to be known. Through this section, the problem statement of the phase synchronization is first discussed. After that, the direct phase synchronization method is described. The proposed PM phase synchronization aided by the CCSK and NB-LDPC association is then illustrated in detail. Finally, a conclusion is drawn up.

5.2.1 Problem statement

In the wireless communication process, detection and time synchronization is not the end of the story. There still exists the phase offset which has a big impact on the generation of the LLRs which is fed to the NB-decoders.

Assuming perfect time and frequency synchronization, the generation of the Log-Likelihood vectors, required by the NB-LDPC decoder, is based on the computation that uses the real part of each of the symbols correlation factor $\boldsymbol{L}_{kq}(i)$ in (3.21) (see [89] for more details). The subscript kq, which represents the index at each symbol level, can be replaced by k for the clarity and simplicity of notations. So, the first step for generating the LLRs is by finding the real part of each of the elements of $\boldsymbol{L}_k$

$$L_k^R(i) = \mathcal{R}(L_k(i)). \tag{5.19}$$

where $\mathcal{R}(x)$ represents the real part of complex x. Then, from (5.19) the LLRs sent to the decoder are normalized as

$$L_k^N(i) = \frac{1}{\sigma^2}(M_k^R - L_k(i)), \tag{5.20}$$

with M_k^R defined as the maximum value of the vector $\boldsymbol{L}_k^R$, i.e.

$$M_k^R = \max\{L_k^R(i), i = 0, 1, \ldots, q-1\}. \tag{5.21}$$

Once the LLRs are computed, the decoding process can start. Assume that a residual phase error ξ affect the correlation vector $\boldsymbol{L}_k = (L_k(i))_{i=0,1,\ldots,q-1}$. The real part of $\boldsymbol{L}_k$ have an amplitude reduced by a factor $\cos(\xi)$ compared to the magnitude $|\boldsymbol{L}_k|$. This amplitude degradation is translated directly in a $10\log_{10}(\cos(\xi)^2)$ dB loss of signal energy, and thus the SNR. Fig. 5.17 shows the degradation of SNR as a function of angle ξ in radian for

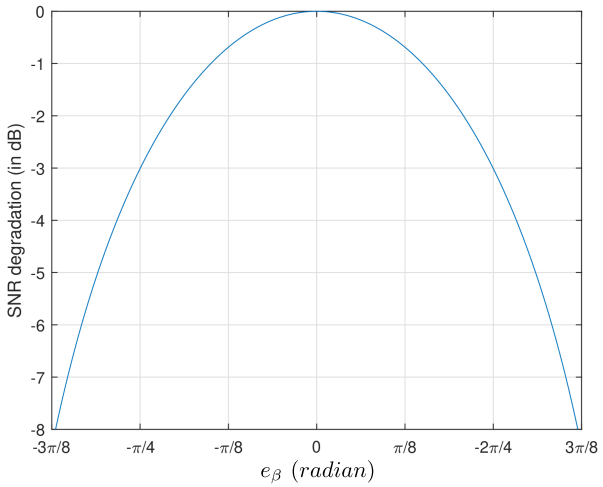


Figure 5.17 – Degradation of SNR (in dB) due to the phase estimation error ξ.

more clarification.

It is also possible to express the phase error ξ as the function of SNR, $\xi(\text{snr}) = \arccos\sqrt{10^{\text{SNR}/10}}$, i.e. to determine the angle required for a given SNR degradation, as shown in Fig. 5.18.

To set the idea, $\xi = 0, \pi/64, \pi/16, \pi/9.3$ and $\pi/4$ generates an SNR degradation of 0 dB, 0.01 dB, 0.16 dB, 0.5 dB and 3 dB, respectively. When $\xi = \pi/2$, no more signal is received. The frequency synchronization task should thus maintains the residual phase error close to zero not to impact significantly the receiver performance.

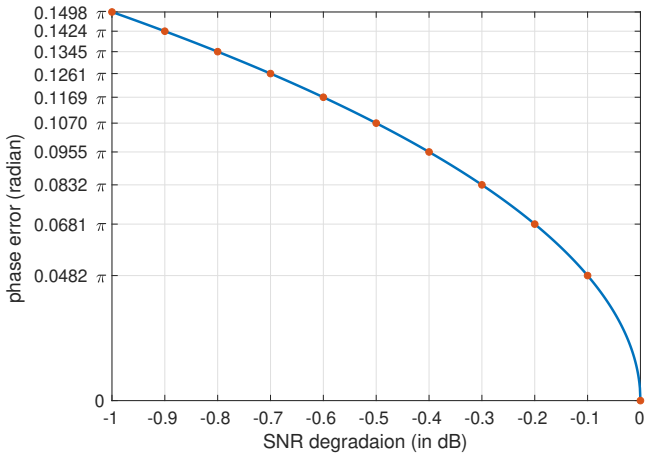


Figure 5.18 – Error phase required for a given level of SNR degradation.

5.2.2 Phase offset in the QCSP frame: theoretical study

Before starting, the output of equation (5.7) is redefined taking into consideration the time synchronization process and the over-modulation sequence. Thus, for the correct decision, i.e. when $d_k = c_k$, γ_k is defined as

$$L_k(c_k) = e^{j(\omega k+\varphi)} \frac{\sin(\pi f_r q)}{\sin(\pi f_r)} + Z_k. \tag{5.22}$$

In the absence of noise (i.e. $Z_k = 0$), the phase of $L_k(c_k)$ is equal to $\Theta_k = \omega k + \varphi$. Note that Θ_k is the equation of a straight line in the phase domain with initial value φ and a slope $\omega = 2\pi f_r q$, as shown in the black solid line of Fig. 5.19. However, in case of wrong decision, we have

$$\gamma_k = e^{j(\omega k+\phi)} \sum_{\ell=0}^{q-1} e^{j2\pi f_r \ell} P_{c_k}(\ell) P_{d_k}(\ell) + Z_k. \tag{5.23}$$

Since by construction, $< P_{d_k}, P_{c_k} >$ is close to zero and considering also that $|q f_r| \ll 1$ (a reasonable assumption for $q = 64$ and $|f_r| < 10^{-3}$), then the first term of (5.23) can be considered as negligible. In that case, γ_k has a random phase given by the phase of the noise term Z_k of (5.23).

Let $\mathbf{\Gamma}$ denotes the vector containing the γ_k values that are coming either from (5.22) or (5.23), i.e. $\mathbf{\Gamma} = (\gamma_k)_{k=0,1,\ldots,N-1}$. Let us define $\mathbf{\Psi} = (\Psi_k)_{k=0,1,\ldots,N-1}$, where Ψ_k is the

phase of γ_k for $k \in 0, 1, \ldots, N-1$. We can notice from (5.22) that Ψ_k is modeled as

$$\Psi_k = \Theta_k + \xi_k, \tag{5.24}$$

with ξ_k is the phase error between the exact phase Θ_k of $L_k(c_k)$ in the absence of noise and the measured phase Ψ_k of $L_k(d_k)$.

Fig. 5.19 illustrates with an example the impact of the channel on the phase of $N = 60$ detected symbols of a QCSP frame. The channel is an AWGN with SNR of -10 dB, the residual frequency offset is $f_r = -3.8507 \times 10^{-5}$ and the initial phase is $\phi = 0.5152$. Two different sets of phases can be observed.

1. The phase Ψ_k of correctly detected CCSK symbols ($d_k = c_k$) that are correlated with Θ_k, and represented by the points of the gray area in Fig. 5.19.
2. The phases Ψ_k of wrongly detected CCSK symbols ($d_k \neq c_k$) represented by the shaded red circles. According to (5.23), those phases are independent from the real phase Θ_k.

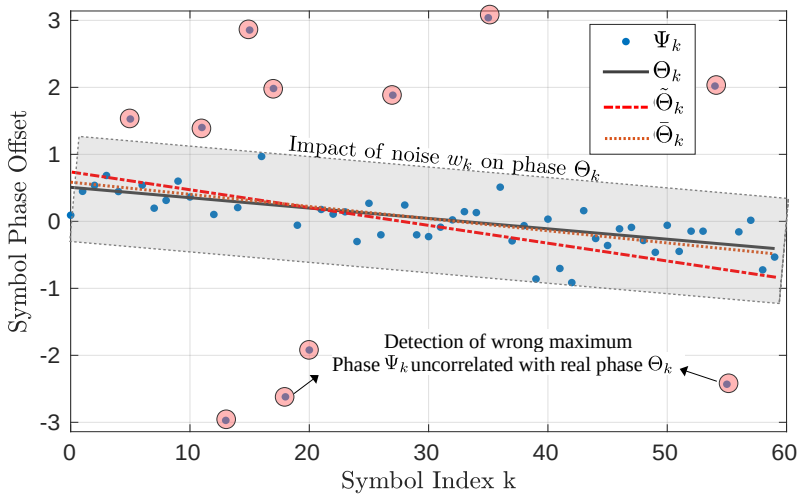


Figure 5.19 – Phase offsets effect due to the channel Ψ_k, real phase offsets Θ_k, direct method estimation $\tilde{\Theta}_k$ and parametric method $\bar{\Theta}_k$.

Therefore, the PDF of the phase error f_ξ can be expressed as the composition of the following two distributions:

1. Distribution of the phase error of correctly detected CCSK symbols f_{ξ_c}: the phase error distribution of these symbols can be modeled as the PDF of the normal

distribution $N(0, \rho)$; see Fig. 5.20 (blue curves), where ρ can be numerically calculated, i.e.

$$f_{\xi_c}(x) = (1/\rho\sqrt{2\pi}) \exp(-\frac{x^2}{2\rho^2}).^{1} \tag{5.25}$$

2. Distribution of the phase error of wrongly detected CCSK symbols f_{ξ_w}, $d_k \neq c_k$: the distribution of the phase error of these symbols can be modeled as the PDF of the uniform distribution over the interval $[-\pi, \pi]$, i.e.

$$f_{\xi_w}(x) = \frac{1}{2\pi} \quad \text{for} \quad x \in [-\pi, \pi], \tag{5.26}$$

and 0 otherwise, see Fig. 5.20 (red curves).

Let us assume that $\mathcal{P}_0$ is the probability of error in the CCSK demodulation. Thus, $f_\xi(x)$ is equal to $f_{\xi_c}(x)$ with a probability of $1-\mathcal{P}_0$ and is equal to $f_{\xi_c}(x)$ with a probability $\mathcal{P}_0$, thus, in average,

$$\begin{aligned} f_\xi(x) &= (1-\mathcal{P}_0) f_{\xi_c} + \mathcal{P}_0 f_{\xi_w} \\ &= \frac{1-\mathcal{P}_0}{\rho\sqrt{2\pi}} \exp(-\frac{x^2}{2\rho^2}) + \frac{\mathcal{P}_0}{2\pi}, \end{aligned} \tag{5.27}$$

when $x \in [-\pi, \pi]$, 0 otherwise. This distribution is shown in Fig. 5.20b.

As an empirical verification of the previous theoretical analysis, a MC simulation of 10^6 CCSK symbols c_k at SNR = -10 dB has been considered. Fig. 5.20a illustrates the experimental phase distribution $f_\xi(x)$ (black curve) as well as its components $(1-\mathcal{P}_0) f_{\xi_c}$ (blue curve) and $\mathcal{P}_0 f_{\xi_w}$ (red curve). Fig. 5.20b illustrates the theoretical counterparts.

5.2.3 Phase synchronization with Direct Method (DM)

In this section, we present the Direct Method to estimate the phase offset. According to (5.22), the vector $\mathbf{\Gamma}$ is a pure sinusoidal vector of length N, with unknown frequency $\omega/(2\pi)$, unknown initial phase φ that is corrupted by an AWGN. The estimation of ω and φ is a classical problem of signal processing. First, ω is already estimated as $\tilde{\omega}$ thanks to the near maximum likelihood estimator using Fast Fourier Transform given in [111].

1. Formally, $f_{\xi_c}(x)$ should be defined over $[-\infty + \infty]$; since $\rho << \pi$, $f_{\xi_c}(x)$ will be given at 0 outside $[-\pi, \pi]$ and the required normalization factor is approximated to 1.

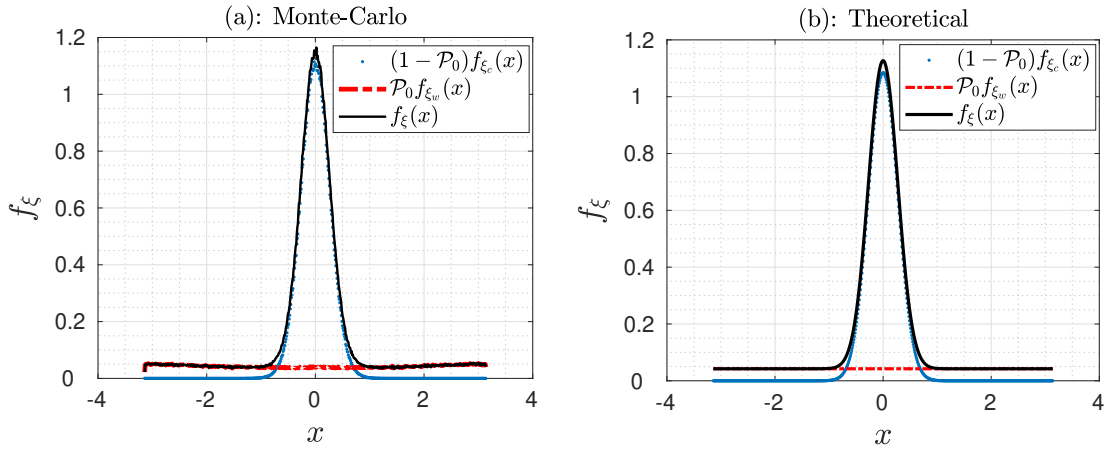


Figure 5.20 – Distribution of ξ_k at -10 dB where $\mathcal{P}_0 = 0.2375$ and $\rho = 0.32$, (a): MC-simulation, (b): Theoretical.

Then, from the estimated value $\tilde{\omega}$, the estimate of the initial phase $\tilde{\varphi}$ is given as

$$\tilde{\varphi} = \text{Phase}\left(\sum_{k=0}^{N-1} \gamma_k e^{-j\tilde{\omega}k}\right). \tag{5.28}$$

From $\tilde{\omega}$ and $\tilde{\varphi}$, the phase of the k^{th} symbol is estimated as $\tilde{\Theta}_k = \tilde{\omega}k + \tilde{\varphi}$. The values of $\tilde{\Theta}_k,\ k = 0, 1, \ldots, N-1$ are shown by the red dashed line in Fig. 5.19. One can note that, by construction, the difference between the estimated phase $\tilde{\Theta}_k$ and the real phase Θ_k is maximized for $k = 0$ and/or $k = N-1$.

The GA method for the frequency and initial phase estimation is similar to the DM method, except that $\mathbf{\Gamma} = (L_k(d_k))_{k=0,1,\ldots,N-1}$ is replaced by the vector $(L_k(c_k))_{k=0,1,\ldots,N-1}$, i.e. all decisions are assumed to be correct. The GA method gives the optimal estimations $\hat{\omega}$ and $\hat{\varphi}$ of ω and φ respectively. The corresponding phase for the k^{th} symbol is thus $\hat{\Theta}_k = \hat{\omega}k + \hat{\varphi}$.

To compare the performance of the DM and GA method, 10^4 QCSP frames of length $N = 60$ over GF(64) have been transmitted at an SNR of -10 dB, with each frame affected by a random frequency offset between $[-10^{-3}, 10^{-3}]$ and a random phase offset. For each received frame, ω and φ are estimated. Fig. 5.21a shows a 2D plot where each point corresponds to a received frame. The y-axis shows the initial estimation error $\Theta_0 - \hat{\Theta}_0$ obtained from the received QCSP frame with the GA method. Similarly, the x-axis represents the corresponding final estimation error $\Theta_{N-1} - \hat{\Theta}_{N-1}$. Fig. 5.21b gives the same result for the DM method. In those figures, the black circle is the boundary for

phase error equals 0.336 radian which corresponds to an SNR degradation smaller than 0.5 dB inside the circle, greater outside.

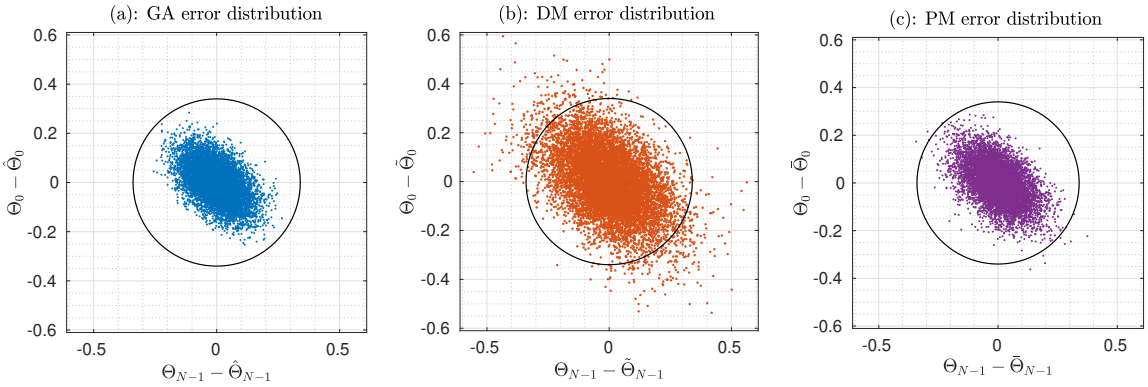


Figure 5.21 – QCSP phase estimation errors distribution for different methods. The circle corresponds to the boundary of SNR degradation above 0.5 dB.

