MIMO Underwater Acoustic Communications over Time-varying Channels: from Theory to Practice

Candidate

Andrea Petroni,

Doctorate cycle XXX

Supervisor

Prof. Mauro Biagi

Acknowledgement

I would first like to thank my supervisor Prof. Biagi for his support through the last three years. He has been a guide during my research activity, continuously motivating me and always giving me the right advise whenerver I needed.

Besides, I would also to thank Prof. Scarano, member of my research group, who willingly followed my study passing his knowledge to me.

Last, but not the least, I thank my labmates and PhD colleagues who I spent each day of the last three years with. Their presence and friendship have been fundamental to let me making the most of this time.

Contents

List of Abbreviations

ACK	Acknowledge
APS	Acoustic Passive Sensor
ARQ	Automatic Repeat reQuest
AUV	Autonomous Underwater Vehicle
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CIR	Channel Impulse Response
CSE	Chip Spaced Equalizer
CSI	Channel State Information
CTS	Clear-to-Send
DCRC	Data Communication Recognition Classifier
DFE	Decision Feedback Equalizer
DFT	Discrete Fourier Transform
DW	Deep Water
FA	False Alarm
FBF	Feedback Filter
FEC	Forward Error Correction
FFF	Feedforward Filter
FFT	Fast Fourier Transform
FSE	Fractionally Spaced Equalizer
GBm	Go-Back-m
IPI	Interpulse Interference
ISI	Intersymbol Interference
JSW	Juggling Stop-and-Wait
LDPC	Low-Density Parity-Check
LMS	Least Mean Square
LOS	Line Of Sight
MD	Missing Detection
MIMO	Multiple-Input Multiple-Output
MLD	Maximum Likelihood Detection
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
MW	Medium Water
NACK	Negative Acknowledge
NARC	Natural/Artificial Recognition Classifier
NSR	Noise-to-Signal Ratio

OFDM	Orthogonal Frequency Division Modulation
OOK	On-Off Keying
OSH	Optimal Separating hyperplane
OSI	Open Systems Interconnection
OSTBC	Orthogonal Space-Time Block Coding
PAM	Pulse Amplitude Modulation
PPM	Pulse Position Modulation
PTS	Prepare-to-Send
QAM	Quadrature Amplitude Modulation
QTO-STBC	Quasi Trace-Orthogonal Space-Time Block Coding
RF	Radio Frequency
RLS	Recursive Least Squares
RTS	Request-to-Send
RTT	Round Trip Time
SAS	Synthetic Aperture Sonar
SEP	Symbol Error Probability
SER	Symbol Error Rate
SIFT	Scale-Invariant Feature Transform
SINR	Signal-to-interference-plus-noise Ratio
SISO	Single-Input Single-Output
SM	Spatial Multiplexing
SNR	Signal-to-noise Ratio
SOFAR	SOund Fixing and Ranging
SONAR	SOund NAvigation and Ranging
SR	Selective Repeat
SSC	Submarine Signal Company
STBC	Space-Time Block Coding
SVM	Support Vector Machine
SW	Shallow Water
S&W	Stop-and-Wait
TDMA	Time Division Multiple Access
TFS	Time-Frequency Slot
TO-STBC	Trace-Orthogonal Space-Time Block Coding
USR	Underwater Selective Repeat
UWA	Underwater Acoustic
UWAN	Underwater Acoustic Network
VQ	Vector Quantization
WHOI	Woods Hole Oceanographic Institution
WVT	Wigner-Ville Transform
ZF	Zero Forcing

Chapter 1

Introduction

Despite more than 70% of our planet surface is covered by water, today the underwater world can still be considered largely unknown. Rivers, lakes, seas and oceans have always been a fundamental resource for human life development, but at the same time they have often represented natural obstacles very hard to surmount. The most impressive example is probably given by the ocean, whose vastness severely limited geographical explorations and discoveries for tens of centuries. Anyway, the growing curiosity about what happens below the water surface has gradually led man to *immerse* in this unknown environment, trying to overcome its inaccessibility and figure out its secrets.

Underwater investigation and exploring have been increasingly supported by technology, advanced over time for different purposes (military, commercial, scientific). In this regard, providing a communication link between remote users has been recognised as one of the main issues to be addressed. The first significant solutions derived from the radio-frequency world, subject of study since the 19th century. Unfortunately both wired and wireless RF inspired signal propagation strategies were not evaluated as succesful. The first, since considering the deployment of meters (up to kilometers) of cable in depth, were too costly and difficult, while the latter ones did not offer good performance in terms of communication range due to signal attenuation.

An alternative way, examined with particular interest from the beginning of the 20th century, has been that one offered by acoustics. Actually, the study of sound and its propagation through different media has been an intriguing topic since the Old World Age, hence the attempt of messaging underwater has seemed to be a great opportunity to convey theoretical principles in a real application. In addition, not only humans but also marine animals use acoustic waves to communicate, even over several kilometers distances as demonstrated by whales. So, since already existing in nature, acoustic communications have been considered as potentially successfull,

furthermore representing an effective trade-off between feasibility and performance, especially if compared to the other electromagnetic signals-based methods.

Communication over RF channels has been extensively investigated so as to become a mature technology. The thorough knowledge about OSI (Open Systems Interconnection) model physical layer issues has allowed the researchers attention to be drawn to the upper layers. Following this direction, the recent advances in technology in this field have been accomplished mainly due to novelties in networks managing rather than to enhancements in the signal propagation study. Moving to acoustics, unfortunately this approach results to be failing if applied in the underwater scenario, as the major challenges rise indeed from physics matters. The underwater environment is varied and variable, so understanding the mechanisms that govern the propagation of sound in water is a key element for the design of a well-performing communication system. In this sense, the physical layer has therefore regained the centrality that has been diminished in other contexts.

The underwater acoustic communications can be adopted in a wide range of applications. The best-known are coastal monitoring, target detection, AUVs (Autonomous Underwater Vehicles) remote control, tsunami alarm, environmental data collection and transmission. Those ones are very specific activities, so the devices to be employed must sometimes meet very strict requirements. In this regard, the solutions commercially available provide good performance (that are paid in terms of high costs). On the other hand, the fact that hardware and software are usually copyrighted leads to have a *closed* system. Having reconfigurable devices is instead an opportunity to match the technology with the environment features and variations, especially in real-time applications.

Considering the comments outlined above, the following dissertation deals with the design of an acoustic communication system. The preliminary theoretical analysis regarding physical layer concerns, such as signal propagation and channel behavior, represents the starting point from which several proposals regarding the implementation of an underwater acoustic networks (UWAN) are introduced. In particular the context of Multiple-Input Multiple-Output (MIMO) communications is investigated, presenting several solutions about transmission schemes and receiver implementation. Furthermore some strategies for access and error control, established at the data link layer level, are proposed. It is worth highlighting that the following has not to be considered as a disjointed discussion about the topics just listed, instead this work aims at presenting practical solutions developed hand in hand with theory, making choices firstly by looking at what nature allows.

The remainder of the manuscript is organized as follows. Chapter 2

presents the history of underwater communications, describing the fields of application where this technology is commonly employed. In Chapter 3 the theory of sound propagation is illustrated, introducing also the concept of Sonar and reporting examples of practical applications. The features of the underwater acoustic channel are detailed in Chapter 4, while Chapter 5 investigates the scenario of MIMO communications. Chapter 6 instead concerns matters related to the OSI data link layer. Finally the conclusion is drawn in the last chapter.

Chapter 2

Underwater acoustics: history and applications

The interest about underwater acoustics is dated back to the Ancient Greece era and it has exponentially increased over time [1]. This curiosity was initially limited only to natural aspects of this phenomenon, then the advances in science allowed researchers to start thinking about sounds as potential means to be used in practical activities. So, getting to the 20th century when the knowledge about sound propagation reached a remarkable degree of maturity, the first acoustic technology-based applications (military above all) flourished [2].

2.1 From the beginning to the 1900s

In the 4th century BC, the Greek philosopher Aristotle was the first one realizing that sound could be heard both in water and in the air. This assumption derived from several observations of natural phenomena, so its discussion rimained mainly limited to theoretical matters. The same conclusion was reached about 2000 years later by Leonardo Da Vinci (1452-1519) who tried to investigate the acoustic propagation with a view to possible purposes for human activities. He wrote down in his notes:

"If you cause your ship to stop and place the head of a long tube in the water and place the other extremity to your ear, you will hear ships at great distances".
[Leonardo Da Vinci, 1490]

Tens of decades after Da Vinci, the english philosopher Francis Bacon (1561-1626) supposed in his work *Historia Naturalis et Experimentalis de Ventis* that sounds arose exactly from water environments and then spread around. Almost during the same years, an important step forward in un-

derstanding the principles of acoustics was made thanks to Galilei's (both Vincenzo and his son Galileo) studies on the properties of vibrating strings. Meanwhile, the French scientist and music theorist M. Mersenne, known to many as the "father of acoustics", published in 1637 *L'Harmonie Universelle* that included the early attempts at measuring the speed of sound in the air. Following this direction, in 1687 Sir Isaac Newton formulated in his masterpiece *Philosophiae Naturalis Principia Mathematica* the first mathematical theory about acoustic propagation. Despite his reasonings were only focused on sounds travelling in air, he believed that the same laws were valid also for the underwater scenario.

The first experiments to prove the feasibility of underwater sound transmission were done during the 18th and 19th century. In 1743 the physicist J.A. Nollet tested how sounds coming from different sources (bells, horns, gunshots, shouts) were perceived by putting his head underwater. This experience let him observe that for instance an alarm clock could be recognized by an underwater listener, but not by those ones in air. More, he inferred that most of sound attenuation occured nearby the water surface. Similar tests were performed later in 1780 by A. Monro II, leading him to suppose that the speed of sound in air and water were almost the same.
However, the most significant results about sound speed measurement came from the experience of the physicist J.D. Colladon and the mathematician J.F.K. Sturm on Lake Geneva in 1826. The setup of the experiment was composed of a bell immersed in water from a boat, used as transmitter, whose vibration produced the ignition of some gunpowder, producing a flash of light. This latter was useful to the scientists, standing in another boat placed 10 miles far and acting as receiver, in order to understand the transmission start for time measurements. Sound propagating through 8°C water was captured by means of a particular trumpet with two ends in the water and in the scientists hear respectively. The achieved result was astonishing since the estimated speed was 1435 m/s, that is just 3 meters per second less than the value currently recognised.
Carrying on towards the 20th century, the potentialities of acoustics were discussed a lot, aiming to find practical applications for this *new* technology. The scarse knowledge about the underwater world, especially the oceans, prompted the researchers to exploit sound for depth measurements and exploring, giving rise to the currently known as echo sounding. Unfortunately, the low reliability of instruments did not provide high performance, so the results were worse than expected. Anyway the Second Industrial Revolution played an important role during this period, speeding up the scientific progress. In addition, the theoretical base of sound propagation was strenghtened thanks to Lord Rayleigh's *The Theory of Sound* published in 1877, considered a milestone of the modern study of acoustics. The spread of new materials as well as the arrival of engine driven boats considerably

broadened the horizons of navigation, so the improvement for what concerned safety and communication among ships was required. Until that time, the coastal monitoring was performed by using lighthouses provided of sirens employed for warning the approaching boats. This method was often unreliable in case of bad weather conditions so, considering the rise of vessels traffic, new more efficient strategies had to be adopted. In this context, the input was given by T. Edison who designed a carbon-granule microphone, suitably covered by a waterproof envelope, to be placed on the hull of the boats. This device was able to detect the warning *messages* emitted by underwater bells positioned under the lightships, letting sailors understand that land was close.

2.2 The World Wars Age and the last decades

Straddling 19th and 20th centuries the scientific spotlight was stolen by the great advances in the radio-frequency field, so the acoustic technology struggled asserting itself. Anyway, the acoustic devices finally entered the market thanks to the Submarine Signal Company (SSC) founded in 1901 by E. Gray and A.J. Mundy who had already patented an underwater electrically driven bell (Fig. 2.1) two years earlier. In 1912, the sink of the
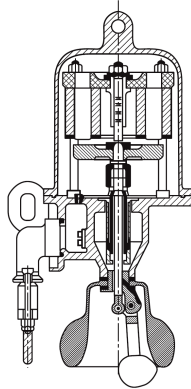
Figure 2.1: Electrically operated bell designed by Gray and Mundy.

Titanic due to an iceberg collision put the problem of navigation safety in the foreground again. Few months after the disaster, the Canadian inventor R. Fessenden joined the Submarine Signal Company and started working on an acoustic transmitter able to overcome the performance of the existing instruments. This device was called Fessenden oscillator (despite it was actually a transducer) and it represented the first model of active sonar. The hardware was composed of a circular metal plate immersed in water and power supplied by means of a copper tube moving in a magnetic field, so the vibration of the plate produced a sequence of tones propagating in

water. Furthermore, the acoustic emission could be tuned in order to realize a Morse code communication. Another considerable novelty characterizing this device was that it could operate in a reverse mode, allowing also the returning echoes to be detected and heard. So, after several tests on iceberg detection and communication between vessels, in 1924 the Fessenden oscillator was launched on the market.

With World War I, broke out in 1914, advancements in underwater acoustics became a real necessity in the military field. The Central Powers line-up could rely on a copious fleet using hydrodynamic echo sounders developed by the German physicist Heinrich Hecht. On the other side, the Allies's formation was not prepared to the submarine warfare at first since the instruments employed for enemies detection were "out-of-date" (they were still designed following the models proposed by Edison many decades before). The United States technology has been mainly provided by the Submarine Signal Company but, as they entered the war, the Naval Experimental Station (New London, CT) was founded so that the US Navy could have an independent research and development laboratory. G.W. Pierce, one of the most honoured members, solved the problem of phase delay related to the binaural detection, that is determining the position of an acoustic source basing on the signal captured by two different hydrophones. In Britain, the research activity for military purposes was controlled by the Board of Invention and Research, joined also by the physicists A.B. Wood and R.W. Boyle who realized one of the first piezoquartz acoustic transducers. The idea of Wood and Boyle was born thanks to the partnership with the French physicist P. Langevin that was working on piezoelectricity. The study of Langevin evolved during the World War I and beyond, leading him to design the prototype of an ultrasonic sonar.

Avoiding to be caught unprepared to a new underwater warfare resulted to be a priority for the great powers already involved in the WWI. The investigation on the marine environment became more accurate, especially for what concerned the in-water acoustics. So, during the '20s and '30s, the study of the relationship between the ocean natural features and sound propagation led to the development of a new research area called acoustical oceanography. Thanks to the new discoveries, remarkable enhancements in echo sounding and bathymetric activities were achieved. Moreover, the advances in electronics were exploited for building better performing devices. In this context, in 1925 G.W. Pierce designed a magnetostrictive transducer operating at 25 kHz that resulted to be more robust and reliable with respect to the previous crystal based instruments. This technology was also employed by the US navy for submarine detection and harbor surveillance as a whole (Fig. 2.2).

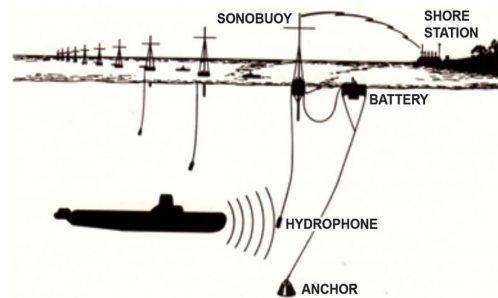The SONAR (SOund NAvigation and Ranging) activities were funda-

Figure 2.2: Naval Research Laboratory's "Herald" harbor security system.

mental in the World War II (1938-1945). At first the German fleet composed of a limited number of submarines did not represent a relevant trouble for the United States and the Allies, in fact the radar-provided airborne convoys and the marine sourveillance systems were sufficient for preventing underwater attacks. However, as the enemy increased its military power, an additional scientific and technological effort was necessary for the US in order to mantain the supremacy during the conflict. In this regard many universities (mainly Harvard, Columbia and the University of California) and other organizations such as the Woods Hole Oceanographic Institution (WHOI) were involved in military issues, supported also by industrial structures. The goal was to encourage the scientific progress not only for academic purposes but also for warfare applications. This strategy proved to be successful since it allowed a remarkable improvement of the underwater environment knowledge. A better understanding of phenomena like sound speed variation, absorption and scattering and other factors such as water temperature, salinity and currents was achieved thanks to many reserches and experiments. Furthermore, the first propagation models based on the Ray Theory (inspired by the theory of light) and Normal Mode were published. The latest theoretical advances led to the design of better performing scanning sonar systems and detection equipments, and also to the production of weapons as acoustic mine and homing torpedo.

Studies about underwater sound propagation continued also after the end of the World War II. One of the most important discoveries along the 1940s and 1950s was made by the geophysicist M. Ewing who was the first to predict and investigate on the SOFAR (Sound Fixing and Ranging) channel. This environment is represented by a horizontal layer of water in the ocean where the speed of sound is minimum and where acoustic waves experience a waveguide propagation. Especially at low frequencies this channel allows the sound to travel hundreds of kilomenters before attenuating, and this fact resulted to be very useful in sonar applications for submarine warfare.

Summarizing, the interest about the underwater world and acoustics has significally increased over the last two millenia. While at first the research was aimed only to scientific and academic scopes, in the last centuries warfare and military issues have represented the driving force behind the advances in underwater acoustics. By the end of the 20th century the research efforts have been mainly moved towards civilian applications such as industry or science. However, the underwater exploration remains fundamental in order to achieve new heights in technology.

2.3 Underwater acoustics in real applications

During the World Wars Age, ocean and coastal surveillance represented the fundamental application where underwater acoustics found use. In this regard, low-frequency technologies were particularly advanced since they provided vessels and submarines detection at long ranges. However, the employment of high-frequency acoustics started to take off thanks to the spread of new both military and civilian activities. This trend was supported by the development of advanced devices providing high spatial resolution, wide bandwidth, small size and relatively low cost. Furthermore, the use of sophisticated signal processing techinques has allowed further enhancements in terms of efficiency. Fig. 2.3 reports an interesting table describing the frequency range of the most diffused underwater acoustic systems.

Frequency (kHz)	0.1	1	10	100	1000
Maximum ranges (km)	1000	100	10	1	0.1
Multibeam sounders					
Sidescan sonars					
Transmission and positioning					
Active military sonars					
Passive military sonars					
Fishery echo sounders & sonars					
Doppler current profilers					
Sediment profilers					
Seismic exploration					

Figure 2.3: Frequency ranges of the main underwater acoustic systems [3].

Active sonars are probably the most known technology in the military field since they can be used for different purposes. In fact, as detailed in the next chapter, by transmitting a signal and then processing the reflections, some information about possible targets (submarines or ships) such as distance and bearing can be inferred. The same operation mode is exploited for sea floor inspection for mine countermeasures procedures as well [4]. An alternative solution is given by the passive sonar-based systems allowing *stealth* target detection since they measure sounds coming from the

surrounding environment without emitting any signal.

The presence of acoustic technology in underwater civilian activities is constantly increasing but already largely diffused thanks to the many scientific and industrial applications where it can be employed. In the context of ocean exploring, gliders and AUVs (Autonomous Underwater Vehicles) are usually equipped with multibeam acoustic sounders and/or sidescan sonars used for bathymetric measurements, sea-bed imaging and sediment profile analysis. Scientific data collection and transmission are performed exploiting acoustic communication networks that are also used for water quality monitoring and early warning against natural events as tsunamis, earthquakes and hurricanes [5]. Sensor networks are employed in the field of industry as well, for instance in the management and control of underwater pipelines structures. Furthermore, the location of hydrocarbon resources is investigated through marine seismic surveying. Finally, acoustic techonlogy turns out to be useful for regulating the shipping activity, optimizing the routes and getting information about vessels traffic.

Chapter 3

Theory of Sound and Sonar principles

Sound propagation can be explained by resorting to many principles of the theory of light since in both cases it comes to waves moving through a medium. Therefore attenuation, reflection, scattering and absorption are the typical occurrencies characterizing the underwater acoustic signal spreading as well. Anyway, as the features of these phenomena depend on both the nature of the source and the propagation medium, underwater acoustics results to be quite different from optics.

3.1 Acoustic waves

In physics, sound is described as a mechanical perturbation causing dilation and compression of particles in a medium. Each particle oscillates around its resting point transferring the energy to the surrounding points. This movement originates a pressure front propagating in a longitudinal direction from the source. The particles vibration is defined as *acoustic pressure*, measured in Pascals (Pa) or more frequently in microPascals (μPa) when dealing with underwater scenarios.

The local properties of the medium, expressed by the density ρ and the compressibility coefficient χ, determine the acoustic wave *propagation velocity* as:

$$c = \sqrt{\frac{1}{\rho\chi}} \qquad [m/s]. \tag{3.1}$$

Several models for calculating the speed of sound in water have been proposed [6], leading to a value approximately equal to $c = 1500$ m/s (the water density is equal to $\rho \simeq 1,03$ Kg/m^3). However in the sea water temperature, salinity and pressure change with depth so the sound speed profile may

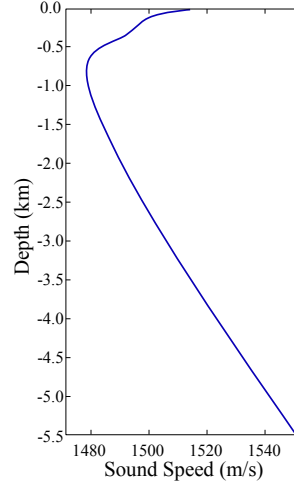present slight variations, as depicted in Fig. 3.1.

Figure 3.1: Underwater sound speed profile as a function of depth.

The acoustic signals employed in the context of communications can be described through a pressure front varying cyclically as a sine wave (Fig. 3.2). This periodic structure is characterized by a frequency f measuring
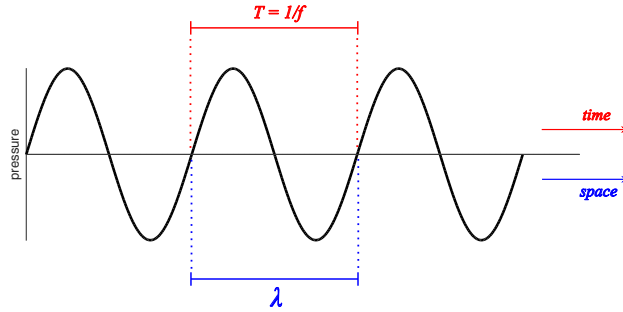
Figure 3.2: Space-time model of an acoustic wave.

the number of vibration cycles per second and expressed in Hertz (Hz). The period T is instead the inverse of frequency ($T = 1/f$) and refers to the duration in seconds of a single vibration cycle. Finally, the spatial description of the signal is given by the wavelength λ, that is essentially the distance covered by a wave over a period T, moving with velocity c. All the physical quantities introduced above are connected by the following relationship:

$$\lambda = cT = \frac{c}{f} \qquad [m]. \tag{3.2}$$

3.1.1 Plane and spherical wave equation

The phenomena related to waves propagation are accurately examined by resorting to the Huygens-Fresnel principle [7]. In the underwater context, the model of an acoustic wave is given by describing the evolution, both in time and space, of a pressure front moving through a fluid medium. For purposes of presentation, this latter is approximated as an *ideal* one, that is homogeneous, isotropic, perfectly elastic and non-dissipative.

Basing on this hypothesis, assume the presence of an expanding spherical source, increasing the local pressure and density. As the spreading of this effect through the medium particles is not instantaneous, the pressure variation in the points far from the source becomes function of the propagation direction. Consider an elementary region with volume V_0 and fluid mass $m = \rho_T V_0$ as in Fig. 3.3, with the fluid density $\rho_T = \rho_0 + \rho$ given by a static component ρ_0 and a variable one ρ due to the mass flow through the volume. The total pressure in the volume is $p_T = p_0 + p$, being p_0 the static ambient pressure and p the acoustic pressure component. As a par-
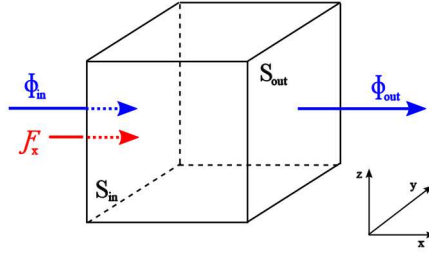
Figure 3.3: Elementary volume of a fluid medium.

ticle moves along the increasing x values direction, p lowers according to a gradient $\partial p / \partial x$, so the corresponding force applied along x is given by:

$$\mathcal{F}_x = -\frac{\partial p}{\partial x} dx dy dz \tag{3.3}$$

Then, having v_x defined as the particle velocity in the x direction, by recalling the Newton's Second Law $\mathcal{F}_x = ma_x$:

$$-\frac{\partial p}{\partial x} dx dy dz = \rho_T dx dy dz \frac{\partial v_x}{\partial t}$$

$$\frac{\partial p}{\partial x} = -\rho_T \frac{\partial v_x}{\partial t} \tag{3.4}$$

Under the hypothesis that density does not change significantly ($\rho \ll \rho_0$) it is possible to assume $\rho_T \simeq \rho_0$. Finally the extension of eq. (3.4) to the three-dimensional case leads to :

$$\nabla p = -\rho_0 \frac{\partial \vec{v}}{\partial t} \tag{3.5}$$

that represents the Euler's first law for point particle motion.

The perturbation coming from the source generates a mass flow passing through the region V_0. Due to the compressibility of the fluid medium the amount of mass entering V_0 may be greater than that one coming out, thus inducing variations in the local density. The mass flow rate per unit area is defined as the product of the fluid density, the area of the considered surface and the velocity component orthogonal to the same surface. By referring to Fig. 3.3, the difference between the flows related to the surfaces S_{in} and S_{out} along the x direction is:

$$\Delta \Phi = \Phi_{out} - \Phi_{in} =$$

$$= (\rho_T v_x dy dz + \frac{\partial \rho_T v_x}{\partial x} dx dy dz) - \rho_T v_x dy dz = \frac{\partial \rho_T v_x}{\partial x} dx dy dz \qquad (3.6)$$

Extended also to the y and z direction and simplifying $\rho_T \simeq \rho_0$, eq. (3.6) becomes:

$$\rho_0 (\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}) dx dy dz = \rho_0 \nabla \cdot (\vec{v}) dx dy dz \qquad (3.7)$$

with ∇ divergence operator. The mass conservation principle states that the variation of the mass contained in a volume must equate the difference between the entering and leaving flow, so:

$$\frac{\partial \rho}{\partial t} dx dy dz = -\rho_0 \nabla \cdot (\vec{v}) dx dy dz$$

$$\frac{\partial \rho}{\partial t} = -\rho_0 \nabla \cdot (\vec{v}) \qquad (3.8)$$

The acoustic phenomenon is assumed to be adiabatic, that is it does not induce temperature variation and the product pV^γ is constant (γ is defined as the ratio between the specific heat for constant pressure and the specific heat for constant volume of an ideal gas). So, considering the conditions before and after the acoustic perturbation, the following constraint is respected:

$$p_0 V_0^\gamma = p V^\gamma \qquad (3.9)$$

From eq. (3.9) it is possible to obtain p as a function of density:

$$p = p_0 \left(\frac{V_0}{V}\right)^\gamma = p_0 \left(\frac{\rho}{\rho_0}\right)^\gamma \qquad (3.10)$$

The derivative in time of eq. (3.10) under the hypothesis of small density variations gives the following Equation of state:

$$\frac{\partial p}{\partial t} = \frac{p_0 \gamma}{\rho_0} \frac{\partial \rho}{\partial t} \qquad (3.11)$$

The combination of eq. (3.8) and eq. (3.11) leads to:

$$\frac{\partial p}{\partial t} = -p_0 \gamma \nabla \cdot \vec{v} \tag{3.12}$$

and, by calculating the time derivative:

$$\frac{\partial^2 p}{\partial t^2} = -p_0 \gamma \nabla \cdot \frac{\partial \vec{v}}{\partial t} \tag{3.13}$$

Finally, by applying the divergence operator to both sides of eq. (3.4) and combinating with eq. (3.13) the following expression is obtained:

$$\nabla^2 p = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \tag{3.14}$$

representing the Wave Equation. ∇^2 is the Laplacian operator, while $c = \sqrt{\gamma \frac{p_0}{\rho_0}}$ is the propagation speed of the wave through the medium. When dealing with a fluid medium, c can be also expressed as a function of $K_s = \gamma p_0$ that is the adiabatic elastic modulus of the fluid. The sinusoidal nature of the acoustic wave allows the writing of eq. (3.14) in the form of the Helmholtz equation [3]:

$$\nabla^2 p + k^2 p = 0 \tag{3.15}$$

where $k = 2\pi f_0 / c = \omega / c$ is the wave number, with f_0 and ω the frequency and the pulsation of the wave, respectively.

By restricting the propagation to the single direction x, eq. (3.15) becomes:

$$\frac{\partial^2 p}{\partial x^2} + \frac{\omega^2}{c^2} p = 0 \tag{3.16}$$

whose solution is a pressure wave with constant amplitude and phase depending on space and time:

$$p(x, t) = P_0 e^{j(\omega t - kx)} \tag{3.17}$$

This kind of wave having its fronts orthogonal to the propagation direction is called *plane wave*.

Another different case is given by instead considering a point source radiating isotropically in the surrounding space. The solution of eq. (3.15) now gives:

$$p(R, t) = \frac{P_0}{R} e^{j(\omega t - kR)} \tag{3.18}$$

representing a *spherical wave*. The wave fronts are not planes as in the previous case but spheres all centered on the point source. Furthermore, the amplitude of the spherical wave is no more constant as in eq. (3.17) but decreases in $1/R$, being R the radial distance from the source.

When dealing with propagation in fluids, the medium particles move in a direction parallel to the propagation. In this context, the waves are usually called compressional waves or longitudinal waves. The propagation through solids allows also the particles to move orthogonally to the propagation direction giving rise to a transverse component called shear wave. In the underwater scenario, the first model is sufficiently accurate to describe the acoustic propagation, so in most of cases the shear wave component can be neglected.

3.1.2   Acoustic impedance and intensity

By resorting to the model of plane wave it is possible to investigate on the relationship between acoustic pressure and particle velocity. Considering a progressive wave moving along the positive x diretion, the variation in time of the particle velocity $v_x = v(x - ct)$ leads to the following expression:

$$\frac{\partial v_x}{\partial t} = -c \frac{\partial v}{\partial t} \tag{3.19}$$

Substituting eq. (3.19) in the one-dimensional Euler's equation (eq. (3.5)) and integrating returns:

$$\frac{\partial p}{\partial x} = \rho_0 c \frac{\partial v}{\partial x}$$

$$p = \mathcal{Z} v \tag{3.20}$$

with $\mathcal{Z} = \rho_0 c$ representing the *characteristic acoustic impedance* of the propagation medium, measured in Rayl (equal to $Kg \cdot s^{-1} \cdot m^{-2}$ in the MKS system of units).

The acoustic impedance turns out to be useful to determine the acoustic intensity I, defined as the mean energy flux per unit of surface and time. Assume a plane wave propagating along the positive x direction, then the instantaneous intensity is given by the product of the pressure and the particle velocity, both depending on time and position [7]:

$$i_x = p(t - x/c) v_x(t - x/c) \tag{3.21}$$

Taking into account eq. (3.20) the intensity becomes function of the pressure only, so:

$$i_x = \frac{[p(t - x/c)]^2}{\rho_0 c} \tag{3.22}$$

It is worth noting that the same expression is valid also for spherical waves, but considering the R direction instead of x. Hence, for a sinusoidal wave the value of i_x is:

$$i_x = \frac{P_0^2 \sin^2(kx - \omega t)}{\rho_0 c} = P_0^2 \frac{1 - \cos[2(kx - \omega t)]}{2\rho_0 c} \tag{3.23}$$

Finally, the integration over time of eq. (3.23) returns the average intensity:

$$I_x = \langle i_x \rangle = \frac{P_0^2}{2\rho_0 c} = \frac{P_{0_{rms}}^2}{\rho_0 c} \qquad [W/m^2] \qquad (3.24)$$

with $P_{0_{rms}} = P_0/\sqrt{2}$ the pressure RMS amplitude.

3.1.3 Measure units: sound pressure level and decibels

Despite pressure is usually quantified in Pascal, due to its large dynamics it is convenient to resort to the logarithmic notation. Once given a reference level p_{ref}, the absolute pressure level is written as:

$$p_{dB} = 20 \log_{10}\left(\frac{p}{p_{ref}}\right) \qquad [dB] \qquad (3.25)$$

Both p and p_{ref} must be expressed according to the same definition, that is they must be both peak-to-peak or RMS values. Furthermore, in underwater acoustics the pressure reference level is $p_{ref} = 1 \ \mu Pa$ while the reference in air is $p_{ref} = 20 \ \mu Pa$. Hence, every time a sound pressure level measured in decibels is reported, it is necessary to specify the reference pressure value.

The logarithmic notation is used also to describe the relation between the acoustic pressure and power radiated by a source. The sound power $\mathcal{P}$ at the distance of 1 meter is given by the acoustic flux orthogonally passing through a surface of area A, so it follows that:

$$10 \log_{10}\left(\frac{\mathcal{P}}{1W}\right) = 10 \log_{10}\left(\frac{IA}{1W}\right) = 10 \log_{10}\left(\frac{I}{1W/m^2}\right) + 10 \log_{10}\left(\frac{A}{1m^2}\right) =$$

$$= 20 \log_{10}\left(\frac{p_{rms}}{1\mu Pa}\right) - 181.8 + 10 \log_{10}\left(\frac{A}{1m^2}\right) \qquad (3.26)$$

where $A = 4\pi$ when dealing with spherically radiating sources.

3.1.4 Reflection and transmission

The progress of a traveling acoustic wave is considerably influenced by the features of the medium. Especially in the underwater scenario where the propagation is constrained by the water surface and floor, reflection and transmission characterize the spreading of waves through the medium. These phenomena occur due to the presence of obstacles in water (fishes, boats, rocks), but most of all they originate by the incidence of waves at the interface between two media. The process is essentially the same presented for electromagnetic waves, therefore the theoretical analysis follows similar steps.

Assuming the incident, reflected and transmitted waves as locally plane, the simplest model of wave reflection/transmission is depicted in Fig. 3.4. The reflected wave propagates in a specular direction with respect to the incident one (that is, $\theta_i = \theta_r$), on the other hand the transmitted wave spreads following the Snell's law:

$$\frac{\sin \theta_i}{c_1} = \frac{\sin \theta_t}{c_2} \tag{3.27}$$

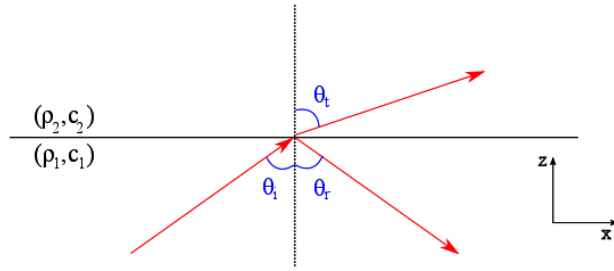Considering the (x,z) plane, the boundary conditions for pressure and ve-

Figure 3.4: Wave propagation according to the Snell's law.

locity normal component at $z = 0$ state that:

$$\begin{cases} p_i + p_r = p_t \\ v_{iz} + v_{rz} = v_{tz} \end{cases} \tag{3.28}$$

For what concerns velocity, the normal component is given by:

$$\begin{cases} v_{iz} = v_i \cos \theta_i \\ v_{rz} = v_r \cos \theta_r = v_r \cos \theta_i \\ v_{tz} = v_t \cos \theta_t \end{cases} \tag{3.29}$$

Furthermore, by recalling eq. (3.20) leads to:

$$\begin{cases} v_{iz} = \dfrac{p_i}{\rho_1 c_1} \cos \theta_i \\ v_{rz} = -\dfrac{p_r}{\rho_1 c_1} \cos \theta_i \\ v_{tz} = \dfrac{p_t}{\rho_2 c_2} \cos \theta_t \end{cases} \tag{3.30}$$

The *reflection* and *transmission coefficient*, defined as $\mathcal{V} = p_r/p_i$ and $\mathcal{W} = p_t/p_i$ respectively and linked by the dependence $\mathcal{W} = 1 + \mathcal{V}$, are introduced to show the relation between incident, reflected and transmitted

waves. Finally, given an incidence angle θ_i, substituting eqs. (3.30) in eqs. (3.28) returns:

$$\begin{cases} \mathcal{V}(\theta_i) = \dfrac{\rho_2 c_2 \cos \theta_i - \rho_1 c_1 \cos \theta_t}{\rho_2 c_2 \cos \theta_i + \rho_1 c_1 \cos \theta_t} \\ \mathcal{W}(\theta_i) = 1 + \mathcal{V}(\theta_i) = \dfrac{2\rho_2 c_2 \cos \theta_i}{\rho_2 c_2 \cos \theta_i + \rho_1 c_1 \cos \theta_t} \end{cases} \qquad (3.31)$$

Eqs. (3.31) are used to obtain the amplitudes of transmitted and reflected waves. Moreover, for $\theta_i \simeq 0$ the reflection coefficient can be expressed as a function of the characteristic acoustic impendances of the media, $\mathcal{Z}_1$ and $\mathcal{Z}_2$ respectively, so:

$$\mathcal{V}(\theta_i \simeq 0) = \frac{\rho_2 c_2 - \rho_1 c_1}{\rho_2 c_2 + \rho_1 c_1} = \frac{\mathcal{Z}_2 - \mathcal{Z}_1}{\mathcal{Z}_2 + \mathcal{Z}_1} \qquad (3.32)$$

If there is a large mismatch between the acoustic impedances, that is $\mathcal{Z}_2 \ll \mathcal{Z}_1$ or $\mathcal{Z}_2 \gg \mathcal{Z}_1$, the incident wave is quasi-totally reflected. This represents the case of *perfect* boundary where $\mathcal{V} \simeq \pm 1$ independently of θ_i.

By considering the reflection surface as ideally plane, the energy loss of the reflected wave would be minimum. However this assumption is never verified in facts since the acoustic transmission/reflection interface is characterized by a microscale roughness that influences the wave propagation when its dimensions are comparable with the signal wavelength. The interface roughness can be expressed by means of the Rayleigh parameter [3]:

$$\mathcal{P} = 2kr \cos \theta_i \qquad (3.33)$$

defined as a function of the incident angle θ_i, the acoustic wavenumber k and the relief amplitude standard deviation r. Hence, depending on the above cited factors, the incident wave scatters in different directions. Part of the signal is reflected without deformation in the specular direction (coherent reflection) while the remainder spreads in the surrounding space, going back also to the source (backscattering). For $\mathcal{P} \leq \pi/2$ the intensity of the signal coherent component can be quantified through the coefficient:

$$\mathcal{V}_{coh} = \mathcal{V} \exp(-\frac{\mathcal{P}^2}{2}) = V \exp(-2k^2 r^2 \cos^2 \theta_i) \qquad (3.34)$$

with $\mathcal{V}$ equal to the reflection coefficient measured for an ideal plane surface. Increasing P, the scattered part of the signal prevails on the coherent one. The scattering at the sea surface is mainly due to the waves, caused by the wind, producing irregularities at the interface. Furthermore, air bubbles created by the water movement represent another phenomenon that may increase the microscale roughness. At the other bound of the medium, the acoustic signal reflection is influenced by the geological composition of the seafloor, usually made of sand, mug and rocks.

3.2   Sonar and transducers

Devices able to emit and/or detect acoustic signals are known as *SONAR* (SOund NAvigation and Ranging). As described in par. 2.3 the activities performed in underwater environment are several and different. Moreover, the features of the employed hardware should be tailored to the requirements of the considered application. In general, sonars fall into two main categories: *passive* and *active*.

Passive sonars (Fig. 3.5a) are realized by acoustic devices able to listen and record sounds originated from external sources. Typical transducers used in passive sonar systems are *hydrophones* (analogous to microphones for out-of-water applications) that convert the front pressure representing the acoustic wave into an electrical signal. Passive sonars find their use in scientific activities such as ocean monitoring and seismic event detection, but also in coastal surveillance and ships/submarines recognition that are typical tasks in the military field.
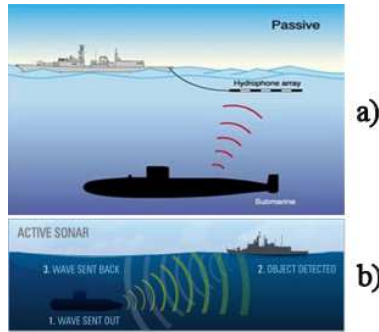

Figure 3.5: Example of passive (a) and active (b) sonar application [8].

Active sonars (Fig. 3.5b) are devices composed of a *projector* emitting a sequence of one or more acoustic pulses and a hydrophone measuring the echoes coming back from the surrounding objects and obstacles. By analyzing the returning echoes it is possible to infer the distance and bearing of the detected target, information that turns out to be essential in navy actions. Furthermore, the features of the active sonars are exploited also for exploring activities, especially in the so called *echo sounding* concerning bathymetry measurements.

The sonar principle of operation is strictly related to the physical phenomena happening underwater. Introducing some parameters representing all those effects described in par. 3.1 allows the formulation of the *sonar equation* that summarizes the relationship between equipment, medium and target [9].

3.2.1 Passive sonar

The passive sonar captures the sound produced by a certain target (Fig. 3.6), whose radiated noise is quantified through the Source Level parameter SL measured at the distance of 1 yd. The signal propagating through the medium experiences absorption and attenuation expressed in terms of Transmission Loss TL. Moreover, a component of background noise, assumed to be isotropic, is summed to the target signal. However, the detected Noise Level NL is reduced by the Directivity Index DI of the sonar transducer (hydrophone). The sum of the parameters introduced above, expressed in
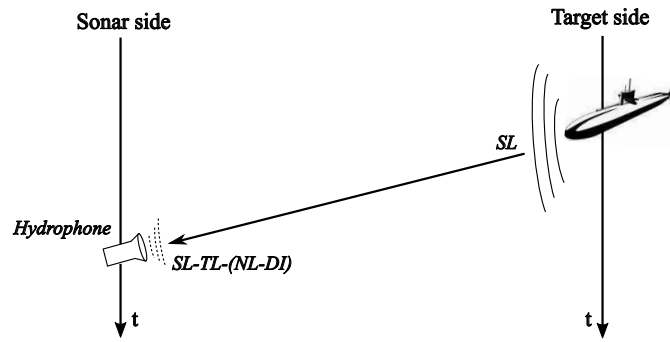

Figure 3.6: Time description of passive sonar signal detection.

decibels, returns the following Signal-to-Noise Level $L_{S/N}$:

$$L_{S/N} = SL - TL - (NL - DI) \qquad [dB] \qquad (3.35)$$

The detection of a target occurs when the $L_{S/N}$ exceeds a certain threshold DT, that is when:

$$SL - TL - (NL - DI) = DT \qquad (3.36)$$

The rearrangement of parameters in eq. (3.36) so that useful signal and noise components are separated leads to obtain the statement:

$$SL - TL = NL - DI + DT \qquad (3.37)$$

representing the passive sonar equation.

3.2.2 Active sonar

The case of active sonar (Fig. 3.7) is rather different from the passive one, since now the radiating source is the sonar itself. The signal emitted by the projector with Source Level SL reaches the target and, after reflection, it comes back to the sonar receiving transducer. Therefore, as the path of the signal doubles with respect to the passive sonar scenario, the Transmission

Loss TL has to be counted twice too. Furthermore, when the sound wave strikes a target, a portion of the incident energy is absorbed while the remaining part is scattered. The ability of the target to reflect energy back in the sonar direction is defined as Target Strenght TS. So considering a back-
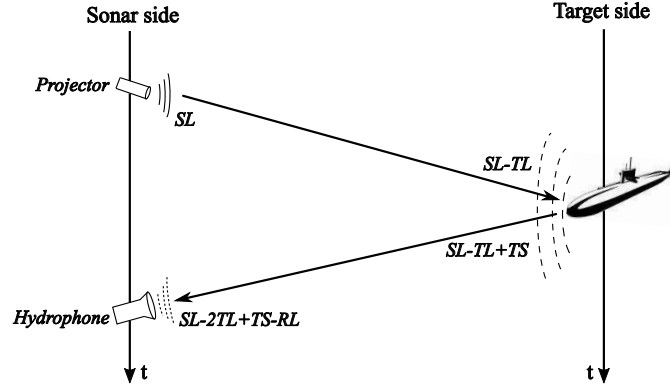

Figure 3.7: Time description of active sonar signal detection.

ground Noise Level NL, the Directivity Index DI and setting a Detection Threshold DT, the active sonar equation is given by:

$$SL - 2TL + TS = NL - DI + DT \qquad (3.38)$$

It is woth noting that the active sonar, as an emitter of acoustic energy, causes an additive self-noise component due to the presence of objects and obstacles (bubbles, rocks, fishes) producing reflections that superimpose over those ones coming from the target. This other noise source is expressed in terms or Reverberation Level RL and it is usually introduced in place of $NL - DI$ in eq. (3.38) that becomes:

$$SL - 2TL + TS = RL + DT \qquad (3.39)$$

Eqs. (3.38) and (3.39) refers to the *monostatic* active sonar where projector and hydrophone are represented by the same transducer. Sometimes transmitting and receiving transducers are instead separated, representing the so called *bistatic* sonar. In this latter case the transmission losses to and from the target are, in general, different, so the accuracy of eqs. (3.38) and (3.39) is limited.

