# RELIABLE UNDERWATER ACOUSTIC VIDEO TRANSMISSION TOWARDS HUMAN-ROBOT DYNAMIC INTERACTION

by

**MEHDI RAHMATI**

**A dissertation submitted to the**

**School of Graduate Studies**

**Rutgers, The State University of New Jersey**

**In partial fulfillment of the requirements**

**For the degree of**

**Doctor of Philosophy**

**Graduate Program in Electrical and Computer Engineering**

**Written under the direction of**

**Professor Dario Pompili**

**And approved by**

____________________________________

____________________________________

____________________________________

____________________________________

**New Brunswick, New Jersey**

**October, 2020**

## ABSTRACT OF THE DISSERTATION

# Reliable Underwater Acoustic Video Transmission Towards Human-Robot Dynamic Interaction

**By MEHDI RAHMATI**

**Dissertation Director:**

**Professor Dario Pompili**

In the past decade, underwater communications have enabled a wide range of applications; there are, however, novel applications and systems—such as coastal multimedia surveillance, oil pipe/bridge inspection, water-quality/marine-pollution monitoring, video monitoring of geological/biological processes from seafloor to air-sea interface, and Underwater Internet of Things (UW IoT)—that require near-real-time multimedia acquisition, classification, and transmission.

Wireless acoustics is the typical physical-layer communication technology for underwater data transmission for distances above a hundred meters; transmitting videos wirelessly underwater using acoustic waves, however, is a very challenging task as the underwater acoustic channel suffers from time-varying attenuation and fading, limited bandwidth, Doppler spreading, high propagation delay, and high bit error rate. For these reasons, state-of-the-art acoustic communication solutions are still mostly focusing on enabling delay-tolerant, low-bandwidth/low-data-rate scalar data transmission or at best low-quality/low-resolution multimedia streaming in the order of few tens of $\mathrm{kbps}$.

On the other hand, while conventional underwater acoustic modems with their fixed-hardware designs hardly meet the data rate and flexibility needed to support video requirements for futuristic multimedia and UW IoT-driven applications, novel algorithms and protocols can be implemented on reconfigurable software-defined architectures so as to perform in-network analysis and/or to transmit a high volume of data to a remote node depending on the environment and system specifications.

For these reasons, the objectives of this research—which led to this doctoral dissertation—were to propose solutions to overcome the limitations of existing acoustic communication techniques and to support robust, reliable, and high-data-rate underwater multimedia transmission. In particular, these objectives were achieved by:

- Developing a new physical-layer solution based on multiple-antenna arrays and Acoustic Vector Sensors (AVSs) and by proposing an underwater acoustic Non-Contiguous Orthogonal Frequency Division Multiplexing (NC-OFDM) technique, called Signal-Space-Frequency Beamforming (SSFB), to boost the data rate for underwater acoustic transmission so as to transfer high-resolution videos.

- Designing a probabilistic Medium Access Control (MAC) solution by introducing a novel underwater Space Division Multiple Access (SDMA) method to share reliably the space among the steered vehicles so as to reduce the acoustic interference in underwater sparse networks.

- Improving the reliability and the quality of multimedia delivery by designing a reliable closed-loop hybrid Automatic Repeat Request (ARQ) coding specifically designed for the harsh underwater environment, and by introducing an efficient and agile collaborative coding strategy to allocate appropriate resources to the communication links based on their status.

- Enhancing the video quality via a cross-layer design for underwater scalable coded videos that are channel compatible, and leveraging the multiplexing-diversity tradeoff in a Multiple Input Multiple Output (MIMO) structure to adjust the video scalability by trading off in real time transmission data rate and reliability according to the user Quality of Service (QoS).

- Presenting a protocol for underwater in-network imagery analysis and monitoring the accumulation of litter and plastic debris at the seafloor using partial information collected by various vehicles around the scene, and using Scalable Video Coded (SVC) multicasting for underwater real-time map reconstruction.

- Proposing a correlation-aware hybrid ARQ technique that leverages the redundancy in the data arising from spatial and temporal correlations of the measured phenomenon; this novel technique can be used in futuristic UW IoT applications with high-density deployed nodes in shallow water.