The gap in performance between the DM and GA methods is clear, where the former is not reliable and introduces a loss of 0.5 dB (and even more than 1 dB in some cases). The next section shows a novel approach that overcomes the problem and allows to get performance close to the GA performance.

5.2.4 Parametric Method (PM) for phase estimation

This section proposes a new method, named PM, for estimating the phase offset. This method is based on two steps. First, a model of the PDF of each phase error is derived. Then, a ML estimator is applied to find the parameters $(\bar{\omega}, \bar{\varphi})$ that "explains the best" the observation.

5.2.4.1 ML estimation method

In statistics, after some observed data is given, ML estimation is a method that estimates the parameters of an assumed probability distribution. This is achieved by maximizing a likelihood function so that, under the assumed statistical model, the observed data is the most probable. The point in the parameter space that maximizes the likelihood function is called the ML estimation. In QCSP system, ML estimator can be applied to

estimate the frequency and phase parameters $(\bar{\omega}, \bar{\varphi})$ of (ω, φ) in (5.22) as the following

$$(\bar{\omega}, \bar{\varphi}) = \arg\max_{\omega,\varphi} \Big(\prod_{k=0}^{N-1} f_{\xi_k}(\Psi_k - \omega k - \varphi) \Big), \tag{5.29}$$

where $\Psi_k = \text{Phase}(\gamma_k)$ as defined in (5.24) and f_{ξ_k} the PDF associated to the phase error of the k^{th} detected symbol. In practice, (5.29) is evaluated on a discrete 2-dimensional grid, one dimension for $\Theta_0 = \varphi$, another for ω. The grid resolution of $\Theta_0 = \varphi$ is $\pi/32$, which corresponds to a maximum error of quantization of $\pm\pi/64$, i.e. a maximum of $10\log_{10}(\cos(\pi/64)^2) = 0.01$ dB of SNR degradation. Since φ varies from $-\pi$ to π, the associated grid contains 64 values. For the grid of ω, it is more convenient to specify the final phase Θ_{N-1}, since $\omega = (\Theta_{N-1} - \Theta_0)/N$ according to (5.28). In the worst cases, $\theta = +\pi$ (respectively $-\pi$) and $f_r = +10^{-3}$ (respectively, -10^{-3}), leading to the final phase Θ_{N-1} inside the interval $[-\Theta^m, \Theta^m]$ with $\Theta^m = \pi + 2\pi(N-1)q f_r = 8.5\pi$. With the same resolution as Θ_0, the Θ_{N-1} grid contains $(2 \times 8.5) \times 32 = 544$ elements. Altogether, the 2-dimensional grid contains $64 \times 544 = 3.5 \times 10^4$ elements. This very high number can be drastically reduced in practice. For example, it is possible to search the optimal solution with a greedy algorithm stating from the parameters given by the DM.

The quality of the PM thus relies on the quality of the PDFs f_{ξ_k}, $k = 0, 1, \ldots, N-1$ derived in (5.27). This PDF depends directly on the two parameters: the probability of detection error $\mathcal{P}_0$ and the variance ρ of the phase error of the correct symbol. In the next two sections, we propose first a parametric estimation based on the CCSK decision process. Then the quality of the parameters' estimation is improved by taking the profit of the information computed thanks to the NB-LDPC code.

5.2.4.2 Dependency of f_ξ on CCSK score ratio R

In (5.11), the parametric R_k is considered a good indicator of the reliability of the decision. For example, $R_k = 0$ means that the decisions d_k and $d_{k,2}$ have same reliability, thus d_k is not a reliable decision. On the contrary, R_k close to one indicates a very reliable decision. This intuitive observation is confirmed by a MC simulation done over 10^8 QCSP frames of length $N = 60$, $q = 64$ at and SNR of -10 dB. The probability of error $\mathcal{P}_0$ is estimated as a function of R. As shown in Fig. 5.22, $\mathcal{P}_0$ decreases from 0.83 when $R = 0$ down to 10^{-5} when R get closed to 0.62. The variance ρ has smaller variation effect when conditioned to R. It decreases from 0.31 when $R = 0$ down to 0.21 for $R = 0.6$.

Also, Fig. 5.23 shows examples of PDF $f_{\xi/R}$ for several values of parameters R, where

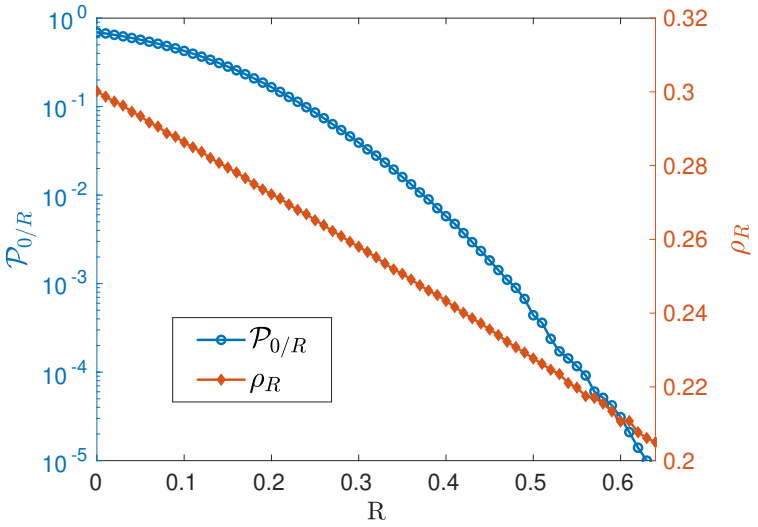


Figure 5.22 – $\mathcal{P}_{0/R}$ and ρ_R as function of R, at SNR = -10 dB.

it clearly shows the effect of the value R on the reliability of the different PDFs. In the sequel, $\mathcal{P}_{0/R}(R)$ and $\rho_R(R)$ denote the PDF parameters values of $f_{\xi/R}$ conditioned to the observed value of R.

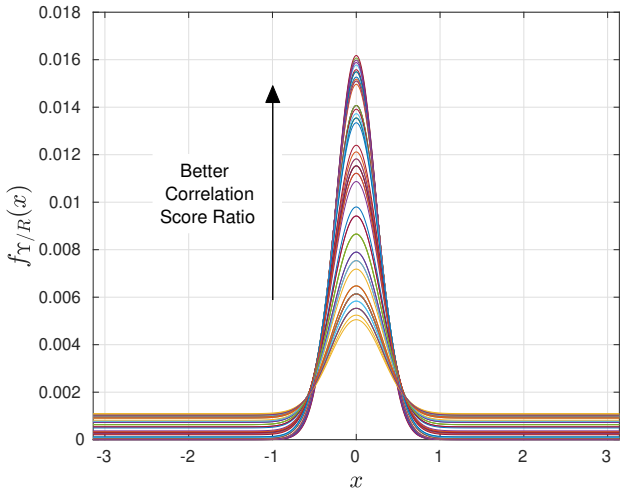


Figure 5.23 – $f_{\xi/R}$ depending on the CCSK ratio R at SNR = -10 dB.

5.2.4.3 Dependency of f_ξ on the NB-LDPC code

The probability of error $\mathcal{P}_{0/R}$ derived in the previous section can be further enhanced thanks to the information given by the NB-LDPC code, giving a new estimate denoted $\mathcal{P}_{0/R-code}$. For a PCM $\boldsymbol{H}$ of $N-K$ rows and N columns, the j^{th} PC equation for a vector $\boldsymbol{D} = (d_0, d_1, \ldots, d_{N-1})$ is defined as in (5.13). The first decoding iteration generates $\mathrm{d}_v = 2$ check to variable messages for each variable node, as in (5.15. To lighten notation, M_1 and M_2 denote the two check to variable node i messages defined in (5.15). Consequently, at each variable node i, there exist three equality tests $(d = M_1), (d = M_2)$ and $(M_1 = M_2)$. These equalities gives in total 8 possible configurations from which it is possible to enhance or decrease the reliability of the decision d, and thus, the estimation of the probability of error $\mathcal{P}_{0/R-code}$. For example, when all the equalities are fulfilled, the local decision is consistent with the code structure and $\mathcal{P}_{0/R-code}$ becomes very low. On the contrary, when $(d \neq M_1), (d \neq M_2)$ and $(M_1 = M_2)$, the local decision is inconsistent with the code structure and $\mathcal{P}_{0/R-code}$ increases. The exact mathematical determination of $\mathcal{P}_{0/R-code}$ as a function the equality test results is illustrated in the probability tree in Fig. 5.24.

Given $\mathcal{P}_d, \mathcal{P}_{M_1}$ and $\mathcal{P}_{M_2}$ are the probability of errors for the messages d, M_1 and M_2 respectively, taken to account the score information. For each variable node v_k have d as an intrinsic information and $M1, M2$ the messages coming from check nodes, we have the following information:

— Two parity checks fulfilled when $d = M_1 = M_2$.
— One parity check fulfilled: $d = M_1 \neq M_2$ or $d = M_2 \neq M_1$.
— No parity checks fulfilled: $d \neq (M_1 = M_2)$ or $d \neq M_1 \neq M_2$.

Consequently, according to Fig. 5.24, the probability of error $\mathcal{P}_{0/R-\mathrm{code}}$ knowing the NB-coding properties can be defined as:

— If two Parity check fulfilled $(d = M_1 = M_2)$,

$$\mathcal{P}_{0/R-\mathrm{code}} = \frac{\mathcal{P}_d \mathcal{P}_{M_1} \mathcal{P}_{M_2}}{\mathcal{P}_d \mathcal{P}_{M_1} \mathcal{P}_{M_2} + 63^2 (1-\mathcal{P}_d)(1-\mathcal{P}_{M_1})(1-\mathcal{P}_{M_2})}. \tag{5.30}$$

— If one parity check fulfilled and $(d = M_1 \neq M_2)$,

$$\mathcal{P}_{0/R-\mathrm{code}} = \frac{63 \times \mathcal{P}_d \mathcal{P}_{M_1}(1-\mathcal{P}_{M_2}) + 62 \times \mathcal{P}_d \mathcal{P}_{M_1} \mathcal{P}_{M_2}}{63 \times \mathcal{P}_d \mathcal{P}_{M_1}(1-\mathcal{P}_{M_2}) + 62 \times \mathcal{P}_d \mathcal{P}_{M_1} \mathcal{P}_{M_2} + 63^2 (1-\mathcal{P}_d)(1-\mathcal{P}_{M_1})\mathcal{P}_{M_2}}. \tag{5.31}$$

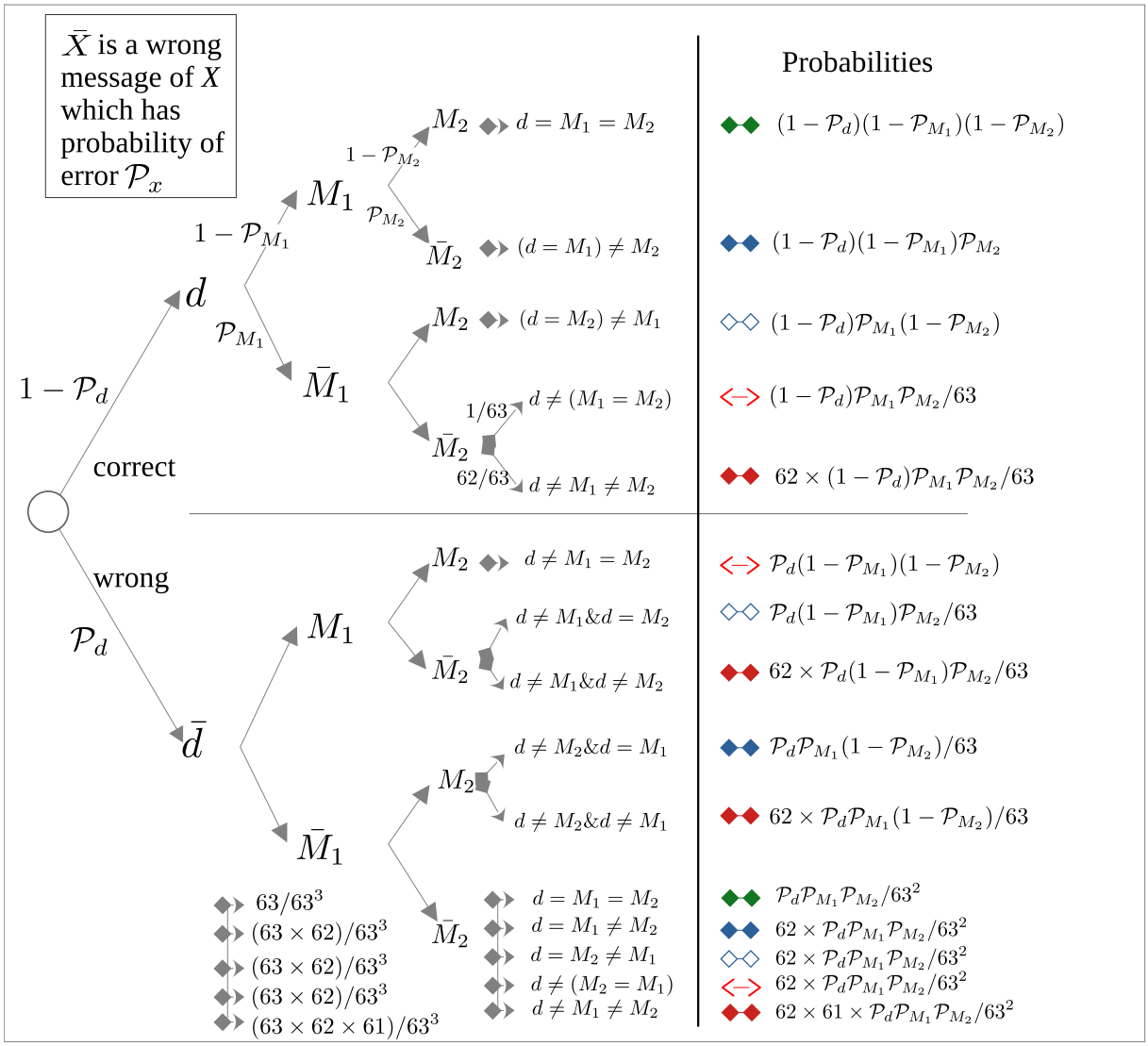


Figure 5.24 – Probability of error for all scenarios of VN for $d_v = 2$.

— If one parity check fulfilled and $(d = M_2 \neq M_1)$,

$$\mathcal{P}_{0/R\text{-code}} = \frac{63 \times \mathcal{P}_d(1-\mathcal{P}_{M_1})\mathcal{P}_{M_2} + 62 \times \mathcal{P}_d\mathcal{P}_{M_1}\mathcal{P}_{M_2}}{63 \times \mathcal{P}_d(1-\mathcal{P}_{M_1})\mathcal{P}_{M_2} + 62 \times \mathcal{P}_d\mathcal{P}_{M_1}\mathcal{P}_{M_2} + 63^2(1-\mathcal{P}_d)\mathcal{P}_{M_1}(1-\mathcal{P}_{M_2})}. \tag{5.32}$$

— No parity checks fulfilled and $d \neq (M_1 = M_2)$,

$$\mathcal{P}_{0/R\text{-code}} = \frac{63^2 \times \mathcal{P}_d(1-\mathcal{P}_{M_1})(1-\mathcal{P}_{M_2}) + 62 \times \mathcal{P}_d\mathcal{P}_{M_1}\mathcal{P}_{M_2}}{63^2 \times \mathcal{P}_d(1-\mathcal{P}_{M_1})(1-\mathcal{P}_{M_2}) + 62 \times \mathcal{P}_d\mathcal{P}_{M_1}\mathcal{P}_{M_2} + 63(1-\mathcal{P}_d)\mathcal{P}_{M_1}\mathcal{P}_{M_2}}. \tag{5.33}$$

— No parity checks fulfilled and $d \neq M_1 \neq M_2$,

$$\frac{63\times 62[\mathcal{P}_d(1-\mathcal{P}_{M_1})\mathcal{P}_{M_2}+\mathcal{P}_d\mathcal{P}_{M_1}(1-\mathcal{P}_{M_2})]+62\times 61\times\mathcal{P}_d\mathcal{P}_{M_1}\mathcal{P}_{M_2}}{63\times 62[\mathcal{P}_d(1-\mathcal{P}_{M_1})\mathcal{P}_{M_2}+\mathcal{P}_d\mathcal{P}_{M_1}(1-\mathcal{P}_{M_2})+(1-\mathcal{P}_d)\mathcal{P}_{M_1}\mathcal{P}_{M_2}]+62\times 61\times\mathcal{P}_d\mathcal{P}_{M_1}\mathcal{P}_{M_2}}. \tag{5.34}$$

Finally, the probability of error $\mathcal{P}_{0/R-\text{code}}$ is updated in the PDF distribution, used for the ML estimation in (5.29), based on the comparison between the different equalities. Fig. 5.25 shows examples of PDF $f_{\xi/R-code}$ for several values of the equality tests to the side of parameters R. It can be figured out that probability of having error close to zero is higher when R is increased and the equality tests are satisfied.