Some parameters as the Source Level SL are employed in the description of both active and passive sonar behavior even if they refer to different elements. Therefore Table 3.1 gathers all the introduced parameters sorted by case (active or passive sonar) and by origin (sonar equipment, medium, target).

Table 3.1: Sonar equations parameters

	Equipment parameters	Medium parameters	Target parameters
Passive Sonar	Hydrophone Directivity Index DI Detection Threshold DT	Transmission Loss TL Background Noise Level NL	Target Source Level SL
Active Sonar	Projector Source Level SL Hydrophone Directivity Index DI Detection Threshold DT Self-Noise Level NL	Transmission Loss TL Background Noise Level NL Reverberation Level RL	Target Strength TS

3.3 Transducers

The performance of projectors and hydrophones, responsible for electro-acoustic signal conversion, represent a key element to consider for the design of efficient underwater transceivers. The most commonly used technology in the field of transducers is made from ceramic hardware exploiting its piezo-electricity properties. Once struck by a pressure front, this kind of material distorts generating a voltage. The process is reversible, so by applying an external voltage the ceramic vibration is induced, resulting in a sound radiated around. An alternative solution to piezoelectric devices are the magneto-strictive transducers. In that case vibrations are driven by a magnetic field applied on a ferromagnetic material causing its shape variation. Being very robust to high amplitude vibrations, the magneto-strictive transducers are mostly employed for low frequency high-power applications. However, they exhibit lower efficiency and a narrower frequency bandwidth with respect to piezoelectric devices. Other specific scenarios request the employment of more complex architectures such as mechanical or hydraulic power sources driving the transducer vibration.

The recent advances in technology have allowed the production of materials providing a higher energy-conversion efficiency, thus improving the performance of transducers. As the amount of energy poured out in water by piezoelectric-based devices is moderate, in the past explosive charges were usually employed to produce high-energy pulses. Scientific measurement campaigns make especially use of this technique. Lately, the investigation on the matching between transducer and water acoustic impedance has led to the development of more powerful transmitting devices.

In particular, this paragraph is devoted to the description of the equipment employed in several tests reported in the next chapters. Therefore the information reported below represents a reference whenever the setup of an experiment is introduced.

Moreover, App. B and App. C report the descritpion of two sonar applications, namely harbor traffic monitoring [10] and bathymetric measurements [11], where this technology find use.

3.3.1 AS-1 hydrophone

The AS-1 hydrophone developed by Aquarian Audio & Scientific [12] is a passive transducer allowing the detection and recording of sounds coming from the surrounding environment (Fig. 3.8). This hardware is designed to provide maximum sensitivity and linear response in the bandwidth of the highest-quality commercial digital sound recording interfaces. The response

Figure 3.8: AS-1 hydrophone and Hydromic pre-amplifier.

is omnidirectional in the human auditory bandwidth, as well as omnidirectional in the horizontal axis at all frequencies (theoretical). The main features of the hydrophone are listed in Table 3.2. The signals acquired by

Table 3.2: AS-1 hydrophone parameters

Linear range	1Hz to 100kHz $\pm$ 2dB
Receiving Sensitivity	-208dBV re 1μPa (40μV / Pascal)
Horizontal Directivity (20kHz)	$\pm$0.2dB
Horizontal Directivity (100kHz)	$\pm$1dB
Vertical Directivity (20kHz)	$\pm$1dB
Vertical Directivity (100kHz)	6dB - 11dB
Operating depth	200m
Output connection	BNC (standard)
Encapsulant	Polyurethane
Operating temperature range	-10°C to +80°C
Cable length	10 m

the hydrophone are directly recorded on a laptop thanks to the use of the Hydromic preamplifier [13]. This device is recognized as a standard USB audio class single channel microphone characterized by a standard BNC input and a USB 2.0 output for the hydrophone and PC connection respectively. The board includes a 250 kHz A/D converter and an integrated 32 bit 80 MHz microcontroller that allow the signal sampling at 192 kHz with 16 bit

per sample.

3.3.2   SAM-1 acoustic modem

The SAM-1 acoustic modem (Fig. 3.9) is a device produced by the Desert
Star Systems LLC [14] that allows the half-duplex transmission and recep-
tion of digital data between underwater stations up to a typical range of 250
meters, with up to 1000 meters in quiet deep water environment. This device
can be connetted and controlled via computer using a serial port connection,
but since it does not have an internal battery an external power supply is
also requested. When used in single channel model, the SAM-1 emits acous-

Figure 3.9: SAM-1 acoustic modem.

tic pulses around a carrier frequency equal to 33,8 kHz, while the use of
other three different carriers is allowed when dealing with frequency hop-
ping mode [15]. Other general hardware specifications are reported in Table
3.3. The data communication features of the SAM-1 modem is based on the
segmentation of the message into 2-character blocks, and then the transmis-
sion of the blocks in sequence. SAM-1 encodes each new data block with an
independent group of 6 acoustic pings. The first ping does not carry data
since it is used for frame synchronization purposes by the receiver. Each of
the remaining five pings represents a 4-bit chunk in the data block and each
4-bit group is mapped according to a 16-PPM (Pulse Position Modulation)
scheme. Hence, the 20 bits of the data block are encoded by six *ping* pulses
contained in a pulse window as shown in Fig. 3.10. However, thanks to sym-
bol re-mapping operations (A.1), it has been also possible to implement 2, 4
and 8-PPM schemes. Pulse position modulation was chosen for the SAM-1
system because it is a very energy efficient modulation format. This means
that there is a specific window of an assigned time duration in which each of
the data pings must occur. As illustrated in Fig. 3.10, a PPM time frame
for SAM-1 contains five fixed duration data ping windows sandwiched with
recovery periods. Each symbol window is evenly divided into 16 slots, and

Table 3.3: SAM-1 acoustic modem specifications

Depth Rating	300 meters max.
Operating temperature	0 - 70 °C
Data I/O	RS-232 serial data link
Sonar transceiver	TX power 183dB@8V - 189db@16V
	4-channel frequency hopping
	TX bandwidth 33.8kHz - 42kHz
	Single channel receiver 33.8 kHz
Sonar transducers	33 - 42 kHz omnidirectional
Sonar modulation	PPM-based single channel RX
Data speed (single channel)	S4: 5 bit/s
	S5: 13 bit/s
	S6: 38 bit/s
Sonar data format	16-bit data + 4 checksum bits

one ping will be transmitted during one of these sub-windows. Which sub-window the ping is in determines the four-bit value that the ping represents. In other words, the value of 0000 is associated with the first sub-window, 0001 with the second, and so on. It is important to note that the pulses are not separated in time and the receiver may suffer for error detection due to multipath. Analytically speaking, the signal received in the i-th ping slot ($i \in \{1, ..., 16\}$) of the k symbol window can be written as:

$$r_{k,i}(t) = s_{k,i}(t - jT_c) * h(t) + n(t) + w(t) \qquad (3.40)$$

where $s_{k,i}(t - jT_c)$ is the symbol j carrying the PPM time shift ($j \in \{0, 1, ..., 15\}$) and T_c is the possible shift. Moreover $h(t)$ represents the channel impulse response and it is responsible for possible reflections (multipath) while $n(t)$ is the acoustic noise and $w(t)$ is the thermal additive white Gaussian noise.

For what concerns the receiving mode, the SAM-1 modem first operates the frame synchronization by resorting to a threshold-based detection mechanism that allows the recognition of the first symbol of each frame (the threshold can be chosen by the user over a set of 99 possible values). The ping detection works according to the following rule:

$$\hat{s}_{k,i} = \arg \max_{j=0,1,...15} y_{k,i}(j) \qquad (3.41)$$

where

$$y_{k,i}(j) = \int_{jT_c}^{jT_c+T_p} r_{k,i}(t)s_{k,i}(t - jT_c)dt \qquad (3.42)$$

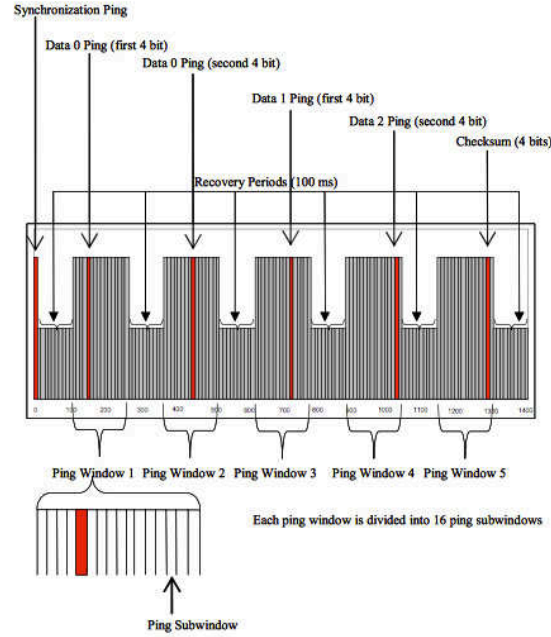T_p being the pulse duration of $s_{k,i}(t)$.

Figure 3.10: SAM-1 frame structure as reported in [15].

3.3.3 Adaptive Detection

Frame synchronization is a fundamental task since it gives the right timing and also the time distance among symbols belonging to the same frame. However, the approach described in eq. (3.41) is really sensible to phase rotation and attenuation as well as multipath. Furthermore, it is worth noting that when the channel is minimum phase, like for example in line-of-sight (LOS) propagation, it is usually expected that the first path presents lower attenuation with respect to the others. This is mainly true in RF applications, while in acoustics it may happen that the first path is comparable in amplitude (i.e., attenuation) with the others. This event occurs when the difference of path length (that is the difference of the length of two propagation rays) is low. In other words, this is equivalent to have a pulse duration larger than the time difference of arrival of two signal replicas.

Concerning signal detection, an improvement in terms of efficiency can be obtained by implementing and adaptive strategy in order to counterbalance the effect of the time-varying channel [16].
In this regard, the key elements to be taken into account are the synchronization, channel knowledge and external interference. The first one gives a reference for the frame transmission, the second one is responsible for interpulse (channel delay spread inducing confusion among symbols) and/or intersymbol (channel delay spread inducing confusion among symbols of dif-

ferent frames) interference while the third one is a colored, and in general sporadic, noise.

3.3.3.1 The role played by synchronization

When the receiver is waiting for the synchronization pulse, it may happen that this latter is attenuated and the detection fails. More reasonably a *confusion* can incur, that is the receiver recognizes some peaks as they are synchronization pings (see Fig. 3.11). Since the shape of the synchronization ping is known, a cross-correlation can be operated as in eq. (3.42), storing amplitude and delay of all the peaks exceeding a threshold. In the non-realistic case of ideal channel only one peak would be recognized, while in a multipath channel scenario M different peaks exceeding the threshold are expected. In the case of LOS channel those values are descending sorted while if NLOS is considered not specific rules are followed. So, for each of the M values it is possible to measure:

$$\tilde{h}(j) = \frac{1}{E_s} \int_{jT_c}^{jT_c+T_p} r_{k,i}(t)s(t)dt \geq \theta_h \qquad (3.43)$$

being θ_h the considered threshold. Finally, the reference time for synchronization is given by:

$$\tilde{\tau}_{sync} = T_s \arg \max_{j=1,...,M} \tilde{h}(j) \qquad (3.44)$$

with T_s the sample time.

3.3.3.2 The role played by channel knowledge

Under the sufficient realistic assumption of channel not changing within a frame, that is, six pings, the information acquired during synchronization phase can be used also for data detection. In fact, since the position of pilots is known within each frame, the receiver is able to evaluate the right symbol despite the possible presence of pulses that are comparable in amplitude. If the channel was ideal the solution would be trivial since $M{=}1$ and only one path is present, so the receiver as in eq. (3.41) will perform very reliably. In the more general case of $M{>}1$ the main component can be due to path overlapping, so a *new temporal reference* for detection is taken by starting from the above estimated $\tilde{\tau}_{sync}$ in eq. (3.44). This is equivalent to consider the conventional 16 receiving windows (as 16-PPM is used), but since the time reference has been changed the rule to follow for symbol detection becomes:

$$\hat{s}_{k,i} = \arg \max_{j=q,q+1,...q+15} y_{k,i}(j) \qquad (3.45)$$

where $q = \tilde{\tau}_{sync}/T_c$. Although this scheme solves the problem of interpulse interference since it realigns the time axis, it fails to work with intersymbol

interference. However, due to the signaling time characterizing the SAM-1 modem and by considering the specific application scenario, it is verified that no intersymbol interferene occurs (among different frames) while interpulse interference is very frequent.

3.3.3.3 The role played by external sounds

External interference is really an issue since, depending on the operating bandwidth, undesired signals can be present at the receiver. An example is depicted in Fig. 3.11, where the data acquired by a hydrophone during some experiments contain not only the signal coming from the SAM-1 modem but also sounds originated by the waves movement and rowing canoeists (the acquired signal has been also band-pass filtered with a bandwidth of 10 kHz). Considering for instance the scenario depicted in Fig. 3.11, the problem
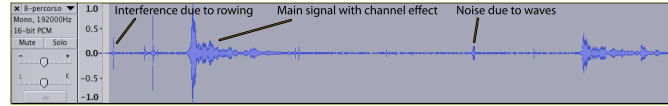

Figure 3.11: Example of data signal plus noise due to canoists.

related to this kind of noise is that the receiver may interpret the first peak as the first synchronization ping so leading to drop all the consecutive data.

3.3.3.4 Experimental results

The proposed adaptive detection scheme (eq. (3.45)) has been tested and compared with those ones achieved when using the direct detection (that characterizes the SAM-1 modem) according to eq. (3.41). The experiments were conducted in a very shallow water reservoir in Rome (41.82821 °, 12.467891°) [11] and consisted of the trasmission of the following text - *Adaptive PPM Acoustic Detection in very Shallow Water Reservoir AMOUR-AQUALAB Project ICT Lab University of Rome "La Sapienza"-*. The performance have been evaluated in terms of bit error rate (BER), reporting the measured values on a conveniently colored Google Earth map. About the colors, those range from green to yellow and red represent a low BER increasing till to values (red) as 0.5. It is possible to note that in Fig. 3.12, after few meters, the color shifts from green, that represents a low BER value, to red. The red region is the largest since it deals with the coverage offered by the modem, moreover the behavior is monotonic with respect to distance thus meaning that by increasing distance, BER grows. The adaptive detection mechanism has been instead realized by using at receiver side the AS-1 hydrophone and the Hydromic pre-amplifier and by processing the
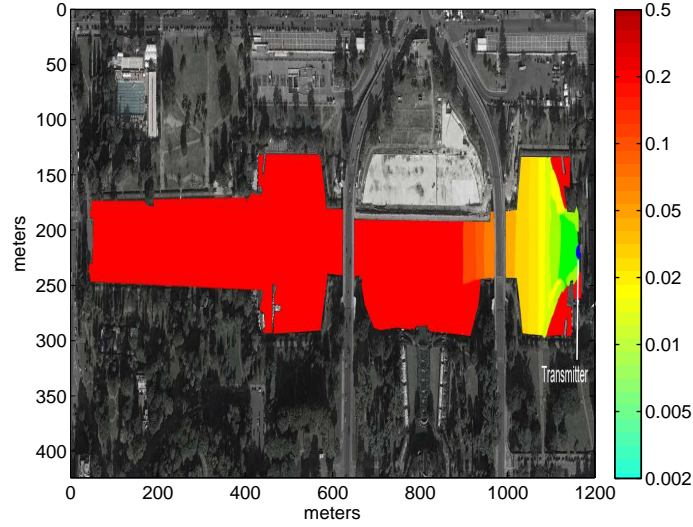
Figure 3.12: BER map for communication with direct detection as in eq. (3.41) with longitudinal transmission.

acquired recordings (in the .wav format) with Matlab software. By comparing the results (Fig. 3.12 and Fig. 3.13) it is possible to argue that the proposed approach is able not only to outperform the modem detection but, moreover, to extend coverage from 100 meters (at BER=0.05) till to 200m. This gain can be explained in two different ways, since two are the elements that allow such a gain when the receiver operates with a hardware/software approach. First, the threshold mechanism to *understand* where the pilot symbol is located allows to synchronize the transmission in a more reliable way. Second, the analysis carried out about interpulse interference allows to reduce BER. To be more precise, this means that widening the coverage, where this latter is defined as a function of the obtained BER, means not only to increase the received signal strength but, mainly, to increase the signal-to-interference-plus-noise ratio (SINR).
A second test has dealt with the positioning of the transmitter on the side of the reservoir (transverse transmission) with the modem and hydrophone located on a small pedalo. The transmitted text was -*Welcome to IEEE Oceans 2015 University of Rome "La Sapienza" Adaptive PPM Acoustic Detection in very Shallow Water Reservoir*-.

As for the previous experiment the performance have been detailed in terms of BER, as a map, both for modem hardware detection and software/hardware one. Fig. 3.14 reports the BER map of modem hardware (direct) detection. The performance rapidly decrease with the increasing
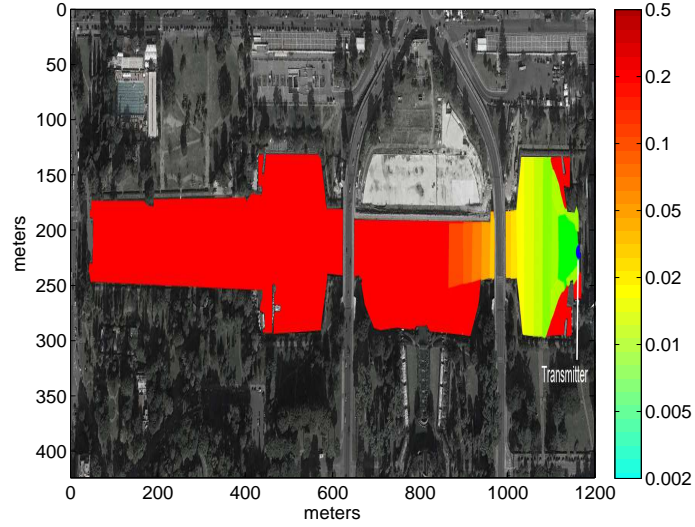
Figure 3.13: BER map for communication with detection as in eq. (3.45) with longitudinal transmission.

distance and at 30 meters the BER achieves 0.1. The behavior is monotonic and the BER rapidly achieves values that are not admissible. As an example, the modem detection recognizes only two ASCII characters out of 156. More interesting is the behavior of the hardware/software detection detailed before. In fact, in Fig. 3.15 it is possible to appreciate that the BER achieved is considerably lower since the color representing BER does not become red so giving evidence of the performance gain. Additionally, the behavior of the BER is not monotonic with the distance. In particular three different zones can be recognized. The first one, closer to the transmitter, is essentially yellow, the second one is green and the third one yellow/green. Since the lake-floor is 1.6 meters there (see [11]), the reflected signals amplitude is comparable with the main path signal and the delay of such a path is of the order of few milliseconds. Therefore, the BER loss is essentially due to interpulse interference. The second zone, the green one, presents very good BER levels since the effect of interpulse interference is mitigated even though the growing distance lowers the signal level with respect to the previous zone. Last, the third zone is yellow/green since the signal level decreases and the effect of attenuation is dominant with respect to the BER behavior.
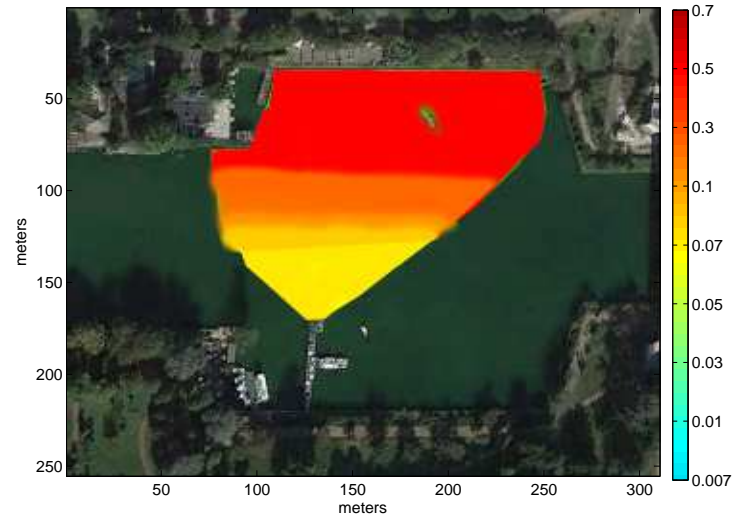
Figure 3.14: BER map for communication with detection as in eq. (3.41) with transverse transmission.
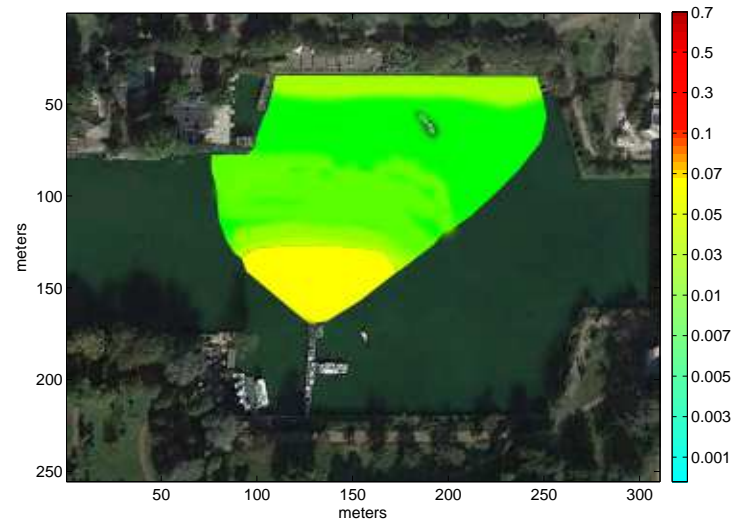

Figure 3.15: BER map for communication with detection as in eq. (3.45) with transverse transmission.

Chapter 4

Underwater acoustic channel

The physical characteristics of sound in association with the properties of
the medium make the acoustic underwater propagation very challenging.
Of course attenuation, reflection and other phenomena described in par. 3.1
are typical of terrestrial communications as well, however, since the speed of
sound is considerably lower than the speed of electromagnetic waves, their
influence on the signal spread results to be more significant. The underwater
acoustic channel features are described below in terms of path loss, noise and
multipath that are function of both frequency and distance. The water depth
is another fundamental element to be taken into account. The reference sce-
narios are usually divided in two categories, namely shallow and deep water
environment respectively. Depth and communication range in addition to
the sound propagation speed characterize the channel time-variability, de-
scribed through both statistical and empirical models. Acoustic propagation
is more reliable at low frequencies so despite the scarce available communica-
tion bandwidth, the underwater systems are anyway considered as wideband
since the fundamental frequency and the bandwidth are comparable.

4.1   Path loss

The path loss represents the attenuation experienced by a signal moving
from the transmitter to the receiver on a single propagation path, so that:

$$P_{rx} = \frac{P_{tx}}{A(d,f)} \tag{4.1}$$

being P_{rx} and P_{tx} the received and transmitted power respectively. As high-
lighted in eq. (4.1), in underwater acoustics the attenuation A is function
not only of the distance d traveled by the wave (as in radio communications)
but also of the frequency f. In detail, the path loss is given by:

$$A(d,f) = d^k a(f)^d \tag{4.2}$$

33

The term d^k models the spreading loss, with $1 \leq k \leq 2$ depending on the geometry of the propagation ($k=1$ for cylindrical spreading and $k=2$ for spherical spreading). The other term $a(f)$ is a coefficient referring instead to the absorption losses due to the transformation of acoustic energy into heat. Eq. (4.2) can be also written in decibels, leading to:

$$10 \log A(d, f) = k \log g + d \log a(f).$$ (4.3)

By doing so, it is possible to express the absorption coefficient by resorting to the following Thorp's empirical formula [17] :

$$10 \log a(f) = 0.11 \frac{f^2}{1 + f^2} + 44 \frac{f^2}{4100 + f^2} + 2.75 \cdot 10^{-4} f^2 + 0.003$$ (4.4)

where f is in kHz and $a(f)$ is given in dB/Km. The values obtained through eq. (4.4) are reliable for frequencies above hundreds of Hz. Fig. 4.1 reports the trend of $a(f)$ showing that asbsorption significantly increases with frequency [18].
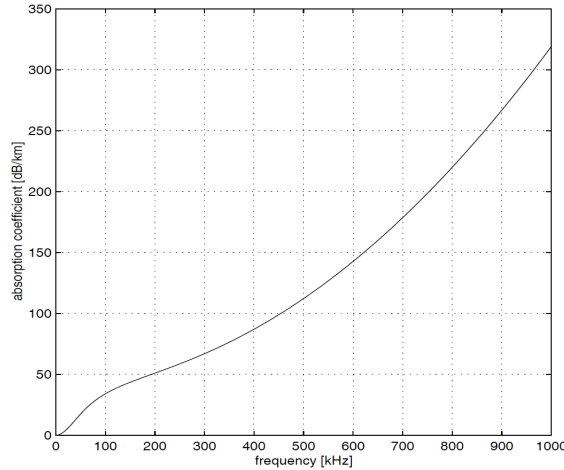

Figure 4.1: Absorption coefficient as a function of frequency.

4.2 Multipath

Underwater communication systems are mainly designed to let the transceivers work in line-of-sight (LOS). However, as the channel is a medium constrained by the water surface and floor, the reflections at the boundaries form multiple secondary paths where the signal radiated by the source travels through (Fig. 4.2). Therefore the receiver collects not only the signal coming from the direct path but also many other replicas, also known as *echoes*. The strength and delay of echoes depend on both the nature of the reflection
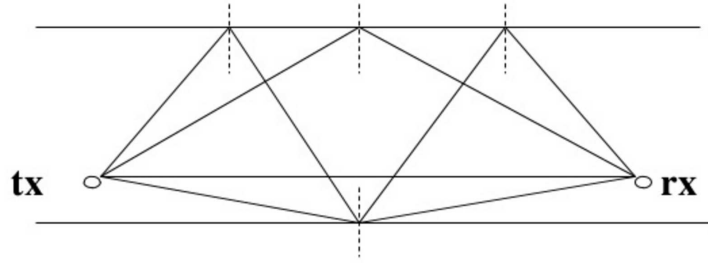
Figure 4.2: The signal reflection at the medium boundaries generates echoes.

(essentially described by the reflection coefficient) and the length of the path they follow to reach the receiver. For instance, in a shallow water scenario direct and secondary paths may be comparable in length so the signals could overlap at the receiver as they arrive almost at the same time. Moreover, the path loss experienced by the signal and its echoes may be similar as well, so the signals amplitudes could be comparable too. When instead dealing with a deep water communications, the length of secondary paths can be significantly greater than the LOS path, so direct signal and echoes arrive separated in time and with a more pronounced difference in strenght.

The multipath effect is also due to the physical characteristics of the channel. Temperature, pressure, salinity variations let water be a layered medium where the speed of sound becomes function of depth (par. 3.1). So the acoustic rays travel following the variation of the medium, giving rise also to scattering phenomena that strengthen the multipath.

4.3   Noise

Noise affecting underwater communications is generally attributed to different causes that lead to define the so called ambient noise. In a deep water scenario the components typically considered are shipping, turbolence, breaking waves and thermal noise. The ambient noise can be modeled through a Gaussian distribution, but it is not considered white. In fact, as shown in Fig. 4.3, at frequencies below 10 Hz the oceanic turbolence results to be dominant. The engine sound of moving vessels is instead identified in the range between 10 Hz and 100 Hz (the intensity of shipping activity is modeled basing on the factor D ranging from 0 to 1). Between 100 Hz and 100 kHz the surface motion due to wind and rain represents the most significant noise source, measured as a function of the wind speed w. Finally, the molecular agitation in the medium becomes non negligible only at frequencies beyond 100 kHz.

The noise characterizing the underwater environment is therefore de-

scribed by the sum of the component introduced above:

$$N(f) = N_{turbolence}(f) + N_{shipping}(f) + N_{wind}(f) + N_{thermal}(f) \qquad (4.5)$$

Each term composing $N(f)$ can be obtained through the following empirical expressions [19]:

$$10 \log N_{turbolence}(f) = 17 - 30 \log f \qquad (4.6a)$$

$$10 \log N_{shipping}(f) = 40 + 20(D - 0.5) + 26 \log f - 60 \log(f + 0.003) \quad (4.6b)$$

$$10 \log N_{wind}(f) = 50 + 7.5w^{1/2} + 20 \log f - 40 \log(f + 0.4) \qquad (4.6c)$$

$$10 \log N_{thermal}(f) = -15 + 20 \log f \qquad (4.6d)$$

Eqs. (4.6) report the power spectral densitity as a function of frequency of each noise source, measured in dB re μPa per Hz (Fig. 4.3).
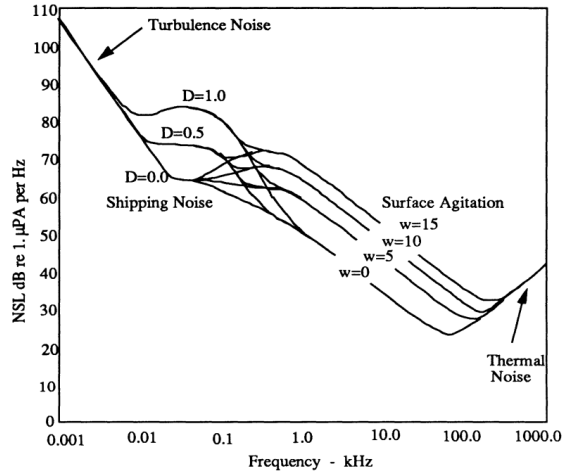

Figure 4.3: Noise spectrum level of deep water ambient noise.

In addition to the ambient noise typical of every scenario, another noise component related to the specific characteristics of the considered site may be present. Common examples are the sound caused by the glaciers cracking in the polar regions or the noise coming from some marine species living in a restricted area. Due to their particular nature, this kind of noise sources can not be generally described as gaussian.

Finally, Fig. 4.4 depicts a composite of possible intermittent and local noise sources characterizing the underwater environment [20].
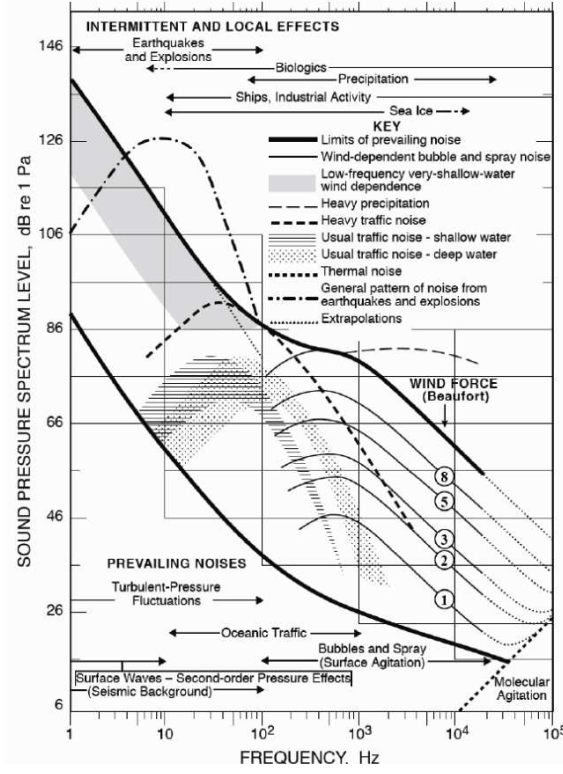
Figure 4.4: Ambient noise components spectra.

4.4 SNR and SINR at the receiver

Path loss and multipath define the model of the underwater acoustic channel, represented by the transfer function [21]:

$$H(d, f) = \sum_{p=0}^{P-1} H_p(d, f) = \sum_{p=0}^{P-1} \Gamma_p / \sqrt{A(d_p, f)} e^{-j2\pi f \tau_p} \qquad (4.7)$$

that takes into account the effect of P different propagation paths, with $p = 0$ being the direct one. Γ_p measures the cumulative reflection loss, while $\tau_p = d_p/c$ is the echo arrival delay. Eq. (4.7) presents each channel path as a low-pass filter for the traveling signal. The inverse Fourier transform of the P contributes returns the overall channel impulse response (CIR):

$$h(t) = \sum_{p=0}^{P-1} h_p(t - \tau_p) \qquad (4.8)$$

that essentially depicts the distribution of the multiple arrivals in time, as shown in Fig. 4.5 reporting an example of CIR estimated during a mea-

surements campaign at the EUR Lake, Rome (coordinates 41.827990° N, 12.467959° E).
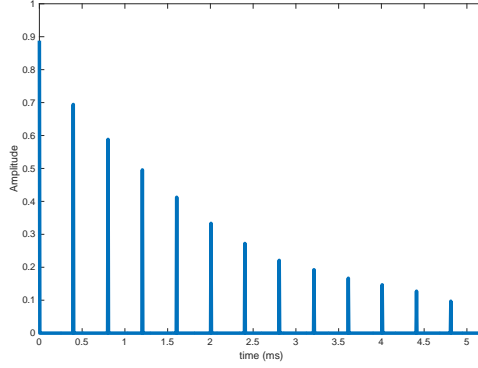

Figure 4.5: Example of channel impulse response measured over a 30 meters link. Water depth was about 3 meters.

As highlighted in eqs. (4.2) and (4.5), both attenuation and noise depend on the frequency. Moreover, path loss is also function of the distance as it refers to the attenuation experienced by a traveling signal. Hence this dependency is reflected in H as well.

The power of the received signal is given by:

$$P_{rx} = P_{tx}|H(d,f)|^2 \qquad (4.9)$$

Under the assumption of single path propagation the channel transfer function can be approximated as:

$$H(d,f) = H_0(d,f) = 1/A(d_0,f) \qquad (4.10)$$

returning the relation in eq. (4.1). In this context it is possible to define the signal-to-noise ratio (SNR) over the receiver noise bandwidth Δf centered in f_0 as:

$$SNR_{\Delta f}(d,f) = \frac{P_{tx}}{A(d,f)N(f)\Delta f} \qquad (4.11)$$

The term $1/AN$ evolves in the frequency domain as described in Fig. 4.6. The curves have been obtained considering a spreading factor $k = 1.5$ for the path loss model while the noise power spectral density has been evaluated setting $D = 0.5$ (shipping activity) and $w = 0$ (no wind). In particular the graph shows that, given the communication distance d, it is possible find an optimal frequency $f_0(d)$ leading to the maximum narrow-band SNR value (Fig. 4.7).

When considering the multipath propagation scenario the approximation made in eq. (4.10) is no more valid since the effect of the secondary paths
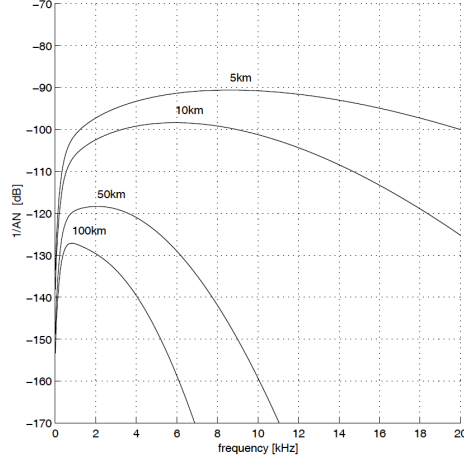
Figure 4.6: Trend of factor $1/A(d, f)N(f)$ for different distances.
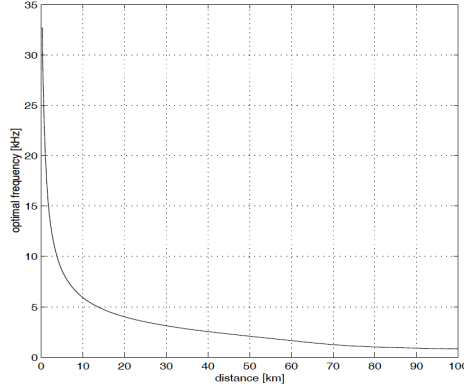
Figure 4.7: Optimal frequency $f_0(d)$ maximizing the narrow-band SNR.

need to be taken into account too. Therefore, the SNR is expressed through its general definition:

$$SNR(d, f) = \frac{P_{tx}|H(d, f)|^2}{N(f)} \qquad (4.12)$$

In the context of acoustic communications, the signal traveling along the direct path is the only one carrying information. The other echoes reaching the receiver are instead considered as disturbance since they cause interference when overlapping with the useful signal. Basing on this concept and on eq. (4.7), the received signal is given by the sum of the above described components:

$$P_{rx} = P_{signal} + P_{interference} = P_{tx}|H_0(d_0, f)|^2 + \sum_{p=1}^{P-1} P_{tx}|H_p(d_p, f)|^2 \quad (4.13)$$

This latest definition is exploited for introducing the signal-to-interference-plus-noise ratio (SINR):

$$SINR(d, f) = \frac{P_{tx}|H_0(d_0, f)|^2}{N(f)\sum_{p=1}^{P-1} P_{tx}|H_p(d_p, f)|^2} \tag{4.14}$$

that represents an additional measure especially used for evaluating the theoretical bounds of the channel capacity in wireless communication systems.

4.5   Channel time-variability

One of the main aspects of the underwater acoustic channel making its modeling very challenging is the time-variability. Natural events like waves, tides and currents induce physical changes in the medium continuously. For instance, the variation of the water surface roughness causes different reflection and scattering phenomena, therefore modifying the multipath propagation profile as well. Since characterized as instantaneus, these occurrencies are considered relevant in the context of communications. On the other hand, circumstances such as seasonal temperature fluctuations are usually neglected since they evolve over a timescale much larger with respect to the communication duration.

However, the channel behaviour remains hard to predict with accuracy due to the random nature of the above cited physical events. Both statistical and deterministic channel models have been proposed, whose effectiveness depends on the specific environment considered. When dealing with shallow water scenarios, signal echoes may be comparable in amplitude with the main path signal, so the Rayleigh fading model could be suitable. Considering instead a deep water environment and assuming the secondary paths to be significantly longer than the line-of-sight link, due to the higher path loss experienced by the echoes, the Rician fading may fit better.

Additionally, channel variability can be also induced by the in-water transmitter and/or receiver motion, resulting in a frequency shift known as Doppler effect (Fig. 4.8). The relevance of this phenomenon is related to the ratio between the communication nodes speed v and the signal propagation velocity in the medium c. It is worth noting that in underwater acoustics the Doppler effect is significantly different from that one experienced in wireless RF communications due to the five orders of magnitude difference between the speed of light and the speed of sound. The problem of frequency spreading is particularly important in multicarrier-based systems where the intercarrier interference can severely degrade the performance.

Changes in the medium and Doppler effect define the evolution of the underwater channel in time and frequency. These properties are summarized in the *channel coherence time*, representing the time window over which the
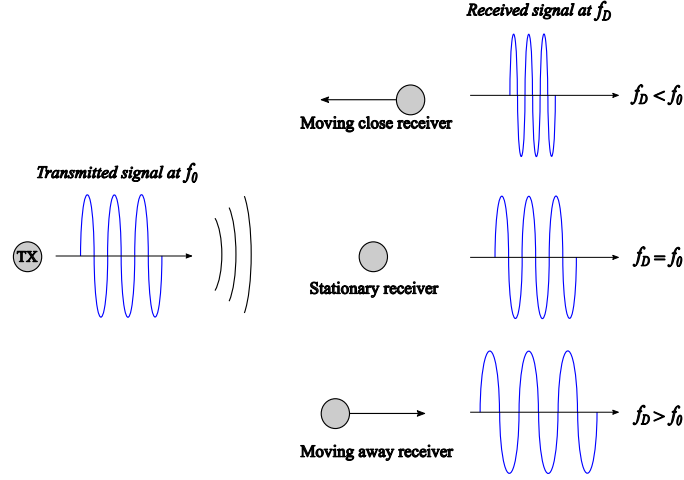
Figure 4.8: Doppler effect for different receiver motions.

channel impulse response is assumed to be not varying. By resorting to the Clarke's model [22], the channel coherence time is given by:

$$T_{coh} = \sqrt{\frac{9}{16\pi f_d^2}} \approx \frac{0.423}{f_d} = \frac{0.423}{a f_0} \tag{4.15}$$

where f_d and f_0 are the Doppler shift and the carrier frequency respectively, while a is the Doppler scaling factor specifying how longer or shorter (it depends on the direction of the receiver movement) the received signal is with respect to the transmitted one. Considering a communication link of length D, the channel can be considered quasi-static if its coherence time is larger than the signal round trip time, that is:

$$T_{coh} > RTT = \frac{2D}{c} \tag{4.16}$$

In other words, the condition in eq. (4.16) holds when:

$$D < D_{coh} = \frac{0.212c}{a f_0} \tag{4.17}$$

As an example, for a transmission at the carrier frequency $f_0 =$30 kHz with $c =$1500 m/s and $a = 1 \cdot 10^{-4}$ the channel can be assumed as not-varying if $D <$106 m.

4.6 Propagation models

The discussion in the previous paragraphs has shown that the physical characteristics of the medium (shallow and deep water) and its time-variability

make the formulation of a general channel model impossible. Furthermore, as the behavior of the channel depends on the frequency as well, different conditions have to be evaluated when dealing with low and high frequency applications. So, despite the notions about the theory of sound are sufficient to basically describe the traveling acoustic waves, the underwater propagation scenarios are usually very complex. Channel modeling can be empirically carried out basing on the information obtained through experiments, otherwise it is possible to build a propagation model by resorting to different techniques that the literature [23] categorizes as follows: Ray Theory, Normal Mode, Multipath Expansion, Fast Field (or Wavenumber Integration) and Parabolic Equation. All the above cited techniques derive from the wave equation, whose solutions are calculated with respect to the frequency domain. Fig. 4.9 reports the description of the domain of applicability of the cited propagation models as a function of both frequency and environment physical characteristics. Moreover, in the last years the devel-

Model type	Applications							
	Shallow water				Deep water			
	Low frequency		High frequency		Low frequency		High frequency	
	RI	RD	RI	RD	RI	RD	RI	RD
Ray theory	○	○	◑	●	◑	◑	●	●
Normal mode	●	◑	●	◑	●	◑	◑	○
Multipath expansion	○	○	◑	◑	◑	◑	●	◑
Fast field	●	◑	●	◑	●	◑	◑	◑
Parabolic equation	◑	●	○	○	◑	●	◑	◑

Low frequency ($<500\,$Hz)   RI: range-independent environment
High frequency ($>500\,$Hz)   RD: range-dependent environment

● Modeling approach is both applicable (physically) and practical (computationally)

◑ Limitations in accuracy or in speed of execution

○ Neither applicable or practical

Figure 4.9: Domain of applicability of underwater acoustic propagation models [24].

opment of software technology has allowed significant advances in channel modeling. In fact, the combination of the basic techinques and the analysis of the increasing amount of data has led to the formulation of dozens of other models [25].

4.7   Channel reciprocity

In underwater acoustic networks with bi-directional communication links, channel state information (CSI) plays an important role since having a good

channel knowledge allows system performance enhancements. Two communicating nodes can use the information obtained by their respective channel estimation in order to adjust the transmission parameters to the medium conditions [26]. But due to channel time-variability, it could happen that forward channel is rather different from the backward one, so the previously performed estimation could be incorrect leading to a communication performance reduction. In RF communication systems channel reciprocity is usually guaranteed by supposing that forward and back transmissions take place within a time window shorter than the channel coherence time [27]. Differences between transmitter and receiver side cause carrier frequency offset, timing offset and sampling clock deviations that may lead to non-reciprocity, but these factors can be compensated by appropriate transceivers calibration. In the field of free-space optics, by several experiments [28] a near-unity correlation between the signal fluctuations measured at both ends of a point-to-point link has been observed. This result, validating channel reciprocity, has allowed to exploit the channel state information to improve the communication performance.

Dealing instead with underwater acoustic communications, the circumstance of channel reciprocity is barely realistic. In fact, since the speed of sound is about five orders of magnitude lower than the speed of radio waves, it may happen that channel coherence time is shorter than signal round trip time, so forward and backward channel could be rather different. In literature, most of works does not address this issue properly, actually very often channel reciprocity is assumed without verifying its acceptability. This could be not convenient for instance when time-reversal method (that is based on channel reciprocity assumption [29]) is employed for improving the performance of a communication link [30]. Although in deep water conditions, where water turbolence and the other natural effects are attenuated, it is licit assuming the channel quasi-static and reciprocal, when dealing with shallow water environments this hypothesis becomes totally inacceptable. The study reported in [31] shows how the ocean environments influence the signal propagation, highlighting that rough sea conditions cause remarkable amplitude and phase fluctuations of the signal. In this regard, the investigation of the multipath effect on the signal propagation returns many information about the channel behavior. In particular, by analyzing two opposite communication directions it is possibile to infer the effectiveness of channel reciprocity [32].