## Acknowledgements

I would like to express my deepest appreciation and warmest gratitude to my Ph.D. advisor, Dr. Dario Pompili, for his constant support, guidance, patience, and encouragement throughout my doctoral studies. Also, I would like to extend my gratitude to Dr. Payman Arabshahi, University of Washington, as well as to Dr. Jingang Yi and Dr. Jorge Ortiz, Rutgers University, for accepting to serve as committee members in my Ph.D. thesis proposal first, and then in the final dissertation defense.

I would like to thank all the Rutgers CPS Lab members and collaborators, Parul Pandey, Zhuoran Qi, Vidyasagar Sadhu, Xueyuan Zhao, Abolfazl Hajisami, Parsa Hosseini, Tuyen Tran, and Ayman Younis. My sincerest thanks go also to all the co-authors of my publications. I am also in debt to the many Rutgers undergraduate and graduate students who helped me in my work, especially Mohammad Nadeem, Agam Modasiya, Karun Kanda, Archana Arjula, Wenjie Chen, Tomasz Brzyzek, Adam Gurney, Seth Karten, and Prince Bose.

I would particularly like to thank the Rutgers Electrical and Computer Engineering (ECE) graduate director, Dr. Zoran Gajic, the ECE graduate program coordinator, Christy Lafferty, and all the other ECE staff for all the help and support throughout my journey at Rutgers.

I would like to thank Dr. Richard Lathrop from the Rutgers Department of Ecology, Evolution & Natural Resources, Sara Malone from the Rutgers Raritan River Consortium (R3C), and Jeffrey Zeszotarski, Rutgers aquatics coordinator, for their help and support during the many field experiments aimed at collecting useful data to compile this thesis.

I am grateful for the financial support from the Rutgers Department of ECE and from the National Science Foundation (NSF), under the awards 1054234, 1739315, and 1763964, which have provided me with the necessary resources to carry on research and finish this dissertation.

Last, but not least, I would like to thank my family for all the love, support, understanding, and encouragement throughout these years of study.

# Dedication

*To my beloved wife and to my family*

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

This chapter provides an overview of the basic concepts, challenges, and applications that motivate this dissertation and explains the objectives and contributions of this research. The key requirements of underwater acoustic communications are discussed in Sect. 1.1. A brief review of the required tools and underwater vehicles to capture and transmit the video are presented in Sect. 1.2. As an application, the need for video capturing and transmission for marine pollution monitoring is introduced in this section. The research objectives, dissertation contributions, and thesis organization are explained in Sect. 1.3.

## 1.1 Underwater Acoustic Communications: Key Requirements

In the past decade, underwater communications and networked-vehicles have enabled a wide range of applications such as oceanographic data collection, ocean pollution assessment, disaster/tsunami prevention, and assisted navigation [95]. Traditionally, monitoring techniques deploy underwater sensors for data recording during the mission which are neither real-time nor interactive (between onshore control systems and the monitoring instruments), and therefore, easily could lead to the complete failure of a monitoring mission [10]. There are, however, novel underwater monitoring applications and systems based on human-robot dynamic interaction—such as coastal and tactical multimedia surveillance, undersea/offshore exploration, oil pipe/bridge inspection, video monitoring of geological/biological processes from the seafloor to the air-sea interface—that require real-time (or near-real-time) multimedia acquisition and classification. These systems should be able to capture multimedia data, store, process, and compress it while it is being transmitted. Underwater multimedia transmission will need to provide support and variable service to these applications with different Quality of Service (QoS) requirements ranging from delay sensitive to delay tolerant and from loss sensitive to loss tolerant [95].

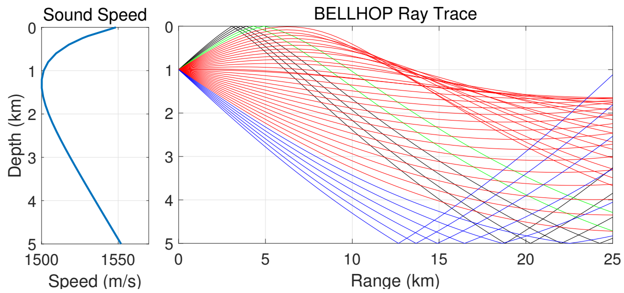



Figure 1.1: Bellhop ray tracing (right box) for a standard sound speed profile (left box) indicating how acoustic beams travel through the underwater acoustic channel when the transmitter is at a depth of $\sim 0.9\ \mathrm{km}$.