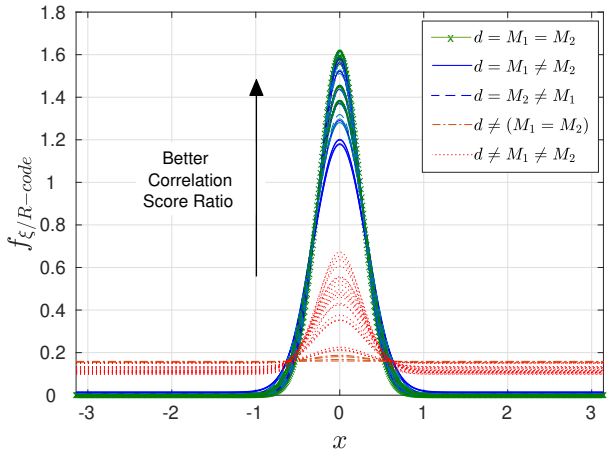


Figure 5.25 – $f_{\xi/R-code}$ depending on the CCSK ratio R and Code-information, at SNR = -10 dB.

This conditioned PDF $f_{\xi/R-code}$ is used to determine parameters $(\bar{\omega}, \bar{\theta})$ in (5.29). Values of $\bar{\Theta}_k,\ k = 0, 1, \ldots, N-1$ are shown as an example by the brown dot line in Fig. 5.19. Also, Fig. 5.21c shows the distribution of the phase offset errors $\bar{\xi}_k$, using the PM phase estimator, for $k = 0$ and $k = N-1$ after transmitting 10^4 QCSP frames in an AWGN channel of SNR = -10 dB and random phase offsets. We can figure out that PM aided by the information of the QCSP frame approaches the GA method.

5.2.5 Simulation results

This section presents the simulation results of the FEC probability of error for a QCSP received frame, of $N = 60$ symbols, exposed to random phase offsets. This is implemented using the different phase synchronization scenarios (no phase synchronization, DM, PM)

and the ideal case of FEC where the frame has no phase offset. The MC simulations are run over an AWGN channel with stopping criteria of 100 miss-corrected frames, NB-LDPC encoder with coding rate $R_c = 1/3$, $q = 64$, initial phase offset randomly distributed in $[-\pi, \pi]$, and frequency shifts considered to be uniform randomly-distributed in the boundary of $\pm 10^{-3}$ (based on what we obtain from time synchronization process). Moreover, the GF(64)-LDPC decoder defined in [114] is implemented using the EMS algorithm with 30 decoding iterations and $n_m = 20$ [115]. It is seen that a NB-LDPC decoder clearly looses its good performance when not caring about the phase offset, and looses up to 0.5 dB when using the DM. However, when using the proposed PM phase estimator (the blue curve) approximately maintains the same performance of the FEC $\mathcal{P}_\epsilon$ as when no phase offset does exist.

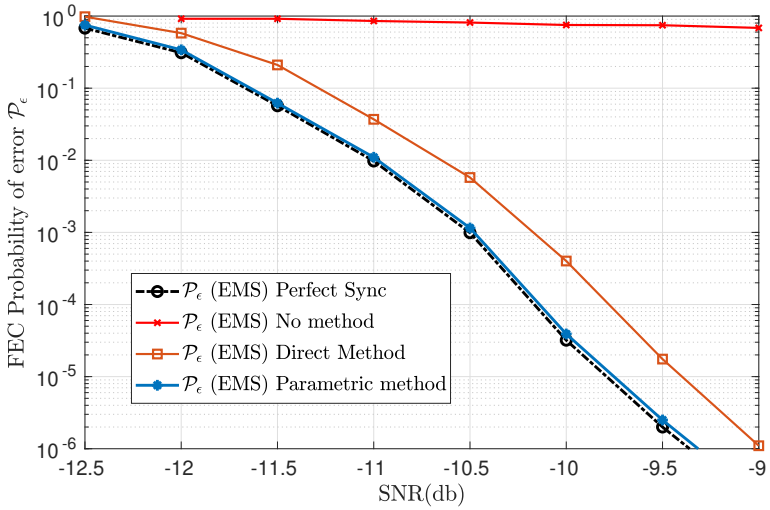


Figure 5.26 – FEC with different scenarios.

5.3 Conclusion

This chapter succeeds to synchronize a QCSP frame at very low SNR (typically -10 dB) without any additional header or footer data, thanks to the CCSK modulation and NB-LDPC coding. First it presented a preamble-less time-synchronization algorithm for QCSP frames operating at -10 dB with a probability of miss-synchronization of the order of 10^{-5}. The proposed algorithm mitigates the time ambiguity in two consecutive steps. The first step is at the symbol level where an OM technique has been proposed and then

enhanced with a weighted version privileging the symbols received with high reliability. The second step is at the chip level by taking advantage of the NB-LDPC code structure. Then, it presented a phase synchronization algorithm for QCSP frames operating also at -10 dB that can help in maintaining the FEC performance as if no phase offset does exist. The proposed PM mitigates the phase ambiguity based on the ML method aided by two side information from the QCSP structure. The first information is from the CCSK demodulation score ratios, where the ML method used a weighted version by privileging the symbols received with high reliability. The second information is from the NB-LDPC code properties, thanks to the code structure.

In the next chapter, we find the normal approximation bound thanks to Polyanskiy's equations in [49] (an estimated Shannon's limit for small packet size). Then we trade off the Detection, synchronization, and FEC probabilities, and find out the performance using the joint probabilities. Consequently, we find out how much a QCSP frame is far from the Polyanskiy's bound.

Chapter 6

QCSP System Performance

Contents

The QCSP frame aims to avoid the use of preambles in short packet transmission for IoT in an unslotted ALOHA protocol. This will increase the spectral efficiency by avoiding the preamble overhead. In previous chapters, we first defined the frame structure, which is based on the association of CCSK modulation and NB-LDPC code. Then, we developed algorithms that can achieve blind detection and self-synchronization for NB coded CCSK short frames, without the use of preambles. In this chapter, we synthesize the previous results with the decoding performance of a NB-LDPC code. Consequently, the QCSP frame is considered a preamble-less frame that can be viewed as a preamble for the detection and synchronization process, and as an encoded codeword carrying the transmitted message.

In the first section, we recall the theoretical Polyanskiy's bound (an estimated Shannon's limit for small packet size) thanks to the normal approximation given for QCSP frame [49]

[116]. This bound is considered as reference with respect to which we see how far we are from the optimal solution. We also give the error correction performance of the NB-LDPC decoder in an ideally synchronized AWGN channel. In the next section, we define the overall probability of reception, i.e. the joint probability of detection, synchronization, and decoding. We start then to define the problem of optimization between those three different probabilities; the robustness of the overall communication performance is directly affected by the weakest one. In the first step, a detection-correction trade-off is analyzed based on the detection results obtained and the error correction rates using the NB-LDPC decoder. Subsequently, we set comparison with up-to-date codes used for short packet transmission. In the next step, we trade off the synchronization performance with the outcome of detection-correction performance. All the probabilities will be affected directly by the QCSP parameters (q, N, R_c, code structure, spreading sequences, etc.). So the question of what is the best way to manage the parameters for having the optimal chain will still be an open question for future work. Finally, a proof of concept is done through the implementation of the proposed detection and synchronization methods on a real-radio system based on SDR transceivers.

6.1 Polyanskiy's bound and CCSK-NB-LDPC decoder

In this section, we recall the error-correction performance and give the upper bound limit that can be reached by a QCSP frame [116] based on Polyanskiy's equation in [49]. This can help us to know how far we are from the optimal performance solution. The maximum achievable coding rate, denoted by R_c^* (defined in (2.2)), for error correction codes with error probability $\mathcal{P}_\epsilon$ can be tightly approximated [49] by

$$R_c^* \approx R - \sqrt{\frac{V}{N}} Q^{-1}(\mathcal{P}_\epsilon), \tag{6.1}$$

where R is the channel capacity (maximum rate achievable in the asymptotic regime), V is the channel dispersion (defined in [49]) and Q^{-1} the inverse Q function where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{u^2}{2}\right) du$. We use the above approximation (known as the *normal approximation*) as a definition of the maximum achievable coding rate in the finite code-length regime. In [49] the channel dispersion parameter is defined as

$$V = H_2(U|Y) - H(U|Y)^2, \tag{6.2}$$

where $H(U|Y)$ is the conditional entropy of the channel input U given the channel output Y, and

$$H_2(U|Y) \triangleq \mathbb{E}_Y \left[-\sum_{s \in \mathbb{Z}_q} \Pi(s) (\log_q(\Pi(s))^2 \right], \tag{6.3}$$

where $\Pi(s) \triangleq \mathcal{P}(U = s|Y)$ denotes the conditional probability distribution of U given Y, i.e. $\Pi(s) = \frac{e^{L(s)}}{\sum_s e^{L(s)}}$. Hence, $H_2(U|Y)$ can be conveniently estimated by MC simulation.

So we can use (6.1) to find the corresponding minimum probability of error that can be reached at each SNR and that is defined as

$$\mathcal{P}_\epsilon^* = Q\left(-\frac{R_c - R}{\sqrt{V/N}} \right). \tag{6.4}$$

Let us consider a QCSP frame over Galois field order 64, with frame of $N = 120$ symbols (as in [49], normal approximation is tight when the blocklength is greater than or equal to 100). So for $R_c = 1/3, 1/2$ and $3/5$ we have 40, 60 and 72 information GF(2^6) symbols respectively. We also assume a perfectly synchronized reception ($\Delta = 0$, $\theta = 0$). Fig. 6.1 (dashed lines) shows the evolution of $\mathcal{P}_\epsilon^*$ as a function of the SNR for several values of R_c (from right $R_c = 3/5$ to left $R_c = 1/3$), over the Galois Field order $q = 64$.

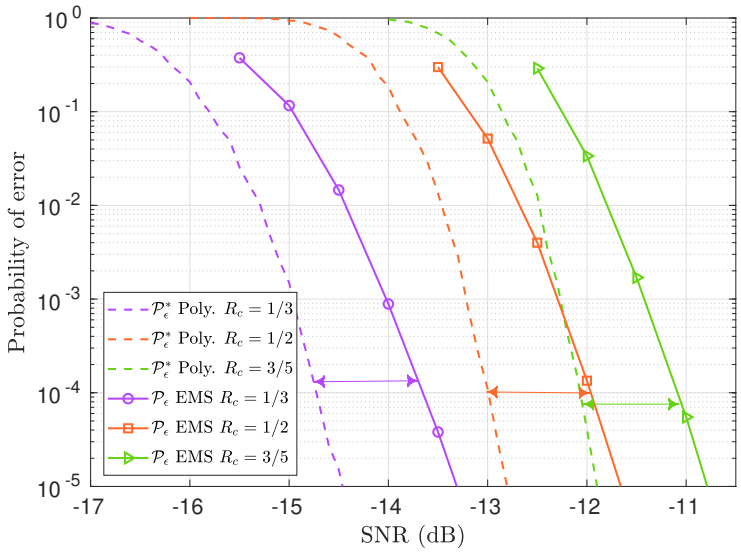


Figure 6.1 – Polyanskiy's bound ($\mathcal{P}_\epsilon^*$), and NB-LDPC (EMS) decoding error ($\mathcal{P}_\epsilon$) for QCSP Frame of $N = 120$ Symbols, and different R_c.

Moreover, the solid curves in Fig. 6.1 represent the probability of error $\mathcal{P}_\epsilon$ obtained

with the GF(2^6)-LDPC code as defined in [114] for the same previous parameters. The decoding algorithm used is the EMS with 30 decoding iterations and $n_m = 20$ (see [115] for the definition of the EMS algorithm and the Parity check matrix being used).

6.2 Overall probability at the receiver side

At very low SNRs, the successful transmission of short frames, as targeted by the NB-code and CCSK association in the QCSP system, is a challenging problem. The overall joint probability of successful transmission in an asynchronous ALOHA system can be expressed as $\mathcal{P} = \mathcal{P}_d \times \mathcal{P}_s \times \mathcal{P}_c$, where $\mathcal{P}_d$ is the probability of detection of the frame, $\mathcal{P}_s$ is the probability of correct estimation of the synchronization parameters, and $\mathcal{P}_c$ is the probability of correction of all transmission errors by the NB-code. This can be expressed as a chain of three connected links as in Fig. 6.2. Thus, the strength of this chain is affected directly by the weakest link between them. For example, let us suppose for a particular case, the probability of correct detection and synchronization is very high, while the probability of correct decoding is low. Consequently, the output of the overall joint probability, in this case, will be low since one of the links of the chain is weak. To conclude, aiming to maximize the probability of successful transmission, we

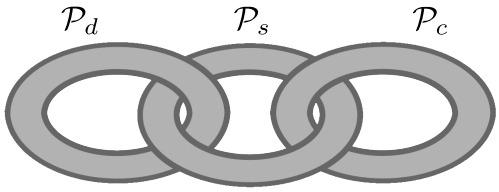


Figure 6.2 – Reception probabilities chain.

must maximize the probability of detection, synchronization, and decoding. For more clarification, Frame Error Rate (FER) is defined as

$$\text{FER} = \mathcal{P}_{\text{md}} + (1 - \mathcal{P}_{\text{md}})\mathcal{P}_{\text{ms}} + (1 - \mathcal{P}_{\text{md}})(1 - \mathcal{P}_{\text{ms}})\mathcal{P}_{\epsilon}, \tag{6.5}$$

where $\mathcal{P}_{\text{md}}$, $\mathcal{P}_{\text{ms}}$ and $\mathcal{P}_{\epsilon}$ are the miss detection, miss synchronization and the error in correction probabilities respectively.

In the following sections, we will discuss this point through a particular example. As for the first stage, we will assume perfect synchronization parameters can be found ($\mathcal{P}_{\text{ms}} = 0$) and study the detection-decoding trade-off for both synchronous and asyn-

chronous AWGN channels. Then, the outcome results will be compared with an up-to-date classical preamble-based frame using modern binary error control code for short packet transmission. After that, we will trade off the blind synchronization algorithm performance with the detection-correction outcome results to study the construction problem of an optimal QCSP frame for short packet transmission.

6.2.1 Detection-correction trade-off

In this part, we follow the strategy proposed by Popovski in [117] that deals with the trade-off between detection and decoding. To give a better illustration, a practical case study is given in Fig. 6.3 that shows the simulation results of the FER of a QCSP frame of $N = 120$ symbols and $q = 64$, in both synchronous and asynchronous complex AWGN channels. FER is considered as the joint effect of miss detection and decoding error probability through a MC simulation. The threshold value U_0 is chosen corresponding to $\mathcal{P}_{\text{fa}} = 10^{-6}$. According to the earlier discussion on detection performance in section 4.3, in the asynchronous channel, we choose to limit the maximum deviation to $q/16 = 4$ chips and the maximum frequency offset to $\theta = \pi/4$ at the receiver side. This is achieved by adjusting the bin size appropriately to $(\delta_\theta = \pi/2, \ell = q/8)$ for uniformly distributed random frequency and time offsets. For the decoding performance, we assume that the synchronization parameters can be perfectly found after the detection process. It is worth noticing here that the MC simulations of the whole system give performances that match (6.5), considering the synchronization process after the detection is perfect ($\mathcal{P}_{\text{ms}} = 0$), i.e. $\text{FER} = \mathcal{P}_{\text{md}} + (1 - \mathcal{P}_{\text{md}})\mathcal{P}_\epsilon$.

As can be seen in Fig. 6.3, for $R_c = 1/2$ and $N = 120$ symbols, the gap between the simulated FERs and the Polyanskiy's bound is around 1.2 dB, i.e., FER = 10^{-4} at SNR = -11.80 dB.

Note also here that using an EMS decoder of a code rate $R_c = 1/3$ is useless since the system will have an over-decoding performance with an overall FER performance curves limited by the detection performance $\mathcal{P}_{\text{md}}$. Symmetrically, if the system uses an EMS decoder of a coding rate in this case of $R_c = 3/5$, we will obtain an over-detection performance, with a FER performance limited by the decoding performance $\mathcal{P}_\epsilon$. For this reason, we have chosen $R_c = 1/2$ as an appropriate rate for the NB-code in the aforementioned scenario.