4.7.1 Multipath estimation

An UWAN composed of N nodes where each communication link is bidirectional is given (Fig. 4.10). Without loss of generality, refer to a single couple of nodes N_i-N_j that represents a simpler point-to-point communication. Signals transmitted from N_i to N_j and vice versa are either synchro-

nization pilots or information carriers. In this case, consider only the first ones since they are usually exploited for channel estimation too. Generally
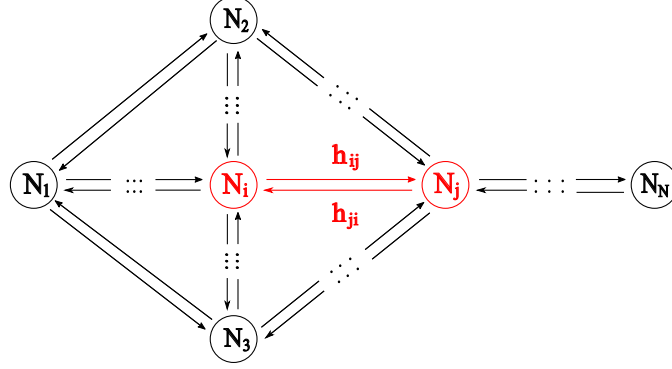

Figure 4.10: N-nodes UWAN with bi-directional links.

speaking, the signals $r_i(t)$ and $r_j(t)$ received by each node are given by:

$$\begin{cases} r_i(t) = s_j(t) * h_{ji}(t) + w_i(t) \\ r_j(t) = s_i(t) * h_{ij}(t) + w_j(t) \end{cases} \tag{4.18}$$

where s_j and s_i are the signals emitted respectively by N_j and N_i, h_{ji} and h_{ij} model the impulse response of the channel for the link from N_j to N_i and from N_i to N_j respectively, and finally w_i and w_j are the noise components superimposing on the useful signal passing through the channel. Since both ends of the link employ the same pilots (whose shape is known) to communicate, $s_j = s_i = s$ is assumed. The noise measured by each node can be expressed as:

$$\begin{cases} w_i(t) = w_T(t) + w_{EXT}^{(i)}(t) \\ w_j(t) = w_T(t) + w_{EXT}^{(j)}(t) \end{cases} \tag{4.19}$$

being w_T the thermal noise component and w_{EXT} that one due to external acoustic sources. The term w_T is statistically independent from the communication direction. On the other hand, w_{EXT} depends on the nature of the source it is generated by, so it is not possible to infer a space-time model of its behavior. For this reason, $w_{EXT}^{(i)}$ and $w_{EXT}^{(j)}$ will generally differ, and consequently also w_i and w_j. Anyway, by assuming the external noise components negligible, the statement $w_i = w_j = w$ can be accepted in order to simplify eq. (4.18), that becomes:

$$\begin{cases} r_i(t) = s(t) * h_{ji}(t) + w(t) \\ r_j(t) = s(t) * h_{ij}(t) + w(t) \end{cases} \tag{4.20}$$

The acoustic channel reciprocity entails that:

$$h_{ji}(t) = h_{ij}(t) \tag{4.21}$$

If property in eq. (4.21) is verified, the shape of the signals received by N_i and N_j in eq. (4.20) should sufficiently look alike, except for an additional component due to external noise fluctuation.

The signal $s(t)$ propagating from a node to the other experiences reflection and scattering at the water surface and floor. The multipath effect can be highlighted by expressing the received signals in eq. (4.20) as the sum of multiple echoes:

$$\begin{cases} r_i(t) = \sum\limits_{m=0}^{M} s(t) * h_{ji}^{(m)}(t) + w(t) = \sum\limits_{m=0}^{M} A_m s(t - T_m) + w(t) \\ \\ r_j(t) = \sum\limits_{k=0}^{K} s(t) * h_{ij}^{(k)}(t) + w(t) = \sum\limits_{k=0}^{K} A_k s(t - T_k) + w(t) \end{cases} \tag{4.22}$$

Superscripts m and k indicate the effect of acoustic propagation on each m-th and k-th signal replica respectively ($m, n = 0$ refers to the main path, the other values to the secondary ones), A_m and A_k represent their amplitude, while T_m and T_k are their arrival delay. Having the echoes amplitude and time distribution comparable in both the communication directions essentially returns that $h_{ji}^{(m)}(t) \approx h_{ij}^{(k)}(t)$ for $m = k$, then the property in eq. (4.21) would be respected. So the problem of channel estimation turns into an arrivals measuring issue.

Since the shape of $s(t)$ is known, the cross-correlation between the signal sent and the received gives:

$$y_{r_c s}(\tau) = \int\limits_{-\infty}^{+\infty} r_c^*(t) s(t + \tau) dt \tag{4.23}$$

with $c = i, j$ depending on the receiving node considered. Then calculating the modulus of $y_{r_c s}$ normalized at the maximum of the auto-correlation of the signal $s(t)$ leads to the final expression:

$$\chi_{r_c s}(\tau) = \frac{\|y_{r_c s}(\tau)\|}{\max\limits_{\tau} y_{ss}(\tau)} \tag{4.24}$$

The function $\chi_{r_c s}(\tau)$ presents several peaks that identify the different signal replicas, each one reaching the receiver with its delay and amplitude as described in eq (4.22) (if the maximum peak is assumed to refer to the main path, the estimation of its arrival delay gives the synchronization for the received information frame). By comparing $\chi_{r_i s}(\tau)$ and $\chi_{r_j s}(\tau)$ it is possible to infer the behaviour of $h_{ji}(t)$ and $h_{ij}(t)$ and hence to understand if the condition of eq. (4.21) can be considered valid.

4.7.2   Tests results

The theoretical concepts previously introduced have been applied for an-
alyzing data coming from a real application. The tests performed have
concerned a bi-directional point-to-point link in a shallow water reservoir in
Rome [11]. Fig. 4.11 reports the positions of the nodes, one placed in P while
the other positioned in three different sites A, B and C, realizing a 30, 70
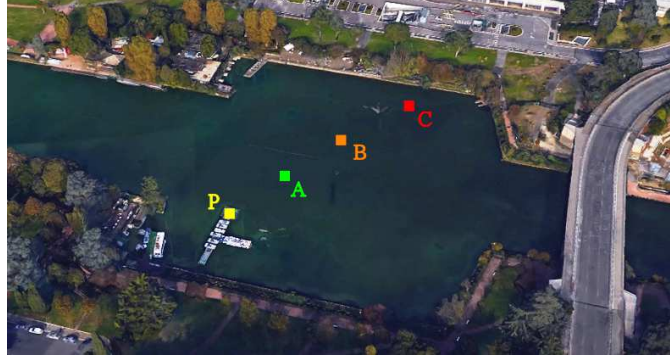and 110 meters communication link respectively. The experiments consisted

Figure 4.11: Map reporting the nodes position P, A, B, C.

in the collection of delay and amplitude of the signal transmitted in both the
communication directions, in order to how the aquired data look like. The
hardware of each node includes a AS-1 hydrophone for receiving and SAM-
1 acoustic modem for transmitting. Since interested in multipath/channel
estimation and not in link performance evaluation, the analysis has been car-
ried out considering only the synchronization ping (whose shape is known)
of each SAM-1 frame (par. 3.3.2).

The first test has involved the P-A node link and consisted in the trans-
mission of a single SAM-1 frame from P to A and later from A to P. By
looking at the signal received by each respective node (Fig. 4.12) it is pos-
sible to appreciate how the channel differently twists the transmitted pulse
(whose noise-free shape is plotted in Fig. A.4 of App. A) depending on the
propagation direction.

In addition to this graphic analysis, the application of eqs. (4.23)-(4.24)
returns quantitative information about channel features. Cross-correlations
$\chi_{r_{As}}(\tau)$ and $\chi_{r_{Ps}}(\tau)$, operated between demodulated transmitted and re-
ceived signal are shown in Fig. 4.13. The local maxima in the cross-
correlation functions represent the various signal replicas reaching the re-
ceiver. The amplitude and delay values exceeding a pre-set threshold (minor
signal components are neglected) have been collected in form of histogram
as drawn in Figs. 4.14 - 4.15. By analyzing the trend of delays in both
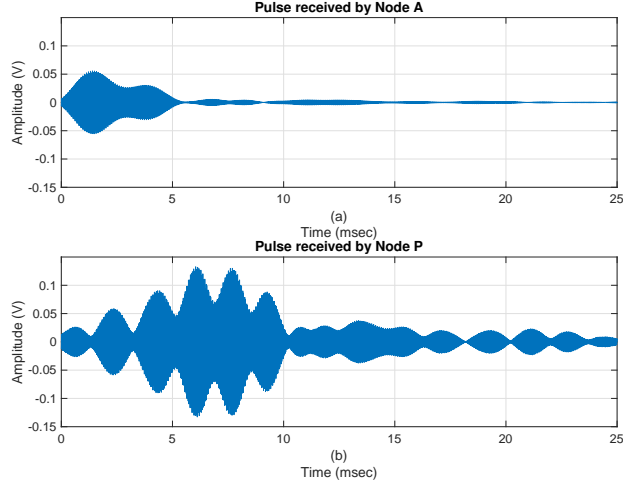cases, note that in the P-A node link most of signal replicas reaches the re-

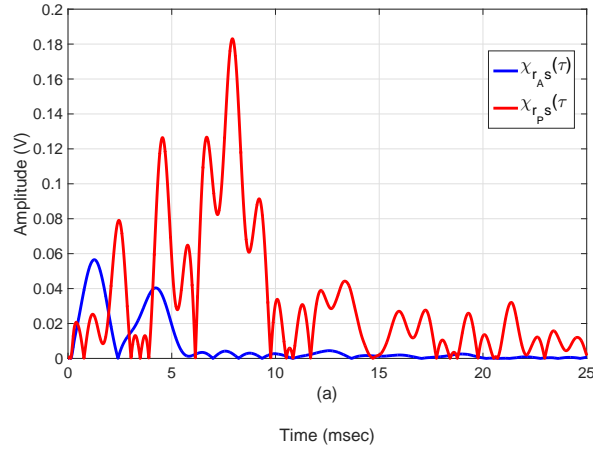Figure 4.12: Pulses respectively received by node A (a) and node P (b).

Figure 4.13: Signals cross-correlation for P→A (a) and A→P (b) link.

ceiver within 5 milliseconds, while the arrivals in the opposite direction path experience delays uniformly distributed up to 25 milliseconds. Concerning instead amplitudes, in the first case nearly all the peaks are less than 60 mV, while in the latter one they range from 10 to 100 mV, with some maxima up to 170 mV. Summarizing, results show that the signal propagating from A to P is affected by a wider delay spread, that one passing from P to A suffers instead from a higher attenuation. Furthermore, this conclusion is confirmed by the fact that in the P to A link the number of identified
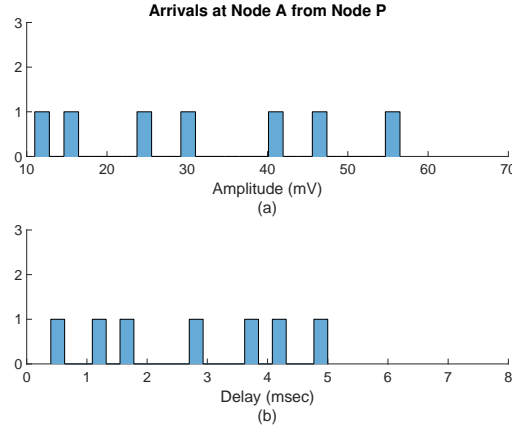
multipaths is 7, while in A to P link it is 28.

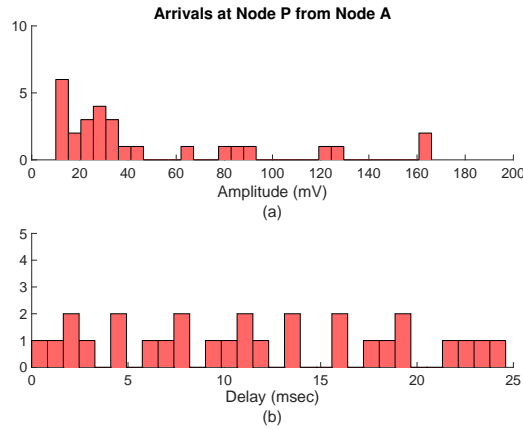Figure 4.14: Histogram of arrivals at Node A from Node P.

Figure 4.15: Histogram of arrivals at Node P from Node A.

The same test has been conducted considering the P-C node link so as to study the effect of communication distance growth on the channel reciprocity. Figs. 4.16-4.17 report the histogram of the resulting measurements. The first aspect to highlight is that the number of signal replicas detected in both propagation directions is 32, that is more than the P-A node link case. This is explained by the fact that, increasing the communication range and considering the low water depth (see [11]), the number of reflections on the surface and on the bottom grows. In this way, the signal experiences a sort of waveguide propagation that leads to a more remarkable multipath effect. Looking at Fig. 4.16, most of signal replicas amplitudes are within 50

mV and their time of arrival is uniformly distributed within 25 milliseconds. Regarding instead the values distributed in Fig. 4.17, many amplitudes are in the range from 50 to 100 mV with a time distribution very similar to the previous case. This means that the channel equally behaves in both direction concerning delay spread but introduces different attenuations.
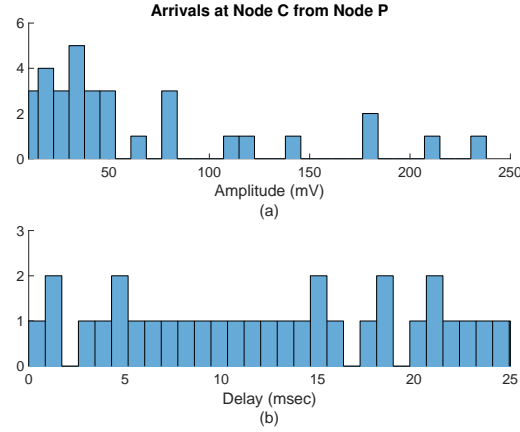

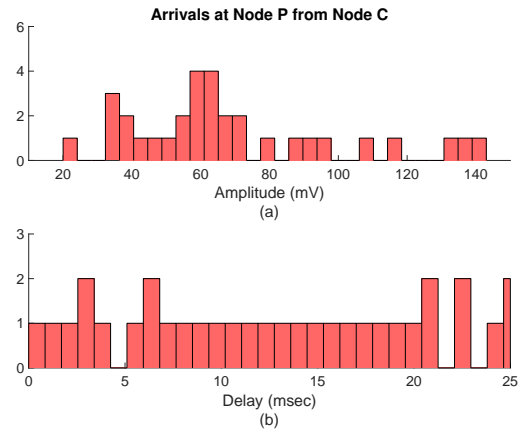Figure 4.16: Histogram of arrivals at Node C from Node P.

Figure 4.17: Histogram of arrivals at Node P from Node C.

The second test has dealt with the analysis of the channel behavior when the communication duration increses. To do so, the three P-A, P-B, P-C node links have been considered, and for each bi-directional communication 30 SAM-1 frames were sent alternating a single frame transmission along a direction and one in the opposite one, and so on. The amount of datea

collected from this experiment is huge, so it has been convenient to show the results in a different fashion. In Table 4.1 the total amount and the average number of signal echoes detected per transmitted pulse is reported. Skimming the values by columns, note that both the reported quantities in-

Propagation Direction	Multipaths Total Amount	Multipaths Average per Pulse
P → A	227	15.13
A → P	379	25.26
P → B	417	27.80
B → P	428	28.53
P → C	380	25.33
C → P	480	32

Table 4.1: Total amount and average of multipaths for different communication links.

crease, demonstrating once again that in a very shallow water environment widening the communication range causes harder multipath phenomenon. More, looking at the cases of P-A and P-C node link, the results obtained from the measurements in the two opposite propagation directions clearly differ, meaning that channel behaviour is not the same. In Table 4.2 and Table 4.3 the collection of replicas amplitudes and delays are respectively listed. Data are gathered depending on different amplitude and delay range, highlighting the percentage of belonging to each set. Even by analyzing this two tables it results that both amplitudes and delays distribution varies from case to case. Furthermore, each one of the three links investigated presents different results for the two opposite propagation directions. This is more evident in links P-A and P-C, while in case P-B the values distributions sufficiently equate.

Propagation Direction	10-60 mV	60-120 mV	120-180 mV	≥180 mV
P → A	67.4%	15.86%	8.81%	7.93%
A → P	58.05%	20.5%8	11.08%	10.29%
P → B	65.95%	16.31%	7.19%	10.55%
B → P	53.5%	19.39%	10.28%	16.82%
P → C	70.26%	15.79%	7.63%	6.32%
C → P	42.92%	51.67%	5.42%	0%

Table 4.2: Signal amplitudes percentage for different ranges.

Propagation Direction	≤7 msec	7-14 msec	14-21 msec	≥21 msec
P → A	51.98%	31.28%	14.1%	2.64%
A → P	34.56%	31.4%8	22.96%	11.08%
P → B	32.37%	27.82%	25.9%	13.91%
B → P	31.31%	30.37%	23.13%	15.19%
P → C	34.21%	31.05%	23.68%	11.05%
C → P	28.13%	25.83%	27.29%	18.75%

Table 4.3: Signal delays percentage for different ranges.

Summarizing, the results coming from the presented analysis suggest that acquiring the channel on one direction does not suffices to infer about the behavior of the channel on the other direction. This difference can be generally justified by the mismatch between channel coherence time and signal round trip time (even if the underwater and radio-frequency channel coherence time were comparable, acoustic waves propagation remains much slower than those EM, so it will always be hard to reach the reciprocity level experienced in a RF scenario), but it is also fundamental to take into account the effect of natural elements. In fact, the reflection angle and scattering (both of sea bed and water surface) depend on the relative position (and direction) of transmitter and receiver. Furthermore, the presence of underwater obstacles or rocks on the seabed as well as water surface profile changes due to wind may cause additional propagation phenomena that could be *different* in the two link directions. Channel reciprocity is far from being verified in underwater communications, especially in shallow water scenarios. Moving to a deeper water environment the effect of natural events at the whater surface would be attenuated, so the features of forward and backward channel may be better matching.

4.8 ISI and channel equalization

The signal distortion induced by the channel and resulting in a large delay spread may lead to have intersymbol interference (ISI). This latter represents the main impairment to face with in wireless communication, and this is especially true in the regard of underwater acoustics where the channel behavior severely degrades the quality of the received signal.

The easiest and cheapest solution to mitigate ISI is introducing a guard time between the transmission of consecutive symbols. The attempt, by doing so, is to avoid that the current symbol delay spread do not fall into the next symbol transmisson time window, hence limiting the signals overlapping. The obvious drawback of this strategy is the rate reduction due to the presence of large *recovery* times making the communication essentially in stand-by.

The alternative solution to the use of guard time is represented by the received signal digital filtering aimed at obtaining the *channel equalization* condition, allowing the symbol detection error probability to be kept below the desired level. However, the introduction of an efficient equalization scheme may significantly increase the receiver complexity, reflecting also on the cost of hardware. This is an important aspect to take care of since devices employed in underwater applications are quite expensive. In this context, as discussed in the following section, it is important to find a solution proving good performance but requesting a reasonable computational cost.

4.8.1 Intersymbol interference

ASK-PAM transmission of a discrete sequence of information $a[k]$ over a real channel is allowed by modulation that essentially consists of identifying each element of $a[k]$ with a real signal. This procedure is realized thanks to the use of a shaping filter $g(t)$ (typically a raised-cosine filter) where the discrete sequence $a[k]$ passes through, returning the continuous signal:

$$x(t) = \sum_{k=-\infty}^{\infty} a[k]g(t - kT_s) \tag{4.25}$$

where $F_s = \frac{1}{T_s}$ is the symbol transmission rate. The shaping operated by $g(t)$ must guarantee no mutual overlapping between consecutive symbols, so that to avoid the so-called intersymbol interference. By sampling $x(t)$ at symbol time T_s the following expression is given:

$$x(nT) = \sum_{k=-\infty}^{\infty} a[k]g(nT_s - kT_s) = a[n]g(0) + \sum_{k=-\infty,k\neq n}^{\infty} a[k]g(nT - kT_s) \tag{4.26}$$

where it is possible to highlight two components, the first referring to the current symbol, the latter one considering the symbols transmitted previously and after. In particular this last term may cause errors at receiver side due to interference with the current symbol. However, if the shaping filter respects the Nyquist criterion, that is:

$$g(kT_{s)} = \begin{cases} 1 & for\ k = 0 \\ 0 & for\ k \neq 0 \end{cases} \tag{4.27}$$

the resulting transmitted signal is:

$$x(nT_s) = a[n]g(0) = a[n] \tag{4.28}$$

By doing so, the received demodulated sequence will correspond to the transmitted one. Moving to the frequency domain, the Nyquist criterion states that:

$$\frac{1}{T_s} \sum_{k=-\infty}^{\infty} G(f - \frac{l}{T_s}) = 1 \tag{4.29}$$

meaning that the replicas of $G(f)$ shifted by $\frac{1}{T_s}$ result in a constant evolution. It is worth underlining that the condition of ISI absence can not be relized if the bandwidth of the spectrum of the shaping filter $G(f)$ is less than the Nyquist bandwidth $f_N = \frac{1}{2T_s}$.

The signal $x(t)$ emitted by the transmitter passes through the communication channel characterized by the impulse response $h(t)$ (Fig. 4.18). The signal $y(t)$ at the output of the channel is expressed by the following
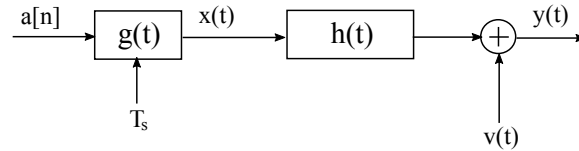

Figure 4.18: Transmission chain including shaping filter at transmitter side and channel.

convolution:

$$y(t) = x(t) * h(t) = \sum_{k=-\infty}^{\infty} a[k] * g(t - kT_s) \tag{4.30}$$

where $c(t) = (g * h)(t)$ collects the cascade of shaping filter and channel impulse response. If the channel has an impulsive nature, the Nyquist features of $g(t)$ remain unaltered, so the signal $y(t)$ can be still considered as ISI-free. On the other hand, in the presence of a frequency-selective channel, that is for instace the underwater acoustic channel , the spectrum of

$C(f) = G(f)H(f)$ does not respect the eq. (4.28) anymore, giving rise to ISI. So, the final expression of the signal at the output of the channel is:

$$y(nT_s) = \sum_{k=-\infty}^{\infty} a[k]c(nT_s - kT_s) = a[n]c(0) + \sum_{k=-\infty, k \neq n}^{\infty} a[k]c(nT_s - kT_s).$$
$$(4.31)$$

It is important to note that, in addition to ISI, the signal passing through the channel can be affected by noise as well. This impairment, described as a white random Gaussian process, zero mean, stationary and ergodic, is summed to the transmitted signal $x(t)$. These features lead to define the so called AWGN (Additive White Gaussian Noise) channel.

4.8.2 Underwater channel equalization

Channel equalization results to be necessary especially when dealing with shallow water scenarios where the multipath effect makes the received signal subject to strong interference. In this regard, suitable techniques are requested in order to allow ISI cancellation. A remarkable part of the works in literature presents solutions based on the implementation of Decision Feedback Equalizers (DFE). This kind of filtering, belonging to the category of non linear equalizers [33], is used when the signal distortion caused by the channel can not reliably mitigated by linear equalizers. In particular, the DFE is characterized by a couple of blocks, the feedforward filter (FFF) and the feedback (FBF) one. The use of this scheme allows not only the cancellation of ISI on the current received symbol, but also the estimation of ISI caused by the current symbol on the future ones, so that it can be substracted out prior to next symbol detection. An example of DFE is reported in [34] where an iterative frequency domain equalization combined with low density parity check (LDPC) decoding is proposed. The fast convergence provided by DFEs is paid in terms of computational cost since the coefficients adaptation concerns two filters, the FFF and the FBF respectively, instead of a single filter as in the case of liner equalization. Furthermore, complexity depends also on the adaptation rule used for filters updating. For instance, the Least Mean Square (LMS) algorithm, the simplest to realize, requests a computational cost proportional to $2N$, while in the Recursive Least Square (RLS) algorithm, faster than LMS, the complexity is proportional to $20N$, being N the total equalizer length considering FFF and FBF.

An example of DFE applied to PPM-based communications, even though not specifically for underwater applications, is reported in [35]. This work considers a chip-oriented transmission, where M chips compose the generic M-PPM symbols. As the process is described as cyclostationary, with $M{=}2$ the chips correlation statistics at even and odd sampling times are different, so the authors propose the implementation of a bank of M DFEs. However, despite each DFE works only once every M chips, the receiver architecture

remains quite complex.

It is worth noting that most of works reported in the literature deal with filtering operated at *chip time*. This choice may lead to time sensitivity reduction that, in addition to the undesired noise amplification effect (as for instance the application of Zero-Forcing could produce), impacts on the system performance. A possible way to overcome this problem is sampling the received signal at rates greater than the nominal one (typically twice the transmitted chip rate) before filtering. This technique is known as fractionally spaced channel equalization [36], and the resulting filter is so-called fractionally spaced equalizer (FSE).

Following this direction, the architecture of a fractionally spaced equalizer adaptively implemented according to a simple LMS rule [37] is presented below. Furthermore, it is shown that the spectral features of PPM modulation result to be particularly suitable for this kind of scheme. The complexity of equalization depends essentially on the number of filter coefficients to be periodically updated, so it is also discussed how the introduction of a time guard between two consecutive chips allows the computational cost reduction.

4.8.3 Spectral features of PPM

In Pulse Position Modulation, the information is carried by the position of the pulse within the symbol time window T_s. Given a modulating signal $m(t)$, the PPM modulated signal is expressed as:

$$x_{PPM}(t) = \sum_k p(t - kT_s - m(kT_s)) = p(t) \sum_k \delta(t - kT_s - m(kT_s)) \quad (4.32)$$

Considering instead digital communications, the signal generated by the PPM modulation of a binary sequence is given by:

$$x_{PPM}(t) = \sum_k p(t - kT_s - \epsilon b_k) = p(t) \sum_k \delta(t - kT_s - \epsilon b_k) \quad (4.33)$$

where ϵ refers to the pulse shift depending on the value of b_k=0,1. In order to avoid the pulses overlapping in the PPM signal, it is sufficient that the following constraint is respected:

$$p(t) = 0 \quad for \quad |t| \geq \frac{T_s}{2} - |m(t)|_{max} \Rightarrow |m(t)|_{max} \leq \frac{T_s}{2} \quad (4.34)$$

As a non-linear transformation is operated on the modulating signal, obtaining the power spectral density of a PPM signal results to be very complex (see [38]). However, under the realistic assumption statistically independent symbols, it is possible to rephrase a particular method for the analytical evaluation of the power spectral density of PAM signals in the context of

pulse position modulation. Basing on [39] the power spectral density of a discrete M-PPM signal (Fig. 4.19) is given by:

$$P_s^{M-PPM}(e^{j\omega}) = \frac{1}{M}\left[1 - \left|\frac{\sin(\omega M/2)}{M\sin(\omega/2)}\right|^2\right] + \frac{2\pi\,\delta(\omega \mod 2\pi)}{M^2} \quad (4.35)$$

Taking into account the zeros of $\dfrac{\sin(\omega M/2)}{M\sin(\omega/2)}$, all the Dirac delta functions $\delta(\omega L \mod 2\pi)$ vanish except for those ones placed in $\omega \mod 2\pi = 0$.
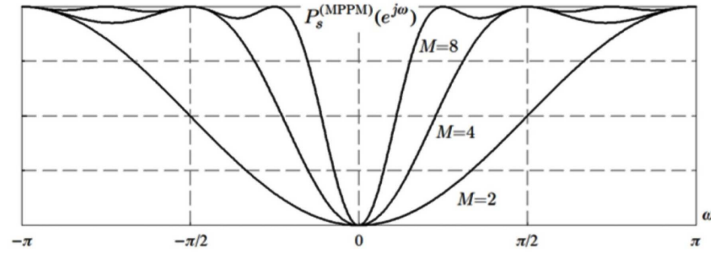

Figure 4.19: Power spectral density of a PPM modulated signal.

The continuous M-PPM signal is obtained through interpolation of the digital one:

$$x^{M-PPM}(t) = \sum_{k=-\infty}^{\infty} x^{M-PPM}[k]g_T(t - kT_c) \quad (4.36)$$

wher T_c is the chip duration. The gaussian shape of the pulse (possibly) placed in a chip is described in Fig. 4.20. Note that in this case the length of the pulse T_p is equal to T_c. Moving to the frequency domain, the power
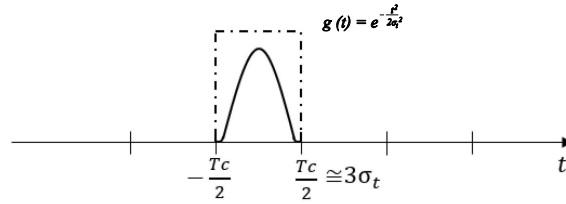

Figure 4.20: Time description of the pulse placed in a chip time window.

spectral density of the signal at the output of the shaping filter is depicted in Fig. 4.21, where it is possible to observe how the spectral structure of the digital PPM is repeated every T_c. If the pulse length is shorter than the chip time window, that is $T_p \leq T_c$, the bandwidth of the signal increases. In fact,
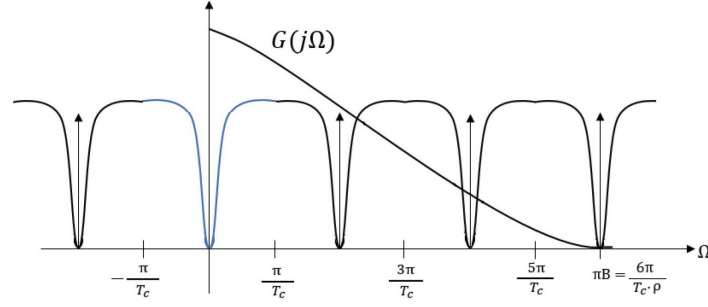
Figure 4.21: Power spectral density of a PPM modulated signal at the output of the shaping filter.

introducing the parameter $\rho = \frac{T_p}{T_c}$ describing the relationship between pulse and chip time, the signal bandwidth is defined as $B = \dfrac{6}{T_c\rho}$ (if T_p decreases, ρ decreases as well and therefore B grows).

Now, consider temporarily the spectrum of a QAM signal obtained using a raised-cosine filter and characterized by a certain roll off factor ρ, as shown in Fig. 4.22. The spectrum structure around $\frac{\pi}{T_s}$ is specular, and this property can be exploited not only for synchronization recovery, but also for equalization since in bandwidth $\frac{2\rho\pi}{T_s}$ the effect of the channel has to be compensated *but not* inverted.
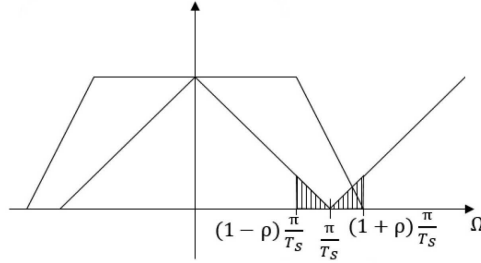

Figure 4.22: Power spectral density of a QAM modulated signal at the output of raised-cosine filter.

Referring back to the PPM case, the spectrum in Fig. 4.21 presents *multiple* bands of width $\frac{2\pi}{T_c}$ and centered in $\frac{2k\pi}{T_c}$ where the channel has to be only compensated. This characteristic demonstrates that PPM is particularly suitable for equalization, therefore reflecting on a remarkable robustness to errors. Furtheremore, the increase of B allows a higher number of spectrum replicas to be contained within the bandwidth, thus reinforcing this property. From a certain point of view, the periodicity shown by the PPM spectrum acts like a sort of frequency repetition coding so that improved robustness to errors is expected.

4.8.4   Fractionally Spaced Equalization model

The proposed fractionally spaced equalization model follows the steps summarized in Fig. 4.23. The sequence $a[n]$ is expanded by Q_{Tx} before entering the shaping filter g_T, so to obtain:

$$a_Q[n] = \begin{cases} a[n/Q] & for \quad n \bmod Q = 0 \\ 0 & for \quad n \bmod Q \neq 0 \end{cases} \tag{4.37}$$

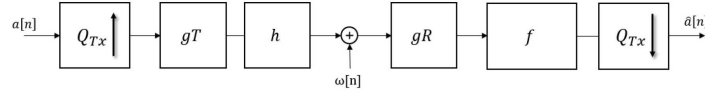In the context of a PPM-based communication, the transmission is con-

Figure 4.23: Fractionally spaced equalization model.

sidered as chip-oriented, that is each element of the sequence $a[n]$ is a chip composing a PPM symbol. Once defined in a discrete time shape $g_T[n] = g_T(nT_s/Q_{Tx})$, the whole filtering chain composed by gT, the channel h and the matching filter g_R at receiver side can be expressed as a unique filter $g[n] = (g_T * h * g_R)[n] = g(nT_s/Q_{Tx})$. Moreover the signal is affected by a noise component expressed by the sequence $w[n] = w(nT_s/Q_{Tx})$, assumed to be statistically independent of the transmitted sequence $a[n]$.

Channel equalization for ISI cancellation is obtained when the filter $f[n]$ in Fig. 4.23 realizes the condition $\hat{a}[n] = a[n]$ when the noise is absent. This result is achieved by satisfying the Nyquist criterion (eq. (4.29)) that, considering the scenario of fractionally spaced equalization [39], becomes:

$$\frac{1}{Q_{Tx}} \sum_{q=0}^{Q_{Tx}-1} F\left(e^{\frac{j\omega - j2\pi q}{Q_{Tx}}}\right) G\left(e^{\frac{j\omega - j2\pi q}{Q_{Tx}}}\right) = 1 \tag{4.38}$$

However it is worth observing that, as the frequency response $G(e^{j\omega})$ may have possible zeros due to some nulls in the channel frequency response $H(j\Omega)$, it will not be possible to completely cancel ISI around those values of ω where $\sum_{q=0}^{Q_{Tx}-1} \left| G\left(e^{\frac{j\omega - j2\pi q}{Q_{Tx}}}\right) \right| = 0$

Finally, a possible drawback of ISI cancellation according to the Nyquist criterion is the noise amplification around those ω where the SNR is high. For this reason, the equalization filter $f[n]$ must be realized following an optimality criterion, such as the Minimum Means Square Error (MMSE) employed in the presented scheme.

4.8.5   Equalizer implementation

The scheme of the FSE for PPM transmission has been implemented via Simulink software as described in Fig. 4.24. The received signal $r[n]$ passes

through a first block providing the sequence buffering, in order to reach chip synchronization, and the automatic gain control. The synchronized sequence $x[n]$ is used both as input of equalization filter f and for filter coefficients update. The block performing this latter operation uses as input $x[n]$, the *old* filter coefficients $f[n]$ and the error between the received sample and the filtered one. Then, following the LMS adaptation rule [40], it generates the updated filter coefficients that, once delayed, are sent to the equalizer f. The whole procedure is ruled by a clock letting the filter updating be done every Q_{Tx} samples. As previously mentioned, the error estimation is carried out according to the MMSE criterion and performed in the block named "Error build" according to the Bussgang approach [41].
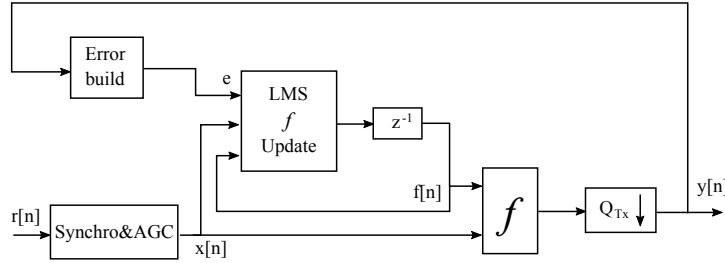
Figure 4.24: Fractionally spaced equalizer scheme.

If the received signal is synchronized at symbol time (recall that each symbol is composed by M chips) the estimator will be named as *synchronous*. In the considered scenario, due to the structure of the PPM symbol, there must be only a single pulse every M chips. This leads to have M possible combinations, that is M different vectors of M chips with probability:

$$P_0 = P_1 = P_2 = \ldots = P_{M-1} = \frac{1}{M} \tag{4.39}$$

On the other hand, when the received signal is not synchronized at symbol time, within a time window of M chips there could be one pulse, two or none. This time the estimator, called as *asynchronous*, must take into account that all the possible combinations (whose number will be grater than that one referring to the synchronous case) are not equiprobable.

In the context of a 2-PPM transmission, Figs. 4.25-4.26-4.27-4.28 describe the behavior of the estimator, both synchronous and asynchronous, for different levels of noise-to-signal ratio (NSR). Refer to the couple of chips entering the estimator as $c[n-1]$ and $c[n]$. When NSR=-7dB (Fig. 4.25) it is possible to observe that with the input chips values close to $c = (0,0)$ and $c = (1,1)$ the synchronous estimator is not able to make a decision because these chips combinations are not allowed. By decreasing noise-to-signal ratio down to NSR=-13dB (Fig. 4.26), the estimator becomes able to perform a *harder* decision. In fact, if the first entering chip results to be slightly

higher than the next one, the estimator considers this difference as due to a higher NSR level, therefore the chip at the output will be estimated as a 1. Vice versa, if the first chip entering is lower than the second, the output chip will be 0.
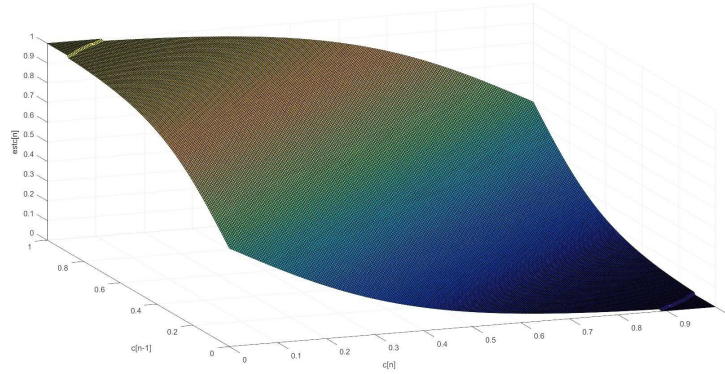

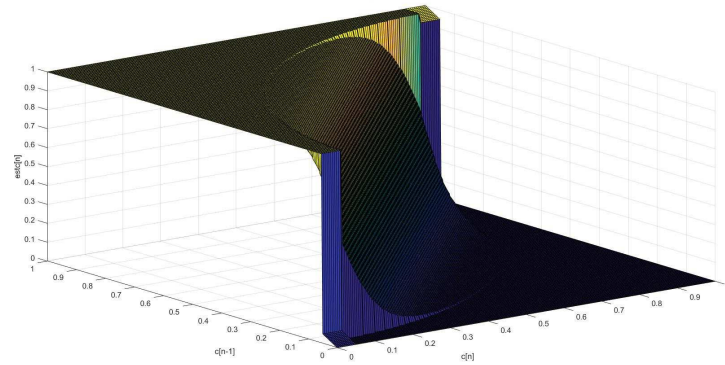Figure 4.25: Behavior of the sychronous estimator for 2-PPM (NSR=-7dB).

Figure 4.26: Behavior of the sychronous estimator for 2-PPM (NSR=-13dB).

Moving to the analysis of the asynchronous case, as the reception of the chips couples $c = (0,0)$ and $c = (1,1)$ is now acceptable, the estimator has quite the same behavior considering NSR=-7dB (Fig. 4.27) and NSR=-13dB (Fig. 4.28). In particular, when NSR=-13dB, the asynchronous estimator is able to work even in the regions of higher indecision (the central part of the graph). In fact, the decrease of NSR reflects on a harder discontinuity when moving from 0 to 1 estimated values.
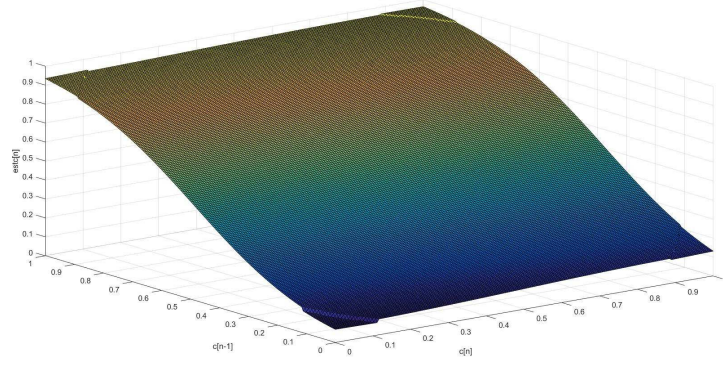
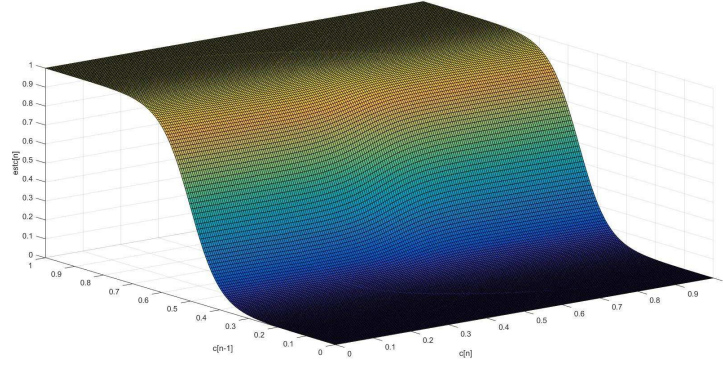Figure 4.27: Behavior of the asychronous estimator for 2-PPM (NSR=-7dB).

Figure 4.28: Behavior of the asychronous estimator for 2-PPM (NSR=-13dB).

4.8.6 Numerical results

The efficiency provided by the proposed LMS-based FSE has been measured through several simulations performed with Simulink software. The equalization scheme has been tested on two different channels taken from the literature, considering a 2-PPM transmission and setting $Q_{Tx} = 9$. The noise-to-signal ratio has been chosesen equal to NSR=-20dB. The performance have been expressed in terms of ISI MSE (Mean Square Error) and the corresponding SEP (Symbol Error Probability), whose average has been measured over 100 Montecarlo runs, each one composed of 50 frames of 512 chips.

The first simulation has considered the channel in [35] used by the authors for testing the performance of a chip-spaced equalizer. Channel impulse and frequency response (that include also the effect of g_T and g_R) are

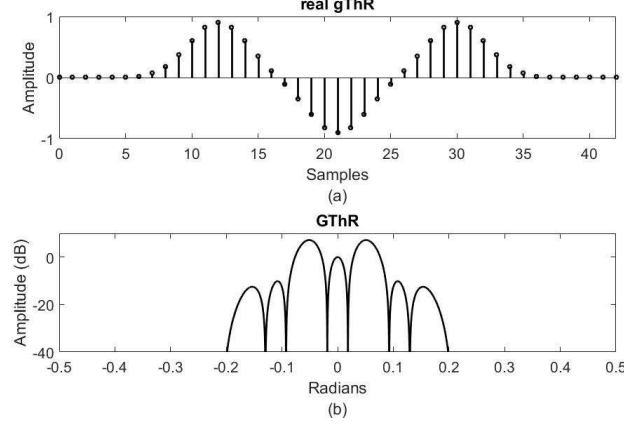reported in Fig. 4.29. The results shown in Fig. 4.30 refer to the proposed

Figure 4.29: Impulse and frequency response for the channel in [35].

equalization scheme implemented both in the chip spaced (CSE) and fractionally spaced (FSE) fashion. The PPM chip structure has been realized considering $\rho = 0.9$ making the pulse length the 90% of the chip duration. By looking at the curves it is possible to observe how the FSE outperforms the CSE, with the symbol error probability reaching a value almost equal to SEP=10^{-11}.
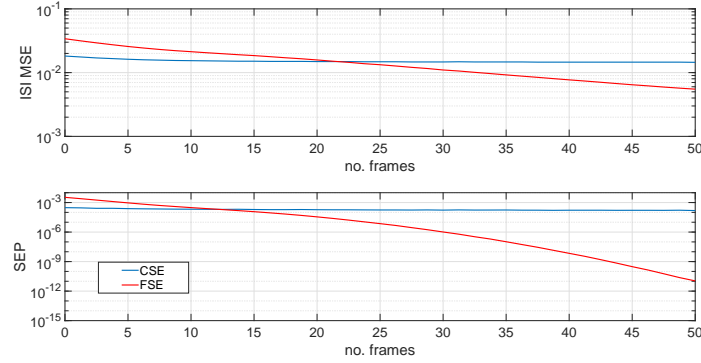

Figure 4.30: ISI MSE and SEP for CSE and FSE.