Efforts have been made at the physical layer to enable video transmission by boosting the data rate and the reliability via wireless acoustic communication technology because of the high medium absorption of Radio Frequency (RF) and of the scattering problem affecting optical wireless communications underwater. However, reliable and high-quality video transmission via acoustic waves is hard to accomplish underwater as the acoustic waves suffer from limited bandwidth, frequency-dependent transmission loss, time variability, large Doppler spreading, high propagation delay, non-Gaussian background noise, etc. [10]. Temperature, salinity, and pressure of the body of water affect the sound speed and change it between $1450\ \mathrm{m/s}$ and $1540\ \mathrm{m/s}$. Therefore, small changes in speed lead to temporally and spatially variability of the acoustic channel and to significant changes in the sound propagation in the ocean [20]. Moreover, sound reflection from the surface and bottom and other objects in the water and also other minor effects (such as sound refraction) form a multipath effect propagation which is the result of spatial variability of sound speed regarding depth and location [136]. The multipath geometry relates to the channel configuration, therefore, vertical channels are specified by little time dispersion, whereas horizontal channels may have extremely long multipath spreads, regarding the water depth [10]. Fig. 1.1 shows a typical sound profile at each certain depth for the deep water which was simulated using Bellhop [97]. As shown in the left figure, the sound velocity becomes minimum at a certain depth due to temperature drop. As the depth increases, the hydrostatic pressure dominates the temperature, and therefore, the sound velocity increases. If the sound source is in the minimum speed depth (or in its vicinity), some parts of the sound energy is trapped in the channel while propagating and do not reach the bottom or surface. Therefore, underwater channel propagates the rays like a waveguide [20].

The underwater acoustic path loss can be modeled as $\mathcal{A}(\ell, f) = \mathcal{A}_0 \ell^k a(f)^\ell$ where $\mathcal{A}(\ell, f)$ is

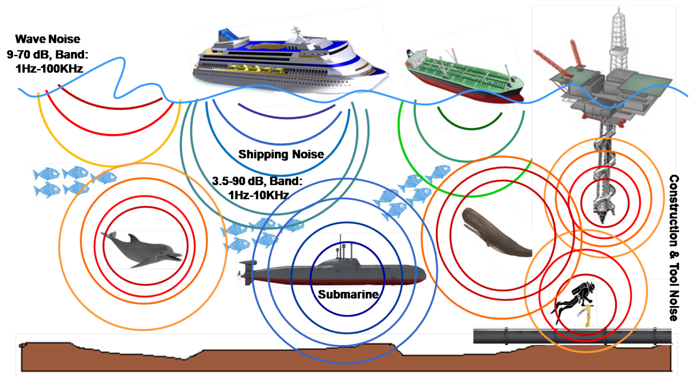



Figure 1.2: Example of underwater dominant man-made and natural acoustic noise sources such as ships, undersea exploration and construction, and their effects on marine life [111].

the experienced path loss (attenuation) on a single path, $\mathcal{A}_0$ is a normalizing constant, $\ell$ [m] is the distance, $f$ [Hz] is a tone of frequency, $k$ is the spreading factor whose value is normally 2, and $a(f)$ is the absorption coefficient [20, pp. 10-12], as $10\log\ a(f) = a'(f) = 0.11\frac{f^2}{1+f^2} + 44\frac{f^2}{4100+f^2} + 2.75\times10^{-4}f^2 + 0.003$ [102, 134]. In this empirical formula, $f$ is in [kHz] and $10\log\ a(f)$ is obtained in [dB/km]. $a(f)$ in [m$^{-1}$] is obtained as $a(f) = 10^{a'(f/1000)/10000}$ where $f$ is in [Hz]. Note that this definition of path loss does not consider the total power. When considering multiple propagation paths, the signal at the receiver is the outcome of several delayed signals of the original signal. The multipath effect (depends on the frequency and transmission range) causes Inter Symbol Interference (ISI) in the underwater acoustic channel which can affect up to hundred of symbols depending on the symbol rate. The mechanism of multipath differs in deep and shallow environments. In shallow water, reflection from the surface and bottom (or other obstacles) define the multipath, while in deep water, ray bending defines this effect based on the sound-speed profile. Acoustic rays bend towards the region of lower acoustic speed which is called "laziness law". The bending effect can be observed at distances of a few kilometers; however, staying within a short/medium range (less than $2$ km), such bending is not notable. Underwater acoustic links can be categorized according to their ranges as very long, long, medium, short, and very short [9], as explained in table 1.1.