So for a given payload, finding the optimal QCSP structure (code rate, q size also) that minimizes FER for obtaining the best detection-correction, is an interesting topic

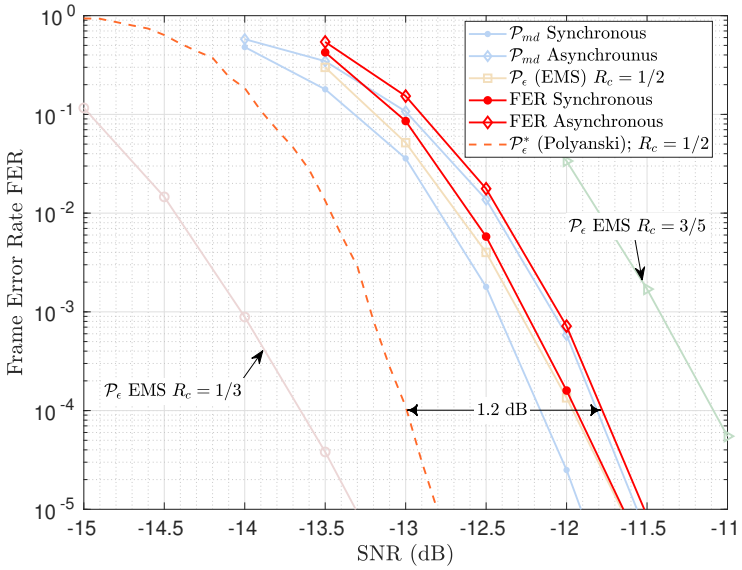


Figure 6.3 – Joint FER due to $\mathcal{P}_{\epsilon}$, $\mathcal{P}_{\text{md}}$ and to $\mathcal{P}_{\text{fa}} = 10^{-6}$ for $N = 120$ symbols, where $R_c = 1/2$, in synchronous and asynchronous CAWGN channel.

that is still an open problem to be addressed in the near future. However, as primarily results in a synchronized channel for a fixed value of N, the method to find the best compromise is rather simple. It consists to assess the performance of miss detection, then selecting the coding rate K/N so that $\mathcal{P}_{\epsilon} \simeq \mathcal{P}_{\text{md}}$.

6.2.2 Comparison with a classical preamble-based frame

In this section we will use the QCSP results obtained in the previous section for $K = 60$ symbols, i.e. $m = 360$ bits of payload, and $R_c = 1/2$ for the comparison. To do this comparison with up-to-date codes, we build an adhoc solution taking elements from the 5G-LDPC 3rd Generation Partnership Project (3GPP) standard in a synchronized channel. The preamble is composed of a length $p = 793$ symbols thanks to Zadoff–Chu sequence. This is the minimum length required to guarantee a probability of misdetection of 10^{-4} with a probability of false alarm of 10^{-6} at SNR of -11.95 dB (result obtained by MC simulation with a fully synchronized preamble, i.e. with perfect time synchronization ($\Delta = 0, f_o = 0$)). For the error correction scheme, the LDPC code, with the rate 1/3 and $k = 360$, of the 3GPP standard is used. This code requires a SNR of 0.2 dB to obtain a FER of 10^{-4} [118]. The transmission of 17 repetitions of encoded frame gives a FER 10^{-4} at SNR of $0.2 - 10\log_{10}(17) = -12.10$ dB. The encoded frame is thus of size

$\frac{360}{2} \times 3 \times 17 = 9180$ QPSK symbols. Subsequently, the total frame length with a classical solution should be equal to $793 + 9180 = 9973$ symbols. To summarize, the size of the QCSP sequence is 7680 ($60 \times 2 \times 64$), while a frame with the classical method requires 9973 symbols. Thus, using the QCSP scheme, reduced the frame size by 22.98% $\approx$ 23% as shown in the schematic of Fig. 6.4. This 23% translates directly into an increase in the wireless channel capacity and energy saving for the wireless sensors.

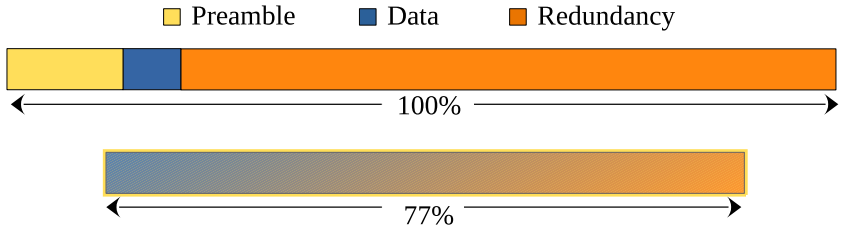


Figure 6.4 – Classical vs. preamble-less proposed approach (QCSP) for transmitting a frame.

Moreover, it is worth mentioning that the QCSP frame has an additional significant advantage compared to the classical frame: the length of the CCSK sequences is 64, whereas the length of the preamble is 793. This implies that the preamble's based frame is 793/64 = 12.39 more sensitive to a frequency offset than the QCSP frame. This ratio of sensitivity is translated directly into the number of parallel filters (frequency bin size δ_θ in the time-frequency grid) required to test the different frequency offset hypotheses.

6.2.3 Detection-synchronization-correction trade-off

In this section, we add the blind synchronization performance to the detection and correction trade-off, where the same frame of $N = 120$ CCSK symbols is considered. The considered coding rate is also $R_c = \frac{1}{2}$ and GF order is $q = 64$. For the detection algorithm, a grid of bin size ($\delta_t = q/8$, $\delta_\theta = \pi/2$) is considered. After applying the blind time-synchronization method discussed in Chapter 5, we can see from Fig. 6.5 that a gap appears between the $\mathcal{P}_{\text{ms}}$ from the first side and $\mathcal{P}_{\text{md}}$ and $\mathcal{P}_{\epsilon}$ from the other side. Consequently, the overall probability will not be balanced, and it will be limited and affected directly by the weakest performance (i.e. the synchronization performance).
It is worth noticing here that all the existing time-synchronization errors are at the chip level, i.e. we have 100 errors from 100 ranging between -4 and 4 chips. Consequently, the problematic step in the synchronization is the VNB method which is then responsible for

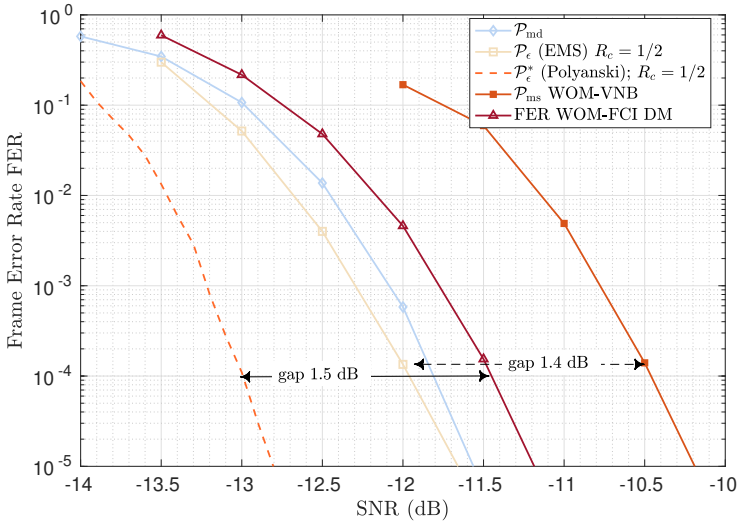


Figure 6.5 – $\mathcal{P}_{\text{ms}}$ from the WOM-VNB blind synchronization method is added.

the synchronization at the chip level and not the WOM.

To remove this gap, we make full decoding iterations at each chip offset hypothesis, then, we choose the hypothesis which obtains a codeword. This method is called Forward Correction Iterations (FCI). Note that the phase estimation is calculated at each of the hypotheses using the DM. The output probabilities show that this gap decreased to less than 0.5 dB from the detection performance. In Fig. 6.5, we show the FER of the received QCSP frame using WOM-FCI in the curve with triangular marks. The QCSP frame can be received with FER = 10^{-4} at -11.5 dB, distanced just by 1.5 dB from Polyanskiy's bound.

We are not deciding that this is the optimal solution. In particular, many questions can be opened to enhance this work and develop it. Should we add a cyclic prefix or special side information of very few chips that help the synchronization algorithm? Should we use a "classical" BPSK spreading sequence or something different? Or, should other NB-decoders as turbo or polar codes give better insights to the synchronization side? This is still really a tricky problem that will be studied soon. In addition to it, approaching the joint detection-synchronization-correction bound utilizing low-complexity coding schemes is also an interesting topic.

6.3 Proof of concept: GNU radio demonstration

Up to this section, the content of the Ph.D. is only theoretical: it contains proposition of algorithms, derivation of theoretical results, and performance estimation through MC simulations. It is now time to verify the theory-practice connection. The objective of this section is thus to show that the QCSP frame can be transmitted and received in a real radio system, even at very low SNR. In this section, although the radio link is real, all the processing is done offline. The implementation of a real-time QCSP system is out of the scope of this Ph.D. report. Nevertheless, in the frame of the QCSP project, another Ph.D. student (Mr. Camille Monière) is working on this task with very interesting preliminary results [19, 119]. In the following sections, the experimental process is first presented. Then, the result of the detection, synchronization, and correction are given and discussed.

6.3.1 Experimental process

The use of Software-Defined Radio (SDR) modules (USRP) for data transmission and reception will be favored to emulate communications between the two IoT type modules. The USRP N210 Kit is chosen [120] as the RF transmitter. It is intended for demanding communication applications requiring this type of rapid development. The datasheet for using USRP N210 is available in [121]. It is supported by GNU Radio software through the emitter and receiving blocks. The USRP source block is used to stream samples from a USRP device (act as the receiver), and the USRP sink block is used to transmit out the samples from a USRP device (acts as a transmitter) [122].

In the following experiment, the transmitter sends a coded and modulated frame many times with a decreasing power level until the loss of power frame detection at the decoder side. These experimental measurements will make it possible to estimate the channel in different environments. In our case, we can accurately estimate the SNR from a very high signal power. At the receiver side, we process the received data offline in Matlab thanks to the proposed detection and synchronization algorithms. These parameters will be re-injected and compared to the theoretical results obtained in order to accurately evaluate the proposed methods. The experimental process is given in Fig. 6.6, and detailed in the next section.

It is worth noticing here that building our own GNU blocks to run in a real-time process is the next step. Currently, the transmitter side is completed, and the GNU

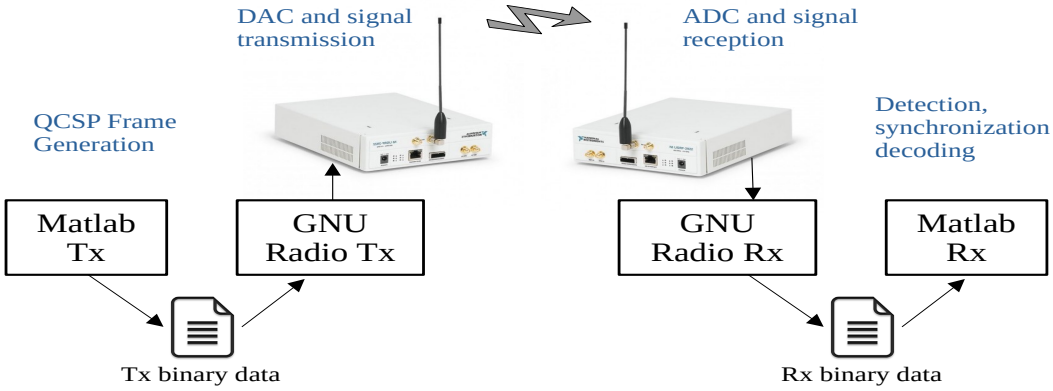


Figure 6.6 – GNU Radio Experimentation.

blocks are realized.

6.3.2 Transmission protocol

We consider here the same QCSP frame F resulted from the previous section where $N = 120$ CCSK modulated symbols and coded by the NB-LDPC code of GF order $q = 64$ chips. F is also over modulated by sequence B as in (3.16). For having the decrease in amplitude, burst of frames of F are included as a super frame SF, which is constructed as

$$SF = \left[F \quad Z \quad \frac{F}{10^{(1/20)}} \quad Z \quad \frac{F}{10^{(2/20)}} \quad Z \quad \ldots \quad Z \quad \frac{F}{10^{(60/20)}} \right], \tag{6.6}$$

where Z is a sequence of zeros of double size of F, i.e. length of Z is $2 \times q \times N$ zeros. By this structuring, it gives a ratio of $60 \times 0.5 = 30$ dB between the first frame amplitude and the last one, or 60 dB in energy. The frame $SF(n), n = 0, ..., 60$ indicates the n^{th} transmitted frame. By construction, $\text{SNR}_{SF(0)}$ (in dB) is the estimated SNR of the first frame which can be found with high accuracy. Then, $\text{SNR}_{SF(n)}$ is the estimated SNR of the n^{th} frame, and given as $\text{SNR}_{SF(0)} - n$ dB. The generated SF is shown in Fig. 6.7.

This is done in Matlab and written in a file as shown in Fig. 6.6. Then, GNU Radio software reads the signal file and transmits it in a real channel. The GNU Radio transmitter (Tx) software chain is shown in Fig. 6.8a, where RF central frequency is 433.92 MHz.

The message is of size 360 bits. It is encoded by a GF(64) NB-LDPC encoder of coding rate 1/2. The chip rate is set to 500 Kchip/s, which corresponds to a raw air binary

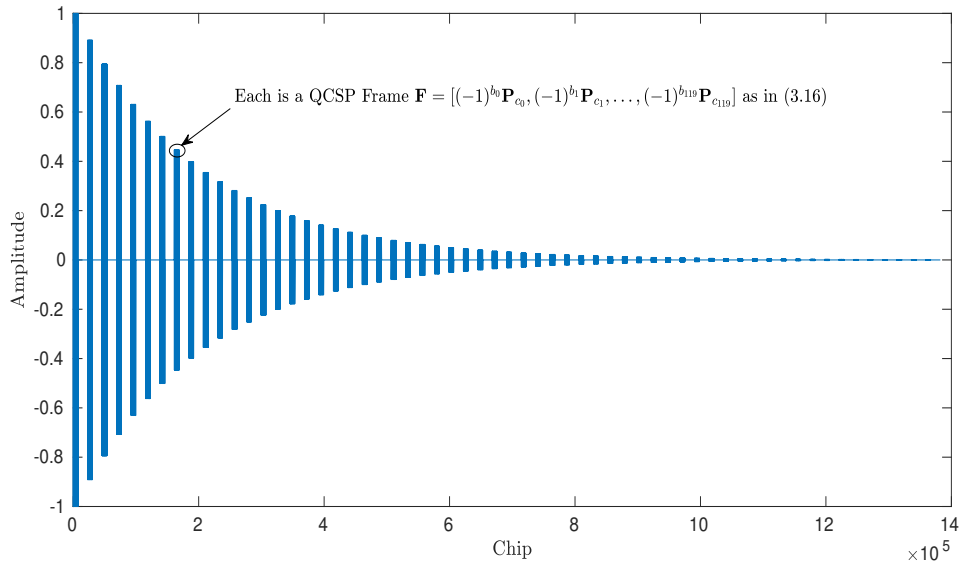
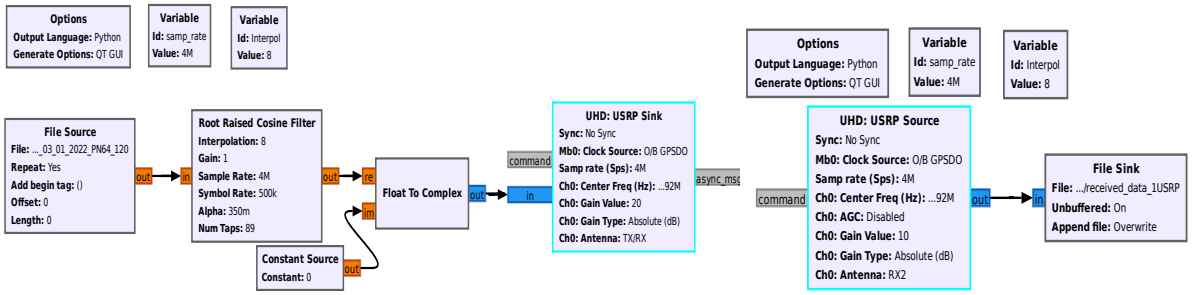


Figure 6.7 – Generated SF.

throughput of $500 \times 6/64 = 46.8$ Kbit/s and an information bit throughput of 23.4 Kbit/s during the QCSP frame transmission. The QCSP frames are over-sampled by a factor $O = 8$ before entering a root raised cosine filter with a roll-off factor of 0.35. The DAC at the emitter side and the ADC at the receiver side are both working at 4 MHz. At the emitter side, the emission power is set to 20 dB on a scale ranging from 0 to 30 dB.



(a) GNU Radio Software: Tx. (b) GNU Radio Software: Rx.

Figure 6.8 – Tx and Rx communication chain in GNU radio software.

At the receiver side, the amplifier is set to 10 dB on a scale ranging from 0 dB to 30 dB. The GNU radio chain is given in Fig. 6.8b. For the target of making more signal processing analysis, the half raised cosine filter and decimation are processed in Matlab.

The scenario was an indoor transmission, with the emitter in a room of the lab and the receiver in another room located on the other side of a corridor. The distance between the doors of the two rooms in the corridor is approximately of 3 m. The minimum free space distance between the two GNU radio modules was approximately 10 m. Both remain still during the experiment.