The second test has aimed at investigating the effect of a larger guard time on the performance of the FSE. In this regard, the channel introduced in [34] has been considered, whose impulse and frequency responce are depicted in Fig. 4.31. The portion of spectrum investigated is equal to 14.778 kHz around the center frequency $f_0 = 33.8$ kHz, so as to consider the working bandwidth of the SAM-1 modem. The simulations has dealt with the case of $\rho = 1$ leading to have the pulse duration equal to the chip time (hence,
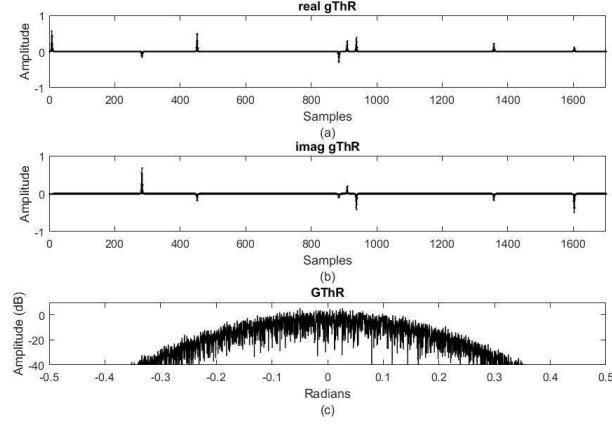
Figure 4.31: Impulse and frequency response for the channel in [34].

no guard interval is used) and $\rho = 0.9$ as in the previous tests. As shown in Fig. 4.32, introducing a guard time within the chip window allows the achievement of slightly better performance in terms of MSE than in the case where $\rho = 1$. This *little* improvement however becomes significant considering the SEP that results in the order of 10^{-33}. Another interesting
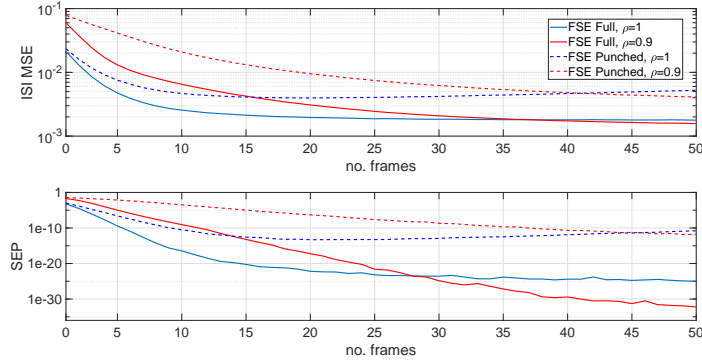

Figure 4.32: ISI MSE and SEP measured for diiferent values of ρ.

aspect to discuss is that when dealing with sparse channels like that one in Fig. 4.31, having a fractionally spaced equalizer may be not convenient since a high number of filter coefficients have to be updated despite some *parts* of the channel do not need to be equalized. In this regard, a possible solution for reducing the complexity requested by the FSE is to *punch* the equalizer support in order to limit the number of coefficients to be updated. Therefore punching the support leads to reduce the computational cost of equalization, but it may also not guarantee the same performance of a full support-based equalizer, as shown in Fig. 4.32.

Summarizing, the spectral properties of pulse position modulation can be exploited in channel equalization for mitigating the effect of ISI. Moreover, better performance can be achieved by introducing a reasonable guard time within the chip window, so as to limit the transmission rate reduction as well.

Chapter 5

MIMO architectures

The discussion about signal propagation presented in the previous chapters has helped in understanding the influence of the underwater channel on the performance of acoustic communications, in terms of both reliability and transmission rate. Those aspects practically represent two sides of the same coin for every communication network, and reaching a trade-off becomes central especially when dealing with highly bandwidth-limited systems. In this regard, realizing Multiple-Input Multiple-Output architectures turns out to be useful since the use of multiple transmit and receive transducers allows significant enhancements of the system performance.

The transmission techniques to be used in the MIMO context are several, each one offering particular benefits about rate or reliability improvement. The choice of the transmission scheme to be employed must be very accurate, otherwise the advantages obtainable may turn into drawbacks. Of course it is important to consider the application scenario considered. The most common activities performed in underwater context such as environmental data collection, coastal surveillance and target detection usually focus only on the reliability of the communication. However, particular services like sea video monitoring generating larger size data need to be supported with transmission rates in the order of kb/s till to Mb/s, that are very hard to achieve without incurring in reliability loss.

However, it is *fundamental* that the system architecture is implemented according to what the nature, that is the medium, allows. In fact, only a thorough knowledge of the channel behavior can correctly suggest the best way to realize a communication system.

The following chapter reports an empirical analysis of how the physical structure of the medium influences MIMO communications. Then, the implementation of some transmission techniques is proposed, comparing the provided performance in terms of BER and rate.

5.1 Channels spatial correlation analysis

When dealing with MIMO communications, high performance can be achieved only if channels are statistically independent, that is spatially uncorrelated. This property depends not only on the spacing among the transducers at transmitter and receiver side, but also on the physical structure of the medium influencing the acoustic propagation. For instance, in the polar regions it is likely to find the water surface covered by ice layers, with this latters representing a boundary to the traveling signal inducing reflection and refraction phenomena that are different from those experienced with fluid sea surface. In this regard, the analysis of the multipath distribution can be used for understenading how the presence of ice at the water surface impacts on the spatial correlation among channels [42].

5.1.1 Reference scenario

Consider a MIMO system composed of $N_{Tx} = 2$ transmitters and $N_{Rx} = 2$ receivers, realizing a 2×2 architecture. This represents a basic MIMO scenario, but the following discussion is valid also for communications where $N_{Tx}, N_{Rx} > 2$. Fig. 5.1 depicts the model under investigation, where the medium is described as a multi-layer structure composed of water and air, eventually separated by an ice layer. The signals generated by the trans-
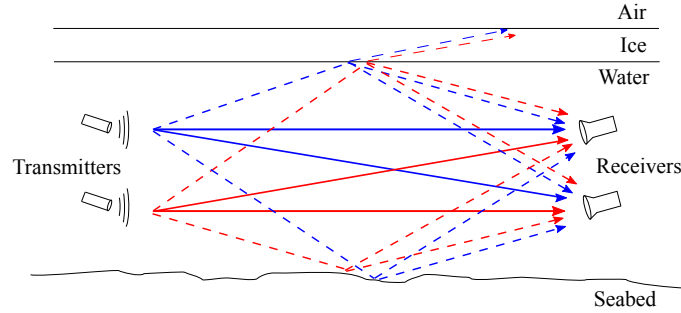

Figure 5.1: 2×2 MIMO reference model.

mitters reach the receiver side both via direct path and via secondary ones induced by the reflections on the floor and off the water surface. In fact, as presented in par. 3.1.4 a propagating acoustic wave incident upon an interface between two media having different acoustic impedance experiences reflection, transmission and scattering phenomena depending on eqs. (3.31)-(3.32)-(3.33).

The scenarios under investigation present the following circumstances:

- the ice layer is not present, so the water-air interface is characterized by a strong impedance contrast (see Table. 5.1) leading to a reflec-

tion coefficient $\mathcal{V}_{w|a}$ close to unity [43]. However, since the fluid water surface is generally non linear due to wind, waves and currents, the reflected wave suffers energy dispersion due to the scattering effect. So, despite the quasi-total reflection of the incident wave, only a part of the reflected signal will propagate coherently in the receivers direction;

- the water surface is covered by an ice canopy where the acoustic waves refect on. Considering the acoustic impedance of ice $\mathcal{Z}_{ice} \simeq 3$, the impedance mismatch at the boundaries with fluid water is not so pronounced. In this case the reflection coefficient is $\mathcal{V}_{w|i} \approx 0.3$ [44], thus meaning that part of the wave is transmitted through the ice layer. Moreover, by assuming the ice layer sufficiently smooth, the energy loss can be considered negligible since the scattering effect is attenuated with respect to the previous case. So the reflected signal will almost fully propagate forward.

Therefore the medium boundary conditions define the nature of multipath, thus influencing the model and relationship among communication channels.

Medium	Speed of sound c (m/s)	Density ρ (Kg/m^3)	Acoustic mmpedance $\mathcal{Z}$ $(Rayl)$
Water	1480	1	1,48
Ice	3840	$0,9 \cdot 10^{-3}$	3,45
Air	340	1,225	415,5

Table 5.1: Acoustic properties of water, ice and air.

5.1.2 Multipath distribution analysis and tests results

According to the 2×2 system architecture previously introduced, the received signal can be written in a discrete-time shape as:

$$\mathbf{Y} = \mathbf{S^T H} + \mathbf{W}, \tag{5.1}$$

where $\mathbf{Y}$ and $\mathbf{S}$ are $[2 \times 1]$ vectors respectively collecting the signal received and transmitted. $\mathbf{W}$ is the $[2 \times 1]$ vector describing the whole disturbance, expressed in terms of thermal and ambient noise. Finally, $\mathbf{H}$ is a $[2 \times 2]$ matrix expressed as:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \tag{5.2}$$

being each element h_{ij}, with $i, j = 1, 2$, the N_t-taps impulse response of the channel (CIR) between the i-th transmitter and the j-th receiving acoustic element.

MIMO communcations result to be efficient if the channels are statistically indepent, that is they are not spatially correlated. This relationship can be inferred by evaluating, tap by tap, the amplitude and delay ratios between the CIRs couples (h_{12}, h_{11}) and (h_{21}, h_{22}). In order to track the channels variations, the CIRs are re-estimated N_e times during the communication, returning N_e values for each measured ratio that allow to evaluate the spatial correlation along time.

So, considering the couple (h_{ij}, h_{ii}), with $i, j = 1, 2$, $i \neq j$, the following $[N_e \times N_t]$ matrix is obtained:

$$
\mathbf{A}_{h_{ij}, h_{ii}} =
\begin{bmatrix}
a_0^{(1)} & \cdots & a_n^{(1)} & \cdots & a_{N_t-1}^{(1)} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
a_0^{(e)} & \cdots & a_n^{(e)} & \cdots & a_{N_t-1}^{(e)} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
a_0^{(N_e)} & \cdots & a_n^{(N_e)} & \cdots & a_{N_t-1}^{(N_e)}
\end{bmatrix}
\tag{5.3}
$$

with $i, j = 1, 2$, $i \neq j$. Each element gathered in $\mathbf{A}_{h_{ij}, h_{ii}}$ represents the amplitude ratio measured for the n-th tap of the e-th estimate of the CIRs (h_{ij}, h_{ii}). Furthermore, the matrix $\mathbf{D}_{h_{ij}, h_{ii}}$, structured as (5.3), is given by collecting the delay ratios.

Analyzing by columns the data in $\mathbf{A}$ and $\mathbf{D}$ it is possible to appreciate how the correlation between channels evolves in time. Having CIRs ratios almost constant demonstrates high spatial correlation between channels, otherwise a uniform distribution of values means that the channels can be considered statistically independent.

The empirical method above proposed has been used in some experiments that took place in a shallow water reservoir in Rome (Villa Ada Lake, 41.932039° N, 12.501497° E), first in October 2016 with mild weather and normal water conditions, and then in January 2017 when the very low temperature caused the water surface freezing, creating a smooth ice layer of about 3 centimeters (Fig. 5.2). In this way it has been possible to consider both the scenarios described above and analyze the effect of ice on the signal propagation. a 2x2 MIMO communication has een realized along a 90 meters link distance (Fig. 5.3). The water depth was about 2 meters.

Two SAM-1 acoustic modems placed 20 centimeters one from the other were used for transmission. At receiver side a couple of 2 meters-spaced AS-1 hydrophones was employed for signal acquisition. The tests consisted in the transmission of a known sequence of symbols represented by the syn-

Figure 5.2: Villa Ada lake (Rome) tests with fluid (Up) and iced (Down) water surface.

Figure 5.3: Environment considered for measurements.

chronization ping of each SAM-1 frame sent (the others remained unused) in order to obtain N_e=30 different measurements per channel, each described by a CIR (estimated through Matlab-based data processing) of N_t=5 taps, that have been used for building the [30 × 5] matrices **A** and **D**.

Focusing on the couple (h_{12}, h_{11}), the distribution of the elements in $\mathbf{A}_{h_{12},h_{11}}$ taken by column is reported in a histogram, with each color corresponding to the n-th tap of the CIRs. In particular Fig. 5.4a-b refer to the scenario with fluid and frozen reservoir surface, respectively. By looking at the results note that, when the water surface is fluid, the channels are scarcely correlated. In fact, the distribution of values in the histogram in Fig. 5.4a is rather omogenous. Differently, regarding the case of frozen surface (Fig. 5.4b) an evident peak in the ratios distribution is present, thus meaning that channels are strictly dependent. Considering data gathered in

$\mathbf{D}_{h_{12},h_{11}}$ the delay ratio results to be approximately constant in both scenarios. The measured values range from 0.98 to 1.03, so the timing of signal replicas arrival at the receiver is similar.

The analysis of the couple (h_{21}, h_{22}), whose results are reported in Fig.
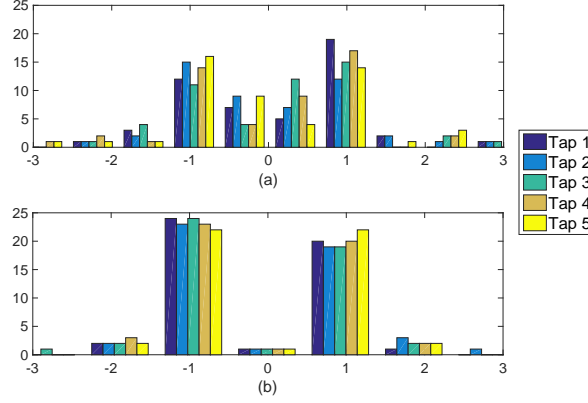

Figure 5.4: Amplitude ratios distribution considering fluid (a) and frozen (b) water surface scenario (channels (h_{12}, h_{11})).

5.5a-b, leads to the same conclusion as for the case of (h_{12}, h_{11}). In fact, the distribution of amplitude ratios is uniform when the water surface is fluid, while the values become more clumped if a frozen interface is present. Anyway, by looking at Fig. 5.5 the difference between fluid and iced water surface scenario is slightly less evident with respect to that one observed in Fig. 5.4 referring to the couple (h_{12}, h_{11}). This is probably due to the fact that, along the path from transmitter 2 to receivers 1 and 2, the roughness of the lake floor and the presence of some rubbles have produced an higher scattering effect making the channels h_{21} and h_{22} less correlated. Concerning the delay ratios, the values reported in $\mathbf{D}_{h_{21},h_{22}}$ are distributed over a narrow interval around the unity, showing a trend comparable with that one measured for the case of (h_{12}, h_{11}).

Summarizing, the presence of an iced water surface layer demonstrates to be an impairment to MIMO underwater acoustic communications, especially in shallow water, since it increases the spatial correlation among channels, adversely influencing the system performance. So, the information obtained through this analysis has to be taken into account when choosing the transmission scheme to be employed. When channels are statistically independent, it is reasonable to use Spatial Multiplexing in order to increase the transmission rate. On the other hand, if channels are strongly correlated it would be more appropriate to employ techniques like Space-Time Block Coding that improve the reliability. In this regard, more details are provided
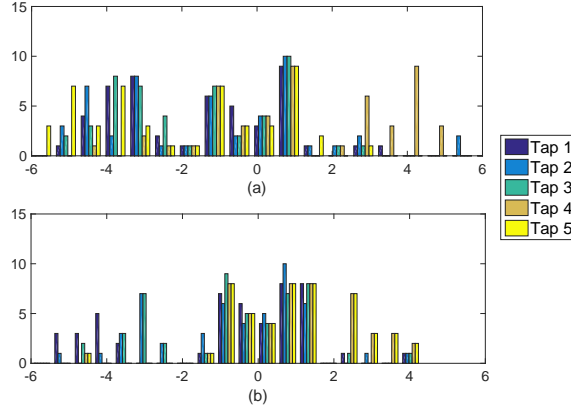
Figure 5.5: Amplitude ratios distribution considering fluid (a) and frozen (b) water surface scenario (channels (h_{12}, h_{11})).

in the next paragraphs.

5.2 Transmission schemes: Spatial Multiplexing vs Space-Time Block Coding

MIMO communications are designed to work using multiple transmitters and receivers. This kind of architecture shows to be useful when dealing with rate improvements issues. In fact, by resorting to spatial multiplexing it is possible to transmit N independent data streams (where N is the number of transmitters) so achieving a rate N-times higher than that one provided by Single-Input Single Output (SISO) communications [45]. Unfortunately this benefit is paid in terms of reliability, since the presence of multiple strems increases the interference among signals that may reflect on higher BER. This is even more true if the channel conditions are not favourable, that is when the channels are strongly correlated [46]. Actually in this latter scenario the first aim is usually to guarantee the reliability of the link, so space-time block coding techniques result to be more suitable since allowing a more robust data decoding with respect to spatial multiplexing.

In this context, a single-carrier MIMO architecture using PPM modulation is proposed below, implemented with both spatial multiplexing (SM) and space-time block coding (STBC) [47]. The perfomance comparison is carried out in terms of achievable rate and BER.

5.2.1 MIMO system model

Consider a MIMO system where N_T and N_R are the number of acoustic transmitters and receivers, respectively (Fig. 5.6). The signal y_i measured at the i-th receiver can be expressed as:

$$y_i(t) = \sum_{j=1}^{N_T} s_j(t) * h_{ji}(t) + w_n(t) \tag{5.4}$$

thus meaning that it collects different signals $s_j(t)$ (with $j = 1, 2, ..., N_T$) coming from each of the N_T transmitters and passing through a channel with impulse response $h_{ji}(t)$, and summed with noise w_n. Since the transmission
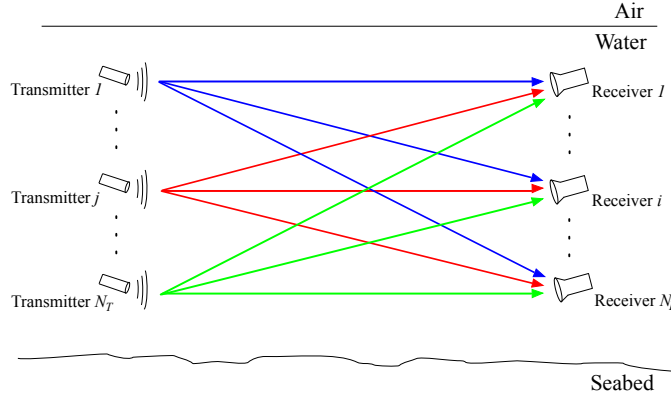

Figure 5.6: MIMO system architecture.

scheme employed is based on a L-PPM modulation, the form of the signal generated by the j-th transmitter is the following:

$$s_j(t) = g(t - kT_c) \tag{5.5}$$

with T_c the PPM chip time and $k = 0, 1, ..., L - 1$ depending on the PPM symbol to be represented.

The whole system can be also described by resorting to the matrix notation as

$$\mathbf{Y} = \mathbf{SH} + \mathbf{W}, \tag{5.6}$$

where $\mathbf{Y}$ and $\mathbf{W}$ are $[L \times N_R]$ matrices collecting respectively the signal received and the noise components, $\mathbf{S}$ matrix is $[L \times N_T]$ and contains the transmitted symbols, and $\mathbf{H}$ is the impulse response $[N_T \times N_R]$ matrix. The m-th matrix $\mathbf{S}_m$, belonging to the set $\mathfrak{S}$ of L^{N_T} elements with $m \leq L^{N_T}$,

represents the space and time distribution of pulses in the form:

$$\mathbf{S}_m = \begin{bmatrix} s_1^{(0)} & s_2^{(0)} & \dots & s_j^{(0)} & \cdots & s_{N_T}^{(0)} \\ s_1^{(1)} & s_2^{(1)} & \dots & s_j^{(1)} & \cdots & s_{N_T}^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_1^{(l)} & s_2^{(l)} & \dots & s_j^{(l)} & \cdots & s_{N_T}^{(l)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_1^{(L-1)} & s_2^{(L-1)} & \dots & s_j^{(L-1)} & \cdots & s_{N_T}^{(L-1)} \end{bmatrix}. \tag{5.7}$$

The symbol generated by the j-th transmitter is reported by column and distributed over a time window $T_{sym} = L \cdot T_c$, so that each element represents one of the L chips composing the PPM symbol. The chips notation follows the rule $(l \bmod L)^1$. Thus the transmitters are distinguished by row. Furthermore, according to the PPM symbol structure, each column vector of the matrix $\mathbf{S}$ is built following two constraints:

$$\sum_{k=0}^{L-1} s_j^{(k)} = 1 \ and \ s_j^{(k)} \in \{0,1\} \tag{5.8}$$

The impulse responses of the channels are instead gathered in the matrix $\mathbf{H}$ shown below:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1i} & \cdots & h_{1N_R} \\ h_{21} & h_{22} & \dots & h_{2i} & \cdots & h_{2N_R} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{j1} & h_{j2} & \dots & h_{ji} & \cdots & h_{jN_R} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{N_T1} & h_{N_T2} & \dots & h_{N_Ti} & \cdots & h_{N_TN_R} \end{bmatrix} \tag{5.9}$$

with h_{ji} referring to the channel between the j-th transmitter and the i-th receiver. During the communication progress the transmitters send some pilot symbols to let the receivers perform the channel estimation, therefore assume as known the $[N_T \times N_R]$ matrix $\tilde{\mathbf{H}}$ with the estimated channel impulse responses. The estimation has to be carried out according to the coherence time of the channel in order to track its variations, hence the matrix $\tilde{\mathbf{H}}$ is periodically updated.

5.2.1.1 Transmission according to Spatial Multiplexing

In the SM case the stream is divided in N_T substreams each modulated by the corresponding acoustic transmitting element. Considering a frame-oriented L-PPM scheme (as that one provided by the SAM-1 modem), the

---

[1]The expression $(.)\bmod L$ outputs the remainder of the integer division between $(\cdot)$ and L.

number of transmittable matrix codewords is L^{N_T}, providing a maximum bit rate equal to:

$$\mathcal{R}_{SM} = \frac{N_s \log_2(L^{N_T})}{T_f} \qquad (5.10)$$

being N_s the number of symbols for frame and T_f the frame time length. The unfavorable aspect of SM is the decrease of the distance among codewords inducing a higher number of detection errors, hence rising the BER.

5.2.1.2 Transmission according to Trace Ortgogonal STBC

STBC can be implemented basing on different rules, each on providing a specific set of transmittable codeword matrices. An example is represented by Orthogonal Space-Time Block Coding (OSTBC), where the codeword to be used are only those ones retaining the orthogonality property. Unfortunately, once assigned N_T, L defining the codewords set $\mathfrak{S}$, it results that the number of elements respecting orthogonality is very low. So, despite the robustness to errors increases as the distance among codewords grows, the achievable transmission rate is considerably lower with respect to that one provided by SM where the number of codewords used is maximum.

A trade-off between OSTBC and SM is given in [48] where Trace-Orthogonal (TO) STBC is introduced. Following this solution, the codeword matrix to be transmitted must retain the property of having the zero-trace between all the codewords, that is:

$$\text{trace}\{\mathbf{S}_p^T \mathbf{S}_m\} = 0, \quad p \neq m, \quad p, m \leq L^{N_T} \qquad (5.11)$$

The number of codewords showing this feature is lower than L^{N_T} (SM) but anyway greater than the number of matrices used OSTBC, so finally the performance of the scheme are in the middle between SM and OSTBC (see [48] for details).

5.2.2 Receiver architecture

Regarding transmission, TO-STBC and SM techniques differ only in the number of utilizable codeword matrices, but at receiver side they require the adoption of different mechanisms for signal detection and decoding.

However, both procedures rely on the channel knowledge, whose estimation is performed by periodically sending a pilot codeword $\mathbf{S} = \mathbf{I}$, being $\mathbf{I}$ the identity matrix. By doing so, the received codeword is:

$$\mathbf{Y} = \mathbf{H} + \mathbf{W} \qquad (5.12)$$

that represents with sufficient accuracy the channel, despite the presence of a noise component. So, the channel estimates are essentially given by:

$$\tilde{\mathbf{H}} \equiv \mathbf{Y} = \mathbf{H} + \mathbf{W} \qquad (5.13)$$

The information carried by $\tilde{\mathbf{H}}$ is then exploited in TO-STBC and SM as follows.

5.2.2.1 TO-STBC-based Receiver

For what concerns the TO-STBC it is possible to resort to a spatial ZF (Zero Forcing) as detailed in [48] that leads to obtain the following metric:

$$\mathbf{Z} = \mathbf{Y}\tilde{\mathbf{H}}^{\sharp} \tag{5.14}$$

where $\tilde{\mathbf{H}}^{\sharp}$ is the pseudo-inverse matrix of the estimated channel $\tilde{\mathbf{H}}$. The received codeword $\mathbf{Y}$, with respect to the transmitted one $\mathbf{S}$, has no more the trace-orthogonal property due to the effect of the channel, so spatial ZF results to be helpful in re-obtaining orthogonal traces.

Hence, the decided TO-STBC codeword is that exhibiting the maximum value among all the possible codewords according to:

$$\text{trace}\{\mathbf{S}_p^T\mathbf{Z}\}, \forall k. \tag{5.15}$$

5.2.2.2 SM-based Receiver

When dealing with SM N_T signals are sent simultaneously over $N_T \cdot N_R$ different channels. As shown in eq. (5.4) the signal collected by each receiver is the sum of N_T independent components (with noise) that need to be correctly separated. This implies that at a single receiver in principle N_T pulses are present in each symbol time window T_{sym} (see the example in Fig. 5.7). So a proper threshold must be set for detecting if a pulse is present or not and this is in line with the strategy conventionally used for SISO detection.

However, this approach does not solve the ambiguity since the receiver, after detecting the symbols, has also to understand which transmitter have been sent by. Resorting to the channel estimate turns out to be useful to accomplish this task. Considering the signal received by the i-th receiver, the energy measured on a PPM symbol time window $T_{sym} = LT_c$ is given by $E_{y_i} = \sum_{l=0}^{L-1} E_{y_i}^{(l)}$, that is the sum of the L chips energy components, each one expressed as:

$$E_{y_i}^{(l)} = \int_{lT_c}^{(l+1)T_c} |y_i(t)|^2 dt \tag{5.16}$$

Under the hypothesis of reliably estimated channel, the attenuation of a pulse can give information about which transmitter emitted it. Therefore, basing on $\tilde{\mathbf{H}}$ and eq. (5.4), the obtained energy estimate is $\tilde{E}_{y_i} = \sum_{l=0}^{L-1} \tilde{E}_{y_i}^{(l)}$, being each single component given by:

$$\tilde{E}_{y_i}^{(l)} = \sum_{j=1}^{N_T} \int_{lT_c}^{(l+1)T_c} |s_j(t) * \tilde{h}_{ji}(t)|^2 dt \tag{5.17}$$

The term $s_j(t)$ can have L different shapes depending on the symbol that represents, so basing on eq. (5.5) and considering al the N_T emitted symbols it follows that:

$$\tilde{E}_{y_i}^{(l)} = \sum_{j=1}^{N_T} \int_{lT_c}^{(l+1)T_c} |g(t - k_j T_c) * \tilde{h}_{ji}(t)|^2 dt \qquad (5.18)$$

with $k_j = 0, 1, ..., L - 1$ depending on the k-th symbol sent by the j-th transmitter. Finally, the estimate of the N_T transmitted symbols is carried out according to the Maximum Likelihood criterion that leads to obtain:

$$\{\hat{s}_1, \ldots, \hat{s}_j, \ldots, \hat{s}_{N_T}\} = \operatorname*{argmin}_{k_1, \ldots, k_{N_T}} \Delta_{i,k}$$

$$\Delta_{i,k} = \sum_{l=0}^{L-1} |E_{y_i} - \tilde{E}_{y_i}|^2 \qquad (5.19)$$

with $\Delta_{i,k}$ collecting $N_R \cdot L^{N_T}$ metrics.

It is worth noting that it is possible that two or more transmitting elements emit the same PPM symbol. This means that multiple pulses are present in the same slot, so in general not all the L PPM slots could be *filled*. In this case, a threshold mechanism for detection could be not reliable since it does not allow to understand how many pulses are placed in the same slot. By instead resorting to the method above described, through the combination of channels estimates it is possible to have a finer decision on the transmitted symbols, even though the detection complexity for each receiver increases with L^{N_T}.

5.2.3 Performance comparison

Some experiments have been conducted in order to evaluate the performace of the implemented SM and TO-STBC. The reference environment and the communication setup have been the same already used for the measurements reported in par. 5.1.2, so a 2x2 MIMO system has been realized over a 90 meters link (Fig. 5.3). According to the SAM-1 modem features, a 16-PPM scheme has been used. Furthermore, by resorting to the symbol re-mapping mechanism (App. A.1) it has been also possible to perform 2-PPM, 4-PPM, 8-PPM.

An example of signal captured at receiver side by a hydrophone is given in Fig. 5.7. The figure reports a 16-PPM symbol time window where it is possible to find two pulses representing two different symbols, each one emitted by one of the two transmitters. The time axis is divided in 16 sub-windows in order to underline the 16 slots composing the PPM symbol. In this case, the receiver recognizes symbol 2 and symbol 7 (counted using the $(l \bmod L)$ rule), but only by resorting to the mechanisms described in par.

5.2.2 it is possible to understand which source they come from, respectively. The test consisted of the transmission of a text file composed by 960 bits,
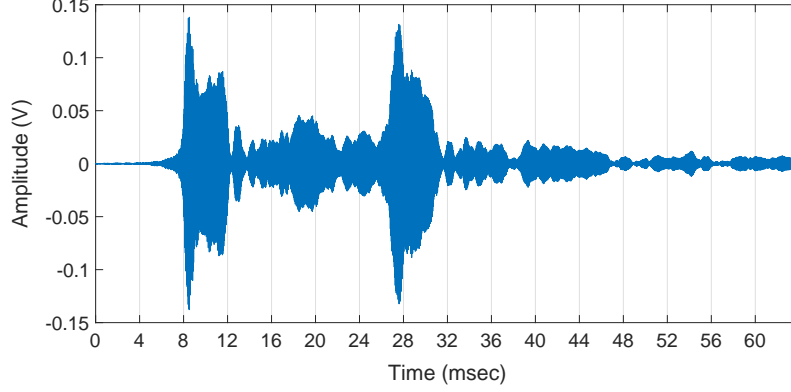

Figure 5.7: Example of signal recorded by a hydrophone.

evaluating the system performance in terms of BER considering the different modulation schemes. Furthermore, since the experiments were conducted when the water surface was both fluid and frozen, it has been possible to infer if the conclusion drawn in par. 5.1.2 about the effect of ice on the channels spatial correlation were correct. It is fundamental highlighting that no time-domain equalization to reduce inter-symbol interference (ISI) and inter-pulse interference (IPI) has been adopted, so to show the effect of the spatial domain processing.

By looking at the results reported in Table 5.2, note how TO-STBC is able to considerably outperform SM in terms of BER. The better performance of TO-STBC with respect to the SM case is mainly due to the use of a limited alphabet of transmittable codewords. In this way, the larger spacing among codewords makes the detection and decoding mechanism more reliable.

An interesting aspect to highlight is that the performance of both TO-STBC and SM increase as the PPM order increases. This fact can be justified by considering that, once set the chip time T_c, the L-PPM symbol time T_{sym} grows with L. The strong multipath effect characterizing the underwater acoustic channel causes very large signal delay spread, inducing ISI that impairs the reliability especially when T_{sym} is short. Since $T_{sym}^{(16ppm)} > T_{sym}^{(2ppm)}$, the delay spread has sufficient time to attenuate without superimposing on the next received symbol temporal window. This limits the ISI, so the BER is lower when the modulation order grows.

Finally, it can be observed that the performance of both TO-STBC and SM get worse when the reservoir surface is frozen. This result perfectly matches that one carried out by the analysis in par. 5.1, hence confirming that when the water surface is frozen the spatial correlation among channels

Transmission	BER	
Scheme	(fluid water surface)	(iced water surface)
SM 2-PPM	0.201	0.208
TO-STBC 2-PPM	0.082	0.050
SM 4-PPM	0.191	0.195
TO-STBC 4-PPM	0.028	0.039
SM 8-PPM	0.112	0.166
TO-STBC 8-PPM	< 0.001	0.019
SM 16-PPM	0.154	0.110
TO-STBC 16-PPM	< 0.001	< 0.001

Table 5.2: Communication BER for different modulation formats

increases causing the system performance reduction in terms of reliability.

The low BER guaranteed by TO-STBC is on the other hand paid in terms of data rate. In fact, SM allows higher rates since, differently from TO-STBC, it has no constraints to meet about the transmittable codewords. The rate provided by both TO-STBC and SM as a function of the PPM constellation are reported in Fig. 5.8. The unit of measurement employed
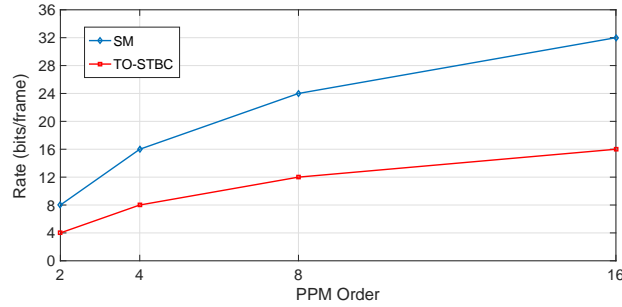

Figure 5.8: Rate comparison between TO-STBC and SM for different PPM constellations considering a 2x2 MIMO architecture.

is *bits/frame* since the results coming from our tests strictly depend on the features of the SAM-1 modem employed. Using other devices, that is having a different frame structure, the rate numerical values change but anyway the trend of curves referring to TO-STBC and SM remains the same respect to one another.

5.2.4 A step forward performance optimization: Quasi Trace-Orthogonal Space-Time Block Coding

The insights about TO-STBC as a potential technique for rate improvement have been validated through the experimental results presented in the

previous paragraph. The transmission rate enhancement with respect to OSTBC is evident [48], while a high level of robustness to errors is mantained. This fact suggests that a modest *sacrifice* in terms of reliability may bring significant benefits for rate, allowing in this context a further step towards the performance of SM. However, it is worth noting that the efficiency loss is *attenuated* by the nature of PPM scheme. In fact, PPM is known to be a bandwidth inefficient modulation providing lower bit rate in comparison with other schemes, but on the other hand it shows higher robustness to ISI. Taking into account this latter aspect, the adoption of a more relaxed condition than orthogonality or trace-orthogonality in STBC can lead to a rate improvement, ensuring reliability at the same time. In this context, it is possbile to introduce a new property about the trace, defined as Quasi Trace-Orthogonal (QTO) condition, characterized by the following constraint to respect:

$$\text{trace}\{\mathbf{S}_p^T \mathbf{S}_m\} \leq \xi, \quad p \neq m, \quad p, m \leq L^{N_T} \tag{5.20}$$

where the parameter ξ is:

$$0 \leq \xi < \lfloor log_2 N_T \rfloor \tag{5.21}$$

The condition in eq. (5.2.4) states that the transmittable codewords are those ones retaining the property of having the trace less than ξ. As ξ depends on N_T, eq. (5.2.4) represents a less strict constraint than eq. (5.2.1.2) since for $N_T \geq 4$ the trace can assume values greater than 0 (for $N_T = 2, 3$ eq. (5.2.4) and eq. (5.2.1.2) are essentially the same). This fact returns that the number of transmittable codewords in QTO-STBC is greater than in the case of TO-STBC, hence leading to a higher transmission rate. Table 5.3 reports the size of the codewords vocabulary for different PPM modulation orders and transmitters employed, and considering the four transmission schemes cited in the current chapter.

Table 5.3: Number of transmittable codewords as a function of scheme, transmitters and modulation order.

N_T	2				3				4			
L-PPM	2-PPM	4-PPM	8-PPM	16-PPM	2-PPM	4-PPM	8-PPM	16-PPM	2-PPM	4-PPM	8-PPM	16-PPM
O-STBC	2	4	8	16	2	4	8	16	2	4	8	16
TO-STBC	2	4	8	16	2	4	8	16	2	4	8	16
QTO-STBC	2	4	8	16	2	4	8	16	2	16	55	256
SM	4	16	64	256	8	64	512	4096	16	256	4096	65536

For what concerns signal detection and codeword decoding, the procedures to be followed are the same discussed when introducing the receiver architecture for TO-STBC. In fact, the difference between TO and QTO relies only in the number of transmittable codewords.

The performance provided by the use of QTO-STBC have been evaluated both in terms of rate and BER, considering a 4x4 MIMO architecture
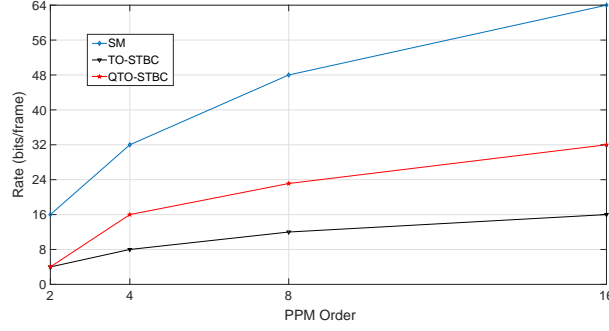
Figure 5.9: Rate comparison between TO-STBC, QTO-STBC and SM for different PPM constellations considering a 4x4 MIMO architecture.

(that is $N_T = 4$) in order to highlight how the proposed technique is able to outperform TO-STBC and approaching SM. The experiments consisted of a 4-PPM based data transmission, considering the same scenario as in Fig. 5.3 (only with fluid water surface). In this case, transmitter and receiver side have been equipped with an horizontal array of 4 SAM-1 modems and 4 hydrophones, respectively. Fig. 5.9 shows the achievable transmission rate of TO-STBC, QTO-STBC and SM, expressed in bits/frame according to the features of the SAM-1 modem. The rate is reported as a funtion of the modulation scheme, that is for 2-PPM, 4-PPM, 8-PPM and 16-PPM. By looking at the trends, it is possible to observe how QTO-STBC is in the middle of TO-STBC and SM. Regarding BER, the results obtained by the tests are collected in Table 5.4, and even in this case QTO-STBC is between TO-STBC and SM. Moreover, the following aspects can be underlined. The

4-PPM Scheme	BER
TO-STBC	< 0.010
QTO-STBC	0.168
SM	0.218

Table 5.4: Communication BER for 4x4 MIMO system.

results in Table 5.4 and Table 5.2 refer to a 4x4 and 2x2 MIMO system respectively. Therefore, as in the first case interference and signals overlapping are stronger, it is reasonable that the BER values in Table 5.4 are higher than those in Table 5.2. By looking at Table 5.3 it can be noticed that in MIMO 2x2 4-PPM SM and in MIMO 4x4 QTO-STBC the number of transmittable codewords is 16 in both cases. Despite the same transmission rate, MIMO 4x4 QTO-STBC outperforms MIMO 2x2 SM in terms of BER, hence

demonstrating the higher reliability provided by QTO-STBC with respect
to SM. Furthermore, as expected, the BER measured for TO-STBC is lower
than the BER for QTO-STBC, but this fact is paid in terms of rate since
the number of transmittable codewords in MIMO 4x4 4-PPM TO-STBC is
lower and equal to 4. Finally, it can be observed that for TO-STBC the
number of used codewords is the same both in MIMO 2x2 and MIMO 4x4,
so the resulting rate is the same but also the measured BER is comparable.

Summarizing, it is worth remarking that for the tests described above
equalization has not been considered. Therefore, by resorting to a well-
performing equalization scheme may lead the measured values of BER to be
all rescaled downward.

Chapter 6

Error control and access

Although the system architecture at the physical layer is designed in order to provide performance as good as possible, it is anyway unlikely that the resulting communication is completely error free. This problem is also handled at the data link layer, where the cooperation among network nodes is exploited for implementing error control strategies making data transmission more reliable. In this context, especially in the underwater scenario, it is also important to understand which is the most suitable portion of spectrum where the transmission may take place, so that the interference with other communications or the impairments caused by the nature of the channel are minimized. Therefore a mechanism for spectrum sensing and access results to be fundamental to achieve high efficiency.

Following this direction, in the current chapter presents an ARQ-based (Automatic Repeat on reQuest) strategy for error detection and correction, while a restatement of the cognitive paradigm is proposed for managing the spectrum access in underwater acoustic networks.

6.1 ARQ protocols for error control

Bi-directional underwater acoustic communications usually request the transmission of some control information among nodes in order to govern the data flow. Attenuation, multipath and noise affecting the signal propagation may lead to errors during the data detection phase, weakening the integrity of the received information. In this regard, the effects induced by the underwater acoustic channel can be, sometimes, partially counterbalanced with the use of coding, that is Forward Error Correction (FEC) and Automatic Repeat reQuest (ARQ), but in some cases ARQ is more reliable than FEC especially in the presence of long error bursts [49]. However, both approaches consider the introduction of overhead information for different purposes. ARQ uses overhead in order to perform error detection and FEC uses overhead for error correction. Moreover, it is worth noting that FEC codes are not always

able to correct errors, so it could happen that the received information is anyway erroneous or incomplete. When instead using ARQ, the integrity of the data is guaranteed by the fact that each frame is retransmitted until it is correctly received.

ARQ is usually implemented in a point-to-point link according to three different schemes that are Stop-and-Wait (S&W), Go-Back-m (GBm) and Selective Repeat (SR). When using S&W, the transmitter sends a frame and waits for a positive acknowledgement (ACK) by the receiver indicating right frame detection, before transmitting the next one. If no ACK message is received within a pre-defined time window, the packet is retransmitted. When GBm is considered, the transmitter can send a sequence of frames, up to M, while waiting for ACKs. The receiver replies with an ACK, reporting also the frame number, that indicates that all the frames until the current one have been detected without errors, so if the ACK referring to the m-th frame is not received the broadcaster node will retransmit frames from the mth to the last one sent. In case of SR, the receiver, by using different ACKs and negative ACKs (NACKs), notifies the transmitter which are the corrupted frames that need to be retransmitted.

ARQ techniques are not new in underwater acoustics. In fact, in [50] a particular ARQ scheme has been implemented by exploiting the long propagation delay in a multi-hop acoustic channel. By choosing the packet size so that transmission time remains smaller with respect to propagation delay it is possible to establish concurrent bidirectional transmissions between nodes, and the backward overhearing is used as a somewhat implicit ACK. In this way overhead and transmission latency are reduced, and enhancement in power saving is also obtained.

A strategy for continuous ARQ is described in [51]. Even in that case, by considering the underwater propagation delay, a transmission scheme based on multiple data packets and ACKs cross-like manner dispatch has been proposed. Simulations results show that the throughput can be improved with respect to the simple S&W.

Furthermore, in [52], two Hybrid ARQ schemes have been presented in order to tackle the problem of error correction capability decrease at low Signal-to-Noise Ratio (SNR). By studying the impact of underwater acoustic channel features on the communication, a particular error control mechanism has been proposed. That strategy allows an improvement in terms of reliability, if compared to simple S&W. The obtained results have been then exploited in [53] to analyze the performance of the proposed schemes in a multi-user UWAN scenario.

An ARQ scheme based on both S&W and Selective Repeat basics has been introduced in [54]. The transmission of the first frame is operated according to the S&W mode so allowing, through the ACK, to measure the round trip time (RTT). If RTT is sufficiently long, the scheme then switches into a SR mode where multiple frames can be sent before receiving an ACK.

In [55] the analysis of the underwater acoustic channel and the variability of its parameters has led to the design of a variable window GBm scheme. With this approach the time window for feedback signalling is dimensioned according to the length of the variable signal propagation delay. This allows to achieve a higher efficiency in terms of channel utilization and error correction than that one given by the simple GBm scheme.

Last, an alternative to ARQ is instead given by [56]. Due to long propagation delays, the feedback waiting time could be very long, so communication through underwater acoustic channels becomes inefficient. The system performs a coded data packets transmission, while feedback signal is used by the receiver to communicate information about channel variation to the transmitter, so that it can adjust the transmit power and packet size.

All the above mentioned works from the literature tackle the issue of optimizing ARQ schemes *without* directly consider the physical layer issues, and, most of them, focus on *point-to-point* links. The contributions that deal with multiple nodes consider them as relays in a multi-hop scenario. Moreover, the analyses have been carried out only by means of simulations, so there are no data coming from field experiments.

Basing on the above considerations, the following section presents the adaptation of some typical rules of the three main ARQ strategies, Stop&Wait, Go-Back-m and Selective Repeat respectively, by considering only the transmission of an (explicit) NACK whenever bit errors are detected in a received frame. In this regard, the NACK message is represented by a single pulse sent on a sub-bandwidth whose center frequency depends on the index of the error-affected frame. Hence the NACK spectrum does not depend on the node that transmits it, therby avoiding the need to implement multi-user detection techniques. The reference communication scenario is a *broadcast* acoustic sensor network, very often used in underwater applications (but largely unexplored in the literature) such as data collection activities where several sensors placed underwater are controlled by a master node that has usually to *(i)* communicate instructions and parameters about the measuring activity to the sensors, *(ii)* carry out the firmware update of the devices remotely (so without removing them from their place), *(iii)* check the presence and safety of potential divers by sending periodically some broadcast *alive* messages. In particular the structure presented is characterized by one transmitter and multiple receivers that employ separated bands for feedforward transmission and feedback signalling. A frequency domain mechanism for NACKs detection is employed. Using the proposed approach grants several benefits with respect to the standard ARQ protocols and FEC that are summarized in Table 6.1.