Recently, deep-sea audio recording by researchers in the National Oceanic and Atmospheric Administration (NOAA) and their partners has revealed that the ocean is not a quiet place. Instead

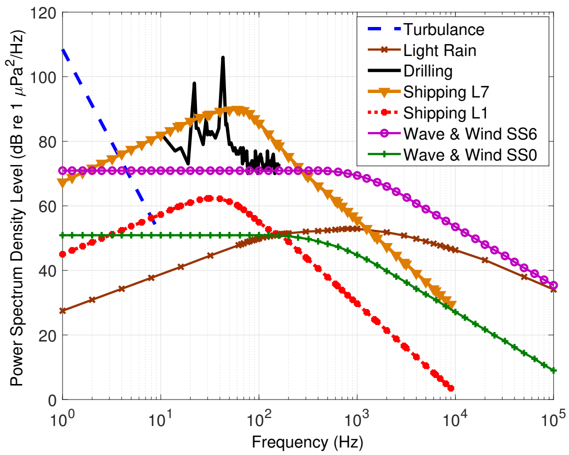



Figure 1.3: Power Spectral Density (PSD) of acoustic noise sources, as reported in [53] and [119], shipping noise based on the traffic/activities, and on the Sea States (SS).

Table 1.1: Typical Range and Bandwidth of Underwater Channels.

| | **Range** (km) | **Bandwidth** (kHz) |
|---|---|---|
| Very long | $1000$ | $< 1$ |
| Long | $10 - 100$ | $2 - 5$ |
| Medium | $1 - 10$ | $\approx 10$ |
| Short | $0.1 - 1$ | $20 - 50$ |
| Very short | $< 0.1$ | $> 100$ |

of finding silence, even the deepest part of the world's ocean, at the bottom of the Mariana Trench[1] with a depth of more than $36,000$ feet, is an incredibly noisy place [56]. Shallow water and coastal regions are also not noise free and are occupied with several noise sources such as impulsive colored noise created by large populaces of snapping shrimp inhabiting these regions Generally, two main categories of noise sources can be recognized; natural anthropogenic sources, as shown in Fig. 1.2. Natural noise is a phenomenon already present in the environment and marine animals are almost adapted to it; however, its effect on marine mammals can be investigated [106]. Sources such as seismic [47], wave, and rain are often high power and occupy the same frequency band as those used by marine animals, as reported in Fig. 1.3.

Current state-of-the-art acoustic communication solutions are still mostly focusing on enabling

[1]The Mariana Trench is located in the western Pacific Ocean, to the east of the Mariana Islands.

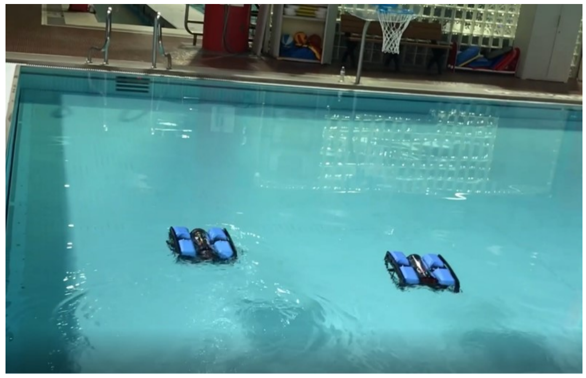


Figure 1.4: Experiments for testing vehicle coordination at Sonny Werblin Recreation Center, Rutgers University, New Jersey.

delay-tolerant, low-bandwidth/low-data-rate scalar data transmission or at best low-quality/low-resolution multimedia streaming in the order of few tens of kbps. Multimedia transmission requires super-effective compression and transmission techniques to reach the required criteria. While traditional commercial acoustic modems with their fixed-hardware designs hardly meet the required data-rate, reliability, and flexibility to support multimedia transmission, other solutions based on adaptive, open source, and reconfigurable architectures employing Software-Defined Acoustic Modems (SDAM) should be utilized.