6.3.3 Data offline processing

The first step at the receiver side is to capture received data and write it in a file as shown in Fig. 6.6. After that, we read the data in Matlab, and plot the received samples of the *SF* as in Fig. 6.9. We can figure out that the first frame is of very high SNR, and

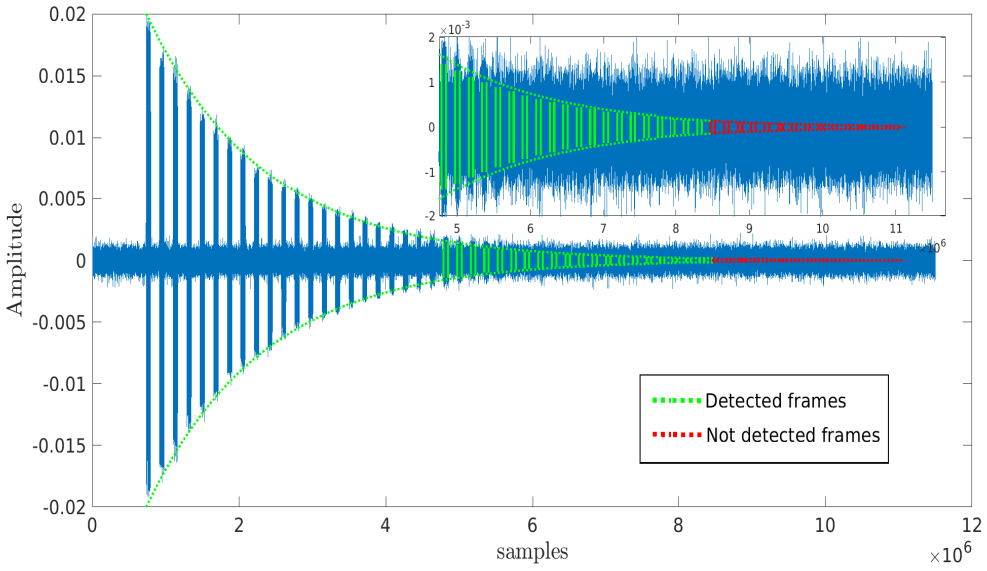


Figure 6.9 – Received SF

continues to have some frames dipped in noise, giving very low SNRs. The green color represents the well-detected frames at the receiver side, and the red color represents the miss detected ones.

The SNR is first accurately estimated from the very high signal power samples and the samples including just noise. The mean energy of the noise σ^2 is first estimated by the received samples before the arrival of the super frame. Then, the mean energy of the first frame $E_{SF(0)}$ is computed. This energy is equal to the sum of the signal energy $E'_{SF(0)}$ and the noise energy σ^2, i.e. $E_{SF(0)} = E'_{SF(0)} + \sigma^2$. Thus, $E'_{SF(0)} = E_{SF(0)} - \sigma^2$. Consequently,

the SNR of the first QCSP frame $SF(0)$ can thus be estimated as:

$$SNR_{SF(0)} = 10\log_{10}\left(\frac{E_{SF(0)} - \sigma^2}{\sigma^2}\right). \tag{6.7}$$

Same can be applied for all the high signal energy frames. As a result, the estimated SNR of the first frame is $\text{SNR}_{SF(0)} \approx 29.41$ dB, the second of $\text{SNR}_{SF(1)} \approx 28.40$ dB, third by $\text{SNR}_{SF(2)} \approx 27.25$ dB, etc. So, successive frames have SNR decreasing by 1 dB as expected theoretically from the structure of SF in (6.6). Finally, the SNR 43^{th} frame in SF can be roughly estimated by successive subtractions to have $\text{SNR}_{SF(42)} \approx$ -12.55 dB. This is shown in Table 6.1.

6.3.4 Output of the detection filter

The detection filter is applied over a grid of bin size ($\delta_t = q/8$, $\delta_\theta = \pi/2$) as defined in section 4.3. Fig. 6.10 shows the output of the detection filter of the received SF as function of time where the optimal frequency bin is considered, i.e. the frequency bin corresponding to the max score function. For the sake of figures-show simplicity, we show

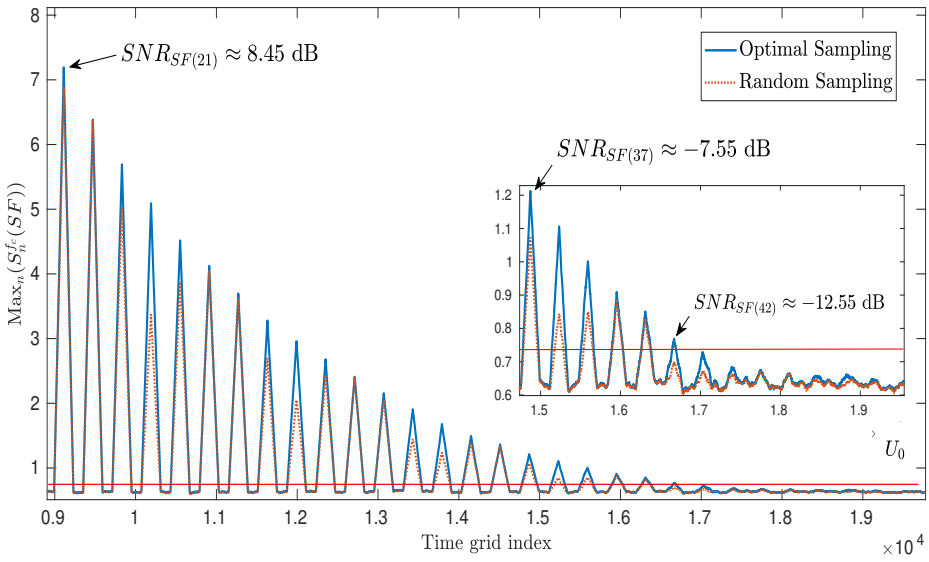


Figure 6.10 – Output of detection filter as function of time bins.

the detection output starting from the 22^{th} frame which is corresponding to $\text{SNR}_{SF(21)} \approx$ 8.45 dB. Since interpolation (oversampling by 8) is considered before the root raised

cosine filter (see Fig. 6.8b), we apply the decimation after the root raised cosine filter at the receiver side in Matlab processing. The blue solid curve in Fig. 6.10 represents the optimal sampling precision. This is assumed as in section 3.4.3 by testing in parallel all the hypotheses of the 8 different oversampling values and by keeping the best one; we can always manage to get the optimal decision. However, the dot orange curve represents the random sampling hypothesis. This drift is due to the precision errors of the oscillator and is solved by taking the maximum different sampling hypotheses. The red line is the threshold value U_0. From these primary results, it is clear that 43 frames F of the SF are detected (correlation score output above the $U_0 = 0.72$). As a result, the SNR where the final detection occurred is approximately -12.55 dB, which can confirm and give a proof of concept of the detection results reached in the theoretical part.

Fig. 6.11 shows the detection filter output of the samples corresponding to the 41^{th} frame as function of the frequency bin value, knowing that the optimal time bin is already maximized. We can notice that the coarse rotation is $\hat{\theta}_c = \pi$, which corresponds to coarse

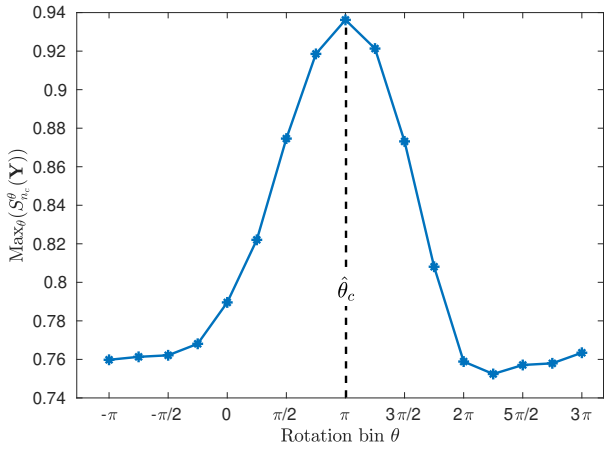


Figure 6.11 – Output of detection filter as function of frequency bins.

frequency offset between both the emitter and transmitter is $f_c = 0.0078$ Hz.

6.3.5 Output of synchronization and decoding block

In this section, we consider the blind synchronization algorithm discussed in chapter 5 to synchronize the data which is detected from the received SF. After that, we use the NB-LDPC decoder defined in section 6.1 for the frame error correction.

Table 6.1 sums-up the out-coming results. The first two columns show the frame index

and the SNR estimations. At the emission side, if we set the index n of the first sample of $SF(0)$ to 0, then the first index of $SF(1)$ will be equal to $8qN$ (the size of $SF(0)$) plus $16qN$ (the size of the null sequence). Thus, the first sample index of $SF(1)$ is equal to $24qN$. Similarly, the index of the ℓ^{th} frame is thus $24\ell qN$. The clocks of the emitter and the receiver have only a relative precision. At the receiver side, the actual number of samples between the first sample of $SF(0)$ and the first sample of $SF(n)$ may differ slightly from $24\ell qN$. This difference is given in the third column of Table 6.1. One can see that the relative precision between the two clocks is 105 samples among 42 Frames, i.e. $42 \times 24qN$ which gives a relative precision of 8 ppm (means parts per million). This corresponds to the typical precision of a good crystal oscillator.
The detection block answer of a frame existence or not is shown in the fourth column.
The time synchronization parameters are then presented. First, let us recall that r and s represent the estimation of the time offsets after the finer grid detection (as discussed in (5.2)) at chip and symbol level, respectively. The value of r is found by the VNB method thanks to the NB-LDPC structure, and s using the WOM method thanks to the CCSK modulation and over-modulation. We can distinguish that for SNR $\leq$ 10 dB the synchronization by simply multi hypothesis detection searches will not work, and generate time estimation errors. Consequently, the proposed time synchronization method comes to solve this problem. At SNR = -12.5 dB for example, the WOM-VNB method succeeded to time-synchronize the 42^{th} frame where it is shifted by 258 chips, i.e. $e = sq+r = 4\times64+2 = 258$. After that, Table 6.1 shows the frequency estimation evolution starting from the detection process when it estimates the coarse frequency offset f_c, till the finer frequency offset $\tilde{f_c}$ discussed in section 5.1.2.4. It is obtained that all the coarse frequency estimations are detected correctly in the bin corresponding to f_c, even at the very low SNR in this example. Some frequency estimation errors occurred while the estimation of the finer frequency offset $\hat{f_o}$ using a grid of smaller frequency bin resolution. This is corrected by the residual frequency estimation thanks to the DM. It is noticed that the frequency offset is increasing successively, this is due to the local oscillator jitter, i.e. instability of the carrier frequency generators at the emitter side and the receiver side. The next column shows the correct estimations of the initial phase offset using the DM.
The final two columns show the CCSK demodulation and the NB-LDPC decoding respectively. We can figure out that for SNR $\leq$ -7.55, we start to have some errors from the CCSK demodulation. The interesting results is that for SNR = 12.5 dB, we are able to decode the 43^{th} frame which corresponds to $\text{SNR}_{SF(42)} = -12.55$ dB.

Table 6.1 – Processing of the received ***SF*** based on the algorithms QCSP proposed algorithms for detection, synchronization and decoding.

SF **index** *SF*(i)	**SNR** (dB)	**Clock jitter effect** (samples)	**Detect or No**	**Time Sync**		**Freq and phase Sync.**				**Decoding**	
				$\hat{r}$ (chips)	$\hat{s}$ (symbols)	$\widehat{f}_c$ (Hz)	$\widehat{f}_o$ (Hz)	$\widetilde{f}_o$ (Hz)	$\hat{\phi}$ (radian)	CCSK errors	NB-LDPC Is Codeword
0	29.41	0	Yes	0	0	0.007812	0.008057	0.008013	2.951911	0	Yes
1	28.40	2	Yes	0	0	0.007812	0.008057	0.008014	0.603615	0	Yes
2	27.26	3	Yes	0	0	0.007812	0.008057	0.008015	-1.588478	0	Yes
3	26.44	6	Yes	0	0	0.007812	0.008057	0.008016	2.668108	0	Yes
21	8.45	11	Yes	0	0	0.007812	0.008057	0.008018	-1.005000	0	Yes
29	0.45	71	Yes	0	0	0.007812	0.008057	0.008022	1.870936	0	Yes
30	-0.55	74	Yes	0	0	0.007812	0.008057	0.008023	0.914991	0	Yes
31	-1.55	76	Yes	0	0	0.007812	0.008057	0.008024	-0.060232	0	Yes
32	-2.55	79	Yes	0	0	0.007812	0.008057	0.008023	-0.955860	0	Yes
33	-3.55	82	Yes	0	0	0.007812	0.008057	0.008022	-1.889084	0	Yes
34	-4.55	84	Yes	0	0	0.007812	0.008057	0.008024	-2.880322	0	Yes
35	-5.55	87	Yes	0	0	0.007812	0.008057	0.008024	2.504311	0	Yes
36	-6.55	90	Yes	0	0	0.007812	0.008057	0.008023	1.620808	0	Yes
37	-7.55	92	Yes	0	0	0.007812	0.007812	0.008023	0.694619	1	Yes
38	-8.55	95	Yes	0	0	0.007812	0.007812	0.008023	-0.222060	1	Yes
39	-9.55	97	Yes	0	0	0.007812	0.008301	0.008022	-1.066529	9	Yes
40	-10.55	99	Yes	1	0	0.007812	0.008301	0.008022	-1.970423	12	Yes
41	-11.55	102	Yes	3	1	0.007812	0.008057	0.008023	-2.876781	29	Yes
42	-12.55	105	Yes	2	4	0.007812	0.007812	0.008024	2.448245	43	Yes
No detection No decoding											

It is worth noticing here that we have used the included MEX file functions in Matlab to be able to call the NB-LDPC decoder which is designed and implemented in C language by the Lab-STICC group. Consequently, the NB-LDPC MEX function behaves just like a Matlab script or function able to decode an NB-LDPC codeword, however, it takes more time than C. To sum up, we have a full Matlab working chain for the QCSP system model. In the frame of the QCSP project, the detection block is already written in C and is working in real-time [19, 119].

Phase and frequency offsets illustration

This section shows the effect of frequency and initial offsets on the real and imaginary part of the samples even on a very high SNR received frame. In Fig. 6.12, we show the scatterplot of the received frame $SF(1)$ of SNR= 28.4 dB, i.e. the imaginary part as function of the real part for all samples in the frame. Before starting the synchronization

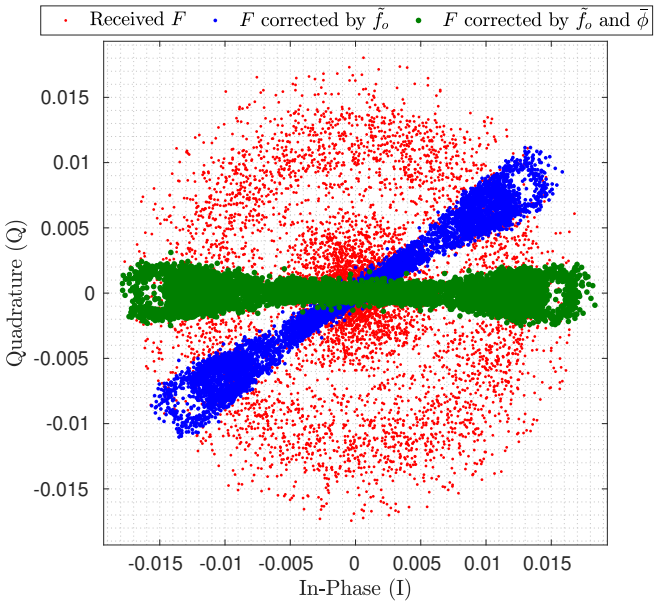


Figure 6.12 – Scatterplot of real-imaginary parts of received $SF(1)$ respectively before synchronization, after frequency synchronization and after phase-frequency synchronization.

(red points), we can notice the phase (ϕ) and frequency f_o effect as in (3.18). Each point in Fig. 6.12 is rotated by a constant value due to ϕ; and each angle of points in the constellation also changes linearly over time by $\theta = 2\pi f_o q$, so the points in the scatter plot shift radially. In the blue points, which are representing the received frame after

estimating the frequency offset by $\tilde{f_o}$, the linear change of phase over time is eliminated and still has just the constant rotation. Finally, the phase offset is estimated by $\hat{\phi}$ in the green dots. Note that this phase offset does not affect the detection method because the score function of the detection algorithm is based on the magnitude of the correlation.

Another representation of the phase and frequency effect is by showing the real and imaginary amplitude of samples as function of time, as in Fig. 6.13.

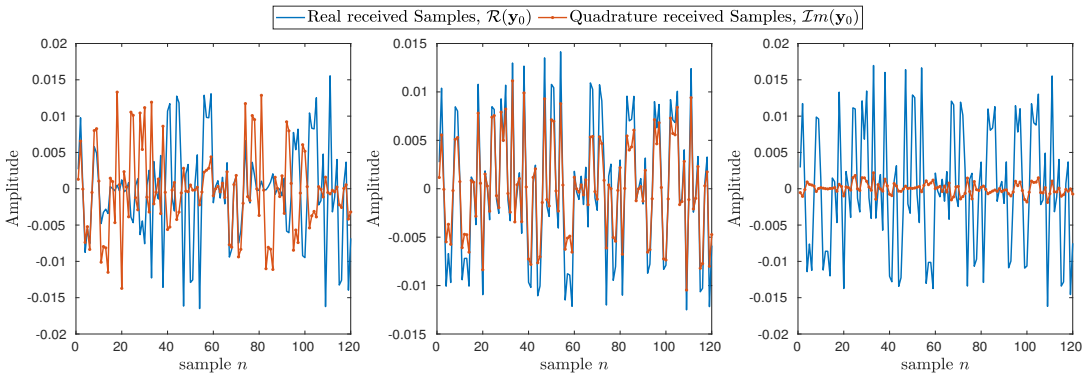


Figure 6.13 – Magnitude of real and imaginary parts as function of time, respectively before synchronization, after frequency synchronization and after phase-frequency synchronization.