Table 6.1: Proposed feedback scheme compared with standard ARQ protocols and FEC

STANDARD ARQ PROTOCOLS	PROPOSED MECHANISM	FORWARD ERROR CORRECTION
✓ Received data integrity is guaranteed		✗ FEC codes could not be able to correct all the errors
✗ Communication throughput reduced due to feedback messaging and frame retransmission		✓ No feedback signalling and no frame retransmission
✗ Too much feedback signalling (an ACK/NACK for each user) saturate the transmission resources	✓ Explicit-NACK/Implicit-ACK use allows signalling reduction if compared to explicit ACK-based schemes	
✗ Possible errors in ACK and NACK detection	✓ NACK recognition performed by means of a simple energy detector	
✗ Multi-user detection is needed for colliding signals	✓ Due to the network structure, no multi-user detection is needed	✓ No multi-user detection is requested
✗ Frame numbering necessary to avoid ambiguity	✓ The index of a corrupted frame equals to that of the sub-bandwidth where the corresponding NACK is sent over	---
✓ Few error detection control bits		✗ Large error correction control bits are requested for error detection and correction
✗ High system complexity	✓ Low system complexity and latency	

6.1.1 Network architecture

Refer to an underwater acoustic network (UWAN) composed of a single transmitting node and N receivers. Considering the case of broadcaster-to-node transmission, the signal $r_i(t)$ received by the ith receiving node is given by:

$$r_i(t) = s(t) * h_i(t) + w_i(t) \tag{6.1}$$

where $s(t)$ is the signal emitted by the transmitter, $*$ is the convolution operator, $h_i(t)$ models the channel impulse response from the transmitter to the ith receiver and $w_i(t)$ is the colored acoustic noise. Although the acoustic underwater environment leads to time-variant behavior of channel impulse response (CIR), consider the transmission of a single frame in (6.1.1) and assume that the CIR does not change during the frame emission (that is, the channel coherence time is longer than the frame emission time). In the field tests it has been experienced that this assumption sometimes is not verified due to, for example, wind intensity and/or direction changes. This occurrence induces errors at the receiver, that from the ARQ point are reflected on by a higher number of retransmissions. From a system perspective, once set the transceiver transmission power P_{tx}, the network coverage area can be inferred by considering the target Bit Error Rate and the attenuation, so that noise and channel effect have to be taken into account too. Consequently, the maximum distance d_{max} where receivers can be placed is determined by using proper models for channel attenuation and, in case, using also a margin able to assure a minimum received power. On the other hand, we can proceed reversely by evaluating the transmission power necessary to guarantee the coverage of an assigned area. In order to handle this kind of problems it is possible to resort to the signal propagation model

proposed by Urick [9] for evaluating the coverage of the network.

The information to be sent is organized in frames of length L bits; the frames are transmitted by the source node at rate R bit/s, so the data emission period T_e is simply given by:

$$T_e = \frac{L}{R} \quad [s] \tag{6.2}$$

The speed of sound in water is approximately equal to $c = 1500$ m/s and the frame arrives at the ith receiver after the following propagation time:

$$\tau_{p_i} = d_i/c \quad [s] \tag{6.3}$$

Then, according to a general ARQ scheme, the receiver generates a feedback signal whose time length is T_N seconds and sends it back to the transmitter. The feedback signal experiences the same propagation delay τ_{p_i}.

The model of the UWAN described above is shown in Fig. 6.1. The blue arrows refer to data transmission from the source towards the receivers, while the red ones represent the channel where potential feedback signals pass through.
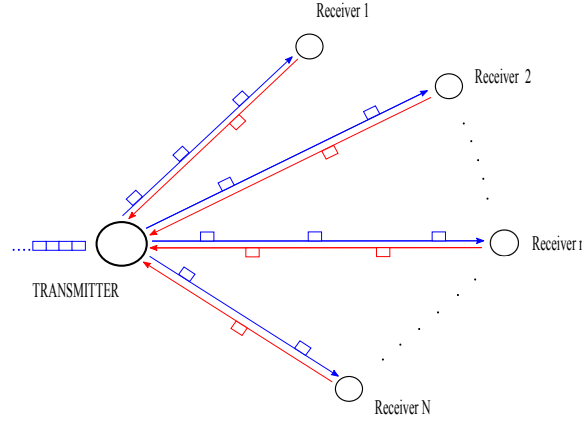
Figure 6.1: Representation of a broadcast network characterized by one transmitting node and N receivers.

6.1.2 Proposed Feedback for ARQ schemes

One of the peculiarities of the proposed ARQ strategy is that the available bandwidth B_{tot} is split into three parts:

$$B_{tot} = B_{tx} + B_{fb} + B_{guard} \quad [Hz] \tag{6.4}$$

where B_{tx} is reserved for forward data transmission and B_{fb} is used for feedback (B_{tx} and B_{fb} can be eventually separated by a guard band B_{guard}).

Depending on the Q-ary modulation used at physical layer, the transmission rate for non-orthogonal constellations is:

$$R = \frac{1}{T_{sym}} \log_2 Q \quad [bit/s] \tag{6.5}$$

where T_{sym} is the symbol transmission time defined as $T_p + T_g$, being T_p the length of the acoustic pulse physically representing a symbol emitted by the source and T_g a guard time used for separating the transmission of a symbol from the next one. Then the parameter β is introduced in order to obtain the relationship $T_{sym} = \beta T_p = \beta/B_{tx}$. The term β defines the ratio between T_{sym} and T_p, so it essentially allows to regulate the length of the guard time that is $T_g = (\beta - 1)T_p$. The time window T_g is set up to counterbalance the effect of multipath in place of using channel equalization mechanism that is, in principle, well performing, even though it increases both the hardware and the computational complexity of the system. However, the proposed scheme is fully compatible with any receiver architecture. The rate in (6.5) is valid for Pulse Amplitude Modulation (PAM), Quadrature Amplitude Modulation (QAM), and also On-Off Keying (OOK). On the other hand, for orthogonal modulations like for example Pulse Position Modulation (PPM), and power efficient modulations in general, the rate is given by:

$$R = \frac{1}{T_{sym}Q} \log_2 Q \quad [bit/s]. \tag{6.6}$$

Concerning the feedback signal, the idea is that each receiver sends back a NACK message whenever it detects a corrupted frame and *does not send back any signal* when the frame detection is error free. The motivation for this choice is twofold. First, using explicit ACK means that each node must communicate its own feedback, frame by frame, to the broadcaster. This induces a lot of signalling in the water and, furthermore, requires that the broadcaster must apply multi-user detection techniques or orthogonal access for feedback signalling so leading to higher system latency and complexity. Second, in a multi-user broadcast scenario it suffices that at least one node detects as corrupted by errors one frame to require retransmission to all nodes.

In order to maintain the NACK structure *as simple as possible* without incurring in multi-user detection or other complex techniques, the feedback bandwidth B_{fb} is partitioned in M sub-bands (Fig. 6.2) to let identify the index of frame(s) affected by errors with a corresponding sub-band. Hence, the feedback bandwidth can be expressed as:

$$B_{fb} = \sum_{m=0}^{M-1} \Delta f_m \quad [Hz] \tag{6.7}$$

with M equal to the number of frames that can be sent without waiting for feedback. The frame sequence length is equal to M. In this way if the

receiver detects an error on the ℓ-th frame, it will send the NACK by using a signal whose bandwidth is centered around the ℓ-th (out of M) sub-carrier. Moreover, using this mechanism will reflect also on the available bandwidth B_{tx}. In fact depending on the value of M, B_{tx} equates $B_{tot} - B_{guard} - M\Delta_f$. Regarding the detection on the forward channel, that is, from the broad-
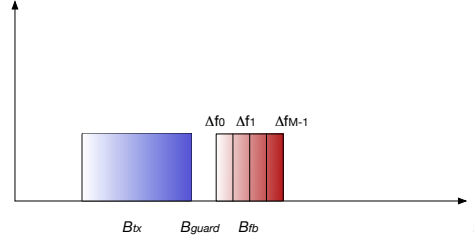

Figure 6.2: Graphical representation of the transmission bandwidth B_{tx} and feedback bandwidth B_{fb}.

caster to the ith receiver, a very simple detection scheme is implemented. Specifically, following the features of the SAM-1 modem, consider that the frame presents, as first pulse, a synchronization ping $s_p(t)$ whose shape is known at each receiver. Hence a cross-correlation can be operated at the ith receiver so as to obtain the following expression:

$$\xi_{i_{sync}}(\tau) = \int_0^{T_{sym}} r_i^*(t)s_p(t+\tau)dt. \tag{6.8}$$

and the synchronization is then given by:

$$\tilde{\tau}_{sync} = T_c \text{arg max } j = 1, ..., L|\xi_{i_{sync}}(jT_s)| \tag{6.9}$$

where L is the number of channel paths, T_s is the sample time. Without adding too many details, longer training sequences allow a finer synchronization. Moreover, as just shown, this mechanism makes not only synchronization possible, but also phase recovery and channel estimation. By basing on the previous steps, the detection is performed according to the maximum likelihood detection (MLD) criterion so the symbol detected will be:

$$\hat{s} = \text{arg max } j = 0, 1, ..., Q - 1 \frac{\xi_i(\tau_{sync})}{\xi_{i_{sync}}(\tau_{sync})}. \tag{6.10}$$

where $\xi_i(\tau_{sync})$ is the cross-correlation as in (6.8) where in place of $s_p(t)$ the pulse shape of data signal $s(t)$ is used. About the emission of the NACK, the

generic receiver will use a bandwidth equal to $M\Delta f$ to send this information through the pulse that lasts T_N seconds ($M = 1$ for S&W). Concerning the mechanism for the emission of NACK, when the ℓ-th frame is affected by errors, the receiver sends the following signal:

$$x_F^{(\ell)}(t) = g_F(t) \cos\left[2\pi(f_0 + \ell\Delta f)t\right] \tag{6.11}$$

which means that the spectrum of the pulse shape $g_F(t)$ is centered on $f_0 + \ell\Delta f$ frequency being f_0 related to the frame 0. The corresponding received signal will be:

$$y_F(t) = x_F^{(\ell)}(t) * h_F(t) + w_F(t) \tag{6.12}$$

where $h_F(t)$ is the channel impulse response related to the feedback channel (in general assumed different from the forward one as demonstrated in par. 4.7) and $w_F(t)$ is the acoustic colored noise.

It is mandatory noting that, although each node acts autonomously by sending NACK if needed, the possible partial overlap of signals related to the same frame does not require multi-user detection techniques since, as detailed in the following, the broadcaster needs simply to measure the energy component of the ℓ-th subchannel to recognize if it is a NACK. From this point of view it does not matter which user generated the signal or if it is generated by multiple users. The only point is understanding whether at least one node requires retransmission. Before detailing the different ARQ mechanisms, it is important to recall that generally the channel impulse response is possibly time-variant, thus meaning that it does not remain the same along time. However, when referring to the transmission of a NACK that corresponds to a single pulse, the hypothesis of channel time invariance is largely satisfied.

Once established the features of the signal to be used for feedback, the rules regulating communication between transmitter and receiver and feedback detection must be detailed. In this context, the cases of S&W, GBm and SR are analyzed separately.

6.1.2.1 Stop and Wait

In the S&W protocol, the transmitter cannot send a new frame until it is sure that no receiver has asked for retransmission of the last one sent. The time to wait before sending the next frame depends on the distance of the furthest receiver (d_{max}) that is the maximum distance where a node can be placed. This is shown in the time diagram of Fig. 6.3, where it is possible to appreciate that after the time needed for emitting one frame T_e, the receiver must wait for two-way propagation $2\tau_{p_{max}}$ with respect to the worst case of a node being at maximum distance, and the emission of a NACK that requires T_N seconds. So the timeout can be set to:
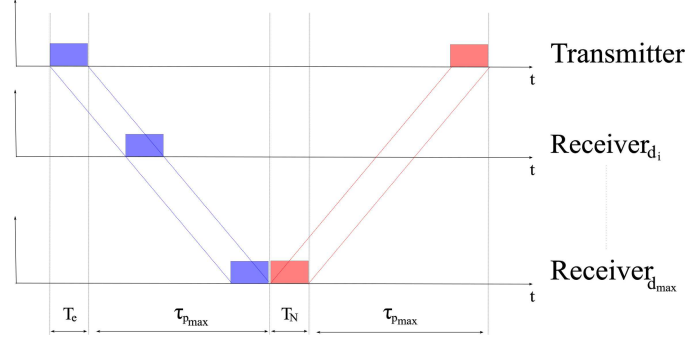
Figure 6.3: Stop&Wait ARQ temporal execution.

$$T_{out} = T_e + 2\tau_{p_{max}} + T_N \quad [s] \tag{6.13}$$

The flowchart of S&W ARQ protocol is shown in Fig. 6.4. At the beginning the transmitter sets the index of frame to be sent to 0 and before sending it the broadcaster stores the frame. Once the frame has been transmitted a timer starts to measure time until the maximum value set to the timeout T_{out}. In the meantime if a NACK has been received, the broadcaster resends the frame, while if no NACK has been detected, the transmitter waits a time equal to T_{out}. Once T_{out} expires the frame is assumed as correctly received and the next frame (with index 1) is stored and sent. The frame numbering follows the update rule $m = (m + 1)\mod 2$ that, for S&W, means to have only 0 and 1 as frame indexes. About the i-th receiver, it sets the index of
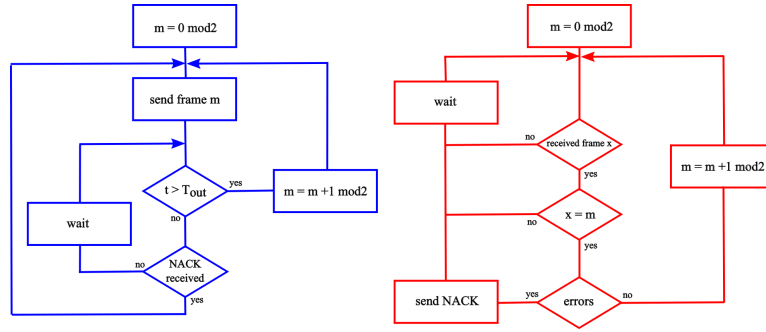
Figure 6.4: Stop&Wait ARQ flowchart. Blue and red diagram refer to transmitter and receiver respectively.

frame to be received to 0 and waits to receive a frame. If it receives a frame with wrong frame index, (e.g., the received frame index is 1), it interprets this occurrence as a frame duplication that can be due, for example, to retransmission required by other nodes. Hence, no action is taken since

the broadcaster will proceed to transmit the next frame (with frame index 0) after the time out if no NACKs are received. If the received frame has the correct frame index the receiver checks if errors are present. If so a NACK is sent otherwise the index of expected frame at the next step is changed according to $m = (m+1) \mod 2$. In order to detect if a NACK has been sent by at least one node, the broadcaster must evaluate the energy component present in the Δf bandwidth via a threshold mechanism. Hence, by measuring the energy as:

$$E_{Y_F} = \int_{T_{out}} |y_F(t)|^2 dt \qquad (6.14)$$

it is possible to infer about the hypothesis of NACK presence of absence by comparing E_{Y_F} with a threshold ϑ_E. Obviously a too low ϑ_E value leads to consider the presence of NACK even when no NACK has been sent. This is false alarm (FA). On the other hand, a too high ϑ_E value induces to miss the presence of a NACK. This is missing detection (MD). It is important to underline here that the threshold setting should *follow* the behavior of the channel. Since dealing with energy components, a quick and easy solution is to consider the above discussions about dimensioning the system for the maximum distance and its relationship with attenuation. This means that the threshold is set up based on the attenuation experienced by the signal coming from the furthest node. This allows the number of missing detection events to be minimized. In fact, while a false alarm induces unnecessary retransmissions, the missing detection leads to information losses and for this reason it is the worst case since it does not allow data integrity. For the proposed broadcast scheme, it suffices the presence of a single NACK to obtain a retransmission.

In the ideal case where no retransmission is needed, the throughput efficiency, defined in [57], is given by the number of data bits per frame transmitted in a time given by eq. 6.13. So it foloows that:

$$\mathcal{T}_{S\&W_{ideal}} = \frac{L(1-\rho)}{T_e + 2\tau_{p_{max}} + T_N} \quad [bit/s] \qquad (6.15)$$

where ρ indicates the percentage of the frame length L used for error detection under the constraint that ρL is an integer. Furthermore, ρL data bits can also be used for symbol synchronization and channel estimation. Considering instead the possible presence of errors and hence retransmissions, expression eq. (6.15) must take into account the error events, so it becomes:

$$\mathcal{T}_{S\&W} = \frac{L(1-\rho)}{T_e + 2\tau_{p_{max}} + T_N} (1 - \sum_{k=1}^{+\infty} P_e^k) \quad [bit/s] \qquad (6.16)$$

where P_e is the error probability for each frame transmission counted k times that is the number of occurred retransmissions. Moreover, once introduced

the Frame Error Rate (FER), that is the total number of retransmitted frames divided by the number of transmitted frames, a method for measuring the throughput efficiency can be:

$$\mathcal{T}_{S\&W} = \frac{L(1-\rho)}{T_e + 2\tau_{p_{max}} + T_N}(1 - FER_{S\&W}) \quad [bit/s] \qquad (6.17)$$

6.1.2.2   Go-Back-M

GBm strategy is a member of continuous ARQ schemes and it allows to send a sequence of M frames without expecting for an ACK. In fact, instead of sending a feedback for each arrived frame, usually the receiver can use a single flag for communicating back that frames up to the m-th of the sequence have been received correctly, so it requests only the retransmission of frames starting from the $(m+1)$-th forward. In the proposed case implicit ACK is assumed as for S&W, hence the receiver transmits back a NACK once it recognizes a frame with errors, asking for the retransmission from that one forward to the last one already transmitted. Concerning NACKs,
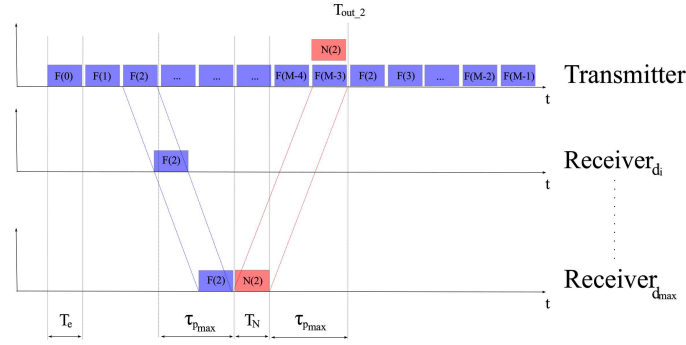
Figure 6.5: GoBack-m ARQ temporal execution. Example of error occurrence at frame 2.

there could be ambiguity of interpretation since the frame numbering follows the $(.)\mathrm{mod}M$ rule. In fact, due to different propagation delays related to receiving nodes at different distances, the broadcaster could not recognize if a NACK refers to a frame of the current or the previous transmitted sequence. This ambiguity is avoided if the m-th frame of the current sequence is sent once the timeout related to the transmission of the possible NACK referring to the m-th frame of the previous sequence has elapsed. The time diagram reported in Fig. 6.5 provides an example. It is possible to appreciate that different nodes receive the same frame in different time instants. This is due to the different distances. However, under the assumption that, for example, frame 0 is affected by errors, the transmission framing must consider the time needed by the possible NACK coming from the most distant destination (in

red in Fig. 6.5) to arrive at the transmitter. Hence, it is not possible to transmit a frame with index 0 before the timer related to the previous 0 frame sequence has expired. So, in order to avoid ambiguity, the sequence length M must be chosen so that:

$$MT_e \geq T_{out} = T_e + 2\tau_{p_{max}} + T_N \qquad (6.18)$$

Substituting eq. (6.1.1) and eq. (6.1.1) leads to have:

$$M \geq \frac{T_{out}}{T_e} = 1 + 2\frac{d_{max}R}{v_p L} + T_N\frac{R}{L} \qquad (6.19)$$

This can also be interpreted in the following way. If the maximum distance increases, M must be increased to avoid ambiguity[1]. This is quite intuitive as increasing the maximum distance means increasing the propagation delay so the NACKs coming from far nodes need more time to propagate. This is obviously a drawback since in case of an error all the sequence must be retransmitted by starting from the first frame not yet positively acknowledged. Besides, this requests a wider bandwidth B_{fb} so reducing B_{tx}.

The flowchart representing how the GBm works is reported in Fig. 6.6. Regarding the tasks performed by the transmitter, at the beginning it initializes the frame index to 0 and sends up to M different frames that can be stored in the transmitter buffer. As for S&W, when the first frame (frame 0) is sent, a timer starts to measure time until to wait to receive a possible NACK. If no NACK related to the first frame is received within timeout, frame 0 is automatically considered as correctly received by all the nodes and thus a new frame with index 0 can be sent since the number of frames that can be sent continuously are set to M. If a NACK has been received and it is related to the frame index ℓ, the transmitter resends all the frames already transmitted starting from the ℓ-th one. Also in this case the index numbering follows the $(.) \bmod M$ rule. From the receiver point of view, the index of the expected frame is initialized and, just in case the frame with the expected index has been received, it evaluates if errors are present. If the frame is error free, the receiver updates the index of the next expected frame. If errors are detected a NACK corresponding to the index of the error-affected frame is sent according to eq. (6.11). As for the S&W, if the frame index is not the expected one, the receiver only waits for the right frame index.
The energy mechanism used in the S&W is slightly changed in GBm since we should measure M different energies each related to the M possible different sub-bands. So the signal $y_F(t)$ is filtered on the m-th subchannel as:

$$y_F^{(m)}(t) = y_F(t) * z_m(t) \qquad (6.20)$$

---

[1]The ambiguity could by solved by allowing gaps between subsequent groups of M packets according to the propagation delay. However, including guard time windows lowers the frame transmission rate and hence the communication performance decreases more.
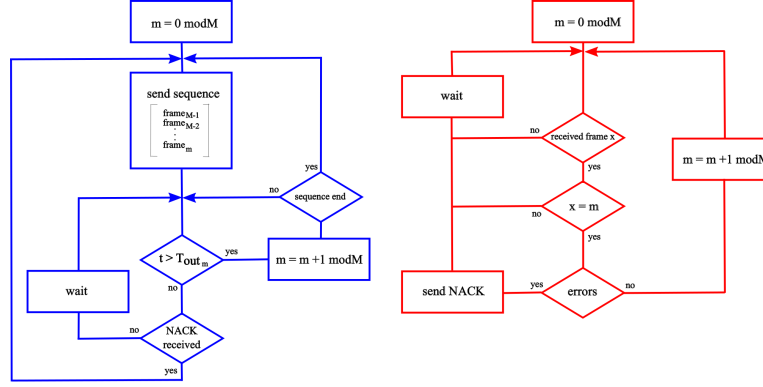
Figure 6.6: Go-Back-m ARQ flowchart. Blue and red diagram refer to transmitter and receiver respectively.

where $z_m(t)$ is a band-pass filter centered around $m\Delta f$, and then the energy is evaluated in the T_{out} time interval as:

$$E_{Y_F}^{(m)} = \int_{T_{out}} |y_F^{(m)}(t)|^2 dt \tag{6.21}$$

Also in this case a threshold mechanism is applied to infer about the presence or absence of the NACK. The main difference with S&W is that, in case of multiple NACKs detected on different frames, the transmitter resend all the frames starting from the oldest one, that is, the frame transmitted and that was not still considered as positively received.

When the channel is error free, the throughput efficiency is given as:

$$\mathcal{T}_{GBm_{ideal}} = \frac{ML(1-\rho)}{MT_e + 2\tau_{p_{\max}} + T_N} \quad [bit/s] \tag{6.22}$$

while in the presence of errors eq. (6.22) becomes:

$$\mathcal{T}_{GBm} = \frac{ML(1-\rho)}{MT_e + 2\tau_{p_{\max}} + T_N} (1 - FER_{GBm}) \quad [bit/s] \tag{6.23}$$

6.1.2.3 Selective Repeat

Selective Repeat also belongs to continuous ARQ family, but differently from GBm, it performs the retransmission only of corrupted frames thanks to the possibility of buffering the received frames. As reported by the frame-time diagram in Fig. 6.7, before using a frame index that has been already used, according to the $(.)\text{mod}M$ rule, the transmitter must wait the time out related to the previously transmitted frame with the same index. The length of the sequence to be sent must be chosen according to the same
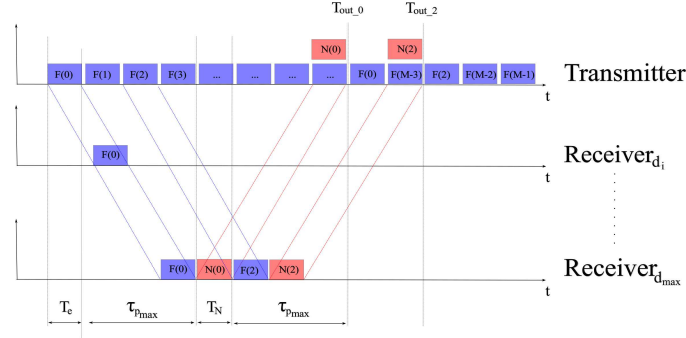
Figure 6.7: Selective Repeat ARQ temporal execution. Error occurrence at frame 0 and 2.

rule in eq. (6.19) previously introduced for GBm case, in order to avoid frames ambiguity as shown in Fig. 6.7. The throughput efficiency derives from eq. (6.16) and, taking into account the number of frames composing the sequence, it follows that:

$$\mathcal{T}_{SR_{ideal}} = \frac{ML(1-\rho)}{MT_e + 2\tau_{p_{\max}} + T_N} \quad [bit/s] \tag{6.24}$$

The mechanism for managing the SR ARQ procedure is essentially the same of GBm and it is reported in Fig. 6.8. The transmission proceeds and the flow is controlled on the basis of the time out so, when time out expires without the presence of a NACK it means that the frame index can be used without ambiguity. The only one difference is the way of managing the error events since, where misdetection occurs, the transmission is simply related to that, or those, frame(s) affected by errors. The mechanism for sending
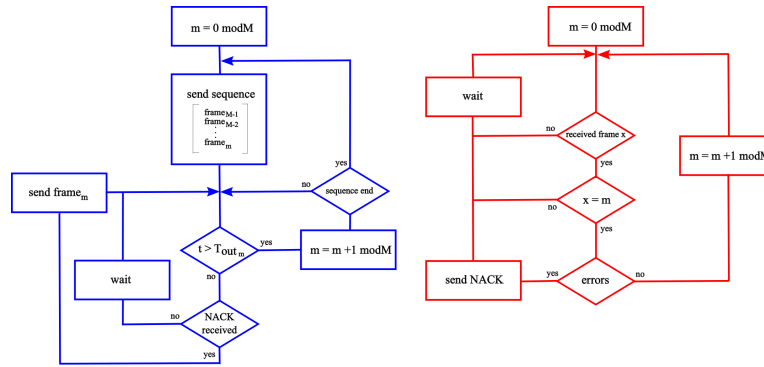
Figure 6.8: Selective Repeat ARQ flowchart. Blue and red diagram refer to transmitter and receiver respectively.

NACK and its detection is absolutely the same, while, in this case, in place of sending all the still unchecked frames, the transmission is limited to those that have been recognized, through NACK, as error affected. This allows to write the throughput efficiency in the following way:

$$\mathcal{T}_{SR} = \frac{ML(1 - \rho)}{MT_e + 2\tau_{p_{max}} + T_N}(1 - FER_{SR}) \quad [bit/s] \tag{6.25}$$

The expression in eq. (6.25) is equal to eq. (6.23) except for the FER. In fact in the SR scheme only the frame corresponding to the received NACK is recognized as error affected, while in GBm protocol all the frames that have been transmitted after the one recognized as affected by errors are considered as corrupted.

An interesting pictorial description of NACKs is that provided in Fig. 6.9 where the time-frequency representation is reported by the means of Wigner-Ville Transform, that is, the NACKs sprectrogram. In the horizontal axis it is possible to find time and on the vertical frequency, while the intensity is represented by the color ranging from blue to red. It is possible to appreciate
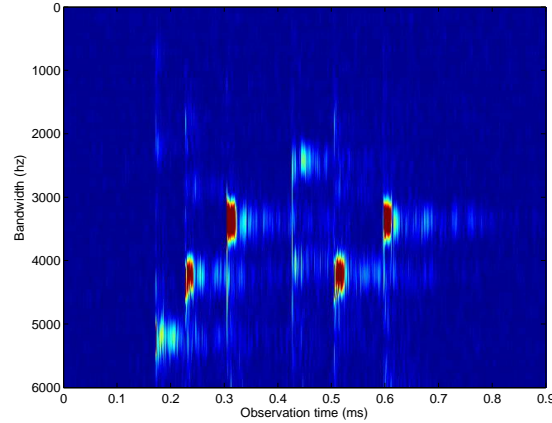
Figure 6.9: Representation of 7 NACKs sent in the water by different users.

7 different NACKs in SR and while the first two are sent by the same user it can be observed that other users (which are more distant from broadcaster) communicate (with higher delay) their NACKs. Hence it is inferred that frames 1 and 5 are detected with errors by a user close to the broadcaster. Frames 2 and 4 are detected with errors by two users with slightly different channels (different colors mean different attenuation). Frame 3 is affected by errors and communicated by a bit more distant user while two other users communicate their NACKs related to frames 2 and 5.

6.1.3 Numerical results

The results achievable with ARQ schemes are discussed by considering both computer simulations based on MATLAB software and real tests performed in a reservoir. The comparison between ARQ with FEC implemented with block coding (whose code rate is specified whenever needed) is also reported. The decoding procedure works according to the Maximum Likelihood (ML) criterion. Furthermore the effect of different modulation formats, namely PAM, PPM, OOK and QAM hyas been taken into account. Each node considered in the experiments was equipped with a AS-1 hydrophone and a SAM-1 modem. Specifically, as described in App. A, the pulse emitted by the SAM-1 modem lasts 406μs, so leading to have $\beta = 773.35$. Moreover, for each SAM-1 frame a factor $\rho = 0.2$ was considered.

In the computer simulations a noise-free pulse shape directly acquired by the SAM-1 modem (App. A) has been used. Using parameters taken from a real device is very useful since it allows a more truthful comparison between simulations and field experiments results. For what concerns the modeling of the channel, a set of 1200 channels acquired on a reservoir have been employed.

The signals received by the hydrophone have been filtered in order to eliminate the noise out of the transmission bandwidth. This gave also the possibility to appreciate the effect of delay spread, that is of the order of $40ms$. Furthermore, Table 6.2 reports the values of delay spread measured for different channel lengths. By looking at the results, it is evident that if the communication distance increases the number of paths grows, leading to a larger delay spread. Basing on those values, measuring the channel coherence bandwidth as the reciprocal of the delay spread lead to estimate a value that ranges from 15 Hz to 20 Hz. Since the transmission bandwidth is about 2.5 KHz around the carrier frequency of 33.8 kHz, the channel results far from being flat. This fact is clearly inferable by looking at the frequency response magnitude (Fig. 6.10) related to the one of the measured channels

Table 6.2: Signal delay spread for different communication links

Delay Spread (ms)	15m. Link	30m Link	60m Link	120m Link
Max.	67.7	75.5	80.7	86.7
min.	23.4	36.5	42.7	51.1
rms	49.6	55.1	63.6	64.5

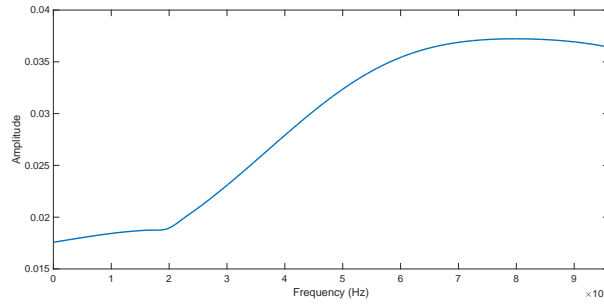The propagation environment induces also a signal Doppler shift. In fact,

Figure 6.10: Frequency response magnitude related to one of the measures channels.

while the main energy component of the transmitted pulse is centered around 33850 Hz, the peak of energy measured for the received signal is usually identified as shifted of few dozens of Hz. Furthermore, the channel coherence time can be calculated from Doppler spread since they are approximately inversely proportional to one another. Table 6.3 reports these features, measured for different communication links. The problem of Doppler shift

Table 6.3: Signal delay spread for different communication links

	15m Link	30m Link	60m Link	120m Link
rms Doppler spread (Hz)	54.1	96.9	73.9	88.1
Channel coherence time (ms)	18.5	10.3	13.5	11.4

has to be considered when organizing the feedback bandwidth, so basing on the values reported in Table 6.3 the ARQ system has been implemented by setting the width Δf of each feedback sub-bandwidth to 100 Hz. It is also worth highlighting that in absence of equalization, especially when dealing with PPM schemes, scarce channel coherence bandwidth and short coherence time do not necessary lead to a severe communication performance decrease. The matter is how much the channel varies rather than how often it does. In fact, if the main tap of the CIR remains almost the same, even if the other taps significantly changed the efficiency of signal detection would not remarkably decrease.

Last, the receiver adopted is the one already detailed par. 3.3.3. Using such a detection mechanism allows the system to be quite simple to implement. More complex receivers including channel equalization may be adopted so leading to better performing receivers. Threshold ϑ_E used for NACK detection is computed on the basis of the most distant node attenu-

ation $A(d_{max})$, leading to a value equal to $\vartheta_E = P/(3A(d_{max}))$ where P is the maximum transmission power by the network nodes, and the term 3 is chosen so as to privilege the minimization of missing detection events.

6.1.3.1   Computer simulation results

Refer to Fig. 6.11 that reports the aerial view of EUR Lake in Rome, reporting the bearing of transmitter T and receiver points A, B, C, D that are positioned at 15m, 30m, 60m, 120m, respectively. At first the performance

Figure 6.11: Map showing the transmitter location T and the receivers' locations, A, B, C, and D.

of the proposed ARQ scheme have been evaluated by considering the effect of β. 4-PAM modulation has been used and the receiver considered was the one placed at point C. Fig. 6.12 reports the throughput efficiency for S&W, GBm and SR. The main aspect to underline is that low values of β induces high FER values so a high number of retransmission is needed. This is essentially due to multipath inducing intersymbol interference so reducing the throughput. Hence, increasing β values leads to reduce the FER. However, this reflects also on increasing emission time so reducing the throughput efficiency. So the *optimal* value is in between. Moreover, as expected, SR outperforms GBm and S&W since with respect to GBm, SR does not retransmit all the frames of the sequence when an error occurred, while the gain offered with respect to S&W can be justified by the transmission of multiple frames (in this case set to M=4).

In Fig. 6.13 the performance have been evaluated considering the transmission of a 500 x 500 pixel image, black and white so each pixel is characterized by only one bit and the total dimension is 25000 bits plus 320 for the header, hence arriving at 25320 bits. The transmission was assumed 2-PAM with $\beta = 1.45$ according to the specifications of some modems in the literature as those produced by Evologics that are also full-duplex [58]. Receiver was placed at point B and the measured peak signal-to-noise ratio (PSNR) is equal to 11.44 dB. The image at the top left corner is the one
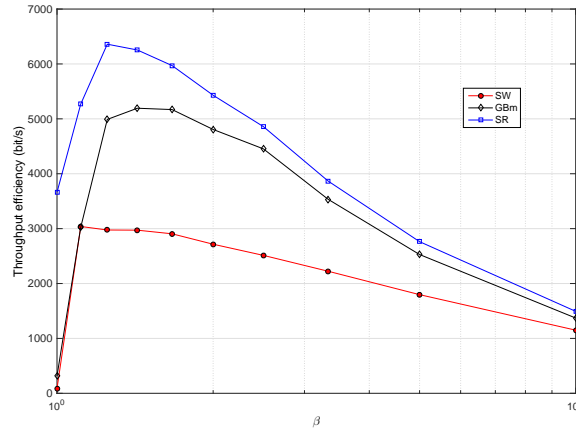
Figure 6.12: Throughput efficiency of S&W, GBm, and SR for 4-PAM at point C.

emitted by the source. It is also the one received by using the SR ARQ and FEC with a code rate of 1/4. It is important to note that the time needed to receive the image with the SR scheme is $6.25s$, while with the use of FEC we need $14.68s$ to complete the data transfer. The image at top right corner
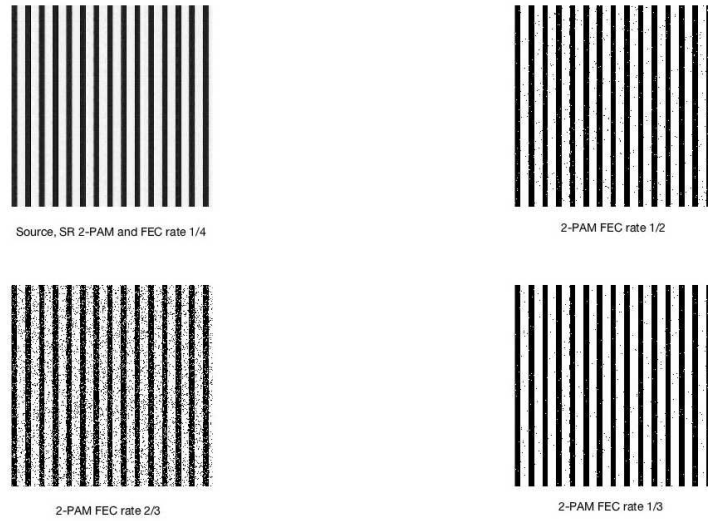

Figure 6.13: Effect of transmission of a 500 x 500 pixel image when SR is considered and also FEC with code rates 1/2, 2/3, 1/3 and 1/4. The measured PSNR is equal to 11.44 dB.

is the one received by using a code rate of 1/2 and it is affected by errors as evident. The time to complete the image transfer is $7.34s$. On the bottom left corner a FEC with code rate 2/3 is used. This is the fastest from the data transfer points of view since it requires $5.5s$ but, at the same time, it presents worse quality. Last, the image at right bottom corner is still (less) error affected and related to the use of a FEC with code rate 1/3. It takes $11s$ to be transferred.

Consider now the *transmission efficiency* as the ratio between the throughput efficiency as in eq. (6.17) and the transmission rate in eq. (6.5), formally defined as $\mathcal{T}_{S\&W}/R$ for S&W. The following simulations are based on all the parameters coming from SAM-1 modem. In Fig. 6.14 reports the simple case of unicast communication where the receiver node is placed in C (the multicast scenario leads to the same performance), in order to emphasize the role played by different modulations (the multicast scenario with respect to modulations will be detailed later). The modulations under investigation
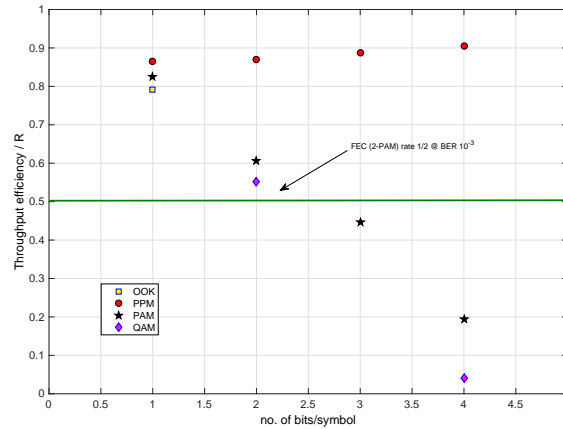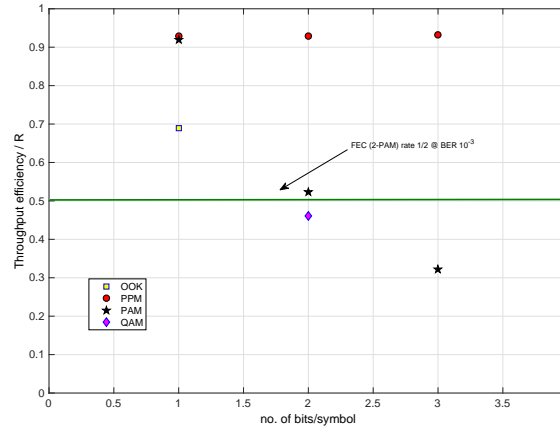

Figure 6.14: Ratio between throughput efficiency and transmission rate for S&W as a function of modulation format. FEC comparison is also included.

are OOK, PPM (from 2 to 16 symbols), PAM (from 2 to 16 symbols, that is 1 to 4 bits), 4-QAM and 16-QAM. The performance of S&W are detailed in terms of the ratio between the throughput efficiency reported in eq. (6.17)) and the transmission rate in eqs. (6.6)-(6.5) depending on the modulation format, orthogonal or non-orthogonal respectively. In other words, on the abscissa it is possible to see the number of bits per symbol (that is $\log_2 Q$) while the ordinate reports the above mentioned transmission efficiency that ranges from 0 to 1. It is possible to appreciate that the best performing is PPM. This result appears a bit counter-intuitive since PPM is bandwidth inefficient. This can be justified by observing that the denominator of eq.

(6.17) contains the term T_e that is the time needed to transmit a frame. So, this term decreases when R increases (the emission time decreases) with respect to propagation delay thus meaning that a lot of time is spent not for forward transmission but for waiting for possible NACKs. This effect reduces the throughput and as a consequence, the modulation formats as PAM or QAM, when the number of bits/symbol increases, become less transmission efficient. It is important to note that for PAM and QAM perfect synchronization is assumed.

The performance of OOK is in line with 2-PAM while QAM does not perform so good with respect to the others even though perfect synchronization here has been assumed. The above modulations jointly with ARQ lead to delivery information without errors, thus meaning that the file received is not affected by errors due to the retransmissions. The results carried out through the use of a FEC scheme, block coding, with code rate equal to $1/2$ are also shown. In that case the transmission efficiency is exactly given by the code rate. However, it is important to note that even though the line for FEC 2-PAM with code rate $1/2$ has an efficiency of 50%, on the other hand it is not error free since the BER is 10^{-3}. It is possible to appreciate that PPM, OOK and 2-PAM are preferable, in terms of efficiency, to FEC.

Moving to the use of the ARQ with the GBm strategy, the obtained results are depicted in Fig. 6.15. The efficiency of ARQ scheme is still
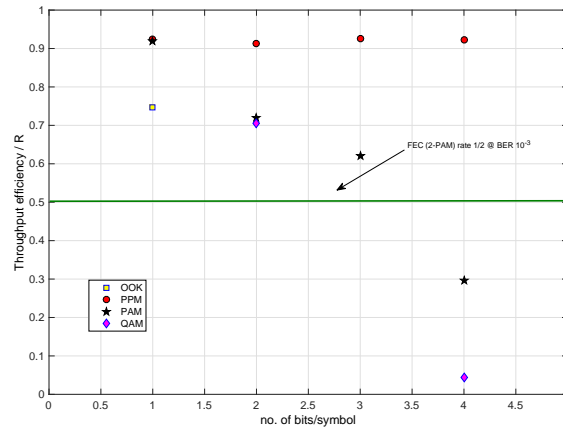

Figure 6.15: Ratio between throughput efficiency and transmission rate for GBm as a function of modulation format. FEC comparison is also included.

evaluated with respect to the modulation format and to the reference value for FEC 2-PAM with code rate $1/2$. Differently from to the S&W case, higher values of efficiency are reached and the reason is that the transmitter can send until to M consecutive frames (M=4 in this case). So the value

assumed by 16-PPM achieves 90% while the performance of 16-PAM and 16-QAM are worse than S&W. This happens since when a frame error occurs GBm requires the retransmission of all the still not acknowledged frames. However, both 4-PAM and 4-QAM present an efficiency that is more close to the line of the FEC. It is possible to highlight that also in this case PPM, OOK and 2-PAM are preferable, in terms of efficiency, to FEC.

Fig. 6.16 reports the same kind of performance analyzed until here for the SR case. While the hierarchy between modulations is the same (PPM is

Figure 6.16: Ratio between throughput efficiency and transmission rate for SR as a function of modulation format. FEC comparison is also included.

the best, 16-QAM the worst), it is interesting to note that the modulations that outperform the FEC case are now PPM, OOK, 2-PAM, 4-PAM, 4-QAM while 8-PAM is close even though worse than FEC. Also in this case the values are better than the previous ones since when an error occurs, the retransmission is limited only to the error affected frames without requesting the retransmission of frames that do not need to be retransmitted as for GBm.

6.1.3.2 Implementation and tests

As mentioned before, both broadcaster and receiving nodes are equipped with a hydrophone and a SAM-1 modem. The placement of transmitting and receiving device within the same node has been selected so that the radiation beam of the modem does not fall into the reception angle of the hydrophone. In this way the interference between the signals received by a node and that transmitted by the same own node is avoided. Furthermore, considering that the forward and feedback bandwidths have been conveniently spaced out and the pass-band filtering operations at the receiver

are well-performing, the potential inter-channel interference is completely removed. Basing on the features of the SAM-1 (see par. 3.3.2), 16-PPM modulation has been tested, but through the control of the source it has been possible to implement also 2,4,and 8-PPM as well as FEC. The other modulation formats, OOK, QAM and PAM, have been emulated following the procedure detailed in App. A.2. The scenario considered is that one already depicted in Fig. 6.11 with the transmitter in the point T of the map in and three receivers in the points B, C and D. The value of M is equal to 4. Table 6.4 reports the throughput efficiency in bit/s for the three ARQ schemes combined with the above considered modulation formats, a FEC scheme for each one of the modulation formats with code rate of 1/2 and 2/3. It is important noting that the data related to OOK, PAM and QAM are emulated while the data referred to PPM are measured on field tests. Moreover, concerning the simulations the channel realizations have been taken from recordings. By observing Table 6.4 it is possible to appreciate

Table 6.4: Throughput efficiency of different schemes including ARQ and FEC.