## 1.2 Video Transmission Using Semi-autonomous Underwater Vehicles

As it was mentioned in the previous section, many futuristic time-critical applications such as multimedia coastal and tactical surveillance, offshore exploration, sea floor mapping, submarine volcanism and hydrothermal vent studies require multimedia data to be retrieved and processed reliably in real time while it is being transmitted for video acquisition and classification [95]. In most underwater missions, Remotely Operated Vehicles (ROVs) are key instruments to support such interactive monitoring applications as they can be equipped with underwater cameras and capture multimedia data from places where humans cannot easily/safely go; however, current underwater vehicles are often tethered to the supporting ship by a cable or have to rise periodically to the surface to communicate with a remote station via RF waves. Tethering is a serious limitation in the development of underwater systems for future applications involving multiple underwater vehicles as it constrains the maneuverability and range of the vehicles engaged in the mission, which run the risk to get tangled and compromise the mission itself. Resurfacing periodically, as shown in Fig. 1.4, on the other hand, does not guarantee interactivity—which is key in real-time applications—and leads to energy/time inefficiencies. Recent research initiatives on human-robot interaction [79, 143] have received considerable attention from both industry and academia in the past few years by developing, analyzing, and validating a proof-of-concept paradigm to maximize the system performance and efficiency of semi-autonomous robots to collect underwater data. To this end, humans/domain experts have used a console to issue low-level commands, where the set of commands is a queue of actions that should be executed consecutively when a robot goes underwater. This approach, however, does not give any flexibility to the humans to refine or cancel the commands that were issued in the beginning of the mission based on the live multimedia streaming from the vehicle. For these reasons, efforts have been made at the physical layer to enable efficient video transmission solutions via wireless acoustic communication technology because of the high medium absorption of RF and of the scattering problem affecting optical wireless communications underwater.

We envision that, in many of the future interactive monitoring applications, humans will participate dynamically in the fine tuning of the system in a closed-loop fashion based on the Quality of Experience (QoE) of the received multimedia stream. Therefore, a dynamic interaction between a human and robots will be essential in which the human/application expert is included in the loop to take dynamic decisions regarding the mission based on the incoming multimedia stream and informs the robots by issuing fresh commands. These commands can include instructions such as identifying Regions of Interest (RoI) from the video feed, allocating underwater vehicles to different regions of interest, changing physical-layer parameters to improve quality of video feed, performing specific tasks, configuring themselves into a suitable mesh topology given the mission requirements, etc.

*The efforts in this dissertation will revolve around the argument that shifting abruptly from a ROV mode—where low-level commands are imparted to the vehicle—to an Autonomous Underwater Vehicle (AUV) mode—where the vehicle(s) are fully intelligent and autonomous—is neither realistic nor practical; rather, in light of the usual gradual evolution and deployment of new technology, we envision a semi-autonomous mode where high-level/mission-oriented commands are imparted by a human in the loop, who may take advantage of domain expertise for decision making. To enable this vision, a software-defined underwater acoustic structure is needed including several mobile nodes and a team of Hybrid Unmanned Air/Underwater Vehicles (HUA/UV).*

(a) (b)

Figure 1.5: An experiment conducted in (a) the Raritan Canal; (b) the Raritan River, New Jersey.

**Pollution Monitoring as an Application of Underwater Video Transmission:** Marine pollution such as litter and debris, both beached and floating objects/liquids, are one of the most serious and fast growing environmental threats to the oceans and seafloors. The negative impacts of this pollution on the environment and on human and marine life are unquestionable. Marine litter develops from various sources and causes a wide range of environmental safety and health issues. The slow degradation rate of marine litter items, combined with the growing quantity of debris collection, is leading to ocean pollution. When the debris, such as plastic, degrades over time, it turns into micro- and then nano-plastics, which is then consumed by fish and eventually by humans.

According to recent studies [37], around 640,000 tonnes of gear is lost in the ocean annually. Lost nets create a huge threat to marine life as they trap and kill at least 136,000 seals, sea lions, and whales. According to the survey conducted by the United Nations Educational, Scientific and Cultural Organization (UNESCO) [147] over $80\%$ of marine pollution comes from land-based activities. From plastic bags to pesticides, most of the waste produced on land eventually reaches the oceans. Rivers carry the litter with their currents to the seas and are one of the main sources of litter entering the seas. There is litter spread widely throughout the seafloor, but its distribution is usually patchy with densities from $1$ up to around $200$ items per each $10\ \mathrm{m}$, as reported for the Messina Strait's channel—one of the geologically active areas of the Central Mediterranean Sea [93].