To conclude, setting emission power to 20 dB gives a budget link margin of 29.41 dB (the SNR of the first received frame) to reach -12.5 dB (the SNR of the last received frame), thus around 40 dB approximately. Fully exploiting this margin would allow, in an open area, to multiply by a factor of 100 the emitter/receiver distance, i.e. from 10 m to 1 km.

6.4 Conclusion

In this chapter, we determined how the overall FER of a QCSP frame is related to the detection, synchronization, and correction probabilities. We defined the complex optimization problem that links frame parameters to the decoder complexity and transmission performance. This problem can be summarized as follows: Assuming a payload of m bits to be transmitted at a given SNR, what is the minimum QCSP frame length that allows a reliable transmission in an unslotted random access channel?

The main degrees of freedom are the choice of GF size q and the NB-LDPC coding rate $R_c = K/N$, with $K = \lceil \frac{m}{\log_2(q)} \rceil$ is the number of GF(q) symbols required to represent

the m-bits of information. The other degree of freedom is the choice of the spreading sequence $\boldsymbol{P}_0$, the choice of the OM sequence $\boldsymbol{B}$, the choice of the NB-LDPC code, and finally, the acceptable computational complexity at the receiver side. The optimization problem has not been solved yet. It is still an open investigation area. However, for a payload of $m = 360$ bits, we propose a solution that outperforms the state of the art. The proposed QCSP frame is constructed over GF(64) with a code rate of 1/2, giving in total a frame length of $N = 120$ symbols. Comparing the proposed QCSP frame with a classical preamble-based frame using up-to-date codes for short packet transmission, we found that the QCSP frame reduced the frame size by 23%. This reduction can be translated directly into an increase in the wireless network capacity and energy saving. For this frame, the required SNR for a transmission with a FER of 10^{-4} is -11.8 dB, which is only 1.5 dB apart from the Polyanskiy's bound in the AWGN channel.

Finally, we present the result of an experimental transmission using the proposed QCSP frame with GNU radio modules. The experimentation results are fully consistent with the theoretical model and validate the transmission of QCSP frame are SNR up to -12 dB of SNR in an indoor environment.

It is worth noticing here that our goal is to put this QCSP structure on the track of blind preamble-less short packet transmission, and we are fully aware that more optimizations can be added in the near future. In the following section, we will conclude the developed work, and more ideas will be proposed to enhance the overall processes to be as close as possible to the optimal bound derived and formulated by Polyanskiy.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

In this thesis, the main focus is to propose, develop and evaluate blind detection and synchronization algorithms for short data packet transmission in LPWA networks without the use of a preamble dedicated for that purpose. Consequently, this increases the spectral efficiency of transmission in an unslotted wireless channel by avoiding the preamble overhead. The ambition of the project is to work on the emergence of NB codes combined with a CCSK modulation. This coded modulation scheme, called CCSK-NB-code, can be easily implemented and provides several advantages compared to the state of the art waveforms. The whole frame can be considered either as a preamble sequence to perform detection and synchronization or as a noisy codeword to perform the non-binary error-correcting process. Owing to this structure, the QCSP frame offers the capability of blind detection and self-synchronization without any additional overhead. In the following, a summary of the main contributions and conclusions of the thesis is provided.

Chapter 3, starts by presenting the overall communication chain of the QCSP system model. The needed definitions as the FEC techniques are recalled along with the theory of finite fields, NB-LDPC codes, and the CCSK modulation. A novel idea is proposed, which is the over-modulation that aims to help the synchronization procedure. Finally, the channel model as well as the time, frequency, and phase offsets have been formulated and described.

In chapter 4, the proposed preamble-less detection method is discussed, it is based on a metric called score function. This metric is calculated through the correlation of the incoming sequence of samples with pre-defined CCSK sequences. Then, a derivation of a formal performance model of the frame detection algorithm is formulated, and it is empirically verified by a MC simulation. Consequently, an analysis of the frame detection algorithm has been studied where it gives insight on the impact of each parameter on

the detection performance according to the QCSP frame structure (size and GF order), the time, and frequency offsets. The simulation results showed that a reliable detection ($\mathcal{P}_{\text{md}} = 10^{-4}$ and $\mathcal{P}_{\text{fa}} = 10^{-6}$) can be obtained at very low SNRs. In a particular example: (i.e., $N = 120$ symbols and $q = 64$ chips), the reliable detection can be reached at -7.5 dB where a frequency rotation uniformly distributed between $[-\pi, \pi]$ and a chip delay $[-q/2, q/2]$. This performance can be enhanced to reach -11.78 dB by using parallel filters, however, this will be on the price of an increase of factor 32 of the complexity. One should not be afraid by this complexity increases since, in [19, 119], a real-time software detection algorithm is presented.

In chapter 5, the first part discusses the proposed self-time synchronized algorithm for QCSP frames operating at very low SNR with a probability of miss-synchronization of the order of 10^{-6}. The proposed algorithm mitigates the time ambiguity in two consecutive steps. The first step is at the symbol level where an OM technique has been proposed and then enhanced with a weighted version privileging the symbols received with high reliability. The second step is at the chip level by taking advantage of the NB-LDPC code structure.

The second part of chapter 5 proposes a phase and frequency synchronization technique. Two methods are proposed. The first, called DM, is a direct estimation of the parameters of a noisy sinusoidal. The second, called PM, is based on the ML estimation using a distribution parameterized by the CCSK demodulator and the NB-LDPC decoder. The FEC results showed that the performance of the proposed phase synchronized frame approximately maintains the same performance as when a Genius-Aided estimation is used, or when no phase offset exists. This is achieved at very low SNRs also.

In chapter 6, global performance analysis in terms of the FER of the proposed QCSP system is presented. This FER has been expressed in terms of the probabilities of miss detection, error correction, and miss synchronization. These probabilities have been traded-off to find the most suitable system parameters (Galois field order q, message length K, code rate R_c, spreading sequence, etc). A comparison with a classical preamble-based frame similar to that used in the 3GPP standard has been considered, where it was shown that a frame size reduction of 23 % can be obtained with the QCSP system. Finally, a proof of concept using SDR implementation has been considered where the conducted experiments on the proposed detection and synchronization methods validated the theoretical results.

7.2 Future work

The work presented in this report does not close the topic of efficient Blind transmission (no preamble) using the QCSP system model. There are still several developments tasks to be considered.

- All the work deals only with the AWGN channel; future work is going to extend the investigation to multi-user detection in the context of IoT multi-user access.
- Using the over-modulation for the detection algorithm.
- Replacing the binary spreading and OM sequence by CAZAC sequence.
- Testing the association of CCSK with other types of NB-codes (Turbo and polar codes) and proposes new ideas thanks to their structures.
- The discussion of the detection, correction, and synchronization approach in the last section of results opens an interesting theoretical question regarding the optimal frame structure to fulfill the requirement of an application with the minimum energy cost at the transmission side.
- Designing architectures that deal with performance-complexity trade-off, is one of the most important tasks that will be studied in the near future.
- Improving the design of the SDR demonstrator using GNU radio with real-time reception for a set of sensors.

To conclude, we believe that the QCSP scheme can be useful in many applications. It could compete with existing solutions such as LoRA, Sigfox, and NB-LTE solutions in a LPWAN. It could be also used to establish a communication link in an ALOHA protocol between a terminal and a communication infrastructure (constellation of low earth orbital satellites, a base station of a mobile network, etc.).

List of Publications

QCSP Project Website: https://qcsp.univ-ubs.fr/.
The results obtained through this PhD have been spread in the scientific community through the following publications,

Patent

— E. Boutillon and K. Saied, "A method for a transmitter to transmit a signal to a receiver in a communication system, and its corresponding receiving method", July 2021.

Journals

— K. Saied, A. Al Ghouwayel, and E. Boutillon, "Quasi Cyclic Short Packet for Asynchronous Preamble-Less Transmission in Very LowSNRs", *Submitted to IEEE Transaction journal on Wireless Communication.* (Major Revision)

— C. Moniere, K. Saied, B. Legal, and E. Boutillon, extention of "Time sliding window for the detection of CCSK frames", *to be submitted to IEEE Open Journal of the Computer Society (OJCS).*

Conferences

— K. Saied, A. Al Ghouwayel, and E. Boutillon, "Blind Time-Synchronization of CCSK Short Frames", *in The 17th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob21)*, Oct. 2021, Bologna, Italy.

— C. Moniere, K. Saied, B. Legal, and E. Boutillon, "Time sliding window for the detection of CCSK frames", *in the IEEE Workshop on Signal Processing Systems (SiPS'2021)*, Oct. 2021, Combria, Portugal.

— K. Saied, A. Al Ghouwayel, and E. Boutillon, "Phase Synchronization for Non-Binary Coded CCSK Short Frames", *submitted to the 2022 IEEE 95th Vehicular Technology Conference: VTC2022-Spring.* (submitted)

Deliverables to ANR

— K. Saied and E. Boutillon."Blind Detection Algorithm for QCSP Frames". [Online]. Available: https://qcsp.univ-ubs.fr/wp-content/uploads/2022/01/QCSP_Detection-1.pdf

— K. Saied and E. Boutillon."Blind Synchronization Algorithm for QCSP Frames". [Online]. Available: https://qcsp.univ-ubs.fr/wp-content/uploads/2022/01/QCSP_Synchronization.pdf

Chapter 8

Appendices

8.1 Different LPWA network protocols

Through the licensed and unlicensed frequency bandwidth, various LPWA network technologies have been emerged. Among these technologies, SigFox, LoRa, and NB-IoT, which are currently the leading emergent technologies with many technical differences. In this section, emerging property techniques and some aspects of the different IoT technologies (SigFox, LoRa and NB-IoT) will be discussed briefly and highlighted, thanks to [38].

SigFox

Being both a company and an LPWAN network operator, the SigFox technology was created in 2010 by the startup SigFox (Toulouse, France). Although it is still an abstract owing to the cooperation with many network operators, SigFox functions and merchandises its own IoT solution in 31 countries via using its proprietary base stations equipped with cognitive software-defined radios in order to interface them to the back end servers using an IP-based network. Thus, the end devices associated with these base stations deploying binary phase-shift keying (BPSK) modulation in an ultra-narrow band (100 Hz) sub- GHZ ISM (Industrial, Scientific, and Medical radio) band carrier. SigFox utilizes unlicensed ISM bands, for instance, 868 MHz in Europe, 915 MHz in North America, and 433 MHz in Asia. Moreover, under the deployment of the ultra-narrow band, SigFox effectively employs the frequency bandwidth and encounters very low noise levels, provoking very low power expense, high receiver sensitivity, and low-cost antenna design at the consumption of maximum throughput of only 100 bps. Additionally, just 140 messages per day are restricted over the uplink with 12 bytes for its extreme payload length. On the other hand, the downlink is constrained to four messages per day with eight bytes of its maximum payload length. Consequently, this implies that the affirmation of each uplink message is not supported. Furthermore, the reliability of the uplink communication is verified through time and frequency diversity as well as transmission duplication. As a result, each end-device message is transferred various times (three by default) along different frequency channels. For this reason, in Europe for instance, the band between 868.180 MHz and 868.220

MHz is partitioned into 400 orthogonal 100 Hz channels (among them 40 channels are saved and not utilized) [1]. Therefore, the messages of the base stations are transferred via a frequency channel which is randomly chosen by the end device because all channels can simultaneously receive messages for the base stations, leading to the simplification of the end device layout and the reduction of its cost.

LoRa

LoRa (Long Range) is a physical layer technology that modulates the signals in sub-GHz ISM band using a proprietary spread spectrum technique. Like SigFox, LoRa deploys unlicensed ISM bands, i.e., 868 MHz in Europe, 915 MHz in North America, and 433 MHz in Asia. although LoRa, which was institutionalized by LoRa-Alliance, is still an abstract in various countries owing to the investment of many mobile operators (e.g., Bouygues and Orange in France, KPN in Netherlands, and Fastnet in South Africa), it is implemented in 42 countries. A narrow-band signal is spread over a wider channel bandwidth by the Chirp Spread Spectrum (CSS), which in its turn supplies the bidirectional communication, leading to a signal that is difficult to be detected or jammed because of its low noise level and high interference resilience [123]. Moreover, Lora employs six spreading factors (SF7 to SF12) to adjust the convenient data rate and range tradeoff. As the spreading factor increases, longer range will be reached at the price of lower data rate, and vice versa. Additionally, depending on the spreading factor and channel bandwidth, the LORA data rate is between 300 bps and 50 kbps, where Lora base stations simultaneously receive the transferred messages using different spreading factors [124]. Consequently, all the base stations in the range receive each message transmitted by an end device via LORAWAN, where every message is maximum pay loaded by 243 bytes. Thus, LORAWAN enhances the successfully received messages ratio by demolishing this excess reception. But, attaining such aspect acquires various base stations in the neighborhood, which may raise the cost of the employed network.

The various needs of wide range IoT applications are addressed using many classes of end devices which are provided by LORAWAN.

— *Bidirectional end devices (class A)*: As shown in Fig. 8.1, class A end devices permit bidirectional communications where two short downlink receive windows pursue each end device's uplink transmission. Moreover, based on its own communication requirements with small variation on a random time basis, the transmissions lot is organized by the end device. Furthermore, after the end device has sent an uplink message, the applications that just need short downlink communication keep this class-A operation as the lowest power end-device system. Thus, downlink communications at any other time will have to wait until the next uplink message of the end device.

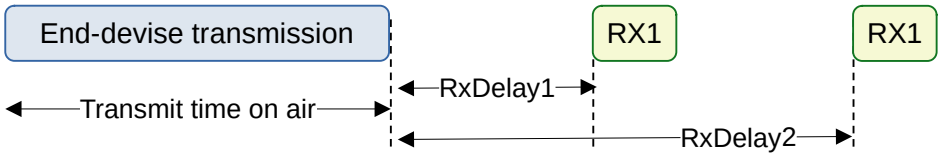


Figure 8.1 – Bidirectional communication between end-device and base station for LoRaWAN class A.

— *Bidirectional end devices with scheduled receives slots (class B)*: class B devices allow an additional receive windows at planned times accompanied with the random receive windows of class A, where end devices receive a time-synchronized beacon from the base station in order to open receive windows at the scheduled time. Consequently, when the end device is listening, the network server recognizes that.
— *Bidirectional end devices with maximal receive slots (class C)*: class C end devices open receive windows in a continuous manner, and only close at excessive energy consumption.

LoRa-Alliance is still updating the specifications of the next version of LoRaWAN with modern aspects and features of roaming, class B clarification, and the temporary switching between class A and class C.

NB-IoT

NB-IoT is a Narrow Band technology dedicated in development 13 of the 3GPP in June 2016, with the ability to coexist with GSM (global system for mobile communications) and LTE (long-term evolution) under licensed frequency bands (e.g., 700 MHz, 800 MHz, and 900 MHz),

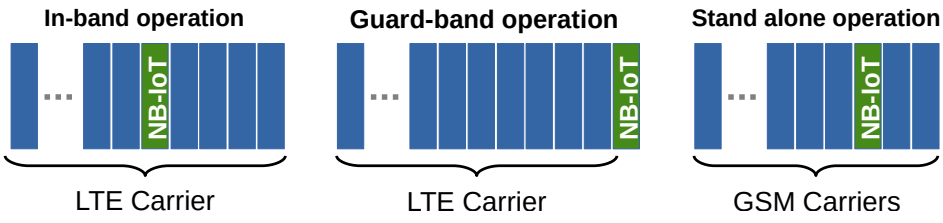


Figure 8.2 – Operation modes for NB-IoT.

in addition of occupying a frequency band width of 200 KHz, which corresponds to one resource block in GSM and LTE transmission [125]. Consequently, as shown in Fig. 8.2, the following operation modes are available through this frequency band selection:

— Stand-alone operation: the currently utilized frequencies bands is a possible scenario
— Guard-band operation: deploying the inactive resource blocks within an LTE carrier's guard band.
— In-Band operation: employing resource blocks within LTE carrier.

For the stand-alone operation, Fig. 8.2 in its right part (GSM carriers) reflects an illustration to specify that the operation is possible in NB-IoT utilization. Actually, the 3GPP provokes the coordination of NB-IoT with the LTE cellular networks, where beside the existing LTE infrastructure; just a software update can assist NB-IoT. Additionally, the LTE protocol is the base of the NB-IoT communication protocol, where the latter sharply restricts the functions of LTE protocol and improves them as required for IoT applications. For instance, the LTE backend framework is utilized to broadcast information that is valid for all end devices within a cell. with size relative to its existence, the communication back end system is kept to a minimum although it acquires resources and expends consumes battery power from each end device via optimizing to little infrequent data messages and averting the unneeded aspects for the IoT goal, e.g., assessments to evaluate the channel quality, carrier aggregation, and dual connectivity. Hence, the end devices are cost-efficient because it only needs low battery energy. Therefore, being constructed on the well-established LTE infrastructure, NB-IoT technology can be classified a modern air interface from the perspective of the protocol stack by permitting connectivity of up to 100 K end devices per cell with the potential for scaling up the capacity by adding more NB-IoT carriers. Single-carrier frequency division with various access (FDMA) in the uplink and orthogonal FDMA (OFDMA) in the downlink are utilized by NB IoT via deploying the quadrature phase shift keying modulation (QPSK) [126], where 200 kbps for the downlink and 20 kbps for the uplink restrict the data rate with 1600-byte maximum payload size for each message. Hence, when transmitting 200 bytes per day on average, 10 years of battery lifetime can be accomplished by IoT technology.