Modulation Format	S&W	GBm	SR	FEC (1/2)	FEC (2/3)
OOK	2.093	1.114	1.709	1.592	1.061
2-PPM	1.214	1.263	1.249	0.796	0.530
4-PPM	1.225	1.268	1.248	0.796	0.530
8-PPM	0.931	0.953	0.943	0.597	0.398
16-PPM	0.632	0.636	0.626	0.398	0.265
2-PAM	2.384	2.504	2.505	1.592	1.061
4-PAM	3.290	2.031	3.309	3.184	2.123
8-PAM	3.222	1.629	3.207	4.777	3.184
16-PAM	1.797	0.696	1.848	6.369	4.246
4-QAM	2.521	1.655	2.526	3.184	2.123
16-QAM	0.3939	0.211	0.411	6.369	4.246

that for OOK the best performance are guaranteed by S&W since the sporadic transmission of OOK well fits with the peculiarities of the S&W. For what concerns PPM, since this is a well performing modulation format in terms of error rate, GBm is preferable since low retransmissions are needed so the problem of retransmitting several frames (up to M) due to errors on one frame rarely occurs. When PAM/QAM is considered GBm suffers as previously described, so SR is the best performing among ARQ schemes. The performance of FEC are also shown and it is possible to appreciate how it outperforms ARQ for 16-PAM and 16-QAM when considering a code rate 1/2 and 2/3. However, even though error correction aims at reducing the number of errors, the residual error probability is in this case of the order of

10^{-2} that leads to conclude that the rate gain is not practical in real uses.

An additive performance measure of the proposed scheme is that related to false alarm (FA) versus missing detection (MD) rates of the proposed scheme with GBm, Juggling Stop-and-Wait (JSW) [59] and Underwater Selective Repeat (USR) [60]. Simulations results are reported in Table 6.5 and refer to the transmission of 1000 frames. It is important to underline that FAR measures the percentage of unnecessary retransmissions that lead to a throughput efficiency reduction, while MDR impacts on data integrity. The simulations consider the number of nodes in the network when 8-PAM is used and the maximum node distance is set to 120 meters. When only one node is considered all the systems compared do not present errors among the 1000 frame sent. This is essentially a unicast transmission where the effect of access is neglected. When the number of nodes is 2 the proposed scheme still does not exhibit errors while both JSW and USR start to perform worse. The reason is that JSW does not consider any form of multiple access so *frame collisions* (that is interference at the physical layer) affect ACK and NACK detection. The performance worsening of USR is justified by the CSMA protocol. In fact, even though the channel is sensed, this kind of protocol badly suites with underwater due to long propagation delay. For instance, it is possible that the channel is sensed as free by a node willing to send its ACK or NACK but, due to the propagation delay, it happens that an ACK or NACK has already been sent by another node. So both signals

Table 6.5: Comparison of FAR and MDR for the proposed ARQ scheme (GBm), JSW and USR.

no. of node	MDR			FAR		
	proposed	USR	JSW	proposed	USR	JSW
1	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$
2	$< 10^{-3}$	$6.2 \cdot 10^{-3}$	$7.1 \cdot 10^{-3}$	$< 10^{-3}$	$2.2 \cdot 10^{-3}$	$3.2 \cdot 10^{-3}$
3	$< 10^{-3}$	$1.4 \cdot 10^{-2}$	$3.6 \cdot 10^{-2}$	$< 10^{-3}$	$9 \cdot 10^{-3}$	$1.4 \cdot 10^{-2}$
4	$< 10^{-3}$	$1.1 \cdot 10^{-1}$	$2.3 \cdot 10^{-1}$	$< 10^{-3}$	$7.2 \cdot 10^{-2}$	$9.1 \cdot 10^{-2}$

overlap at the receiver (broadcaster). This effect becomes more important in terms of performance worsening when the number of user increases because the number of possible collisions grows. This is not true in the case of the proposed scheme since the ACK is not sent (so less feedback is sent) and when NACK is transmitted the possible overlapping of signals can be fruitful since it increases the energy detected.

6.2 Cognitive access for underwater acoustic communications

The cognitive paradigm aims at redefining a communication device so as to let it adapt its operating mode to the surrounding communication environment [61]. Among the various topics related to the cognitive approach, the problem of dynamic spectrum access is becoming a problem of paramount importance tackled by the scientific community [62][63][64][65]. The possibility to allow a cognitive node to opportunistically detect and exploit unused spectrum holes may be used in distributed systems such as sensors networks to devise a device-centric interference control as well as a distributed access technique. The former approach allows non licensed users (referred to as *secondary* users) to transmit over licensed bands unused by licensed nodes (referred to as *primary* users). The overall process consists of a preliminary spectrum sensing, where the unused spectrum holes are detected, and spectrum access, where the node tests whether the available detected spectrum is enough to meet its performance constraints. In the latter approach, no primary user is present and all the nodes are hierarchically equivalent, i.e., no preemption policy is envisaged. In such a scenario, the operations of spectrum sensing and access are performed in order to realize a distributed interference controlled medium access control. This framework is also suited to represent specific environments where few regulations or licensing procedures are present, such as for the case of underwater acoustic (UWA) communications.

While a huge amount of works have been devoted to cognitive access in wireless radio communication system [62]-[66], still only few contributions have been proposed to adapt the idea of cognitive access to an underwater acoustic communication system. Nonetheless the advantages of a cognitive access scheme could be profitably exploited in an UWA environment to both face the intrinsic acoustic spectrum scarcity and realize easily self-configurable sensor networks. As for the case of most of the techniques originally developed for radio communications, also the framework of cognitive access can be revisited in UWA and great attention should be paid to the intrinsic differences among the Hertzian and the acoustic medium in order to correctly recast the problem of cognitive access so that its solution results both feasible and efficient. A short survey on open issues and appealing perspective of applying the cognitive framework in UWA is described in [67].

The cognitive paradigm for underwater communication has been presented in [68] where different applications of the intelligence of the nodes in an underwater scenario are presented. The authors of [68] analyze the capacity gain attainable by resorting to either a dynamic or an opportunistic access in different propagation scenarios (deep, medium and shallow water).

In [69] the authors propose a Fast Fourier Transform (FFT) based cognitive access protocol to accommodate for spectrum sensing and spectrum access in UWA. The overall procedure is based on a FFT step to identify available and occupied channels through the use of an energy detection stage based on the assumption that the presence of a non negligible received power in a given channel stands as an indicator of an active transmission. In [70] a channel allocation procedure is devised relying on user location knowledge: in a nutshell, due to higher attenuation affecting high frequency signaling, the authors propose to assign high frequency sub-channels the pairs of nodes topologically closer to each other, leaving less attenuated sub-channels to users located at a larger distance. The preliminary identification of available channels is still performed by an energy detection stage.

In all the above referenced works, the task of spectrum sensing is performed under the assumption that whenever a non-negligible power level is received in a selected sub-channel, then a communication is supposed to be active on that very sub-channel. In other words, it is implicitly assumed that *each source of interference is equivalent to a communication taking place.* This conjecture is inherited from the cognitive radio framework, and it is completely justified in the specific application scenario of wireless radio communications, where electromagnetic interference is almost always due to communicating devices. On the other hand, the same assumption is, as a matter of fact, not completely true in underwater communications. It is in fact quite likely that the presence of received acoustic energy may *not* be originated by a transmitting device, as a lot of interfering sources such as mammals, artificial signals (engines), air bubbles, are present in the underwater environment.

Stemming on these considerations, it is easily understood how a direct rethinking of the cognitive access approaches originally designed for wireless radio systems in UWA may result in a non efficient access performance, as it will tend to avoid transmissions in available sub-channels because of the misinterpretation of the interfering source.

Such a consideration paves the way to a different vision for realizing a cognitive access when dealing with UWA communication systems. The role played by interference is different since its characteristics depend on the source nature (vessels, mammals) and, sometimes, cannot be predicted. This is the reason why several works in the literature analyzed the features of acoustic sources (see for example [71][72][73]) and other contributions aimed at mitigating or suppressing such interference (see [74][75][76]). However, estimation and tracking of interference, as well as its cancellation, may require a rate reduction for signal analysis and proper adaptive filter design.

This suggests that a more interesting mechanism can be used by the nodes without increasing excessively the processing required for the acous-

tic devices. The idea is to setup the underwater sensor network with nodes exhibiting the following properties [77]: awareness (about environment and application); ability to be smart (take decisions); learning (understand and classify signals); adaptivity (to different context specifications); reliability (so as to give to the service the required performances) and efficiency (utilizing the minimum necessary resources).

In particular, the following section discusses a new way to look at cognitive access in UWA communication systems by suitably coupling the energy detection step, usually envisaged in the spectrum sensing phase, with a classification step utilized to discriminate active transmissions from natural and artificial source of interference [78]. Such a classification phase will allow the cognitive nodes to understand not only *where* (in time and frequency domains) the interference is present, but also *the nature* of the interference prior to perform an access. In this regard, a test an image processing based procedure to detect and classify the presence of interfering sources in an underwater acoustic communication system is introduced. The procedure, whose first results based on correlation properties have been presented in [79], starts by analyzing the Wigner-Ville Transform (WVT) of the received signal by means of classical image processing tools. Specifically, an edge detection step is performed to detect the regions (in time and frequency) where an interfering signal is present. Successively, the detected pattern is analyzed to infer the nature of the emitting source. As the emission of different acoustic source will be characterized by different patterns in the WVT, the problem of inferring the nature of the interference is in turn recast as a problem of image classification.

6.2.1 Problem formulation

Consider an active sensor network where all the nodes (in a number of N_u) are able to sense the channel/interference and transmit in a time and frequency *grid* composed by time and frequency slots. This choice is justified by the possibly long delay spread of the channels that can induce performance degradation in terms of bit error rate. Generally speaking this may be compensated with equalization and or multi-carrier solutions. Channel equalization and Orthogonal Frequency Division Modulation (OFDM) are quite different. The first one suffers from imperfect estimates and its performance depends on the the unequalized taps of the channel. On the other hand OFDM suffers at transmission side if bit loading is performed and if the channel attenuation (usually represented the sub-channels attenuation) is poorly estimated [80]. However, this latter is more robust with respect to the case of single carrier modulation and channel equalization. The following discussion is presented considering the multi-carrier approach.

Assume the transmission to be slotted, thus meaning that the the access is organized in frames of time length T_f that is lower than coherence time,

so the channel can be considered time-invariant. Moreover, assume to have in a frame N_t time slots and N_f frequency sub-channels so the generic slot is identified in time and frequency domain as Time Frequency Slot (TFS). Before proceeding, it is important to anticipate two important features of this approach. First, no network synchronization is required, thus meaning that the only synchronization needed is between transmitter and receiver, so, in principle, different nodes can experience different interference profiles due to relative positions. Second, consider two different situations. One is related to the possibility of accessing on a single TFS within a frame This is justified by the need to evaluate the access probability. The other possibility is linked to the use of multiple TFSs and in the numerical results section the sum rate achievable by the network is shown.

In the most general case, the signal $y_k^{(j)}(t)$ received at the generic k-th cognitive node on the j-th (out of N_f) frequency sub-channel is given by the superimposition of two contributions. The first term accounts for the signals $s_u^{(j)}(t)$, received by the other (possible) active users on the same sub-channels, convolved with the channel impulse response $h_{u,k}^{(j)}(t)$ representing the communication channel between the generic u-th (out of N_u) active node and the k-th node on the j-th sub-channel. Generally speaking, the term $h_{u,k}^{(j)}(t)$ should be time-variant and written as $h_{u,k}^{(j)}(t, \tau)$, since dealing with the model related to the time-frequency slot, the dependency from τ can be omitted. The second term represents the emissions $\chi_m^{(j)}(t)$ coming from natural/artificial sources of interference possibly present in the propagation scenario; also in this case the model accounts for the effect of the channel $\eta_{m,k}^{(j)}(t)$ representing the generic link between the m-th (out of N_M) interference source and the k-th cognitive node on the j-th sub-channel. Stemming on this assumption $y_k^{(j)}(t)$ can be written as follows:

$$y_k^{(j)}(t) = \sum_{u=1}^{N_u} s_u^{(j)}(t) * h_{u,k}^{(j)}(t) + \sum_{m=1}^{N_M} \chi_m^{(j)}(t) * \eta_{m,k}^{(j)}(t) + n(t), \qquad (6.26)$$

where $n(t)$ represents the ambient/thermal noise that is generally colored on the whole bandwidth but usually white on the j-th frequency sub-channel. The overall signal $y_k(t)$ received by the k-th cognitive node can be then written as the sum of the $y_k^{(j)}(t)$ signals received in the different sub-channels, that is:

$$y_k(t) = \sum_{j=1}^{N_f} y_k^{(j)}(t). \qquad (6.27)$$

The signal model in eq. (6.26) is general and accounts both for the case where all the nodes employ the same numerical modulation technique, i.e., an OFDM/OFDMA scheme, and for the case where the nodes belong to different networks with different transmissions schemes. Specifically, the former

case refers to an application scenario where all the nodes are hierarchically at the same level (no primary users are present) while the latter may refer to the case of primary users employing different transmission techniques with respect to (w.r.t.) the secondary users.

Dealing with the effect of the acoustic interference, when the delay spread of channel related to the link from the interference source to reference receiver is limited to few TFS (along time), the interference does not spread heavily over the time so its influence is limited to one or few TFSs. On the other hand, when the acoustic sound propagates with a $\eta_{m,k}^{(j)}(t)$ presenting a long delay spread, the energy is spread among several TFSs and this, in turn, may affect the performance since it reduces accessing opportunities.

From the model in eq. (6.26), the amount of interference $I_k^{Rx}(i,j)$ sensed by the k-th user in the j-th sub-channel and during the i-th time slot (i.e., on the (i,j) TFS) accounts for both the interference due to active communications $I_{k,N_u}^{Rx}(i,j)$ and for the one due to natural/artificial interference $I_{k,N_M}^{Rx}(i,j)$, as well as for the noise power $N^{Rx}(i,j)$. Specifically, it can be written that:

$$I_k^{Rx}(i,j) = I_{k,N_u}^{Rx}(i,j) + I_{k,N_M}^{Rx}(i,j) + N^{Rx}(i,j). \tag{6.28}$$

The terms $I_{k,N_u}^{Rx}(i,j)$ and $I_{k,N_M}^{Rx}(i,j)$ are defined as follows, according to the model in eq. (6.26):

$$
\begin{aligned}
I_{k,N_u}^{Rx}(i,j) &= \sum_{u=1}^{N_u} P_u^{(i,j)} |h_{u,k}^{(i,j)}|^2 \\
I_{k,N_M}^{Rx}(i,j) &= \sum_{m=1}^{N_M} P_m^{(i,j)} |\eta_{m,k}^{(i,j)}|^2
\end{aligned}
\tag{6.29}
$$

where $P_u^{(i,j)}$ and $P_m^{(i,j)}$ are the powers of the signals $s_u(t)$ and $\chi_m(t)$ emitted respectively by the u-th user and the m-th source of interference and sensed by sensing node[2] in the TFS (i,j) while $h_{u,k}^{(i,j)}$ and $\eta_{m,k}^{(i,j)}$ are the channel gains of the (i,j)-th TFS measured at the k-th terminal and due to the u-th and m-th communication and natural interference sources respectively. About the channel model, several approaches are available in literature (see i.e. [81], [82]). It is important to stress, at this stage, that a conventional cognitive sensing (like that performed in all the cognitive radio systems as in [66]) only takes care of the whole $I_k^{Rx}(i,j)$ without distinguish if it is due to $I_{k,N_u}^{Rx}(i,j)$ or $I_{k,N_M}^{Rx}(i,j)$.

Consider a scenario where a pair of secondary nodes tries to access the medium. The access procedure is distributed so that there is no centralized

---

[2]The indexes (i,j) refer to the position of TFS from the sensing point of view under the implicit and obvious assumption that it is generated some time slots before.

controller as for conventional radio cognitive schemes. Besides, all the secondary nodes are considered at the same priority so that there are not nodes that are privileged for the access. Finally, if a node is already connected, i.e., it has already accessed the medium, it continues its transmission until the end (unless primary nodes start transmission in the same time and frequency slots). This last assumption assures that new secondary nodes cannot replace other secondary nodes already connected. It is important to highlight that a TFS used for communication (for example $i = 3, j = 5$ thus meaning 3rd time slot and 5th frequency slot) is the same within each frame till the end of communication.

The access problem formulation for the k-th node concerns the maximization of the access probability under the constraint that no interference is caused over other already active communications (either originated by a primary or a secondary user). A possible high level (of abstraction) problem may be:

$$\text{maximize} \quad \text{Pr}\{access\} \quad \text{s.t.} \quad coexistence.$$

Formally speaking the above problem can be described as:

$$\max_{P^{(k)}} \sum_{i,j} Pr\left\{C_{ij}^{(k)} > C^\star | i, j\right\} Pr\{i, j\} \tag{6.30a}$$

$$s.t. \ P_{i,j}^{(k)} \leq P_{max}(1 - D(I_{k,N_u}^{Rx}(i, j))), \tag{6.30b}$$

$Pr\{i, j\}$ being the probability of trying to access the (i, j)-th TFS. The goal is then to maximize the probability to perform the access in the TFS (i, j), that is the probability that the capacity $C_{ij}^{(k)}$ of the TFS (i, j) is over a performance, application dependent, minimum value $C^\star$. The maximization constraint takes care of both *i)* establishing a bound on the maximum transmitted power, by means of the constant P_{max} and *ii)* assuring coexistence among secondary and primary users which is, in turn, provided by the function $D(\cdot)$. This last is a binary indicator function defined assuming unitary value when its argument exceeds a given threshold, and zero value otherwise. At a first sight, the use of $D(\cdot)$ operator for managing coexistence may appear a limitation. However, since the access to the medium with coexistence constraints can be interpreted as a procedure that allows transmission if no harmful interference is produced against already established connections, each possible transmitting node has only two possibilities: transmit or not to transmit, that is, a difficult approach for managing interference.

Now, by assuming that no one of the $N_t N_f$ possible TFSs is privileged for trying to access with respect to others, a uniform distribution for $Pr\{i, j\}$ can be posed as:

$$Pr\{i, j\} = \frac{1}{N_f N_t} \quad i = 1, ..., N_t, \ j = 1, ..., N_f \tag{6.31}$$

so that the problem in eq. (6.30) can be rewritten as:

$$\max_{P^{(k)}} \frac{1}{N_f N_t} \sum_{i,j} Pr\left\{C_{ij}^{(k)} > C^\star | i, j\right\} \tag{6.32a}$$

$$s.t. \ P_{i,j}^{(k)} \leq P_{max}(1 - D(I_{k,N_u}^{Rx}(i,j))). \tag{6.32b}$$

Since focusing on the probability of accessing the medium on the (i, j)-th TFS, the maximum for all i, j must be found, so rewriting the problem in the following shape leads to:

$$\max_{P_{i,j}^{(k)}} Pr\left\{\log_2\left(1 + \frac{P_{i,j}^{(k)}|h_{k^{Tx},k^{Rx}}^{(i,j)}|^2}{\Gamma(\mathcal{N}_0 B + I_k^{Rx}(i,j))}\right) > C^\star\right\} \quad \forall i, j \tag{6.33a}$$

$$s.t. \ P_{i,j}^{(k)} \leq P_{max}(1 - D(I_{k,N_u}^{Rx}(i,j))). \tag{6.33b}$$

About the above problem, $h_{k^{Tx},k^{Rx}}^{(i,j)}$ is the channel modeling the link from the k-th transmitter to k-th receiver by using the (i, j)-th TFS, $\mathcal{N}_0 B$ is the noise power, while $I_k^{Rx}(i, j)$ is the measured interference, and Γ is the gap approximation to guarantee a target BER [83] starting from the required Symbol Error Rate (SER), that can be expressed as:

$$\Gamma = \frac{1}{3}\left[Q^{-1}\left(\frac{SER}{4}\right)\right]^2, \tag{6.34}$$

with the position:

$$Q(x) = \int_x^\infty \frac{e^{-\zeta^2/2}}{\sqrt{2\pi}} d\zeta. \tag{6.35}$$

It is worth noting that depending on the channel conditions and impairments, Doppler, delay spread and so forth, the above model may not suffice to account for all these disturbing effects. However, for TFS with bandwidth within the coherence band, formulation as in eq. (6.33) holds. Moreover, for imperfect synchronization and/or channel estimation, setting higher values of Γ can partially compensate the loss, even though it may lead to lose transmission opportunity since the maximum power does not guarantee the minimum capacity.

Once chosen a TFS (namely $j^\star$) by the U-th user (the mechanism for selecting a channel will be detailed in the implementation issues section), the link model is:

$$y_k^{(j^\star)}(t) = s_U^{(j^\star)}(t) * h_{U,k}^{(j^\star)}(t) + \sum_{m=1}^{N_M} \chi_m^{(j^\star)}(t) * \eta_{m,k}^{(j^\star)}(t) + n(t). \tag{6.36}$$

The conventional optimization procedures, coming from the cognitive radio field in the framework of wireless networks, are based on forbidding the

transmission whenever a non negligible interference is present so as to assure coexistence among primary and secondary users. The assumption beyond this approach is that when (relevant) interference is sensed over a TFS, then a communication is supposed active on that TFS. This, in turn, is equivalent to consider the problem in eq. (6.33) with the following constraint:

$$P_{i,j}^{(k)} \leq P_{max}(1 - D(I_k^{Rx}(i,j))).$$

As previously anticipated, the framework of underwater cognitive network is a bit different since it is possible to experience interference even if it is not due to communication sources. This leads to observe that, differently from conventional cognitive optimization tools, the constraint on the optimization problem in eq. (6.33) should be correctly considered by evaluating the indicator function $D(\cdot)$ only on the portion of sensed interference actually due to active communications:

$$P_{i,j}^{(k)} \leq P_{max}(1 - D(I_{k,N_u}^{Rx}(i,j))).$$

As it will be clear by the numerical results , applying a conventional sensing like those in [84]-[85] may lead to lower the access opportunities as some time-frequency slots, in principle available for transmission, are considered as unavailable.

Solving the problem in eq. (6.33) corresponds to maximize the signal-to-interference-noise-ratio (SINR) without inducing interference. In the cognitive framework two are the possible choices. A node can transmit in a TFS at full power (so maximizing SINR, that is capacity) or avoid to transmit since the channel is heavily attenuated and/or too much interference is present. Transmitting when interference is high means, in principle, to induce interference on a receiver. Since, as previously specified not all the interference is due to an active communication, a node can decide to transmit if it *recognizes* the interference is not due to a communication.

This suggests that a mechanism able to understand if data communication (interference) is active on each TFS can effectively maximize access probability thus avoiding to violate coexistence (transmission operated on channels already used for communication purposes) and missing transmission opportunities (transmission avoided on some TFSs since natural/artificial interference is present). So, implicitly, the maximization problem can be solved by recognizing if a transmission is active or not on a specific TFS.

As a result, the key point in the approach herein described is to discriminate to which extent the contribution due to active communications $I_{k,N_u}^{Rx}(i,j)$ affects the interference $I_k^{Rx}(i,j)$ sensed within a time slot. Under this vision, the access problem can be reshaped as the estimation of the contribution of $I_{k,N_u}^{Rx}(i,j)$ starting from its noisy observation $I_k^{Rx}(i,j)$ according to the model in eq. (6.28). Different estimation criteria can be exploited:

for instance, in the case of ML estimation one would obtain the estimate $\hat{I}_{k,N_u}^{Rx}(i,j)$ of $I_{k,N_u}^{Rx}(i,j)$ as:

$$\hat{I}_{k,N_u}^{Rx}(i,j) = \arg \max_{I_{k,N_u}^{Rx}(i,j)} p\left(I_{k,N_u}^{Rx}(i,j)|I_k^{Rx}(i,j)\right), \qquad (6.37)$$

$p\left(I_{k,N_u}^{Rx}(i,j)|I_k^{Rx}(i,j)\right)$ being the pdf of the interference generated by communications sources conditioned to the value of measured interference $I_k^{Rx}(i,j)$. A complementary procedure would encompass the estimation of $I_{k,N_M}^{Rx}(i,j)$, starting from $I_k^{Rx}(i,j)$ and thus obtaining $\hat{I}_{k,N_u}^{Rx}(i,j)$ from the model in (6.28) according to the following relationship:

$$\hat{I}_{k,N_M}^{Rx}(i,j) = \arg \max_{I_{k,N_M}^{Rx}(i,j)} p\left(I_{k,N_M}^{Rx}(i,j)|I_k^{Rx}(i,j)\right). \qquad (6.38)$$

Once obtained a reliable estimate of $I_{k,N_u}^{Rx}(i,j)$, the problem in eq. (6.33) could be solved. Unfortunately, due to the additive relation in eq. (6.28) comprising the noisy superimposition of two components, both the ML problems in eq. (6.37) and eq. (6.38) would result in biased estimates for the quantity of interest, so that their value could not be employed in solving eq. (6.33).[3]

To overcome this problem, a different approach is proposed. The idea of understating the nature of the interference falls in the binary hypothesis testing problem since, for each TFS, the interference level can be due to acoustic natural/artificial sources (noise of different kinds) or to a communication active in the network. This implicitly suggests that possible merit figures to evaluate the performance can be the false alarm probability as well as the missing detection one. In order to proceed with interference nature understanding, a 2-dimension (time and frequency) description of the interference by using the WVT is exploited, so as to obtain an image and move from an estimation problem to a classification one.

6.2.2 WVT Based Interference Discovery And Classification

A possible way to analyze the nature of the interference is observing the time and frequency properties of the interfering signals. The use of the WVT [86][87][88] allows to show the spectral components of a signal with respect to the time variable and therefore to have a two-dimensional description of the perceived interference available. According to [86], the expression for the WVT of the signal in eq. (6.26) is given by:

$$Y_k^{(WV)}(u,\xi) = \int_{T_f} y_k(u + \frac{\vartheta}{2})y_k^*(u - \frac{\vartheta}{2})e^{-j\vartheta\xi}d\vartheta. \qquad (6.39)$$

---

[3]In fact, referring for the sake of simplicity only ot the problem in eq. (6.37), the estimate $\hat{I}_{k,N_u}^{Rx}(i,j)$ will exhibit a bias affected by the unknown value of $I_{k,N_M}^{Rx}(i,j)$.

The WVT application returns a snapshot of the energy sensed over a given bandwidth for a selected number of consecutive time intervals short enough to assume stationarity of the received signal. Such a time/frequency representation turns quite useful when dealing with cognitive access scheme. Consider, as an example, a scenario encompassing a primary (or secondary) user accessing a selected portion of the available spectrum in a TDMA fashion; the WVT allows to promptly identify the periodicity of the emission while an energy-detection based spectrum sensing, performed only during a single time interval will just detect the presence of a transmission on the said bandwidth, thus failing to recognize such spectrum portion is not occupied for the whole frame duration, but only for a given time slot. Besides, because of its intrinsic bi-dimensional nature, known tools from image processing community can be exploited in support of WVT.

Possible hardware tools for implementing such a scheme are the so-called FFT chips that, time by time, output the FFT samples, allowing to store different FFT vectors composing the WVT image at restrained computational complexity.

In this framework, the WVT provides the receiver with an image of the perceived interference that can be further processed to *extract* significant information. To elaborate, once the k-th cognitive node has sensed the medium for a given time interval, it can proceed to evaluate, through a *double sampling*, the time and frequency samples $Y_k^{(WV)}(m,n)$ of $Y_k^{(WV)}(u,\xi)$ by simply applying FFT at each time sample interval. Successively a partitioning in $N_t N_f$ slots is operated so as to obtain $N_t N_f$ sub-images $Y_{(i,j),k}^{(WV)}(m,n)$, each one pertaining to a different (i,j) TFS.

A classical cognitive access scheme considers the level of interference and decides if accessing the medium or not by basing on a threshold θ that in general is difficult to set since some a priori information, such as the channel impulse response and its attenuation, may be unavailable. Within the WVT based framework herein described, such an approach is equivalent to compare the values of each sub-image $Y_{(i,j),k}^{(WV)}(m,n)$ with a threshold θ so as to individuate the free slots and the occupied ones. According to a thresholding procedure, the time frequency slots available for transmission are those respecting the condition on the threshold θ:

$$\sum_m \sum_n Y_{(i,j),k}^{(WV)}(m,n) < \theta. \tag{6.40}$$

Failing to meet this relationship means that the TFS is interfered. As described in detail above, such a process may be mislead, in the underwater scenario, by the fact that a perceived high energy level is not always originated by an active transmission.

In order to understand if the interference is due to an active communication, an image classification procedure is employed. Interestingly, using

image classification to understand which kind of interference faced in the current TFS does not require to tune the above mentioned threshold θ. In fact, in image classification, an image is classified according to its visual content, for example, if it contains a particular object or not. An important application is image retrieval, that is searching through an image dataset to obtain (or retrieve) those images with particular visual content. Such an approach perfectly matches with the problem at hand of determining whether an interference is due to a communication or to a natural/artificial source (as explained in the previous paragraphs, understanding if the interference contained in the image is due to a communication or not, can increase the number of access opportunities).

Different image classification techniques can be envisaged to solve the problem at hand. Below, two possible schemes are considered. The first approach, named *Natural/Artificial Recognition Classifier* (NARC), aims at detecting the presence of a natural interference, while the second, referred to as *Data Communication Recognition Classifier* (DCRC), is built for detecting the presence of active communications. Both of the schemes share a quite conventional features extraction process [89] and rely on a classical Support Vector Machine based classifier [90].

In the NARC a training dataset containing typical sub-images representing the WVT of interference due to mammals, ships and so forth is required and it tackles the problem in eq. (6.38)). The DCRC approach tackles the problem in eq. (6.37) and requires as training set a collection of sub-images arising from typical transmission techniques (single carrier or multi-carrier).

Before a brief description of the classification techniques herein employed, it is worth remarking that in line of principle, any procedure driven to discriminate among interference due to communications and to natural/artificial interference can be put to good use in this framework. The approach here employed stands just as a possible realization of a more general operative procedure aimed at discriminating the sensed interference. The classification procedure herein employed consists essentially of 3 different stages, where the first two are performed *off-line* relying on selected recordings of data transmissions or interference sounds.

6.2.3 Training set preparation and features extraction

The training data consist of N_{ds} bi-dimensional sequences (images) of size $X_1 \times X_2$ for each category related to typical sounds (mammal emissions, engine sound, ice-breaking, earthquake for NARC and OFDM/single carrier for DCRC). The feature vector extracted by each sequence in the training set consists of Scale-Invariant Feature Transform (SIFT) (see [89]) computed on a regular grid across the image ('dense SIFT') and vector quantized into visual words. The frequency of each visual word is then recorded in a histogram for each tile of a spatial tiling as shown in the Fig. 6.17. The final

feature vector for the image is a concatenation of these histograms. This process is summarized in Fig. 6.17. For more details about the analysis of SIFT refer to [89]. Here only its core operations are reviewed. Specifically the SIFT procedure is based on the following steps:

1. *Scale-space extrema detection.* The first stage of computation searches over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian functions to identify potential interest points that are invariant to scale and orientation.

2. *Keypoint localization.* At each candidate location, a detailed model is fit to determine location and scale. Keypoints are selected based on measures of their stability.

3. *Orientation assignment.* One or more orientations are assigned to each keypoint location based on local image gradient directions. All future operations are performed on image data that have been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations.

4. *Keypoint descriptor.* The local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation that allows significant levels of local shape distortion and change in illumination.
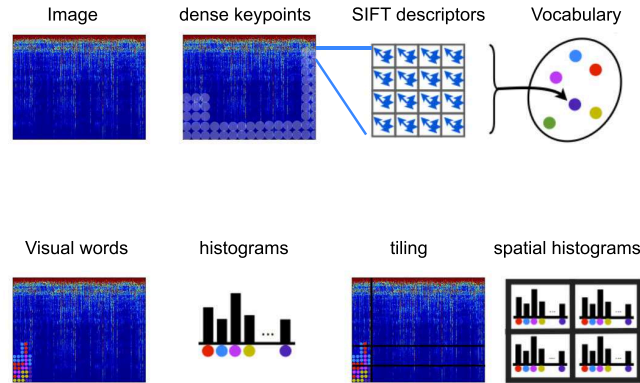
Figure 6.17: Different steps for image classification.

About the first stage, i.e., scale-space extrema detection, it is aimed at detecting keypoints by using a cascade filtering approach that utilizes efficient

algorithms to identify candidate locations that are then examined in further detail. The keypoint detection step is to identify locations and scales that can be repeatedly assigned under differing views of the same object. Detecting locations that are invariant to scale change of the image can be accomplished by searching for stable features across all possible scales, using a continuous function of scale known as scale space. Starting from the original image $Y_{(i,j),k}^{(WV)}(m,n)$, a Gaussian-filtered one can be defined as:

$$L(m,n,\rho) = G(m,n,\rho) \otimes Y_{(i,j),k}^{(WV)}(m,n), \qquad (6.41)$$

with

$$G(m,n,\rho) = \frac{1}{2\pi\rho^2} e^{-(m^2+n^2)/2\rho^2}. \qquad (6.42)$$

and $\otimes$ being the two-dimensional convolution operator. To efficiently detect stable keypoint locations in scale space, Lowe [89] proposed using scale-space extrema in the difference-of-Gaussian function convolved with the image, $\Delta(m,n,\rho)$, which can be computed from the difference of two nearby scales separated by a constant multiplicative factor c:

$$\Delta(m,n,\rho) = (G(m,n,c\rho) - G(m,n,\rho)) \otimes I(m,n)$$

$$= L(m,n,c\rho) - L(m,n,\rho). \qquad (6.43)$$

In order to detect the local maxima and minima of $\Delta(m,n,\rho)$, each sample point is compared to its eight neighbors in the current image and nine neighbors in the scale above and below. It is selected only if it is larger than all of these neighbors or smaller than all of them. The cost of this check is reasonably low due to the fact that most sample points will be eliminated following the first few checks.

Regarding keypoint localization step, once a keypoint candidate has been found by comparing a pixel to its neighbors, the next step is to perform a detailed fit to the nearby data for location, scale, and ratio of principal curvatures. This information allows points to be rejected that have low contrast (and are therefore sensitive to noise) or are poorly localized along an edge.

Dealing with orientation assignment stage, by assigning a consistent orientation to each keypoint based on local image properties, this latter can be represented relative to this orientation and therefore achieve invariance to image rotation. Based on a set of experimentations with a number of procedures to assign a local orientation, the following approach was found to give the most stable results. The scale of the keypoint is used to select the Gaussian smoothed image, L, with the closest scale, so that all computations are performed in a scale-invariant manner. For each image sample $L(m,n)$ at this scale, the gradient magnitude $\Omega(m,n)$ and orientation $\Theta(m,n)$, are

precomputed using pixel differences as shown below:

$$\Omega(m,n) = \sqrt{(L(m+1,n) - L(m-1,n))^2 + (L(m,n+1) - L(m,n-1))^2}$$
$$\tag{6.44}$$

$$\Theta(m,n) = \tan^{-1}\left(\frac{L(m,n+1) - L(m,n-1)}{L(m+1,n) - L(m-1,n)}\right) \tag{6.45}$$

Finally, in regards to keypoint descriptor, the image gradient magnitudes and orientations are sampled around the keypoint location, using the scale of the keypoint to select the level of Gaussian blur for the image. To achieve orientation invariance, the coordinates of the descriptor and the gradient orientations are rotated relative to the keypoint orientation.

To move from SIFT description to vocabulary creation, it has been used the vector quantization (VQ) technique which clusters the keypoint descriptors in their feature space into a (usually) large number of clusters using the K-means clustering algorithm and encodes each keypoint by the index of the cluster to which it belongs (see [91] for a detailed description of the procedure). In the case under investigation the feature vector contains 24 elements. Each cluster is described as a visual word that represents a specific local pattern shared by the keypoints in that cluster. Thus, the clustering process generates a visual word vocabulary describing different local patterns in images. The number of clusters determines the size of the vocabulary, which can vary from hundreds to over tens of thousands (in this approach this number is limited to few tens for DCRC and few hundreds for NARC). Mapping the keypoints to visual words, each image can be represented as a group of visual words. This last representation can be converted into a visual word vector similar to the term vector of a document (see [91]). The visual word vector may contain the presence or absence information of each visual word in the image, the count of each visual word (i.e., the number of keypoints in the corresponding cluster), or the count weighted by other factors. Then by basing on occurrence of features a histogram $\mathbf{x}$ is built.

6.2.3.1 Training Stage

The interference training images will be used as the positives, and the background images (only noise) as the negatives. The classifier is a linear Support Vector Machine (SVM) (see [90]). This classifier, for this application, present low complexity and, as it appears in the numerical results section, presents reliable performance. Starting from the definition of the set composed by $(\mathbf{x}_i, A_i)$ gathering the examples $\mathbf{x}_i \in \mathcal{R}^g$, g being the dimension of the input space and $\mathbf{x}_i$ the histogram vector belonging to the class $\mathcal{A} = \{1, -1\}$ (or equivalent $\mathcal{A} = \{\mathcal{A}+, \mathcal{A}-\}$. Generally speaking, the goal is to define a hyperplane which divides the set of examples such that all the points with the same label are on the same side of the hyperplane. This is equivalent to

looking for $\boldsymbol{\omega}$ and γ values so that:

$$\mathcal{A}_i(\boldsymbol{\omega}'\mathbf{x}_i + \gamma) > 0, \quad i = 1, ..., N_{ds}N_{cat} \tag{6.46}$$

Under the assumption, in this particular case of linearly separable set, it is always possible to rescale $\boldsymbol{\omega}$ and γ so to have:

$$\min_{1 \leq i \leq N_{ds}N_{cat}} \mathcal{A}_i(\boldsymbol{\omega}'\mathbf{x}_i + \gamma) \geq 1, \quad i = 1, ..., N_{ds}N_{cat} \tag{6.47}$$

so that the distance between the closest point to the hyperplane is $1/||\boldsymbol{\omega}||$, where n_{cat} is the number of images categories. Among the separating hyperplanes, the one for which the distance to the closest point is maximal is called separating surface or optimal separating hyperplane (OSH). Since the distance to the closest point is $1/||\boldsymbol{\omega}||$, finding the OSH amounts to minimizing $||\boldsymbol{\omega}||^2$ under the constraint $\mathcal{A}_i(\boldsymbol{\omega}'\mathbf{x}_i + \gamma) \geq 1, \quad i = 1, ..., N_{ds}N_{cat}$. The quantity $2/||\boldsymbol{\omega}||$ conventionally called the margin, and thus the OSH is the separating hyperplane which maximizes the margin. The margin can be seen as a measure of the generalization ability: the larger the margin, the better the generalization is expected to be [92][93]. Since $||\boldsymbol{\omega}||^2$ is convex, minimizing it under linear constraints shown above, can be achieved with Lagrange multipliers.

Being $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_{N_{ds}N_{cat}})$ the Lagrange multipliers associated with constraints above, the optimization problem amounts to maximizing

$$W(\boldsymbol{\alpha}) = \sum_{i=1}^{N_{ds}N_{cat}} \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j \mathcal{A}_i \mathcal{A}_j \mathbf{x}_i' \mathbf{x}_j \tag{6.48}$$

with $\alpha_i > 0$ and under constraint $\sum_{i=1}^{N_{ds}N_{cat}} \alpha_i \mathcal{A}_i = 0$. This can be achieved by the use of standard quadratic programming methods [94].

Once the vector $\boldsymbol{\alpha}^\star = (\alpha_1^\star, ..., \alpha_{N_{ds}N_{cat}}^\star)$ solution of the maximization problem has been found, the hyperplane decision function can thus be written asN

$$f(\mathbf{x}) = \text{sgn}\left(\sum_{i=1}^{N_{ds}N_{cat}} \alpha_i^\star \mathcal{A}_i \mathbf{x}_i' \mathbf{x} + \gamma\right) \tag{6.49}$$

The scheme in Fig. 6.18 well shown the behavior (on a two dimension space) of classification and the geometrical definition of margin, OSH.

6.2.3.2 Testing Stage

After a signal acquisition, the one-dimensional signal is transformed via the WVT and then a TFS partitioning is obtained via image segmentation. The first two steps detailed before are applied to each of the image representing the TFS in a frame in order to classify the interference over each frame, so the above mentioned process (including SIFT, visual words organization,
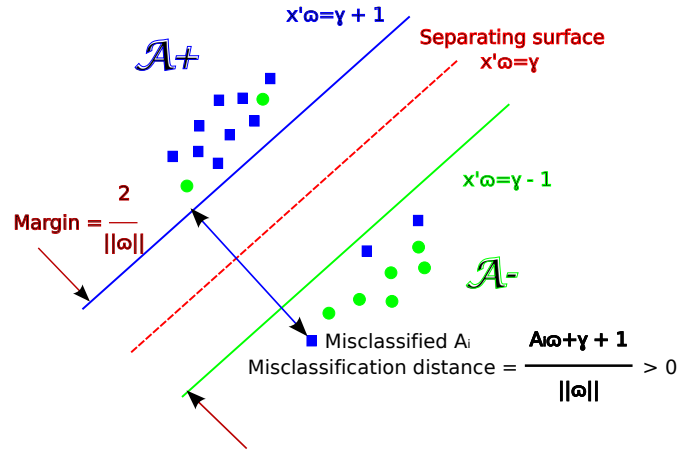
Figure 6.18: Diagram showing the classification geometry.

histogram and tiling) is run newly in order to compare it with the data already stored. This is done through a neural network.

In this regard, as above anticipated, two different approaches have been considered. NARC requires several typical underwater disturbances are stored in terms of imaged (WVT). These images gather interference signals as bubbles, earthquakes, ice-breaking, mammals, ship engines and so forth. This suggests, implicitly, that a memory should be dedicated to this kind of storage and, more, a wide class of interference signals must be taken into account for an accurate classification. DCRC has stored a smaller set of signals that are the typical used for communication (single-carrier, multi-carrier and sonar-like, this latter is not communication in strict sense but it is generated for acquiring some information). Even though some of the above mentioned noise sources can be considered as background noise, several noise components are however evaluated in order to perform a deep analysis and resort to a more complete and realistic scenario.

By looking at Fig. 6.18, it is easy to note that a misclassification is possible, that is, a TFS is classified with a specific kind of interference where this is absent. In details, for NARC the $\mathcal{A}+$ space represents the natural/artificial interference while in the DCRC the data communication is in $\mathcal{A}+$. When a misclassification occurs, a false alarm or missing detection event happens. It is important to stress that in the NARC it is not mandatory to recognize if the interference is due to whales or ships (or other natural/artificial sources) since the only one goal is to separate (from a classifier point of view) the communication sources from other interfering ones. It is not the same for the DCRC case.

6.2.3.3 On the Classification of Unknown Signals

At this stage, a discussion about how the approach herein presented deals with unknown signals is in order. In fact, during the online classification procedure a cognitive node may happen to sense a signal that has not be accounted for in the construction of the training set. Of course, no matter how large is the employed training set, during online operation there will always be the chance to sense a signal whose nature is unknown. Besides, it is important to stress that also a signal that, in line of principle, is known, that is a signal who has been included in the training set, may be affected by the channel to such an extent that the classification procedure may fail to recognize it. Both the cases of unknown signals/interferers and of mis-detected signals due to channel impairments affect the classification performance as an increased probability of miss/probability of false alarm, depending on which strategy (NARC/DCRC) is employed. For instance, an unknown natural interference will not be recognized by the NARC; as a consequence the unknown signal will be erroneously classified as an active communication thus resulting in a less favorable percentage of access. On the contrary, an unknown communication signal will not be recognized by the DCRC, resulting in a possible coexistence issue.

Summing up, the more complete the training set for the NARC/DCRC, the higher the performance of the access schemes. Of course a suitable trade-off between computational complexity for classifier training and expected performance has to be correctly taken in to account when designing the samples to be included in the training set.