To address this issue, using only static sensors attached to fixed monitoring stations with predefined configurations is not a real-time and efficient solution for data collection as the phenomenon of interest may occur sporadically and propagate spatially through the water bodies. Deploying a team of Autonomous and Semi-Autonomous Underwater Vehicles (AUVs)—which is capable of

(a) (b)

Figure 1.6: Videos of pollution in the Raritan River, New Jersey, when the water is murky and the visibility is low, after a heavy rain or flood.

chasing the phenomenon of interest instead of waiting for the pollution to reach the fixed stations—equipped with cameras and other sensors, can help in detecting the pollution on the seafloor and riverbed, as well as the surface, build a map of the pollution, and therefore, can issue early warnings so to reduce the damage to human and aquatic life.

The Raritan River is a major river of central New Jersey and is a unique laboratory available to Rutgers, i.e., a perfect case study. It is also the New Jersey's largest contiguous wildlife corridor offering refuge to numerous threatened and endangered species [41]. This river has experienced pollution from industrial facilities toxic dumping for over 100 years. The watershed is also impacted by contaminated sites and sewage treatment systems. Pollution from contaminated sites leaks into the river and harms the environment and public health. According to the Environmental Protection Agency (EPA) reports [46], over 16 noxious chemicals and solids were found infecting the section of the Raritan River that borders New Brunswick, NJ. Three of those chemicals—arsenic, benzopyrene, and the pesticide heptachlor epoxide—have the potential to adversely affect the drinking water supply. The watersheds should be monitored regularly to provide usable data about water quality and the overall health of the Raritan watersheds. The Raritan Headwaters Association [4] holds a specific stream monitoring program; based on visual assessment and on manual collection of water samples at each site, they can classify coarsely the sites as excellent, good, fair, or poor. *However, more research should be performed to enable streamlined and improved monitoring of such an important area.* As shown in Fig. 1.5, the vehicles are equipped with multiple on-board sensors and are deployed in the Raritan River for gathering scientific data via collaborative strategies. Fig. 1.6 shows two frames of a captured videos from the bottom of the Raritan River.

## 1.3 Dissertation Contributions and Organization

To overcome the mentioned challenges, the following objectives are proposed in this dissertation:

**(1)** ***To design novel communication solutions for robust, reliable, and high-data rate underwater multimedia transmission in the order of hundreds of*** Kbps ***at the operating ranges of the application of interest*** **[110]**. The solution exploits all the resources in the frequency/signal/space domains, and relies on novel vehicle tracking methods, multiple-antenna structure/coding, and/or Acoustic Vector Sensors (AVSs).[2]

**(2)** ***To design a novel Medium Access Control (MAC) based on a probabilistic Space Division Multiple Access (SDMA) method for short/medium distances in order to achieve a reliable coordination among underwater vehicles [108, 113].*** The innovative probabilistic solution focuses on cancelling or alleviating the inter-vehicle interference while the inherent position uncertainty of vehicles is considered in a sparse underwater mobile network. Spatially separable and non-separable scenarios are studied and the parameters are optimized to minimize the statistical interference.

**(3)** ***To design an efficient and agile collaborative strategy to allocate appropriate resources to the communication links in order to achieve high throughput and reliable underwater acoustic networks for transmitting distributed and large volume of data [107, 114].*** The proposed method adjusts the physical and link-layer parameters collaboratively for a Code Division Multiple Access (CDMA)-based underwater network. An adaptive Hybrid Automatic Repeat Request (HARQ) solution is employed to guarantee reliable communications against errors in poor links.

**(4)** ***To develop a reliable acoustic wireless video transmission technique in the extreme and uncertain underwater environment while optimizing the received video quality and the user's experience [104,115].*** The innovative method is an adaptive solution that is specifically designed for Multi-Input Multi-Output (MIMO)-based Software-Defined Acoustic Modems (SDAMs). Cross-layer techniques utilizing diversity-spatial multiplexing and Unequal Error Protection (UEP) are implemented along with the scalable video compression at the application layer to keep the video distortion under an acceptable threshold and to achieve a high physical-layer throughput.