8.2 Basic sequences generation for CCSK modulation

P_0 **for** $q = 64$

“0111011001011101011001110000010000101110000111011100100001101011”

P_0 **for** $q = 128$

“10000001001001101001111011100001111111000111011000101001011111010101010000
10110111100111001010110011000001101101011101000110010001”

P_0 **for** $q = 256$

“011101010101001001001101101011110010011000100001010010010001101010100001
010101110110111000000110011001110000110010010110001110001001111001110010 1
011110100111001100010110000101100000100100111011110000110000111010100010 1
1111110111111101111011110000000001101110”;

P_0 **for** $q = 512$

“1010000001100100000011000111111100011010100110101011010000101111100000011 1
110101011010110001000101000001010110000010110110110110001011101000110011 0
100000000010111101110110010000001101111101111111000111010100110001000001 0
111011000111010010110010000111100100101011110100011010001110000111100011 1
100110010101000101011110011000111011011001000001100010000100000110100100 0
101110011111111101010100011100011110011110011001110010010111101110010010 0
100111100101110011000110111101100111111011110110000000111110110010101001 0
00”;

P_0 **for** $q = 1024$

“100001001001101000111100100001001011010111001010110100010101111101010011
000000001111111011100011000110101110111110101100000000011100010110100100 0
100010101100000010001011010010001001001100110001111010001010110011000111 01
010110110111111000110111011001110110010111110011010100011010100001010111 0
001111101110001111111111011100001001010101000001100100100011100011100101 0
100011100001100011001000010011000000000000010101001101111011110100101001 11
011101010010110001100111000001111100111000000110000010001011010101101000 1

101000000101111111111011001101100101100111010001010001110010110010100001010
101101101001001001000111000111001101011100111010000101001101001110000011000
001010110000100010101111101011011100011000101100101111111100000101111101000
000010011011100001010010101000111010100111111101001111110100100111011000000
000010111111111010011110101011110001001101101111101001001100110111001010100
100110100010101100101110000111101001110110110110010100011001011100001101100
100100111111110001101100101100100010001011011011100111010100000111010000001”

P_0 **for** $q = 2048$

“000101010000011011011010110001101111101011011111010111111111010001011001
011000010100011110110000011001000111101001010011001001001101011110000010100
111011100100111101100011100010000100011101111101100101011011111011000010100
000001110011100011011101000100011010000001001110100011110010011011100010000
010010011010010001111111111011011110010110010110000000110001101100110110100
011100010110000010000000101100010010100110111100101110010110111000000101100
100111010001101111011000000011100101011101100110000101100101000101101101100
001101110101111101100111101100010001101010011010011001101100010101111101110
111100000011110001101001000000010001100110101111111010110001101110101111000
111111111101100100000111111111010001100100100101110111111001110010011000100
000010100010111010101101111110010010011010010101011010000000111100001111100
101100011000101001111101011010001000000110011110000000011001111010001001100
110101100011101011111001001101100001000000101010000000011110001100100101100
010101011101000001011001100101000110010111010100110011101100101101001010100
000111010001101101010011001000001000100001000101110010011011100111001011100
110110011111100000100001010110000000110101011011101011110010001100000100100
111111011000011001100111010000101101111111010000101000101111101110011111010
001011001100011010000000011000111110001000110101010111010100101011010010100
101100000010010111001010001011110000110110101001110010011000001100010101100
001010110100111000111101000000111010100110111101000011011100111111011010110
110110000110101000111101101101100010101011110111111011101010000111001010000
101101101010010110000001000010100110010000001010100001000001101000010111100
110001110010010111001111011010110110000110000011101110000100010000111111100
000011010011100110110111111011000101101010010001001011000111110000110110000
001000101100110010110001110110000011111011111100111101011111100101011011100
000100010110110111001111000010000011111011100011111101000100111110010101000
000100000111001101110010110000100001100010001110011110001111010001011010000

00100101011100111010001011001010000010110000011110000101101101111010001100
01011".

8.3 OM sequences

OM sequence for $N = 60$

B: "11110101000111000000001000101100110110000100110111000101010010"

OM sequence for $N = 120$

B: "01010011110011001111010100010110101101010000001110001110100000000010010010011011110110100011001100010111000111011011101010"

BIBLIOGRAPHY

1. Raza, U., Kulkarni, P. & Sooriyabandara, M., Low Power Wide Area Networks: An Overview, *IEEE Communications Surveys Tutorials* **19,** 855–873 (2017).
2. Akyildiz, I. F., Weilian Su, Sankarasubramaniam, Y. & Cayirci, E., A survey on sensor networks, *IEEE Communications Magazine* **40,** 102–114 (2002).
3. Ratasuk, R., Vejlgaard, B., Mangalvedhe, N. & Ghosh, A., *NB-IoT system for M2M communication in 2016 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)* (2016), 428–432.
4. Wang, Y. E. *et al.,* A Primer on 3GPP Narrowband Internet of Things, *IEEE Communications Magazine* **55,** 117–123 (2017).
5. Berrou, C., Glavieux, A. & Thitimajshima, P., *Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1 in Proceedings of ICC '93 - IEEE International Conference on Communications* **2** (1993), 1064–1070 vol.2.
6. Lora-Alliance, *LoRaWAN TM 101 A Technical Introduction* Technical Marketing Workgroup 1.0, November 2015, `http://www.lora-alliance.org`.
7. Lavric, Alexandru and Petrariu, Adrian I. and Popa, Valentin, Long Range SigFox Communication Protocol Scalability Analysis Under Large-Scale, High-Density Conditions, *IEEE Access* **7,** 35816–35825 (2019).
8. Durisi, G., Koch, T. & Popovski, P., Toward Massive, Ultrareliable, and Low-Latency Wireless Communication With Short Packets, *Proceedings of the IEEE* **104,** 1711–1726 (Sept. 2016).
9. Polyanskiy, Y., Asynchronous Communication: Exact Synchronization, Universality, and Dispersion, *IEEE Transactions on Information Theory* **59,** 1256–1270 (Mar. 2013).
10. Abassi, O., Conde-Canencia, L., Mansour, M. & Boutillon, E., *Non-Binary Low-Density Parity-Check coded Cyclic Code-Shift Keying in 2013 IEEE Wireless Communications and Networking Conference (WCNC)* (Apr. 2013), 3890–3894.

11. Wong, A. Y. .-. & Leung, V. C. M., *Code-phase-shift keying: a power and bandwidth efficient spread spectrum signaling technique for wireless local area network applications in CCECE '97. Canadian Conference on Electrical and Computer Engineering. Engineering Innovation: Voyage of Discovery. Conference Proceedings* **2** (May 1997), 478–481 vol.2.

12. Dillard, G. M., Reuter, M., Zeiddler, J. & Zeidler, B., Cyclic code shift keying: a low probability of intercept communication technique, *IEEE Transactions on Aerospace and Electronic Systems* **39,** 786–798, ISSN: 1557-9603 (July 2003).

13. Voicila, A., Declercq, D., Verdier, F., Fossorier, M. & Urard, P., Low-complexity decoding for non-binary LDPC codes in high order fields, *IEEE Transactions on Communications* **58,** 1365–1375, ISSN: 1558-0857 (May 2010).

14. Liva, G., Paolini, E., Matuz, B., Scalise, S. & Chiani, M., Short Turbo Codes over High Order Fields, *IEEE Transactions on Communications* **61,** 2201–2211, ISSN: 1558-0857 (June 2013).

15. Zhou, R., Le Bidan, R., Pyndiah, R. & Goalic, A., Low-Complexity High-Rate Reed-Solomon Block Turbo Codes, *IEEE Transactions on Communications* **55,** 1656–1660, ISSN: 1558-0857 (Sept. 2007).

16. Mori, R. & Tanaka, T., *Non-binary polar codes using Reed-Solomon codes and algebraic geometry codes in 2010 IEEE Information Theory Workshop* (Aug. 2010), 1–5.

17. Pfletschinger, S. & Declercq, D., *Getting Closer to MIMO Capacity with Non-Binary Codes and Spatial Multiplexing in 2010 IEEE Global Telecommunications Conference GLOBECOM 2010* (Dec. 2010), 1–5.

18. Saied, K., Ghouwayel, A. & Boutillon, E., Short Frame Transmission at Very Low SNRs, *Submitted to IEEE Trans. on Wireless Communication (major revision)* (Jan. 2022).

19. Monière, C., Saied, K., Legal, B. & Boutillon, E., *Time sliding window for the detection of CCSK frames in accepted to the IEEE Workshop on Signal Processing Systems (SiPS'2021)* (2021), 1–6.

20. Saied, K., Ghouwayel, A. & Boutillon, E., *Time-Synchronization of CCSK Short Frames in 17th International Conference on Wireless and Mobile Computing, Networking and Communications WiMob2021* (Bologna, Italy, Oct. 2021), `https://hal.archives-ouvertes.fr/hal-03404770`.

21. Saied, K., Ghouwayel, A. & Boutillon, E., *Phase Synchronization for NB-LDPC Coded CCSK Short Frames in Submitted to the 2022 IEEE Vehicular Technology Conference VTC2022* (Helsinki, Finland, 2022).

22. Hyder, M. & Mahata, K., Zadoff–Chu Sequence Design for Random Access Initial Uplink Synchronization in LTE-Like Systems, *IEEE Transactions on Wireless Communications* **16,** 503–511 (2017).

23. Malik Muhammad Usman Gul, Sungeun Lee & Xiaoli Ma, *Robust synchronization for OFDM employing Zadoff-Chu sequence in 2012 46th Annual Conference on Information Sciences and Systems (CISS)* (2012), 1–6.

24. Li, Y., Li, D., Cui, W. & Zhang, R., *Research based on OSI model in 2011 IEEE 3rd International Conference on Communication Software and Networks* (2011), 554–557.

25. O'Shea, T. J., Roy, T., West, N. & Hilburn, B. C., *Physical Layer Communications System Design Over-the-Air Using Adversarial Networks in 2018 26th European Signal Processing Conference (EUSIPCO)* (2018), 529–532.

26. 4G LTE World Coverage Map, `http://www.worldtimezone.com/4g.html`.

27. 5G White Paper, *NGMN 5G Initiative* `https://www.ngmn.org/work-programme/5g-white-paper.html`.

28. Palattella, M. *et al.,* Internet of Things in the 5G Era: Enablers, Architecture and Business Models, *IEEE Journal on Selected Areas in Communications* **34,** 1–1 (Mar. 2016).

29. Miraz, M. H., Ali, M., Excell, P. S. & Picking, R., *A review on Internet of Things (IoT), Internet of Everything (IoE) and Internet of Nano Things (IoNT) in 2015 Internet Technologies and Applications (ITA)* (Sept. 2015), 219–224.

30. Dogra, A., Jha, R. K. & Jain, S., A Survey on Beyond 5G Network With the Advent of 6G: Architecture and Emerging Technologies, *IEEE Access* **9,** 67512–67547 (2021).

31. Durisi, G., Koch, T. & Popovski, P., Toward Massive, Ultrareliable, and Low-Latency Wireless Communication With Short Packets, *Proceedings of the IEEE* **104,** 1711–1726 (2016).

32. Poor, H. V., Goldenbaum, M. & Yang, W., *Fundamentals for IoT networks: secure and low-latency communications in* (Jan. 2019), 362–364, ISBN: 9781450360944.

33. Ajaykumar, N. & Sarvagya, M., *Secure and energy efficient routing protocol in wireless sensor network: A survey in 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (2017), 2313–2322.

34. Dhurandher, S. K., Misra, S., Obaidat, M. S. & Gupta, N., *QDV: A Quality-of-Security-Based Distance Vector Routing Protocol for Wireless Sensor Networks Using Ant Colony Optimization in 2008 IEEE International Conference on Wireless and Mobile Computing, Networking and Communications* (2008), 598–602.

35. Kang, B., Kim, D. & Choo, H., Internet of Everything: A Large-Scale Autonomic IoT Gateway, *IEEE Transactions on Multi-Scale Computing Systems* **3,** 206–214 (July 2017).

36. Evans, D., *How the Next Evolution of the Internet Is Changing Everything in* (2011).

37. Raza, U., Kulkarni, P. & Sooriyabandara, M., Low Power Wide Area Networks: An Overview, *IEEE Communications Surveys Tutorials* **19,** 855–873 (2017).

38. Mekki, K., Bajic, E., Chaxel, F. & Meyer, F., A comparative study of LPWAN technologies for large-scale IoT deployment, **5,** 1–7 (Mar. 2019).

39. Centenaro, M., Vangelista, L., Zanella, A. & Zorzi, M., Long-range communications in unlicensed bands: the rising stars in the IoT and smart city scenarios, *IEEE Wireless Communications* **23,** 60–67, ISSN: 1558-0687 (Oct. 2016).

40. Patel, D. & Won, M., *Experimental Study on Low Power Wide Area Networks (LPWAN) for Mobile Internet of Things in 2017 IEEE 85th Vehicular Technology Conference (VTC Spring)* (2017), 1–5.

41. Baharudin, A. M. & Yan, W., *Long-range wireless sensor networks for geo-location tracking: Design and evaluation in 2016 International Electronics Symposium (IES)* (2016), 76–80.

42. Guibene, W. *et al., Evaluation of LPWAN Technologies for Smart Cities: River Monitoring Use-Case in* (Mar. 2017), 1–5.

43. Sinha, R. S., Wei, Y. & Hwang, S.-H., A survey on LPWA technology: LoRa and NB-IoT, *ICT Express* **3,** 14–21, ISSN: 2405-9595, `https://www.sciencedirect.com/science/article/pii/S2405959517300061` (2017).

44. Bockelmann, C. *et al.,* Towards Massive Connectivity Support for Scalable mMTC Communications in 5G Networks, *IEEE Access* **6,** 28969–28992 (2018).

45. Abramson, N. M., *THE ALOHA SYSTEM: another alternative for computer communications in AFIPS '70 (Fall)* (1970).

46. Kotaba, R. *et al.,* How to Identify and Authenticate Users in Massive Unsourced Random Access, *IEEE Communications Letters* **25,** 3795–3799 (2021).

47. Chen, X., Chen, T.-Y. & Guo, D., Capacity of Gaussian Many-Access Channels, *IEEE Transactions on Information Theory* **63,** 3516–3539 (2017).

48. Shannon, C. E., A mathematical theory of communication, *The Bell System Technical Journal* **27,** 379–423 (1948).

49. Polyanskiy, Y., Poor, H. V. & Verdu, S., Channel Coding Rate in the Finite Blocklength Regime, *IEEE Transactions on Information Theory* **56,** 2307–2359, ISSN: 1557-9654 (May 2010).

50. Erseghe, T., Coding in the Finite-Blocklength Regime: Bounds Based on Laplace Integrals and Their Asymptotic Approximations, *IEEE Transactions on Information Theory* **62,** 6854–6883 (2016).

51. Lancho, A., Östman, J., Durisi, G., Koch, T. & Vazquez-Vilar, G., *Saddlepoint Approximations for Rayleigh Block-Fading Channels* Apr. 2019.

52. Chiani, M. & Martini, M. G., Analysis of Optimum Frame Synchronization Based on Periodically Embedded Sync Words, *IEEE Transactions on Communications* **55,** 2056–2060 (2007).

53. Kim, S., Joo, K. & Lim, Y., *A delay-robust random access preamble detection algorithm for LTE system in 2012 IEEE Radio and Wireless Symposium* (2012), 75–78.

54. Wu, G., Hu, S. & Li, S., *Low complexity time-frequency synchronization for transform domain communications systems in 2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)* (2015), 1002–1006.

55. Ye, Z., Duan, C., Orlik, P. V., Zhang, J. & Abouzeid, A. A., A Synchronization Design for UWB-Based Wireless Multimedia Systems, *IEEE Transactions on Broadcasting* **56,** 211–225 (2010).

56. Schmidl, T. M. & Cox, D. C., Robust frequency and timing synchronization for OFDM, *IEEE Transactions on Communications* **45,** 1613–1621 (1997).

57. Zhu, D. & Heath, R. W., *Directional timing synchronization in wideband millimeter wave cellular systems with low-resolution ADCs in 2017 51st Asilomar Conference on Signals, Systems, and Computers* (2017), 37–41.

58. Schlüter, M., Dörpinghaus, M. & Fettweis, G. P., Bounds on Phase, Frequency, and Timing Synchronization in Fully Digital Receivers With 1-bit Quantization and Oversampling, *IEEE Transactions on Communications* **68,** 6499–6513 (2020).