6.2.3.4 Implementation issues: computational cost and protocol

Regarding the requirements at different levels needed for implementing the proposed scheme, it is important to underline that both transmitter and receiver (in the following transceiver since a node can operate as transmitter and receiver), should be able to transmit on different sub-channels (OFDM-like) and sampling the received (sensed) interference with a sampling time of the order of hundreds on μs. The network synchronization is not required since the only synchronization needed is the one between cognitive pairs willing to communicate each other. About the cost of testing images for evaluating if communication or other sources are operating, it is given by $X_1 X_2 \log_2 X_2 + |\mathcal{S}_d|$, X_1 being the time sub-interval during which the sensing is operated, X_2 the number of points representing FFT (recall that WVT can be performed as short-term FFT), while $|\mathcal{S}_d|$ is the cardinality of the set $\mathcal{S}_d$ collecting the SIFT descriptor. In the proposed scheme $X_1 = 50$, $X_2 = 50$, $|\mathcal{S}_d| = 24$ for the single image. So by considering CPU processing times nowadays available, such a computational burden is in-line with the low bit-rate requested, that is, the processing time is several order of magni-
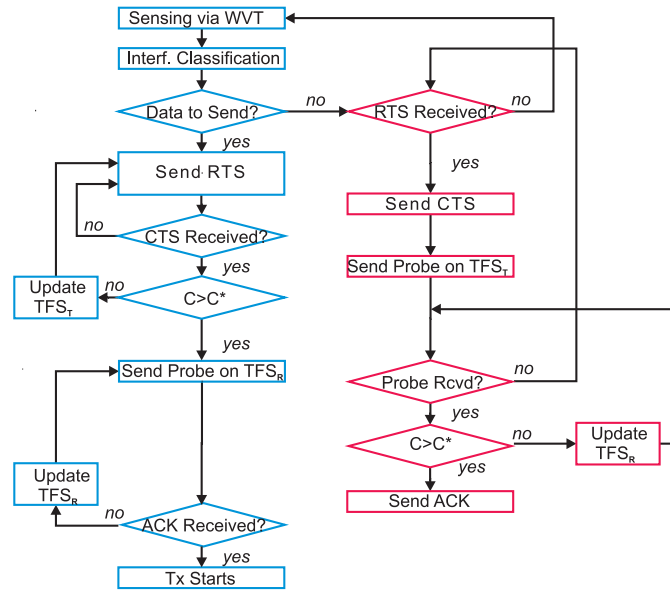
Figure 6.19: Proposed access protocol representing a modified version of [95].

tudes lower than the transmission one. Possible firmware updates may also increase, during maintenance, the dimension of the training set by including new interference images so as to reduce miss-classification probability.

In Fig. 6.19 a reference protocol that can be implemented by the transceiver is reported by proposing a modified version of that described in [95] where Prepare-to-Send (PTS) / Request-to-Send (RTS) / Clear-to-Sent (CTS) mechanism has been adopted in an orthogonal frequency-based access by applying sensing on all the sub-channels. More, the access is operated on the free slot(s) with the receiver that must decode all the streams to understand if it is the destination of the packet. The modified version here proposed considers a two-step communication protocol in order to avoid to occupy the channel (for signaling purposes), it uses a common (separated) channel for RTS/CTS mechanism and, more, it indicates the TFS where to probe the channel since it is not granted, in general, that an available TFS presents low attenuation. This aspect has been not taken into account in [95]. The left part of the Fig. 6.19 refers to the transmitting functions of the transceiver, while the right part to the receiving ones. The first step is sensing and interference classification, that is the main goal of this work. If the cognitive node has data to send, it sends the RTS (containing its address, the destination address and a randomly sorted listed of TFSs where it can receive data) and waits for CTS (containing the source address, destination address and a randomly sorted listed of TFSs where the receiver can receive

data). The choice to randomly sort the TFS is tied to lower the probability that two nodes tries to access to the same TFS. Moreover, if the transmitter does not receive CTS it sends RTS once more. If the CTS has been received (together with a probing signal on the selected TFS_T) the transmitter evaluates the capacity (through the expression given in eq. (6.33)) and if it does not meet the constraint, it drops TFS_T from the list and send a new RTS by updating the list of TFS otherwise it probes the receiver on the selected channel (namely TFS_R) and waits for the ACK. If this latter is received then communication takes place, otherwise transmitter probes once more on the next TFS_R in the list communicated by the receiver. Moreover, the use of RTS-CTS allows the transmitter to evaluate distance, that is, propagation delay toward the receiver by measuring the time needed for receiving CTS, so synchronization procedure is easier and can implemented according to the ML criterion via an Early-Gate late scheme [96], Ch.4.

About the functionalities to be operated by the transceiver in receiving stage, these have been already partially described. When a RTS has been received on the common channel, probing on the TFS_T indicated by the transmitter is operated, jointly with the communication of the list of channels to be used for probing by the transmitter. If the received probe meets the required capacity, then and ACK is sent otherwise the receiver expects to receive a new probe on the next preferred TFS_R.

6.2.4 Numerical results

The numerical results have been obtained by merging real data acquisitions along with emulated ones obtained by MATLAB simulations. In order to pictorially show where and when the access can take place, 4 pairs of transmit-receive nodes sharing a bandwidth of 24 KHz are considered, with the possibility of operating single carrier transmission or OFDM over 64 sub-channels (320 Hz band for each sub-channel plus 48 Hz guard bands) with a maximum power per node of 1 Watt. The coherence time is 1 second and that is the frame length and the slot time for sensing is 3 ms. This value is in line with the sampling capabilities of hydrophones available in the market working with a sample frequency of 300kHz and allows to consider multipath effect in WVT in an accurate way. The communication and interference signals in eq. (6.26) are available thanks to field recordings, so no models are used for interference, and the channel effect has been emulated in medium water (MW) as well as in deep water (DW), while in shallow water (SW) the effect has been already taken into account by recordings. The channel model in MW and DW follows the Bellhop model with the tool of wave3qD (see [82],[97] and also [98], Ch.3, for ray-tracing like techniques) according to a bathymetric map considering different depth levels for emulating propagation environments ranging from 400 to 1000 meters. The Bellhop model is used to obtain channel impulse response for each source-

sensing node pair, so including interference propagation. Hence, through Bellhop model, several channel impulse responses are given. Each channel impulse response is characterized by delays and taps. Once applied WVT, the time and frequency representation of the received signals are obtained.

The recorded sound sources are related to ships' engines, sonar signals emitted by buoy (primary user), whales sounds, air bubbles and earthquake since these are generated at different depths. Interference signals representing mammals emissions, wind sound and ship engine noises are obtained by field recording performed in the Mediterranean Sea in front of Fosso Chiarone, Italy, with a bathymetry of maximum 200m. The sounds generated by whales are referred to the *Globicephala melas* while the sound of diesel engines is related to a maximum propulsion energy of 120kW taken from 3 ships posed at different distances (see [99]). About sonar, bubbles and earthquake, these are taken in the ocean in front of Portugal through a proper measurements campaign [100]. If not differently specified, the nodes are equally uniformly distributed over an area of 400m x 400m with the above mentioned depth. So, the results obtained are averaged on 40 topologies. In other words, for each simulation, the position of all the nodes is randomly extracted in order to consider both cases where interference sources are close to receiver and/or transmitter and when are very distant. About nodes movement, direction have been randomly chosen in the interval $[0°, 359°]$ while speed is uniformly (randomly) selected in the interval $[0, 5]$ km/h. It is really important to highlight that if the transmission of a nodes pair operates on the (i, j)-th TFS, the effect of their interference, due to nodes movement can be detected on a different TFS due to Doppler shift, for example $(i, j + 1)$. This does not affect the sensing from an access point of view since what drives the access is the interference detected so, the receiver will be able to receive on the (i, j)-th TFS and not on $(i, j + 1)$-th one.

Before detailing the numerical results, some graphical examples are shown highlighting how the procedure works with respect to the conventional approach aiming at sensing the channel without distinguishing the nature of interference.

In Fig. 6.20(a) the whole signal (confined for sake of significant representation to 1.4 KHz), represented in time and frequency domains via the WVT, is reported. For sake of representation, here a 256 sub-channels setup has been considered (80 Hz bandwidth and 12 Hz guard bands). It is possible to recognize the contribution of bubbles and earthquakes at low frequencies less than 200 Hz. The images for the classifiers representing TFSs are 50×50 pixels. The Figs. 6.20(b)-6.20(c) show respectively the access granted and not possible due to coexistence issues basing on the interference profile in Fig. 6.20(a) when a constraint of $C^\star = 3$ bit/TFS is considered. The TFSs are also depicted to show the access opportunities. It is interesting to note that, in the specific scenario where several source

(a)

(b)

(c)
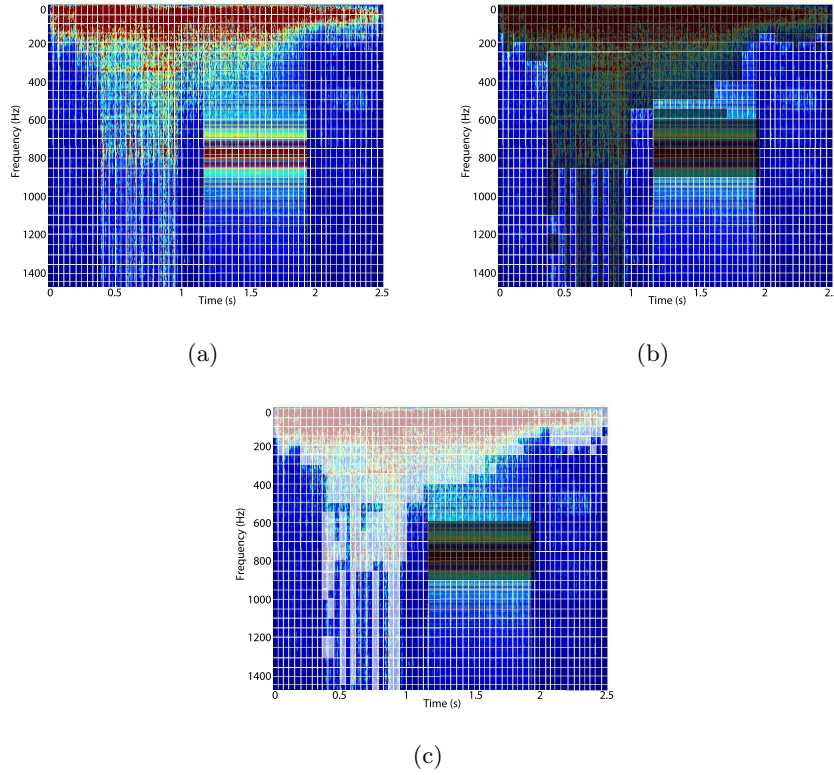
Figure 6.20: Interference and access opportunities with high interference level scenario for classic cognitive and proposed one. (a) Interference. (b) Classic CR access. Dark slots mean no transmission due to coexistence. (c) NARC access. Whited slots mean no transmission due to high interference.

of interference are present (earthquake, bubbles, communication), the conventional cognitive access does not occur in the black TFSs. The reason, in this sense, is simply the coexistence without taking care, as previously mentioned, of the nature of interference. NARC recognizes the different nature of interference so the access is not granted in the communication case (black zones in Fig. 6.20(c)) and also where interference is high and not sufficiently low to guarantee the minimum capacity (white zones). By comparing Fig. 6.20(b) and Fig. 6.20(c) it is clear that the last one allows more access opportunities thus maximizing the access probability.

In Fig. 6.21 the image of interference, the access opportunities in a conventional cognitive radio approach and the proposed one are compared as for the previous case depicted in Fig. 6.20 with the difference of lower interference by natural sources. The only one communication source here is a sonar. Both the approaches do not allow any node to transmit over the

TFSs interested by sonar emission. It appears quite clear that the conven-

(a)   (b)

(c)
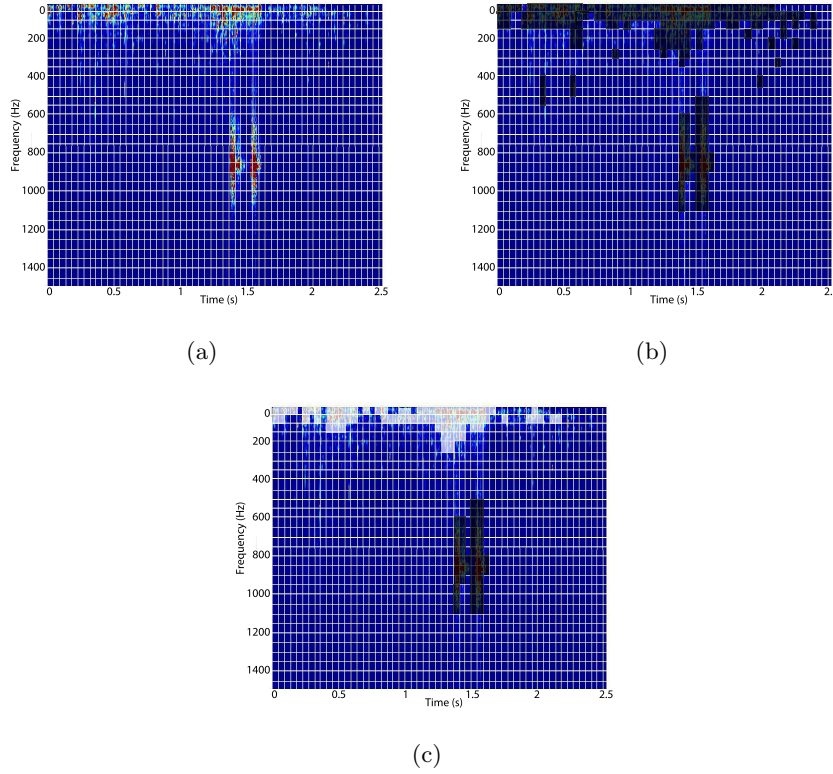
Figure 6.21: Interference and access opportunities with low interference level scenario for classical cognitive and proposed one. (a) Interference. (b) Classic CR access. Dark slots mean no transmission due to coexistence. (c) NARC access. Whited slots mean no transmission due to high interference.

tional cognitive access does not allow transmission opportunities for several TFS, approximately all those at very low frequencies (see Fig. 6.21(b)) The classification-based ones are access due to interference level (whited slots in Fig. 6.21(b)) while it is successful for several TFSs at low frequencies.

Because of its ability in discriminating non-communicating interferers from communication sources, the herein proposed approach exhibits a gain in percentage of access with respect to the classical approach, when non-communicating interfering sources are active in the propagation scenario under concern. To assess such a gain several simulations have been run under different interference scenarios so as to show that the higher the number of non-communicating interferers, the higher the gain attained by the herein proposed scheme. The selected scenarios differ in terms of both number and nature of the active interfering sources.

Table 6.6 reports the average percentage of accessing pairs attained by the NARC, DCRC and the classical scheme. From now consider TFS that has a time duration of 30ms and a bandwidth of 320Hz for the numerical results. Results in Table 6.6 have been obtained after running simulations with different positions of the nodes and of the interference sources. A 320 Hz sub-channel bandwidth has been set up; a successful access is performed when a cognitive pair is able to establish a connection at 2 Kb/s (equivalent to $C^\star = 6$ bit/TFS) with BER of 10^{-6}. The first column in Table 6.6 summarizes the specific interference scenarios. It is worth remarking here, for the sake of clarity, that both the NARC and DCRC approaches, as well as the classical one, do not need any *a priori* information about the number/kind of interferers, so that the data reported in the first column of Table 6.6 should just be considered as a description of a specific simulation scenario, and are not to be intended as some kind of input required by the access scheme.

Results in Table 6.6 clearly show how an approach able to recognize natural/artificial from communication-based interference allows to increase the number of cognitive pairs able to establish a connection, expecially in those scenarios with several non-communicating interferers, where a gain of about 20% of average accessing pairs is observed. The NARC approach,

Table 6.6: Average percentage of accessing nodes with 20 pair of cognitive nodes. Target rate is 2Kbit/s

no./kind of interferer	NARC access %	DCRC access %	access % according to [79]	access % (conventional)
1 ship	99%	99%	99%	92%
2 ships	98.5%	98.5%	97%	88%
2 ships and 1 whale	98.5%	98.2%	96%	86%
3 ships and 1 whale	98%	97.4%	92%	78%
3 ships and 2 whales	98%	96.8%	90%	75%
3 ships, 2 whales and 1 comm	97%	96.8%	82%	65%
3 ships, 2 whales and 2 comms	96.5%	96.8%	79%	65%

driven to recognize natural/artificial interferences, exhibits a slightly better performance. Such a gain is obtained at the cost of a large database that should encompass a large number of natural/artificial interference samples. On the other hand, DCRC still performs well while requiring a more restrained database, encompassing only the WVTs of the transmission techniques employed by the primary and secondary nodes supposedly present in the specific transmission scenario.

Both the proposed schemes outperform that proposed in [79] where, as anticipated before, a cross-correlation-based approach has been adopted to recognize the nature of the interference. Even if less costly with respect to NARC and DCRC since no image classification has been proposed in [79],

the number of non-communications sources increases (e.g., 3 ships, 2 whales and 2 comms) the gain appears more evident since the access percentage is around 96% for NARC/DCRC and 79% for the scheme in [79]. However, when the number of non-communications sources is low, the performance are comparable. It is expected that increasing the target rate will reduce the percentage of both the classified and unclassified methods as shown in the following.

It is interesting to investigate how the value of the target rate $C^\star$ affects the access percentage in combination with the propagation environment. Figs. 6.22-6.23 report the percentage of access opportunities (i.e. the access probability multiplied by 100) of NARC, DCRC and of the conventional cognitive (CC) approach. The scheme depicted in Fig. 6.19 has been also
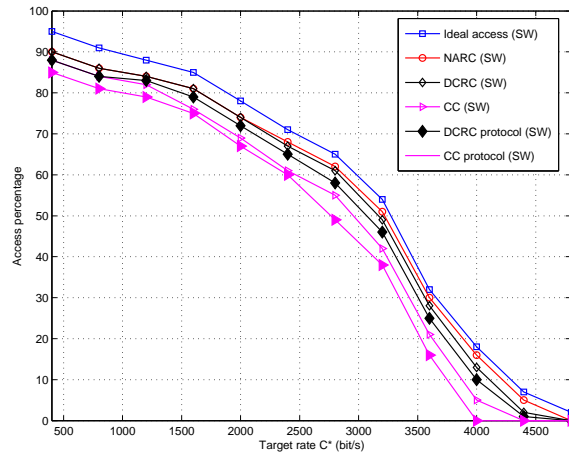
Figure 6.22: Access opportunities for proposed schemes and conventional compared with both the ideal solutions and protocol based ones in Shallow Water.

implemented for DCRC and CC. Here 20 pairs of cognitive nodes are considered, randomly disposed in the water (SW, MW and DW) and for each propagation scenario different ambient noise effects have been considered. In that context, the interference sources are the above detailed whales, ships, wind, bubbles, earthquake, sonar that present different effects at different sea levels due to distance, frequency, salinity and temperature. For the sake of comparison the performance of the ideal access scheme is also reported, that is an oracular approach where all the interference is known a priori (intensity and nature), representing an upper bound for the performance of the cognitive access schemes. The herein described approach outperforms the conventional one at every target rate, while still exhibiting a gap with
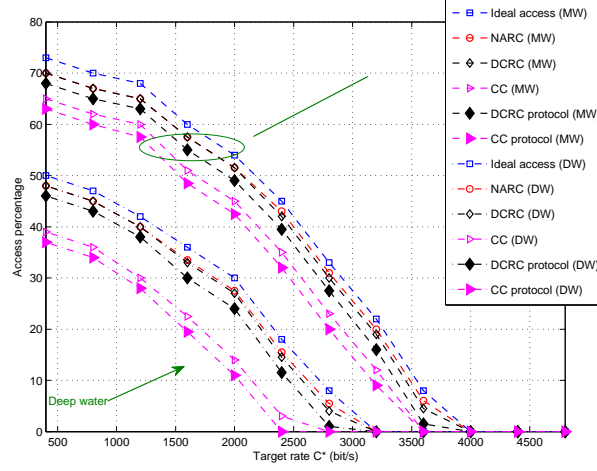
Figure 6.23: Access opportunities for proposed schemes and conventional compared with both the ideal solutions and protocol based ones in Medium-Deep Water.

respect to the ideal case. The gap is mostly present at low target rate values and it is mainly due to the imperfect classification; in fact, a low rate $C^\star$ is easily obtained even if some interference is present in the time frequency slot, so that the performance gap are almost always due to the fact that the NARC/DCRC have misclassified the sensed interference. For high target rate, instead, the gap is restrained; this fact is justified by considering that the access, even in the case of classification affected by errors, is not performed due to the high-quality link requests. The conventional approach that considers all interference as communication one, fails to have performance comparable with the ideal case.

The effect of the protocol implementation has been also reported for the DCRC and CC solutions. It is possible to observe that the access opportunities loss with respect to the ones obtained by resorting only on sensing, is limited since it is due to possible collisions in the RTS-CTS mechanism on common channel and/or to the (random) selection of a (sensed free) TFS unable to guarantee the target rate due to channel attenuation. In this regard, this second effect is emphasized when MW and DW propagation scenarios have been taken into account. Even if the performance hierarchy is the same of SW with ideal case outperforming in order NARC, DCRC, DCRC protocol, CC and CC protocol, the performance degrades in terms of access probability. Moreover, it is interesting to note that DCRC protocol (including collisions in RTS-CTS) outperforms the CC when collisions are not considered for this last case.

Furthermore, the results obtained by the two different schemes have been compared also by considering the effect of sporadicity of the interference. Specifically, in Fig. 6.24 the access opportunities are considered in terms of how much the interference lasts. Note that sporadic interference leads to underuse the spectrum. In fact, if sporadic disturbing signal is sensed, the transmitter probably will avoid to use that TFS even if it will became free in the next frame. The horizontal axis reports the (average) interference time duration in terms of percentage of time frame, so a T_f value means that the interference lasts a frame, while 0.1 T_f means that over the TFSs grid the source is active for 10% of time.
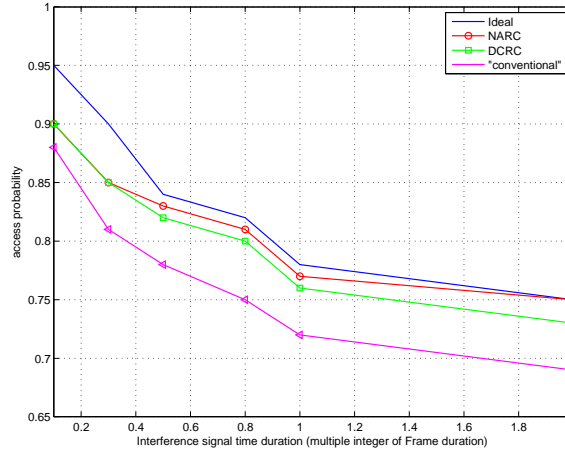
Figure 6.24: Percentage of access opportunities on 100 frames for 12 interference sources.

In this regard, the ideal approach represents an upper bound on access probability while the other plots are related the effect of NARC, DCRC and conventional cognitive procedures respectively. When the interference activity is low, that is, it is really sporadic, access probability is high even if the gap from the ideal case is higher. This is due to the overestimation of interference. In fact, when it is present in some TFSs, the access procedure avoids to access in those places in the next frame. By increasing the time lasting of interference, the access probability decreases both for ideal and non-ideal procedures since the sources occupy more TFSs but, in this case, no overestimate is possible by classification methods while worse performance are achieved by conventional cognitive.

Furthermore, in Table 6.7 the comparison between NARC and DCRC is reported. It is important to stress that false alarm for NARC means to detect a natural/artificial signal when a communication is present and this

is equivalent to missing detection of DCRC. On the other hand the missing detection of NARC means to recognize a communication signal when it is absent and this is equivalent to the false alarm of DCRC. The first column indicates the number of images per category. For NARC $N_{cat} = 12$ so that the whole number of images in the training set, reported in the second column, is $12 \cdot N_{ds}$, while for DCRC $N_{cat} = 3$.

The rationale behind the increased number of categories N_{cat} for NARC is tied to the more variegate behavior of natural/artificial interferences with respect to communication signals. This difference in the value of N_{cat} directly affects both the processing time/power and the required storage memory as well as the performance in terms of access probability. In this respect, results in Table 6.7 show how NARC exhibits a slight better performance at the expense of an increased value of the above referenced complexity figures of merit. So, even if DCRC performs a bit worse with respect to NARC, it requires fewer images to be stored.

Table 6.7: NARC and DCRC Missing Detection (MD) and False Alarm (FA) probabilities for different number of images in the training set.

	NARC				DCRC			
N_{ds}	$N_{ds}N_{cat}$	MD	FA	FA after $C^\star$=3bit/TFS	$N_{ds}N_{cat}$	FA	MD	MD after $C^\star$=3bit/TFS
6	72	0.39	0.15	0.01	18	0.42	0.18	0.011
12	144	0.11	0.02	0.006	36	0.13	0.032	0.0063
18	216	0.07	0.03	0.0021	54	0.09	0.0952	0.0024
24	288	0.0085	0.00082	0.0001	72	0.01	0.0011	0.00085

An interesting point to be stressed is that the MD for NARC and FA related to DCRC represent the missing transmission opportunities, that is the case when a communication is detected erroneously. This reduces the performance quality in terms of access probability. On the other hand the FA of NARC and MD of DCRC deal with the coexistence violation. Now, since a possible transmission is allowed on the basis of achieved capacity, this check avoids (for few images available) to violate the coexistence (due to the higher interference level).

Until now the analysis has presented the gain offered by the proposed scheme in terms of access probability without considering the possible advantage from a transmitting point of view. An interesting performance metric is the sum rate, defined as the sum of the rates of the links active in the network. Analytically speaking the quantity to be evaluated is the following:

$$\mathfrak{R} = \sum_{u=1}^{N_u} R_u. \tag{6.50}$$

where R_u is the transmission rate. As anticipated at the beginning of the work, here the possibility for the users of accessing to more than one TFS within a frame is considered. The maximum of TFSs per frame is set to 3.

The constraint of eq. (6.33)) on capacity C^* influences eq. (6.50)) since it acts on R_u while the interference, i.e. coexistence and the method used for recognizing the nature of disturbing signals, influences the number of slots where access is possible so it can decrease or increase the number of terms to be summed (N_u). In Fig.6.25 the sum rate of the underwater sensor network composed by 20 node pairs is depicted. For the DCRC approach, the presence of 1, 2 and 3 ships has been respectively considered and a case where 3 ships plus one sonar are present. While the performance in the case of the ships are very close (10kb/s sum rate difference as a maximum value), when sonar is present, the impossibility to transmit on some slots reduce the rate till to 56kb/s at SNR equating 30dB. The behavior of a
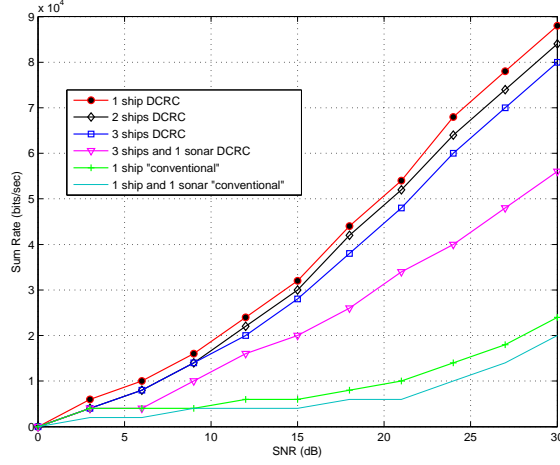
Figure 6.25: Sum rate of the underwater sensor network for average SIR of 10dB.

sensing mechanism, labeled in Fig. 6.25 as "conventional", since it resorts to the energy level of interference despite of its natures, achieves a very low sum rate that is below 30kb/s at SNR=30dB both for 1 ship as a source of interference and 1 ship plus a sonar signal.

Another interesting result is that shown in Fig. 6.26 where the effect on the sum rate induced by the time length of the TFS is reported. The ideal case is considered, which is the best solution, that is assumed to be aware of each information needed for the access, and the proposed DCRC compared also with the "conventional" approach based on energy detection as for [66]. The behavior of each curve is not monotonic. This can be justified by observing that when the time slot is very small (please note that TFS below 1ms are currently not achievable by existing modems as a result of a deep search on vendors websites) even if the access percentage is high,
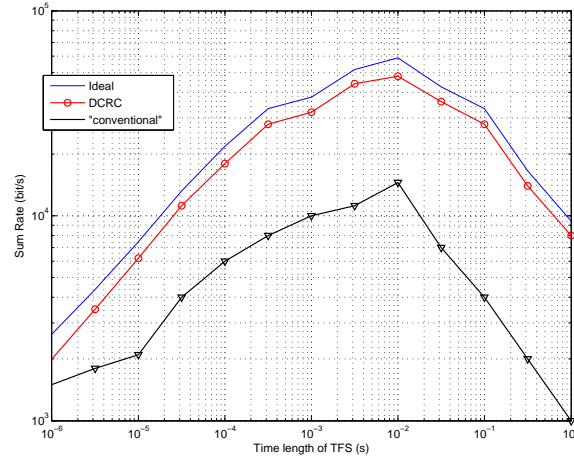
Figure 6.26: Sum rate of the underwater sensor network when different values of slot time are considered.

the resource is available for a very short time. On the other hand when TFS has a time length of the order to 0.1 - 1 sec, the resource is available for long times but for few users. So, the best solution is in the middle and, in this case, obtained for a TFS time length of 1ms for all the three considered schemes. It is worth notice that ideal solution outperforms at maximum value DCRC for 10% while the gain of DCRC with respect the maximum value of the energy detection approach is 300%.

Chapter 7

Conclusion

This dissertation has concerned the design and implementation of underwater acoustic communications. The application scenarios for this kind of technology are several and often requiring specific architectures depending on the activity to be performed. In fact the number of situations in which underwater sensors and/or robots need to be connected is growing, and now several works in the literature consider problems related to the setup and management of underwater acoustic networks. An example is given by the project of Woods Hole Oceanographic Institution, namely Challenger Deep [101], that considers an UWANs represented by a mother ship communicating with sensors/actuators or autonomous underwater vehicles.

The first step of network design is understanding which performance has to be provided. However, the constraints imposed by the considered application must match with those ones dictated by nature. In this regard, the work here proposed has passed through the preliminar study of the underwater acoustic channel as it represents the main source of impairment to communications. Attenuation, multipath and noise are the typical effects occurring during signal propagation, so the information obtained from a reliable model of these phenomena can be used for making the system design appropriate to the characteristics of the medium.

In particular, for what concerns point-to-point communications, the investigation on Multiple-Input Multiple-Output architectures has led to the analysis of many transmission schemes, presenting also the description of the transceivers structure. Then, moving to upper level of the OSI model, solutions for error and medium access control have been proposed in order to improve the efficiency of the whole underwater network.

The main insight given by theoretical and experimental results is that the performance of underwater communications strongly suffers from the channel time-variability. Therefore, realizing a communication working according to *adaptive* rules seems to be the right way to make the possibilities offered by technology complying with the behavior of the medium. The

implementation of underwater networks is somehow made difficult by the fact that the equipment employed is usually composed of commercial devices whose hardware and software design is *not open*, so reconfigurability is not permitted. Recently, the need to overcome these constraints has encouraged the debate about underwater technology challenges. The work by Demirors et al. [102] reports an interesting discussion about the implementation of software-defined UWANs, highlighting how this solution can provide enhancements in terms of software portability, computational capacity, energy efficiency and realtime reconfigurability. Furthermore, the authors propose the architecture of a software-defined acoustic modem and evaluate its performance and capabilities with tank and lake experiments.

Finally, the other aspect to be highlighted is that the network design challenges to be tackled in the underwater scenario are very different from those encountered when dealing with terrestrial communications. Underwater sensors cost, memory, power consumption and deployment, together with the impairments due to the communication medium, are considered fundamental factors more than in out-of-water services [103]. For these reasons, the adoption of protocols deriving from terrestrial applications turns out to be an inefficient solution, so it becomes necessary to build a different architecture where each layer of the network stack [104] is developed by taking into account all those issues related to the underwater environment.

Appendix A

SAM-1 modem symbol re-mapping

As described in par. 3.3.2, The SAM-1 acoustic modem is implemented with a 16-PPM orthogonal modulation scheme. This means that the symbol time window is splitted in 16 different slots where the acoustic pulse can be placed in. So 16 different symbols can be sent, each one representing a number of bits $b = 4$. In Table A.1 the list of the 16-PPM symbols is reported. So,

16-PPM	encoded bits
Symbol 1	0000
Symbol 2	0001
Symbol 3	0010
Symbol 4	0011
Symbol 5	0100
Symbol 6	0101
Symbol 7	0110
Symbol 8	0111
Symbol 9	1000
Symbol 10	1001
Symbol 11	1010
Symbol 12	1011
Symbol 13	1100
Symbol 14	1101
Symbol 15	1110
Symbol 16	1111

Table A.1: 16-PPM symbol list

considering a bit stream to be sent, by using the 16-PPM scheme the information is organized in 4-bits groups each one corresponding to one of the

16 symbols here listed.

A.1   PPM Re-Mapping

This section describes how to move to 8-PPM, 4-PPM and 2-PPM by using the symbol list referring to the 16-PPM scheme provided by the SAM-1 modem, hence without the need of modifying the software.
Considering a simple case of communication link between two network nodes, an example of bit stream to be transferred is the following:

$$0\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 0\ ...$$

Depending on the L-PPM scheme employed, the stream is organized in b-bits blocks, each one encoded in a symbol.

A.1.1   8-PPM

In 8-PPM, a number of bits $b = 3$ is encoded in each of the 8 symbols allowed by this scheme. In order to realize this ste of symbols, 8 elements are taken from the 16-PPM symbols alphabet, that are symbol 1, 3, 5, 7, 9 , 11, 13 and 15 reported in Table A.1. It is worth noting that in this way 4 bits corresponding to a 16-PPM symbol are used for transmitting 3 bits of the data stream, so there is 1 bit transmitted that is useless and has to be discarded at the receiver side. In particular, the first 3 bits of each 16-PPM symbol represent the corresponding 8-PPM symbol, while the fourth bit is redundant (third column of Table A.2). So each 3-bits group of the original bit stream is expressed through these 8 symbols, as shown in Fig. A.1.
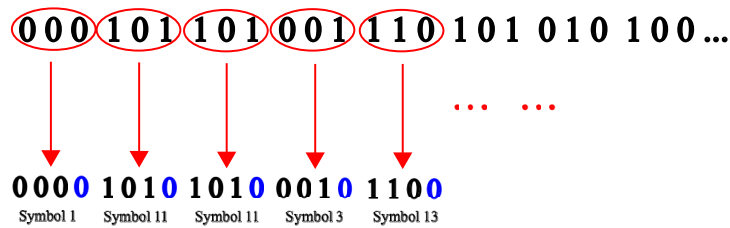
Figure A.1: Re-mapping from 16 to 8 PPM.

Then at the receiver side, the received bit stream has hence to be organized in groups of 4 bits where the first three are the original data, while the fourth, being redudant information, is discarded. Proceeding with this operation, the original data stream is finally obtained.

A.1.2   4-PPM

The 4-PPM scheme provides a number $b = 2$ bits encoded in each one of the 4 possible symbols. These latters, selected from the 16-PPM symbols alphabet, are respectively symbol 1, 6, 11 and 16. In this case the first two bits represent the 4-PPM symbol while the latter two are redundant (fourth column of Table A.2). So this time the bit stream to be sent is organized in blocks composed of two bits and then remapped as shown in Fig.A.2.
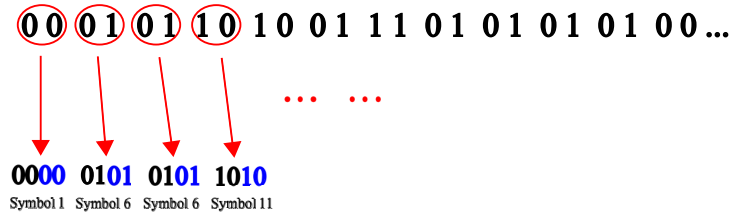
Figure A.2: Re-mapping from 16 to 4 PPM.

Then the receiver will select the two data bits and delete the redundancy bits.

A.1.3   2-PPM

In 2-PPM transmission scheme, only 2 symbols of the 16-PPM list are used, that is symbol 1 and 16. Now only the first bit of each symbol represents the information, while the other three are useless (fifth column of Table A.2). So the remapping procedure is that depicted in the following Fig.A.3.
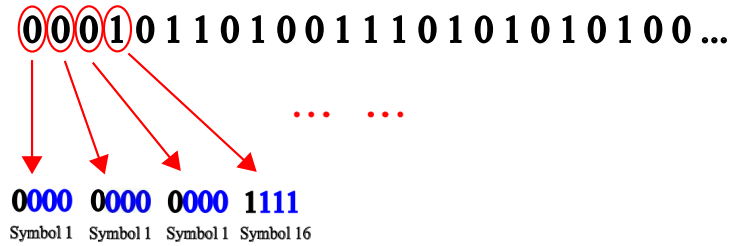
Figure A.3: Re-mapping from 16 to 2 PPM.

The de-mapping operations at the receiver are the same described for the above cases, except for the fact that there is only one bit to *save* every four ones.

Reading Table A.2 column by column it can be appreciated that the 8-PPM, 4-PPM and 2-PPM symbols have been chosen in order to be maximally equally-spaced one from the other. In this way the transmission

16-PPM Alphabet	16 PPM	16 to 8 PPM	16 to 4 PPM	16 to 2 PPM
Symbol 1	0000	0000	0000	0000
Symbol 2	0001	-	-	-
Symbol 3	0010	0010	-	-
Symbol 4	0011	-	-	-
Symbol 5	0100	0100	-	-
Symbol 6	0101	-	0101	-
Symbol 7	0110	0110	-	-
Symbol 8	0111	-	-	-
Symbol 9	1000	1000	-	-
Symbol 10	1001	-	-	-
Symbol 11	1010	1010	1010	-
Symbol 12	1011	-	-	-
Symbol 13	1100	1100	-	-
Symbol 14	1101	-	-	-
Symbol 15	1110	1110	-	-
Symbol 16	1111	-	1111	1111

Table A.2: Re-mapped symbols list

results to be more robust since the space distance between symbols is set for minimizing the intersymbol interference. Obviously, the drawback of re-mapping is the obligation to transmit exceeding bits, thus leading to the communication throughput decrease. In particular, given a N information bit stream to be sent, the number of bits actually transmitted is $\frac{4}{3}N$, $2N$ and $4N$ for the case of 8-PPM, 4-PPM and 2-PPM re-mapping respectively.

A.2 PAM, QAM, and OOK schemes

Since the SAM-1 modem is not equipped with transmit power control, it has not been possible to directly implement Pulse Amplitude Modulation (PAM) and Quadrature Amplitude Modulation (QAM) modulations for real applications. However, as the noise-free shape of the acoustic pulse generated by the modem has been measured (Fig. A.4), an emulation framework has been built using this *real* feature. The purely simulated part concerns the generation of different amplitudes for PAM and QAM symbols, and also the detection mechanism to use at receiver side.
Following this direction, 4-QAM, 16-QAM, 2-PAM, 4-PAM, 8-PAM and 16-PAM can be emulated and exploited in the simulations.

For what concerns the On-Off Keying (OOK), since this modulation substantially consists in the transmission of a pulse or nothing, it has been

sufficient to use the acoustic pulse template multiplied in amplitude by 1 or 0 in order to have the two possible OOK symbols. Differently from PAM and QAM, the implementation of the OOK modulation on the SAM-1 modem has been realized. In fact, by first transmitting a sequence composed of the symbol 1 of the 16-PPM alphabet and then by offline multiplying by 0 or 1 the received symbols an OOK modulated data stream is obtained.
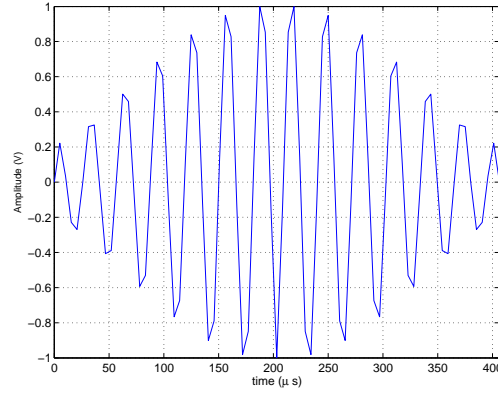
Figure A.4: Template of the acoustic pulse emitted by the SAM-1 modem.

Appendix B

Passive sonar application: vessels traffic estimation

Underwater surveillance in naval bases or harbor environments is an important task to be accomplished. It can allow statistics acquisition in an automatic fashion without involving human presence registering ships entrances and leavings. It is an effective alternative to active radars that increase electromagnetic pollution even though approaches based on opportunistic target detection are essentially based on passive radars (see [105]). Today, the most used automatic systems for detection of naval presences are active sonar systems where an active modem/sonar emits some acoustic signals and process the reflections of objects so as to understand if a vessel is present or not. However, the performance of an active sonar can be disturbed by the sounds coming from ships' engine noise. This justifies the research on the use of passive sensor systems as a complement to active sonar systems or alternative to active sensors. By applying signal processing algorithms to the sounds generated in the above detailed environments, passive sensors performance can be improved. This problem has been investigated in the literature where several sensor architectures [106][107] and techniques [108] have been proposed.

In this context, an image processing tool for detecting the presence of ships and their trajectory is presented below [10]. In particular, the proposed approach is able to detect when ships enter the harbor or leave. To do so, the signal acquired is represented on a time-frequency base (that is the Wigner-Ville Transform) and image classification techniques are applied to extract information about the source of the recorded sound.

B.1 System model and signal analysis

Assume the presence of (at least) two acoustic passive sensors (APSs). One is able to capture the signals coming from the harbor and the other is able to acquire signals from the open sea in front of the harbor. The model for the signal received by the i-th hydrophone is:

$$y_i(t) = \sum_{l=1}^{U} v_l(t) * h_{li}(t) + \sum_{k=1}^{M} \chi_k(t) * \eta_{ki}(t) + n(t) \qquad \text{(B.1)}$$

where $v_l(t)$ is the (possible) ship engine sound generated by the l-th vessel out of U and $h_{li}(t)$ is the effect on acoustic propagation. Similarly, $\chi_k(t)$ is related to other acoustic sources (in a number of M) whose propagation is modeled as $\eta_{ki}(t)$. This latter is a disturbing signal for the vessels traffic analysis. Last $n(t)$ is the additive generally colored Gaussian noise.

A possible strategy for understanding the nature of the signals (that is, if the sound is generated by a vessel or not) is observing the time and frequency properties of the acquired sounds. By means of the WVT [86], it is possible to infer the spectral components of a signal with respect to the time variable and therefore to obtain a two-dimensional description, that is, an image. In fact, the WVT allows to acquire a snapshot of the energy sensed over a given bandwidth for a selected number of consecutive time intervals short enough to assume stationarity of the received signal. Such a time/frequency representation turns quite useful when dealing with engine recognizing. So, according to [86], the expression for the WVT of the signal in eq. (B.1) is given by:

$$Y_i^{(WV)}(u, \xi) = \int_{T_O} y_i(u + \frac{\vartheta}{2}) y_i^*(u - \frac{\vartheta}{2}) e^{-j\vartheta\xi} d\vartheta \qquad \text{(B.2)}$$

being T_O the observation time. So, from the image representation of a vessel engine some information can be acquired. For example, by considering acquisition by multiple hydrophones, the direction of the vessel can be inferred. The signals classification is obtained by resorting to the Support Vector Machine [109].

Specifically, in order to reduce the cost of the analysis and avoid to manage large files, the acquisition can be partitioned into frames of T_O seconds. The choice of T_O value must be performed according to ambiguity issues. In fact, since two different frames may refer to the same vessel, T_O should be greater than the time needed with an assigned minimum speed to cover the whole harbour-sea distance.

Another possible cost reduction strategy can be to select the signal portion containing sufficiently high energy levels. So, an edge detection step is performed to detect the regions (in time and frequency) where a sufficiently strong signal is present. Successively, the detected pattern is analyzed to

infer the features of the sound. As the emission of different sources (vessels, mammals, shrimps) will be characterized by different patterns in the WVT, the problem of inferring the characteristics of the transient is in turn recast as a problem of image classification. In Fig. B.1 it is possible to recognize the presence of three different ships. The first one is characterized by the presence of a very low frequency components with high energy (red). After, the presence of a second ship having higher frequency components can be recognized. Finally, a third vessel is present and this latter is characterized by a very low frequency occupancy.
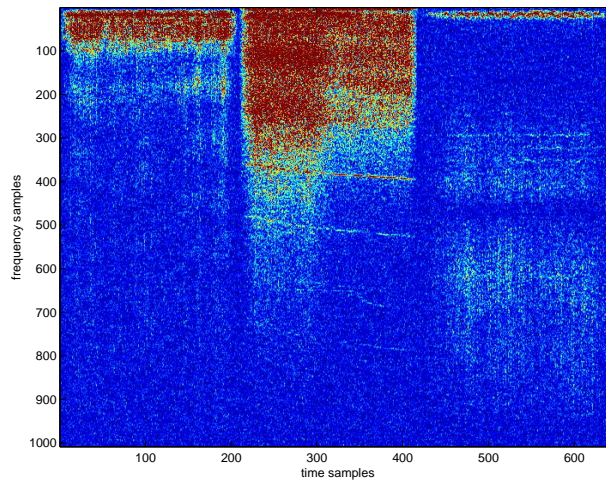
Figure B.1: Spectrogram of three different ships.