---

[2]While a regular hydrophone can only measure the scalar acoustic pressure, an Acoustic Vector Sensor (AVS)—which is formed by a hydrophone and three accelerometers—measures all the three components of the acoustic particle motion (i.e., pressure, velocity vector, and direction) and can estimate the angle of arrival of the acoustic wave in addition to its scalar pressure value. AVSs can be constructed using a variety of technologies from mechanically or optically based on Micro-Electro-Mechanical Systems (MEMS).

**(5)** ***To design a system for monitoring the accumulation of litter and plastic debris at the bottom of rivers using a team of Autonomous Underwater Vehicles (AUVs) that exchange the recorded video in order to reconstruct the map of regions of interest [112].*** The solution focuses on in-network scalable underwater video sharing between AUVs and develops a framework for underwater imagery analysis using partial information collected by various vehicles around the scene. This method maximizes Quality of Service (QoS) via an innovative multicasting scalable coded video, while achieving the maximum Quality of Experience (QoE) for the scene reconstruction.

**(6)** ***To develop a reliable and persistent water monitoring technique in smart Underwater Internet of Things (UW IoT) [116].*** The solution realizes a reliable correlation-based HARQ to transmit data between the buoys and the fusion center such as a drone. This method leverages the correlation of the data to avoid costly retransmissions in a chaotic Direct Sequence Spread Spectrum (DSSS) system that guarantees secure buoy-drone transmissions.

The solutions that will be proposed in this dissertation will impact several areas of research since it has a strong multidisciplinary component that involves a nexus of ideas from sensor technology, communications, coding, networking, algorithms, statistical inference, and dynamical systems. The novel methods and algorithms will have a wide applicability in the areas of science and technology that concern $(i)$ real-time multimedia transmission of coordinated robots and $(ii)$ the study of dynamic interaction of such robots with their environment and the humans.

**Dissertation Organization:** This dissertation will be geared towards enabling our vision of reliable and high-speed underwater acoustic multimedia transmission towards dynamic human-robot interaction. This rest of this dissertation is organized as follows.

**Chapter 2** describes the work in physical layer by exploiting multiple-antenna arrays and AVSs. A novel solution is proposed to boost the data rate for underwater acoustic transmission so as to transfer high-resolution video underwater.

**Chapter 3** follows a novel probabilistic approach to design an efficient MAC-layer solution to share reliably the space among the steered vehicles so as to reduce the acoustic interference.

**Chapter 4** discusses the data transmission reliability by applying a robust closed-loop hybrid Automatic Repeat Request (ARQ) coding technique based on the structure of the nodes and the possible collaboration between them in the field.

**Chapter 5** proposes an adaptive cross-layer solution for transmitting underwater scalable-coded video using a MIMO-based reconfigurable Software-Defined Acoustic Modem (SDAM). Multiplexing diversity tradeoff is navigated in order to balance the transmission data rate and the reliability.

**Chapter 6** introduces a novel framework for underwater imagery analysis and multicasting using scalable coded video and partial information collected by various vehicles around a scene in order to reconstruct the map of environment.

**Chapter 7** proposes a novel architecture for UW-IoT and a correlation-aware hybrid ARQ that leverages the correlation in the data to avoid costly retransmissions and thereby enable timely reconstruction of the phenomenon.

**Chapter 8** summarizes the main contributions of this dissertation and presents suggestions on future research directions that will push the state-of-the-art in underwater video transmission.

# Chapter 2

# Signal-Space-Frequency Beamforming for Underwater Acoustic Video Transmission

In this chapter, a hybrid solution that is capable of transmitting at high data rates underwater via acoustic waves at *short/medium distances* is proposed. This physical-layer signaling solution introduces a novel method, called Signal-Space-Frequency Beamforming (SSFB), for a multiple-hydrophone structure where each one consists of Uniform Circular Array (UCA) hydrophones (mounted on an underwater vehicle) to steer the beam in both azimuth and elevation planes; then, an array of Acoustic Vector Sensors (AVS)[1] are mounted on the surface buoy. Detection is performed based on the beam spatial separation and direction of arrival angles' estimation. Simulation results confirm that this solution outperforms state-of-the-art underwater acoustic transmission techniques, whose data rates are limited only to few tens of kbps.