59. Lui, G. & Tan, H., Frame Synchronization for Gaussian Channels, *IEEE Transactions on Communications* **35,** 818–829 (1987).

60. Liang, Y., Rajan, D. & Eliezer, O., Sequential Frame Synchronization Based on Hypothesis Testing With Unknown Channel State Information, *IEEE Transactions on Communications* **63,** 2972–2984 (2015).

61. Ali, U., Kieffer, M. & Duhamel, P., *Frame Synchronization based on robust header recovery and Bayesian testing in 21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications* (2010), 933–938.

62. Chiani, M. & Martini, M., Practical frame synchronization for data with unknown distribution on AWGN channels, *IEEE Communications Letters* **9,** 456–458 (2005).

63. Chiani, M. & Martini, M., On sequential frame synchronization in AWGN channels, *IEEE Transactions on Communications* **54,** 339–348 (2006).

64. Suwansantisuk, W., Chiani, M. & Win, M. Z., Frame Synchronization for Variable-Length Packets, *IEEE Journal on Selected Areas in Communications* **26,** 52–69 (2008).

65. Elzanaty, A., Koroleva, K., Gritsutenko, S. & Chiani, M., *Frame synchronization for M-ary modulation with phase offsets in 2017 IEEE 17th International Conference on Ubiquitous Wireless Broadband (ICUWB)* (2017), 1–6.

66. Bana, A.-S., Trillingsgaard, K. F., Popovski, P. & de Carvalho, E., *Short Packet Structure for Ultra-Reliable Machine-Type Communication: Tradeoff between Detection and Decoding in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2018), 6608–6612.

67. Sekimoto, T. & Kaneko, H., Group Synchronization for Digital Transmission Systems, *IRE Transactions on Communications Systems* **10,** 381–390 (1962).

68. Massey, J., Optimum Frame Synchronization, *IEEE Transactions on Communications* **20,** 115–119 (1972).

69. Jia, H. & Dodds, D., *Frame synchronization for PSAM in AWGN and Rayleigh fading channels in Canadian Conference on Electrical and Computer Engineering, 2005.* (2005), 44–50.

70. Imad, R., Poulliat, C. & Houcke, S., *Frame Synchronization Techniques for Non-Binary LDPC Codes over GF(q) in* (Jan. 2011), 1–6.

71. Godard, D., Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems, *IEEE Trans. Commun.* **28,** 1867–1875 (1980).

72. Fujita, T., Uchida, D., Fujino, Y., Kagami, O. & Watanabe, K., *A burst modulation/demodulation method for short-packet wireless communication systems in 2008 14th Asia-Pacific Conference on Communications* (2008), 1–5.

73. Azari, A., Popovski, P., Miao, G. & Stefanovic, C., *Grant-Free Radio Access for Short-Packet Communications over 5G Networks in GLOBECOM 2017 - 2017 IEEE Global Communications Conference* (2017), 1–7.

74. Bloessl, B. & Dressler, F., mSync: Physical Layer Frame Synchronization Without Preamble Symbols, *IEEE Transactions on Mobile Computing* **PP,** 1–1 (Feb. 2018).

75. Walk, P., Jung, P., Hassibi, B. & Jafarkhani, H., MOCZ for Blind Short-Packet Communication: Practical Aspects, *IEEE Transactions on Wireless Communications* **19,** 6675–6692 (2020).

76. Rahamim, Y., Freedman, A. & Reichman, A., *Methods for carrier synchronization of short packet turbo coded signals in 2004 IEEE 15th International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE Cat. No.04TH8754)* **3** (2004), 1983–1987 Vol.3.

77. Cassaro, T. & Georghiades, C., Frame synchronization for coded systems over AWGN channels, *IEEE Transactions on Communications* **52,** 484–489 (2004).

78. Matsumoto, W. & Imai, H., *Blind synchronization with enhanced sum-product algorithm for low-density parity-check codes in The 5th International Symposium on Wireless Personal Multimedia Communications* **3** (2002), 966–970 vol.3.

79. Imad, R. & Houcke, S., *Blind frame synchronization and phase offset estimation for coded systems in 2008 IEEE 9th Workshop on Signal Processing Advances in Wireless Communications* (2008), 11–15.

80. Weinberger, N. & Merhav, N., Codeword or Noise? Exact Random Coding Exponents for Joint Detection and Decoding, *IEEE Transactions on Information Theory* **60,** 5077–5094 (2014).

81. Nagaraj, S., Khan, S., Schlegel, C. & Burnashev, M. V., Differential preamble detection in packet-based wireless networks, *IEEE Transactions on Wireless Communications* **8,** 599–607 (2009).

82. Liva, G., Durisi, G., Chiani, M., Ullah, S. S. & Liew, S. C., Short codes with mismatched channel state information: A case study, *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC),* 1–5 (2017).

83. Klaimi, R., Nour, C. A., Douillard, C. & Farah, J., *Low-complexity decoders for non-binary turbo codes in 2018 IEEE 10th International Symposium on Turbo Codes Iterative Information Processing (ISTC)* (Dec. 2018), 1–5.

84. Davey, M. C. & MacKay, D., Low-density parity check codes over GF(q), *IEEE Communications Letters* **2,** 165–167, ISSN: 2373-7891 (June 1998).

85. Lidl, R. & Niederreiter, H., *Introduction to Finite Fields and their Applications* 2nd ed. (Cambridge University Press, 1994).

86. Deschamps, J.-P., *Hardware Implementation of Finite-Field Arithmetic* 1st ed., ISBN: 0071545816 (McGraw-Hill, Inc., USA, 2009).

87. Howie, J. M., *Fields and Galois Field Theory* 1st ed., ISBN: 9781846286278 (Springer, London, 2007).

88. Poulliat, C., Fossorier, M. & Declercq, D., *Design of non binary LDPC codes using their binary image: algebraic properties in 2006 IEEE International Symposium on Information Theory* (2006), 93–97.

89. Abassi, O., Conde-Canencia, L., Mansour, M. & Boutillon, E., *Non-Binary Low-Density Parity-Check coded Cyclic Code-Shift Keying in 2013 IEEE Wireless Communications and Networking Conference (WCNC)* (2013), 3890–3894.

90. Dillard, G., Reuter, M., Zeiddler, J. & Zeidler, B., Cyclic code shift keying: a low probability of intercept communication technique, *IEEE Transactions on Aerospace and Electronic Systems* **39,** 786–798 (2003).

91. Berrou, C., Glavieux, A. & Thitimajshima, P., *Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1 in Proceedings of ICC '93 - IEEE International Conference on Communications* **2** (1993), 1064–1070 vol.2.

92. Moreira, J. & Farrell, P., *Essentials of Error-control Coding in* (2006).

93. Tanner, R., A recursive approach to low complexity codes, *IEEE Transactions on Information Theory* **27,** 533–547 (1981).

94. Chen, J. & Fossorier, M., Density evolution for two improved BP-based decoding algorithms of LDPC codes, *IEEE Communications Letters* **6,** 208–210 (May 2002).

95. Davey, M. & MacKay, D., *Low density parity check codes over GF(q) in 1998 Information Theory Workshop (Cat. No.98EX131)* (1998), 70–71.

96. Barnault, L. & Declercq, D., *Fast decoding algorithm for LDPC over GF(2q) in* (Jan. 2003), 70–73, ISBN: 0-7803-7799-0.

97. Barnault, L. & Declercq, D., *Fast decoding algorithm for LDPC over GF(2/sup q/) in Proceedings 2003 IEEE Information Theory Workshop (Cat. No.03EX674)* (2003), 70–73.

98. Song, H. & Cruz, J. R., Reduced-complexity decoding of Q-ary LDPC codes for magnetic recording, *Magnetics, IEEE Transactions on* **39,** 1081–1087 (Apr. 2003).

99. Declercq, D. & Fossorier, M., Decoding Algorithms for Nonbinary LDPC Codes Over GF(q), *IEEE Transactions on Communications* **55,** 633–643 (2007).

100. Voicila, A., Declercq, D., Verdier, F., Fossorier, M. & Urard, P., *Low-Complexity, Low-Memory EMS Algorithm for Non-Binary LDPC Codes in 2007 IEEE International Conference on Communications* (2007), 671–676.

101. Conde-Canencia, L., Al Ghouwayel, A. & Boutillon, E., Complexity Comparison of Non-Binary LDPC Decoders, *ICT-MobileSummit* (June 2009).

102. Savin, V., *Min-Max decoding for non binary LDPC codes in 2008 IEEE International Symposium on Information Theory* (2008), 960–964.

103. Boutillon, E., *"PN Sequence Generation in QCSP systems"* `https://qcsp.univ-ubs.fr/wp-content/uploads/2021/11/Sequence_PN.pdf`.

104. Kitsos, P., Sklavos, N., Zervas, N. & Koufopavlou, O., *A reconfigurable linear feedback shift register (LFSR) for the Bluetooth system in ICECS 2001. 8th IEEE International Conference on Electronics, Circuits and Systems (Cat. No.01EX483)* **2** (2001), 991–994 vol.2.

105. Sudeepa, K., Aithal, G., Rajinikanth, V. & Satapathy, S. C., Genetic algorithm based key sequence generation for cipher system, *Pattern Recognition Letters* **133,** 341–348, ISSN: 0167-8655, `https://www.sciencedirect.com/science/article/pii/S0167865520300933` (2020).

106. Cho, H., Song, H.-Y., Ahn, J. M. & Lim, D. W., Some new RS-coded orthogonal modulation schemes for future GNSS, *ICT Express* **7,** 530–534, ISSN: 2405-9595, `https://www.sciencedirect.com/science/article/pii/S2405959521000485` (2021).

107. Chauvat, R., Peña, A. G., Anghileri, M., Floch, J.-J. & Paonni, M., *Ultra-Sparse Binary LDPC Codes with CSK Signals for Increased Data Rates in Future GNSS in 2018 9th ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC)* (2018), 1–11.

108. China Satellite Navigation Office, *BeiDou Navigation Satellite System, Signal In Space, Interface Control Document, Open Service Signals* `http://en.beidou.gov.cn/SYSTEMS/Officialdocument/201806/P020180608525871869457.pdf`.

109. Beaulieu, N. C., An infinite series for the computation of the complementary probability distribution function of a sum of independent random variables and its application to the sum of Rayleigh random variables, *IEEE Transactions on Communications* **38,** 1463–1474, ISSN: 1558-0857 (Sept. 1990).

110. Akopian, D., Fast FFT based GPS satellite acquisition methods, *IEE Proceedings - Radar, Sonar and Navigation* **152,** 277–286, ISSN: 1350-2395 (Aug. 2005).

111. Fitz, M., Further results in the fast estimation of a single frequency, *IEEE Transactions on Communications* **42,** 862–864 (1994).

112. K. Saied, E. B., *Blind Synchronization Algorithm for QCSP* https://qcsp.univ-ubs.fr/wp-content/uploads/2022/01/QCSP_Synchronization.pdf.

113. Zhang, X., Cai, F. & Lin, S., Low-Complexity Reliability-Based Message-Passing Decoder Architectures for Non-Binary LDPC Codes, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* **20,** 1938–1950 (2012).

114. (2020) Web site on Non-Binary LDPC. [Online]. Available: http://www-labsticc.univ-ubs.fr/nb-ldpc/.

115. Song, H. & Cruz, J., Reduced-complexity decoding of Q-ary LDPC codes for magnetic recording, *IEEE Transactions on Magnetics* **39,** 1081–1087 (2003).

116. Savin, V., *Non-Binary Polar Codes for Spread-Spectrum Modulations in 2021 11th International Symposium on Topics in Coding (ISTC)* (2021), 1–5.

117. Bana, A.-S., Trillingsgaard, K. F., Popovski, P. & de Carvalho, E., *Short Packet Structure for Ultra-Reliable Machine-Type Communication: Tradeoff between Detection and Decoding in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2018), 6608–6612.

118. 3GPP, *Performance evaluation of LDPC codes for NR eMBB data* Discussion and decision R1-1713740, Version 6.1.4.1.6 (3rd Generation Partnership Project (3GPP), Aug. 2017), https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_90/Docs/R1-1713740.zip.

119. Monière, C., Legal, B. & Boutillon, E., *Efficient Software and Hardware Implementations of a QCSP Communication System in Accepted in The Workshop on Design and Architectures for Signal and Image Processing (DASIP22)* (Budapest, Hungary, 2022).

120. Ettus, Ettus official website: N210. [Online]. Available: https://www.ettus.com/all-products/un210-kit/.

121. USRPs, USRP N210 Datasheet. [Online]. Available: https://www.ettus.com/wp-content/uploads/2019/01/07495_Ettus_N200-210_DS_Flyer_HR_1.pdf.

122. Tucker, D. C. & Tagliarini, G. A., *Prototyping with GNU radio and the USRP - where to begin in IEEE Southeastcon 2009* (2009), 50–54.

123. Reynders, B., Meert, W. & Pollin, S., *Range and coexistence analysis of long range unlicensed communication in 2016 23rd International Conference on Telecommunications (ICT)* (2016), 1–6.

124. Mikhaylov, K., Juha Petaejaejaervi, & Haenninen, T., *Analysis of Capacity and Scalability of the LoRa Low Power Wide Area Network Technology in European Wireless 2016; 22th European Wireless Conference* (2016), 1–6.

125. Wang, Y.-P. E. *et al.,* A Primer on 3GPP Narrowband Internet of Things, *IEEE Communications Magazine* **55,** 117–123 (2017).

126. Adhikary, A., Lin, X. & Wang, Y.-P. E., *Performance Evaluation of NB-IoT Coverage in 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)* (2016), 1–5.

Title: Quasi-Cyclic Short Packet transmission for IoT

Keywords: Short Packets, CCSK, NB-LDPC, Over-modulation, Detection, Synchronization, FEC

Abstract: The efficient transmission of short frames is a prerequisite for the effectiveness of a wireless Internet of Things (IoT) network. In classical systems, each message is preceded by a preamble to help in detecting the arrival of the frame and simplify the demodulation operations. To avoid the bandwidth loss introduced by this preamble, we propose to study a new type of frame called Quasi-Cyclic Short Packet (QCSP). The whole QCSP frame can be considered first as a preamble to simply perform the detection and synchronization functions, then as an encoded codeword to correct the transmission errors. A QCSP frame is based on the association of a Cyclic Code Shift Keying (CCSK) modulation with a non-binary error correction code.
This PhD studies the QCSP frame reception problem by combining theoretical aspects with the definition and evaluation of algorithms. We first studied a detection algorithm adapted to QCSP frames. A theoretical model, validated by Monte-Carlo simulation, allows us to fully characterize the proposed algorithm. Then, we develop different time, frequency and phase synchronization algorithms. In particular, we propose to add an over-modulation of CCSK symbols to remove the time synchronization ambiguity at the symbol level. In addition, the non-binary code structure is also used to help the time and phase synchronizations. We formalized the QCSP frame parameter optimization problem as a trade-off between detection, synchronization and decoding performance. Finally, Software Defined Radio (SDR) modules allow us to experimentally validate the theoretical contributions of the thesis. It is thus possible to transmit 360 bits of information at a very low signal-to-noise ratio (-12 dB) with a transmission time reduced by 23% compared to the use of a classical frame.

Titre : Transmission par Trames Courtes Quasi-Cycliques pour l'Internet des Objets

Mot clés : Trames Courtes, CCSK, NB-LDPC, sur-modulation, Détection, Synchronisation, FEC

Résumé : La transmission efficace de trames courtes conditionne l'efficacité d'un réseau sans fil d'Internet des Objets (IoT). Dans un système classique, chaque message est précédé d'un préambule connu du récepteur pour l'aider à détecter l'arrivée de la trame et simplifier les opérations de démodulation. Pour éviter la perte de bande passante introduite par ce préambule, nous proposons d'étudier un nouveau type de trame appelée Quasi-Cyclic Short Packet (QCSP). L'ensemble d'une trame QCSP peut être considérée d'abord comme un préambule pour réaliser simplement les fonctions de détection et de synchronisation, ensuite comme un message codé pour corriger les erreurs de transmission. Une trame QCSP est constituée de l'association d'une modulation Cyclic Code Shift Keying (CCSK) avec un code de correction d'erreurs non binaire.
Cette thèse étudie le problème de réception de trames QCSP en combinant des aspects théoriques avec la définition et l'évaluation d'algorithmes. Nous avons tout d'abord étudié un algorithme de détection adapté aux trames QCSP. Un modèle théorique, validé par des simulations de Monte-Carlo, nous permet de le caractériser entièrement. Ensuite, nous proposons différents algorithmes de synchronisation en temps et en fréquence. Nous avons en particulier proposé d'ajouter une sur-modulation des symboles CCSK pour supprimer l'ambiguïté temporelle au niveau symbole. De plus, la structure du code non-binaire est aussi utilisée pour aider la synchronisation temporelle et fréquentielle. Nous avons formalisé le problème d'optimisation des paramètres d'une trame QCSP comme un compromis entre performance de détection, de synchronisation et de correction. Enfin, des modules radios nous permettent de valider expérimentalement les contributions théoriques de la thèse. Il est ainsi possible de transmettre 360 bits d'information à très faible rapport signal à bruit (-12 dB) avec une durée de la transmission réduite de 23% comparé à l'utilisation d'une trame classique.