B.2   Traffic analysis and numerical results

The estimation of the the time of arrival of vessels (arrival or leaving) can be performed according to the flowchart of Fig. B.2. Consequently, the traffic probability density function (pdf) of inter-arrivals may be performed off-line. By resorting to Fig. B.2, after signal acquisition, a number of audio recording equating the number of hydrophones is stored. The signal analysis block is a pattern recognition tool operated on images as, for example, suggested in [109] by resorting to support vector machine and it is used for classifying images. In fact, the first question to be answered is about the nature of signal acquired, thus meaning that the system must understand if the captured sound has been generated by a vessel engine or not. If the sound appears to be due to interference the algorithm proceeds to process the next acquisition in the audio file. On the other hand, if the signal is

recognized as a vessel approaching then a new question must be answered. This consideration implicitly suggests that the theoretical structure is the
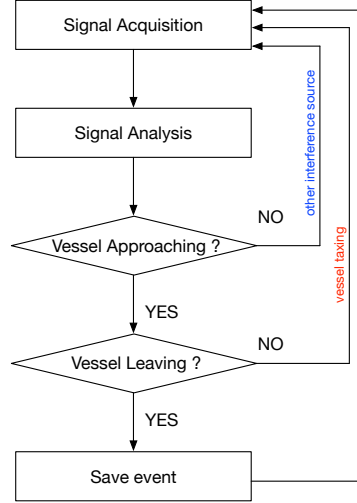
Figure B.2: Flowchart of the scheme used for estimating the vessels traffic.

typical of hypothesis testing. In fact, the events to be considered are interference signals recognized as vessels (false alarm) or vessels recognized as interference (missing detection). False alarm is only due to interference, while missing detection is due to badly classified vessels sound or highly attenuated vessel sound unable to be detected by the hydrophones.

When no missing detection or false alarm incur the right behavior is the following. If a vessel is entering or exiting the harbor area (in this regard, this means that the hydrophone looking at the open sea or at the harbor respectively recognized the engine sound) the second hydrophone is expected to hear the engine noise leaving the area, otherwise there is the possibility of vessel navigating in front of the harbor or inside it (that is *taxiing*). This latter case should not be stored as entrance/leaving the harbor, however it could represent another possible cause of false alarm or missing detection. On the other hand, if the second hydrophone recognizes the presence of an engine's sound leaving, the event is saved, thus meaning that the time stamp is stored. Table B.1 reports the different cases.

In order to test the proposed approach the inter-arrivals have been modeled as an exponential decreasing pdf and this is the basis for generating traffic that is represented by the presence of vessels sound. The λ factor, describing both mean and variance is set to $\lambda = 0.05$ vessels per minute, equivalent to one vessel every 20 minutes. After, the audio recordings of the hydrophones have been off-line analyzed by applying the algorithm previ-

Table B.1: False Alarm and Missing Detection events for different recognition. Bold elements on first column report real situations while first row elements refer to what has been detected.

	interference detected	vessel taxiing detected	vessel detected
no vessel	$\checkmark$	false alarm	false alarm
vessel taxiing	missing detection	$\checkmark$	false alarm
vessel approaching/leaving	missing detection	missing detection	$\checkmark$

ously detailed. The saved events have been finally reported on a histogram so as to show their distribution. The classification is based on a 12 different images related to ship engines sounds. In Fig. B.3 a possible placement of hydrophones is depicted in relationship with the harbor of Civitavecchia, Italy. It is possible to appreciate the four labels that are harbor direction (right side - HDRS), sea direction right side (SDRS), harbor direction left side (HDLS) and sea direction left side (SDLS). In Fig. B.4 the red stroked

Figure B.3: A possible sensors placement.

and yellow filled circle markers show the vessel arrivals according to the Poisson model, while the blue star markers represent the output of the estimation process detailed also in Fig. B.2. The plot represent the traffic among 24 hours by detailing the minutes. It is interesting to note that at

minute 560 the estimation produce a false alarm and the same is at minutes 790 and 910, while no missing detection events have been recognized. In order to understand how the false alarm events reflect on statistics, by
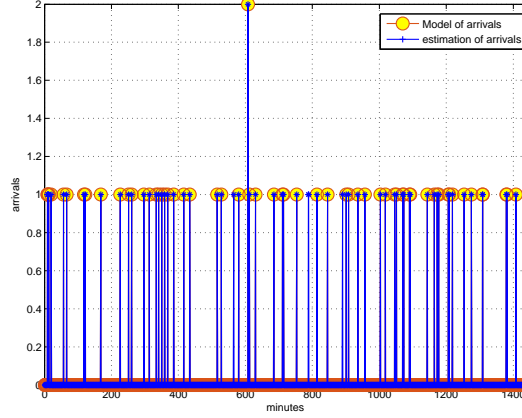
Figure B.4: Vessel arrivals. Model and estimation are reported.

looking at Fig. B.5 reporting both the modeled exponential decreasing pdf and the estimated one it is possible to note how these two plots are really close each other so detailing the reliability of the proposed approach.

Moreover, in Table B.2 the false alarm and missing detection rates (FAR and MDR) are reported considering a different number of employed hydrophones. Missing detection is represented by a not-recognized vessel, while

hydrophones configuration	FAR	MDR
HDRS (1)	0.4	0.2
SDRS (2)	0.04	0.03
HDLS (3)	0.002	0.006
SDLS (4)	0.0015	0.0045

Table B.2: False Alarm Rate and Missing Detection Rate for different hydrophones configurations.

false alarm is due both to interference recognized as vessel or vessel taxiing detected as entering or leaving the harbor. When one hydrophone is considered, it is in the harbor direction (right side - HDRS), when two are used the sea direction right side (SDRS) is added, when three hydrophones are considered the harbor direction left side (HDLS) is taken into account and when using four hydrophones all the directions showed in Fig. B.3 are considered. By observing the results in table B.2 it is possible to appreciate that using two hydrophones in place of one dramatically improves the per-
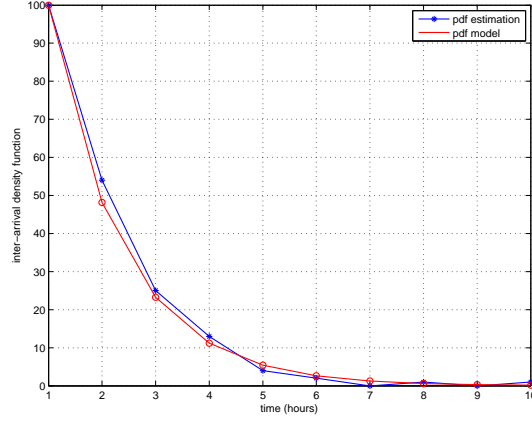
Figure B.5: Probability density function of inter-arrivals model and estimation according to the proposed scheme

formance while the difference in using three or four hydrophones is not so evident.

Summarizing, the proposed tool can be used for traffic analysis and can gather different statistics with reliable performance when multiple hydrophones are used. The computational cost is limited since based on support vector machine and WVT implementation.

Appendix C

Active sonar application: acoustic bathymetric mapping

The use of sonar systems for evaluating bathymetry is a consolidated technique even though it presents some drawbacks. Several applications ranging from *simple* sounding to synthetic aperture sonar (SAS) allow to estimate depth and, especially for SAS, obtain images of the water floor (see [110]-[111]).

In shallow water environments, the swath coverage and performance of sidescan bathymetric systems are influenced by the interference due to multipath. So in [112] a hybrid sonar design is proposed, showing that the combination of multibeam and sidescan technologies helps in suppressing multipath effects and thus extending the swath coverage. Dealing with seafloor mapping, in [113] an interferometry sonar installed on an underwater acoustic vehicle in addition to a side scanning sonar is employed to highly resolve bathymetry mapping. A possible way to increase the accuracy of bathymetry data is described in [114]. By processing the adjacent across-track and along-track beam pairs an improvement of the angle of arrival estimation is obtained. Finally, in [115] the use of high resolution methods and tracking algorithms in order to mitigate the multipath interference effect is discussed.

A typical feature of the above reported approaches is that the transmission and receiver sections are contained in a unique device. From one hand, this is an advantage since only one clock is needed and the hardware is compact. However, assuming that the pulse emitted lasts T_p seconds, it is in general impossible to detect possible reflections due to obstacles at a distance less or equal to $0.5T_p v$ where v is the speed of sound in the water, usually 1500m/s. This leads to conclude that for pulses lasting milliseconds or tens of milliseconds is not possible to detect reflections of object/obstacles

!h]
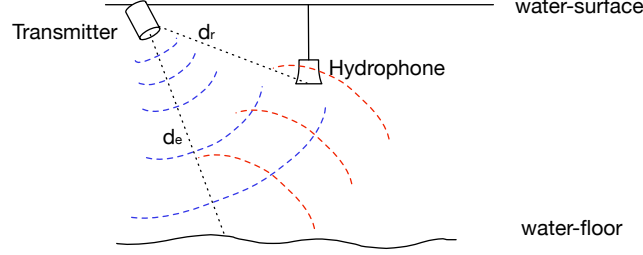
Figure C.1: Example of transmitter-hydrophone propagation due to reflections.

including water-floor at few meters of distance (e.g. $T_p = 10ms$ leads to 7.5 meters distance impossible to detect). Usually this point is not so critical, but it becomes dramatic when a very shallow water environment is considered.

In this regard, a low cost system characterized by a hardware/software integration for the study of the bathymetry of a very shallow water reservoir is presented [11]. The reservoir in Rome (coordinates 41.827990 N, 12.467959 E) has been considered for measurements since it is expected that the depth is less than 5 meters. As previously specified, analyzing so low values of depth requires very large bandwidth and this can be really costly. The following analysis proposes the bathymetric map, as well as the distribution of the reflections (amplitude) so as to try to classify the water-bed.

C.1 Bathymetric analysis

Suppose to transmit a signal $s(t)$ as that generated by SAM-1 modem. For very shallow water bathymetry, where transmitter and receiver are posed at a distance equating d_r, the channel impulse response can be modeled (as for communication purposes) as a two rays channel (see Fig. C.1). The first ray is the direct one and the propagation delay is $\tau_0 = d_r/v$. The depth of the water-floor is the half of the propagation path followed by the signal reflected by the floor, so $\tau_1 = 2d_e/v$. This implicitly means that the time distance between the direct and reflected ray is:

$$\Delta\tau = \tau_1 - \tau_0 = \frac{2d_e - d_r}{v}. \tag{C.1}$$

The above relationship allows to conclude that depth is given by:

$$d_e = \frac{v\Delta\tau + d_r}{2}. \tag{C.2}$$

So, the signal acquired by the hydrophone is characterized by three components as detailed by the following relationship:

$$y(t) = s(t) * [a\delta(t - \tau_0) + b\delta(t - \tau_1)] + n(t) \tag{C.3}$$

where the term $a\delta(t - \tau_0)$ considers the attenuation $(a < 1)$ and delay (τ_0) of the direct path, while $b\delta(t - \tau_1)$ models the effect of reflection. Last $n(t)$ models the additive, generally colored, Gaussian noise. Furthermore, eq. (C.3) can be rewritten as follows by considering eq. (C.1):

$$y(t) = s(t) * [a\delta(t) + b\delta(t - \Delta\tau)] + n(t). \tag{C.4}$$

Moving to the Fourier analysis, by recalling that the convolution in time becomes the product in the frequency domain and, more, $\mathfrak{F}\{\delta(t - \Delta\tau)\} = e^{-j2\pi f \Delta\tau}$, it is possible to write eq. (C.4) as:

$$Y(f) = S(f)[a + be^{-j2\pi f \Delta\tau}] + N(f). \tag{C.5}$$

Note that, following the Fourier theory, the term $n(t)$ is not directly transformable, but a discrete time implementation (via Discrete Fourier Transform - DFT) allows the computation of $Y(K)$, this latter being the sampled version of the received spectrum. So, the effect of the channel in the sampled frequency domain is descrived by the following expression:

$$\tilde{H}(K) = \frac{Y(K)}{S(K)} \tag{C.6}$$

whose power spectral density is given by

$$|\tilde{H}(K)|^2 = a^2 + b^2 + 2\,a\,b\,\cos(2\pi K \Delta\tau). \tag{C.7}$$

It is interesting to observe that in the discrete frequency domain the above expression is a sinusoidal tone, so the distance among two maximum values is the frequency domain period, that is, $(\Delta\tau)^{-1}$. So, the distance among those two maximum (or minimum) values leads to obtain $\Delta\tau$, allowing the evaluation of eq. (C.2) since d_r is known as well as v. By removing the assumption of having only two paths, the analysis carried out till now is, in general, no more valid since it can fail to be robust with respect to several reflections. The model in eq. (C.4) in fact becomes:

$$y(t) = s(t) * \left[a\delta(t) + b\delta(t - \Delta\tau) + \sum_{i=2}^{M} \alpha_i \delta(t - \theta_i)\right] + n(t), \tag{C.8}$$

where α_i terms model the attenuation of the paths reflected by other surrounding objects and θ_i is the relative delay. The summation goes from 2

to M since it is assumed that M-1 reflective components may be present. This last consideration is mostly true when approaching the borders of the reservoir or close to boats and so forth. In this case, a direct application of the Fourier analysis leads to the following expression:

$$Y(f) = S(f) \left[a + be^{-j2\pi f \Delta \tau} + \sum_{i=2}^{M} \alpha_i e^{-j2\pi f \Delta \theta_i} \right] + N(f). \qquad \text{(C.9)}$$

that is not so easy to manage, and it is not so straightforward to arrive at eq. (C.7). This implicitly suggests to follow an alternative approach.

In the simple case of only two paths present, the cross-correlation between the signal transmitted by the modem/SONAR $s(t)$ and the received one $y(t)$ returns, in the discrete time domain:

$$\mu_{ys}(m) = a\mu_s(m) + b\mu_s(m-z) + \mu_{sn}(m) \qquad \text{(C.10)}$$

When noise is low with respect to signal and since $a > b$, a very simple (post) processing can lead to subtract the term $a\mu_s(m)$ where a can be inferred by observing the maximum value of $\mu_{ys}(m)$. In this sense, by computing $\mu_{ys}(m) - \tilde{a}\mu_s(m)$ the following relationship is given:

$$\mu_{ys}^{(b)} = \mu_{ys}(m) - \tilde{a}\mu_s(m) = b\mu_s(m-z) + \mu_{sn}(m) + \mu_\epsilon(m) \qquad \text{(C.11)}$$

being $\mu_\epsilon(m)$ the residual error. Expanding this approach to the case of surrounding objects, that is the case where more than a single secondary paths are present, it is obtained:

$$\mu_{ys}(m) = a\mu_s(m) + b\mu_s(m-z) + \sum_{i=2}^{M} \alpha_i \mu_s(m-q_i) + \mu_{sn}(m). \qquad \text{(C.12)}$$

As for the previous case, an estimated version of $a\mu_s(m)$ canbe subtracted, where the error is simply due to a since $\mu_s(m)$ is perfectly known. This leads to have:

$$\mu_{ys}^{(b)} = \mu_{ys}(m) - \tilde{a}\mu_s(m) =$$

$$\mu_{ys}^{(b)} = \mu_{ys}(m) - \tilde{a}\mu_s(m) =$$

$$b\mu_s(m-z) + \mu_{sn}(m) + \sum_{i=2}^{M} \alpha_i \mu_s(m-q_i) + \mu_\epsilon(m) \qquad \text{(C.13)}$$

A discussion about the relationship between $b\mu_s(m-z)$ and $\sum_{i=2}^{M} \alpha_i \mu_s(m-q_i)$ must be carried out. The performance expected by the proposed scheme strongly depend on b and α_i from one side and z and q_i. If the reflections of the surrounding objects are, from the amplitude point of view, stronger

than b, it is possible that they could be confused with the reflection of lake floor. In this respect, also the relationship between z and q_i is important since if $z >> q_i$ it is reasonable that the estimation of the delay related to the reflection of z does not give rise to ambiguity or errors. If $z > q_i$ and the difference is not so evident, ambiguity is possible even though low error is expected The problem in terms of performance is tied to the case when $z < q_i$ that can incur when being close to the reservoir boundaries and when very close vessels/boats are present. So, the estimated (sampled) delay is given by:

$$\tilde{z} = \arg\max_m \mu_{ys}^{(b)}(m) \qquad (C.14)$$

and the continuous time delay is $\tilde{\tau}_1 = \tilde{z}T_s$. Finally, as previously anticipated, the distance (depth) is given by:

$$\tilde{d}_e = \frac{v\left(\tilde{\tau}_1 - \tau_0\right) + d_r}{2} \qquad (C.15)$$

C.2 Real measurements

The measurement campaign at the EUR Lake in Rome was conducted realizing an active sonar system composed of the SAM-1 modem as projector and the AS-1 hydrophone as receiver. The acquired data (in the .wav format) have been processed with Matlab software. Once organized the lake surface in different sections (a square of 20m x 20m, each one characterized by its center), the estimated bathymetry has been obtained by sending 12 acoustic pings for each area and the depth has been calculated by averaging on different measures. Fig. C.2 the measure of the cross-correlation between the received signal and the emitted acoustic ping is plotted. The main com-
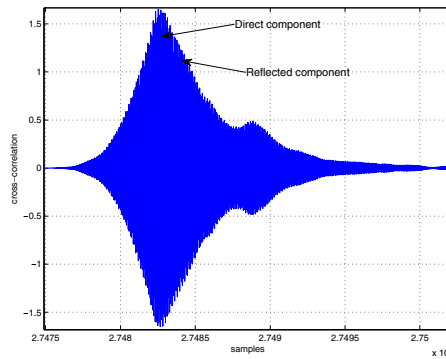
Figure C.2: Cross-correlation of transmitted signal with received one.

ponent is due to the main path (transmitter-receiver) while the second one

is the effect of the reflection. The map reported in Fig. C.3 shows that the measured depth ranges from 1.80 meters to 3.90 with a maximum close to bridges due to concrete reinforces for pillars. This justifies also the shape of
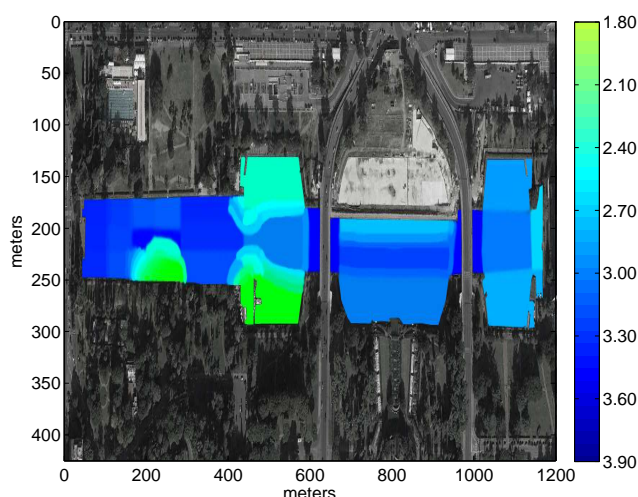
Figure C.3: Bathymetric map for reservoir.

the bathymetric regions in the picture. Spanning the reservoir from left to right it is possible to observe that at the beginning (extreme left) the depth is maximum, that is, a few centimeters less than 4 meters. moving slightly to the right a *small floor step* is present and the depth is around 3.60 meters. Then two different areas are present (around 200m in the horizontal axis). The upper part is still around 3.50 meters while a step of more than 1 meter (leading depth to 2 meters) is present in the lower part. After, the lake presents a depth decrease in the upper and lower part in correspondence of 500m horizontal axis. In fact, while in the central zone (200 m vertical axis) the depth is around 3.20 meters, the upper part presents a smaller depth while the lower part of the reservoir achieves 1.80 meters. This is justified by the presence of a jetty where the pedalos can be rent and start. As anticipated, under the pillars of the road passing on the reservoir the depth is the maximum. After, when the horizontal axis achieves 800 meters, the depth is around 3 meters even though in the upper part is a bit smaller. Then, after the second series of pillars (at 1000 meters on the horizontal axis), the depth decreases since it is between 2.40 and 2.10 meters. It is important to note that the expected error is higher in correspondence of the borders of the reservoir due to reflections by the walls and in the previously detailed jetty area since there some pedalos and small boats were present causing additional reflections.

Finally, the lake-floor composition has been discussed by studying the classification of the attenuation. All the peaks related to reflection (the main component) have been collected ind gathered into a vector. Successively, a histogram (Fig. C.4) has been drawn in order to have a statistical description of the amplitudes of the reflections. As the lake-floor is characterized by asbestos, this analysis helps at investigating the presence and/or absence of possible sediments. By considering the two modes highlighted in the
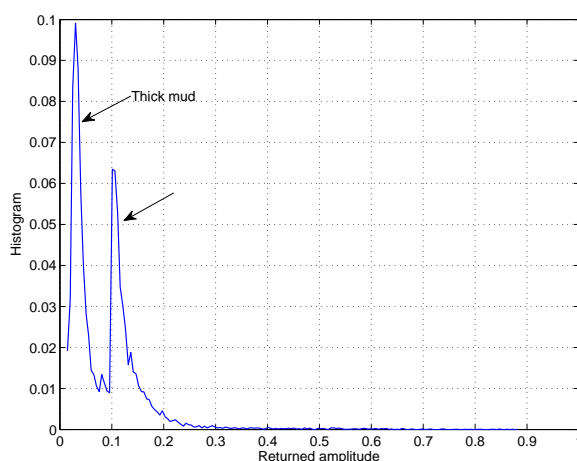
Figure C.4: Histogram of the returned amplitudes, measured with the hydrophone.

histogram, one centered around 0.02 and on around 0.12, and by resorting to the studies in [116][117][118] it has been possible to argue that the floor is characterized by mud essentially of two different kinds. One is thick since the signal is strongly attenuated (and this is the largest part) and one is thin mud. No values higher than 0.8 have been found even though values over 0.7 are related to the part of the lake under the two bridges. However, since the areas spanned by the roads (under the roads) are very small if compared with the whole dimension of the reservoir, the peaks obtained as reflection of the signal under this sections represent a very small percentage of all the acquisitions.

List of Figures

156

List of Tables

Bibliography

[1] L. Bjørnø, "Features of underwater acoustics from aristotle to our time," *Acoustical Physics*, vol. 49, no. 1, pp. 24–30, 2003. [Online]. Available: http://dx.doi.org/10.1134/1.1537384

[2] T. G. Muir and D. L. Bradley, "Underwater acoustics: A brief historical overview through world war II," *Acoustics Today*, vol. 12, no. 3, Fall 2016. [Online]. Available: http://acousticstoday.org/wp-content/uploads/2016/09/Underwater-Acoustics.pdf

[3] X. Lurton, "Chapter 2 - underwater acoustic wave propagation," in *An Introduction to Underwater Acoustics: Principles and Applications*. Springer-Verlag Berlin Heidelberg, 2010. [Online]. Available: http://www.springer.com/gb/book/9783540784807

[4] L. Freitag, M. Grund, C. Von Alt, R. Stokey, and T. Austin, "A shallow water acoustic network for mine countermeasures operations with autonomous underwater vehicles," 10 2017.

[5] P. Kumar, P. Kumar, P. Priyadarshini, and Srija, "Underwater acoustic sensor network for early warning generation," *Oceans*, October 2012.

[6] "Underwater acoustics. technical guides - speed of sound in sea-water," *National Physical Laboratory, Teddington, Middlesex, UK*, 2000. [Online]. Available: http://www.comm-tec.com/Library/Technical_Papers/speedsw.pdf

[7] H. Medwin and C. S. Clay, "Chapter 2 - sound propagation," in *Fundamentals of Acoustical Oceanography*, ser. Applications of Modern Acoustics, H. Medwin and C. S. Clay, Eds. San Diego: Academic Press, 1998, pp. 17 – 69. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780124875708500040

[8] BrightMags, *How does sonar work?* [Online]. Available: http://brightmags.com/how-does-sonar-work/

[9] R. Urick, *Principles Of Underwater Sound.* McGraw-Hill Ryerson, Limited, 1983. [Online]. Available: https://books.google.it/books?id=hfxQAAAAMAAJ

[10] A. Petroni, M. Biagi, S. Colonnese, R. Cusani, and G. Scarano, "Vessels traffic estimation through image processing applied to acquisitions by hydrophones," in *OCEANS 2015 - Genova*, May 2015.

[11] A. Petroni, S. Pergoloni, M. Biagi, S. Colonnese, R. Cusani, and G. Scarano, "Acoustic bathymetric mapping in very shallow water reservoir," in *OCEANS 2015 - MTS/IEEE Washington*, Oct 2015.

[12] Aquarian Audio & Scientific, *AS-1 Hydrophone Specifications.* [Online]. Available: http://www.aquarianaudio.com/as-1-hydrophone.html

[13] Dodotronic, *Hydromic Preamplifier.* [Online]. Available: http://www.dodotronic.com/hydromic/?v=cd32106bcb6d

[14] Desert Star Systems LLC, *SAM-1 Specifications*, 2011. [Online]. Available: http://www.desertstar.com/page/sam-1

[15] G. G. Xie, J. H. Gibson, K. Bektas, "Evaluating the feasibility of establishing full-duplex underwater acoustic channels." The Third Annual Mediterranean Ad Hoc Networking Workshop June 27-30, Bodrum, Turkey, 2004.

[16] A. Petroni, S. Pergoloni, M. Biagi, S. Colonnese, R. Cusani, and G. Scarano, "Adaptive PPM acoustic detection in very shallow water reservoir," in *OCEANS 2015 - MTS/IEEE Washington*, Oct 2015.

[17] L. Berkhovskikh and Y. Lysanov, *Fundamentals of Ocean Acoustics.* Springer, 1982.

[18] M. Stojanovic, "On the relationship between capacity and distance in an underwater acoustic communication channel," in *Proceedings of the 1st ACM International Workshop on Underwater Networks*, ser. WUWNet '06. New York, NY, USA: ACM, 2006, pp. 41–47. [Online]. Available: http://doi.acm.org/10.1145/1161039.1161049

[19] R. Coates, *Underwater acoustic systems*, ser. A Halstead Press book. John Wiley Sons Canada, Limited, 1989.

[20] G. M. Wenz, "Acoustic ambient noise in the ocean: Spectra and sources," *Journal of The Acoustical Society of America*, vol. 34, 1962.

[21] M. Stojanovic, "On the relationship between capacity and distance in an underwater acoustic channel," pp. 41–47, 11 2006.

[22] R. Clarke, "A statistical theory of mobile-radio reception," vol. 47, pp. 957–1000, 07 1968.

[23] P. Etter, *Underwater Acoustic Modeling: Principles, techniques and applications, Second Edition.* Taylor & Francis, 1995. [Online]. Available: https://books.google.it/books?id=WZlImldKm9wC

[24] F. Jensen, "Numerical models of sound propagation in real oceans," in *OCEANS 82*, Sept 1982, pp. 147–154.

[25] P. Etter, "Recent advances in underwater acoustic modelling and simulation," *Journal of Sound and Vibration*, vol. 240, no. 2, pp. 351 – 383, 2001. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022460X00932127

[26] P. Qarabaqi and M. Stojanovic, "Adaptive power control for underwater acoustic communications," in *OCEANS 2011 IEEE - Spain*, June 2011, pp. 1–7.

[27] S. Haile, "Investigation of channel reciprocity for OFDM TDD systems," University of Waterloo, Ontario, Canada, 2009.

[28] R. R. Parenti, J. M. Roth, J. H. Shapiro, F. G. Walther, and J. A. Greco, "Experimental observations of channel reciprocity in single-mode free-space optical links," *Opt. Express*, vol. 20, no. 19, pp. 21 635–21 644, Sep 2012. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-19-21635

[29] M. Fink, D. Cassereau, A. Derode, C. Prada, P. Roux, M. Tanter, J.-L. Thomas, and F. Wu, "Time-reversed acoustics," *Reports on Progress in Physics*, vol. 63, no. 12, p. 1933, 2000. [Online]. Available: http://stacks.iop.org/0034-4885/63/i=12/a=202

[30] P. Gagnon, "Clock synchronization through time-variant underwater acoustic channels," 2012. [Online]. Available: https://calhoun.nps.edu/handle/10945/17368

[31] T. C. Yang, "Characteristics of underwater acoustic communication channels in shallow water," in *OCEANS 2011 IEEE - Spain*, June 2011, pp. 1–8.

[32] A. Petroni, S. Pergoloni, H. L. Ko, T. H. Im, Y. H. Cho, R. Cusani, G. Scarano, and M. Biagi, "Channel reciprocity analysis for bidirectional shallow water acoustic communications," in *OCEANS 2017 - MTS/IEEE Anchorage (to be published)*, Sept 2017.

[33] J. G. Proakis, "Adaptive equalization techniques for acoustic telemetry channels," *IEEE Journal of Oceanic Engineering*, vol. 16, no. 1, pp. 21–31, Jan 1991.

[34] S. Zhao, X. Zhang, and X. Zhang, "Iterative frequency domain equalization combined with LDPC decoding for single-carrier underwater acoustic communications," in *OCEANS 2016 MTS/IEEE Monterey*, Sept 2016, pp. 1–5.

[35] A. G. Klein and C. R. Johnson, "MMSE decision feedback equalization of pulse position modulated signals," in *2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577)*, vol. 5, June 2004, pp. 2648–2652 Vol.5.

[36] G. Ungerboeck, "Fractional tap-spacing equalizer and consequences for clock recovery in data modems," *IEEE Transactions on Communications*, vol. 24, no. 8, pp. 856–864, August 1976.

[37] G. Scarano, A. Petroni, M. Biagi, and R. Cusani, "Second-order statistics driven LMS blind fractionally spaced channel equalization," *IEEE Signal Processing Letters*, vol. 24, no. 2, pp. 161–165, Feb 2017.

[38] M. Di Benedetto, *Comunicazioni Elettriche. Approfondimenti.* Prentice Hall, Pearson, 2008.

[39] G. Scarano, *Elaborazione Statistica dei Segnali.* Independently published, vol. 1-2, 2017. [Online]. Available: https://www.amazon.it/Elaborazione-Statistica-dei-Segnali-Vol-I/dp/1549792679

[40] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proceedings of the IEEE*, vol. 64, no. 8, pp. 1151–1162, Aug 1976.

[41] S. Colonnese, G. Panci, and G. Scarano, "Using SOS in blind bussgang equalization for QAM communication systems," in *IEEE 5th Workshop on Signal Processing Advances in Wireless Communications, 2004.*, July 2004, pp. 415–418.

[42] A. Petroni, S. Pergoloni, H. L. Ko, Y. H. Cho, R. Cusani, G. Scarano, and M. Biagi, "On the MIMO multipath channels spatial correlation in shallow water communications," in *OCEANS 2017 - MTS/IEEE Anchorage (to be published)*, Sept 2017.

[43] O. C. Orji, W. Sollner, and L. J. Gelius, *Sea Surface Reflection Coefficient Estimation*, 2013, pp. 51–55.

[44] G. R. Garrison, R. E. Francois, and T. Wen, "Acoustic reflections from arctic ice at 15-300 kHz," *Acoustical Society of America Journal*, vol. 90, pp. 973–984, Aug. 1991.

[45] B. Li, J. Huang, S. Zhou, K. Ball, M. Stojanovic, L. Freitag, and P. Willett, "MIMO-OFDM for high-rate underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 34, no. 4, pp. 634–644, Oct 2009.

[46] J. Cross, J. Zhang, and Y. R. Zheng, "Impact of spatial correlation of fading channels on the performance of MIMO underwater communication systems," in *2011 - MILCOM 2011 Military Communications Conference*, Nov 2011, pp. 430–434.

[47] A. Petroni, S. Pergoloni, H. L. Ko, T. H. Im, Y. H. Cho, and M. Biagi, "PPM space-time block coding vs. spatial multiplexing for shallow water acoustic links," in *OCEANS 2017 - Aberdeen*, June 2017, pp. 1–4.

[48] M. Biagi, S. Rinauro, S. Colonnese, S. Pergoloni, G. Scarano, and R. Cusani, "Trace-orthogonal Pulse Position Modulation space time block coding for underwater links," in *2014 Oceans - St. John's*, Sept 2014, pp. 1–5.

[49] M. Zorzi, "Some results on error control for burst-error channels under delay constraints," *Vehicular Technology, IEEE Transactions on*, vol. 50, no. 1, pp. 12–24, Jan 2001.

[50] J.-W. Lee, J.-P. Kim, J.-H. Lee, Y. S. Jang, K. C. Dho, K. Son, and H.-S. Cho, "An improved ARQ scheme in underwater acoustic sensor networks," in *OCEANS 2008 - MTS/IEEE Kobe Techno-Ocean*, April 2008.

[51] M. Gao, W.-S. Soh, and M. Tao, "A transmission scheme for continuous ARQ protocols over underwater acoustic channels," in *Communications, 2009. ICC '09. IEEE International Conference on*, June 2009.

[52] J. Yu, H. Chen, L. Xie, and J.-H. Cui, "Performance analysis of hybrid ARQ schemes in underwater acoustic networks," in *Oceans - St. John's, 2014*, Sept 2014.

[53] A. Lin, H. Chen, and L. Xie, "Performance analysis of ARQ protocols in multiuser underwater acoustic networks," in *Oceans - Washington, 2015*, Oct 2015.

[54] S. Azad, P. Casari, F. Guerra, and M. Zorzi, "On ARQ strategies over random access protocols in underwater acoustic networks," in *OCEANS, 2011 IEEE - Spain*, June 2011.

[55] P. Kwatra, "ARQ protocol studies in underwater communication networks," in *Signal Processing and Communication (ICSC), 2013 International Conference on*, Dec 2013, pp. 121–126.

[56] R. Ahmed and M. Stojanovic, "Joint power and rate control for packet coding over fading channels," *IEEE Journal of Oceanic Engineering*, vol. 42, no. 3, pp. 697–710, July 2017.

[57] J. M. Morris, "Throughput performance of data-communication systems using automatic-repeat-request (ARQ) error-control schemes," *Naval Research Lab Washington DC*, September 1977.

[58] Evologics, *R-series Specifications*. Evologics, 2015. [Online]. Available: https://www.evologics.de/en/products/USBL/index.html

[59] M. Gao, W. S. Soh, and M. Tao, "A transmission scheme for continuous ARQ protocols over underwater acoustic channels," in *2009 IEEE International Conference on Communications*, June 2009.

[60] S. Azad, P. Casari, and M. Zorzi, "The underwater selective repeat error control protocol for multiuser acoustic networks: Design and parameter optimization," *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 4866–4877, October 2013.

[61] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, Aug 1999.

[62] M. Zhao, C. c. Yin, and X. j. Wang, "A novel dynamic spectrum access algorithm for cognitive radio networks," *Journal of Communications and Networks*, vol. 15, no. 1, pp. 38–44, Feb 2013.

[63] L. Yang, H. Kim, J. Zhang, M. Chiang, and C. W. Tan, "Pricing-based decentralized spectrum access control in cognitive radio networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 2, pp. 522–535, April 2013.

[64] H. Shokri-Ghadikolaei, F. Sheikholeslami, and M. Nasiri-Kenari, "Distributed multiuser sequential channel sensing schemes in multichannel cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2055–2067, May 2013.

[65] M. Tao and Y. Liu, "Spectrum leasing and cooperative resource allocation in cognitive ofdma networks," *Journal of Communications and Networks*, vol. 15, no. 1, pp. 102–110, Feb 2013.

[66] N. Kundargi and A. Tewfik, "Doubly sequential energy detection for distributed dynamic spectrum access," in *2010 IEEE International Conference on Communications*, May 2010, pp. 1–5.

[67] W. Yonggang, T. Jiansheng, P. Yue, and H. Li, "Underwater communication goes cognitive," in *OCEANS 2008*, Sept 2008, pp. 1–4.

[68] A. O. Bicen, A. B. Sahin, and O. B. Akan, "Spectrum-aware underwater networks: Cognitive acoustic communications," *IEEE Vehicular Technology Magazine*, vol. 7, no. 2, pp. 34–40, June 2012.

[69] D. Torres, Z. Charbiwala, J. Friedman, and M. Srivastava, "Spectrum signaling for cognitive underwater acoustic channel allocation," in *2010 INFOCOM IEEE Conference on Computer Communications Workshops*, March 2010, pp. 1–6.

[70] N. Baldo, P. Casari, and M. Zorzi, "Cognitive spectrum access for underwater acoustic communications," in *ICC Workshops - 2008 IEEE International Conference on Communications Workshops*, May 2008, pp. 518–523.

[71] P. H. Dahl, J. Miller, D. H. Cato, and R. K. Andrew, "Underwater ambient noise," vol. 3, 01 2007.

[72] N. Kottege, F. Kroon, R. Jurdak, and D. Jones, "Classification of underwater broadband bio-acoustics using spectro-temporal features," in *Proceedings of the Seventh ACM International Conference on Underwater Networks and Systems*, ser. WUWNet '12. New York, NY, USA: ACM, 2012, pp. 19:1–19:8. [Online]. Available: http://doi.acm.org/10.1145/2398936.2398961

[73] K. Fristrup and W. A. Watkins, "Marine animal sound classification," vol. 97, 05 1995.

[74] W. Shen, H. Sun, E. Cheng, W. Su, and Y. Zhang, "Narrowband interference suppression in underwater acoustic OFDM system," in *2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR 2010)*, vol. 1, March 2010, pp. 496–499.

[75] X. Guo and J. Zhuo, "Coherent interference suppression for near-field underwater acoustic imaging," in *2012 IEEE 11th International Conference on Signal Processing*, vol. 3, Oct 2012, pp. 2080–2084.

[76] E. Cho, Steve, H. Song, and W. S. Hodgkiss, "Adaptive interference suppression and cancellation for underwater acoustic communications," vol. 129, April 2011.

[77] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, Feb 2005.

[78] M. Biagi, A. Petroni, S. Colonnese, R. Cusani, and G. Scarano, "On rethinking cognitive access for underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. PP, no. 99, 2016.

[79] M. Biagi, S. Rinauro, and R. Cusani, "UWA interference analysis for cognitive access," in *2013 MTS/IEEE OCEANS - Bergen*, June 2013, pp. 1–5.

[80] M. Stojanovic, "Low complexity OFDM detector for underwater acoustic channels," in *OCEANS 2006*, Sept 2006, pp. 1–6.

[81] M. Carmen Domingo, "Overview of channel models for underwater wireless communication networks," vol. 1, pp. 163–182, 09 2008.

[82] M. B Porter, "The Bellhop manual and user's guide: Preliminary draft," 01 2011.

[83] A. Garcia-Armada, "SNR gap approximation for m-psk-based bit loading," *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 57–60, Jan 2006.

[84] M. Naraghi-Pour and T. Ikuma, "Autocorrelation-based spectrum sensing for cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 2, pp. 718–733, Feb 2010.

[85] K. W. Choi, "Adaptive sensing technique to maximize spectrum utilization in cognitive radio," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 2, pp. 992–998, Feb 2010.

[86] J. Ville, *Theorie et Applications de la Notion de Signal Analytique*. Cables et Transmission, 1948.

[87] L. D. Paarmann and M. D. Najar, "Analysis of the wigner-ville transform of periodic signals," in *Proceedings of IEEE-SP International Symposium on Time- Frequency and Time-Scale Analysis*, Oct 1994, pp. 68–71.

[88] J. C. Moss, P. G. Adamopoulos, and J. K. Hammond, "Time-frequency distributions: a modification applied to the pseudo-wigner-ville distribution and the spectrogram," in *IEEE International Symposium on Circuits and Systems,*, May 1989, pp. 1256–1259 vol.2.

[89] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov 2004. [Online]. Available: https://doi.org/10.1023/B:VISI.0000029664.99615.94

[90] B. Schölkopf, C. J. C. Burges, and A. J. Smola, Eds., *Advances in Kernel Methods: Support Vector Learning.* Cambridge, MA, USA: MIT Press, 1999.

[91] J. Yang, Y.-G. Jiang, A. G. Hauptmann, and C.-W. Ngo, "Evaluating bag-of-visual-words representations in scene classification," in *Proceedings of the International Workshop on Workshop on Multimedia Information Retrieval*, ser. MIR '07. New York, NY, USA: ACM, 2007, pp. 197–206. [Online]. Available: http://doi.acm.org/10.1145/1290082.1290111

[92] V. Vapnik, Ed., *Statistical Learning Theory.* New York: John Wiley, 1998.

[93] P. Bartlett and J. Shawe-Taylor, "Advances in kernel methods," B. Schölkopf, C. J. C. Burges, and A. J. Smola, Eds. Cambridge, MA, USA: MIT Press, 1999, ch. Generalization Performance of Support Vector Machines and Other Pattern Classifiers, pp. 43–54. [Online]. Available: http://dl.acm.org/citation.cfm?id=299094.299098

[94] M. B. qnd C.M. Shetty, *Nonlinear Programming.* New York: Wiley Publishing, 1979.

[95] Q. Chen, W. C. Wong, M. Motani, and Y. C. Liang, "MAC protocol design and performance analysis for random access cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 11, pp. 2289–2300, November 2013.

[96] H. V. Poor, *An Introduction to Signal Detection and Estimation (2Nd Ed.).* New York, NY, USA: Springer-Verlag New York, Inc., 1994.

[97] S. M. Reilly and G. Potty, "Verification tests for hybrid gaussian beams in spherical/time coordinates," *Report to the High Fidelity Active Sonar Training (HiFAST) Program*, 10 January 2012.

[98] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics.* Springer, 2011.

[99] "ESONET measurements campaigns." [Online]. Available: http://www.esonet-noe.org

[100] C. Soares, C. Martins, F. Zabel, and A. Silva, "On the applications of a compact autonomous acoustic recorder," in *OCEANS 2011 IEEE - Spain*, June 2011, pp. 1–5.

[101] W. H. O. Institution, "Challenger deep," in *Project, available at http://www.whoi.edu/visualWHOI/ocean-robots-challenger-deep*, Sept 2014.

[102] E. Demirors, G. Sklivanitis, T. Melodia, S. N. Batalama, and D. A. Pados, "Software-defined underwater acoustic networks: toward a high-rate real-time reconfigurable modem," *IEEE Communications Magazine*, vol. 53, no. 11, pp. 64–71, Nov 2015.

[103] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: research challenges," *Ad Hoc Networks*, vol. 3, no. 3, pp. 257 – 279, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1570870505000168

[104] J. Heidemann, M. Stojanovic, and M. Zorzi, "Underwater sensor networks: applications, advances and challenges," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1958, pp. 158–175, 2011. [Online]. Available: http://rsta.royalsocietypublishing.org/content/370/1958/158

[105] M. Biagi, S. Rinauro, S. Colonnese, S. Pergoloni, R. Cusani, and G. Scarano, "Near-sea multi-target opportunistic multiple-input multiple-output detection," in *2014 Oceans - St. John's*, Sept 2014.

[106] A. Sutin, H. Salloum, M. DeLorme, N. Sedunov, A. Sedunov, and M. Tsionskiy, "Stevens passive acoustic system for surface and underwater threat detection," in *2013 IEEE International Conference on Technologies for Homeland Security (HST)*, Nov 2013.

[107] D. Thomson and S. Dosso, "AUV localization in an underwater acoustic positioning system," in *2013 MTS/IEEE OCEANS - Bergen*, June 2013.

[108] W. A. P. van Kleunen, K. C. H. Blom, N. Meratnia, A. B. J. Kokkeler, P. J. M. Havinga, and G. J. M. Smit, "Underwater localization by combining time-of-flight and direction-of-arrival," in *OCEANS 2014 - TAIPEI*, April 2014.

[109] M. Biagi, S. Rinauro, S. Colonnese, R. Cusani, and G. Scarano, "Understanding interference by cognitive nodes: The underwater case," in *2014 Oceans - St. John's*, Sept 2014.

[110] H. V. Thurman, *Introductory Oceanography*, ser. New Jersey, USA. Prentice Hall College, 1997.

[111] M. P. Hayes and P. T. Gough, "Synthetic aperture sonar: A review of current status," *IEEE Journal of Oceanic Engineering*, vol. 34, no. 3, pp. 207–224, July 2009.

[112] D. Alexandrou, D. Pantzartzis, and Z. H. Michalopoulou, "A hybrid sonar design for shallow water bathymetry," in *OCEANS 91 Proceedings*, vol. 1, Oct 1991, pp. 567–571.

[113] A. Asada, T. Ura, H. Koyama, T. Sakamaki, Y. Nose, T. Obara, and K. Nagahashi, "Expanded interferometry and synthetic aperture applied to a side scanning sonar for seafloor bathymetry mapping," in *Proceedings of the 2004 International Symposium on Underwater Technology (IEEE Cat. No.04EX869)*, April 2004, pp. 47–50.

[114] C. Barbu, W. Avera, D. Bibee, M. Harris, and C. Steed, "AQS-20 sonar processing enhancement for bathymetry estimation," in *Proceedings of OCEANS 2005 MTS/IEEE*, Sept 2005, pp. 2025–2029 Vol. 3.

[115] A. A. Saucan, C. Sintes, T. Chonavel, and J. M. L. Caillec, "Enhanced sonar bathymetry tracking in multi-path environment," in *2012 Oceans*, Oct 2012, pp. 1–8.

[116] D. Alexandrou and D. Pantzartzis, "A methodology for acoustic seafloor classification," *IEEE Journal of Oceanic Engineering*, vol. 18, no. 2, pp. 81–86, Apr 1993.

[117] B. Chakraborty, V. Mahale, G. Navelkar, and R. G. P. Desai, "Seafloor classification using acoustic backscatter echo-waveform - artificial neural network applications," in *OCEANS 2006 - Asia Pacific*, May 2006.

[118] J. H. Clarke, "Toward remote seafloor classification using the angular response of acoustic backscattering: a case study from multiple overlapping GLORIA data," *IEEE Journal of Oceanic Engineering*, vol. 19, no. 1, pp. 112–127, Jan 1994.