## 2.1 Overview

Underwater networks enable a wide range of applications such as oceanographic data gathering, pollution monitoring, disaster prevention, and assisted navigation—just to name a few—in which *mostly scalar values* are sensed from the environment and transmitted to an onshore or surface station. However, many futuristic time-critical applications such as multimedia coastal and tactical surveillance, offshore exploration, sea floor mapping, submarine volcanism and hydrothermal vent studies require multimedia data to be retrieved and processed reliably in real time while it is being transmitted for video acquisition and classification [95].

For many of these futuristic applications, transmitting reliably videos underwater is a challenging problem in the environment in which Radio-Frequency (RF) waves are absorbed for distances

[1] Acoustic Vector Sensors (AVS) are hydrophones that are able to capture the acoustic particle velocity/direction of arrival in addition to measuring regular scalar pressure.

above a few tens of meters, optical waves require narrow laser beams and suffer from scattering and ocean wave motion, and acoustic waves—while being able to propagate up to several tens of kilometers—lead to a communication channel that is very dynamic, prone to fading, spectrum limited with passband bandwidths of only a few tens of kHz due to high transmission loss at frequencies above $50\ \mathrm{kHz}$, and affected by a non-Gaussian noise [136].

In most cases, Autonomous/Remotely Operated Underwater Vehicles (AUVs/ROVs) are key enabling instruments to support such futuristic applications as they can be equipped with cameras. However, current underwater vehicles are often tethered to the supporting ship by a high-data-rate fiber cable or have to surface periodically to communicate with a remote onshore station via terrestrial RF waves. Tethering is a serious limitation for the development of underwater systems for multimedia applications involving one or more underwater vehicles as it constrains severely the maneuverability and range of the vehicles, which run the risk to get tangled and compromise their mission. Resurfacing periodically, on the other hand, does not guarantee interactivity, which is key in real-time applications, and leads to energy/time inefficiencies.

**Challenges:** Although acoustic communication is the typical physical-layer technology underwater for distances above a hundred meters, yet, achieving high data rates for video transmission through the acoustic channel is hard to accomplish as acoustic waves suffer from attenuation, limited bandwidth, Doppler spreading, high propagation delay, and time-varying propagation characteristics [109, 136]. For these reasons, state-of-the-art acoustic communication solutions are still mostly focusing on enabling delay-tolerant, low-bandwidth/low-data-rate transmission or at best low-quality/low-resolution multimedia streaming in the order of few tens of $\mathrm{kbps}$.

To achieve higher data rates in the bandwidth-limited underwater acoustic channel, several techniques should be combined together. For example, signal beamforming along with multiple antenna arrays [142] could achieve this goal; however, the main challenge is the position uncertainty of the users, which leads to inaccuracies in the estimation of beam angles—a key piece of information in beamforming—and therefore to overall performance degradation. The problem becomes even worse over time if the vehicle remains underwater for long because of the accumulation of its position error, which leads to non-negligible drifts in the vehicle's position estimation, as attested by many works on underwater localization [27, 70, 141].

**State of the Art:** Over the past few years, researchers have come up with advancements in sensor technology that seem quite promising to overcome the limitations of traditional scalar hydrophones, which detect the acoustic pressure without any directional sensitivity. For example, *Acoustic Vector Sensor (AVS)* array consists of hydrophones that are able to capture the acoustic particle velocity/angle of arrival [128] in addition to measuring regular scalar pressure. This interesting characteristic can be used in determining the position of an acoustic source. Source localization using an array of AVSs is performed in [103] for multipath scenarios. This device can be exploited in a broad range of environments and is constructed using a variety of mechanical, optical, and Micro-Electro-Mechanical Systems (MEMS) technologies [33].

**Contributions:** In this research, we exploit the potential of this recent sensor technology and present a novel AVS-based method to increase the effective data rate of underwater acoustic communications, i.e., to support reliable and high-data-rate video transmissions (in the order of hundreds of $\mathrm{Kbps}$ at the operating ranges of the application of interest, i.e., up to a few kilometers). To achieve this goal, we propose a novel signaling method, called *Signal-Space-Frequency Beamforming (SSFB)*, that makes use of multiple domains to leverage the benefits of AVS. A novel arrangement for the vehicle's antenna is presented, while the surface buoy (receiver) is equipped with AVS hydrophones and detects the signal based on the estimated direction of arrival angles. In addition to modulation, each antenna in a multiple antenna structure and also each subcarrier in Orthogonal Frequency Division Multiplexing (OFDM) system participate in the data rate increase, while inter-antenna-interference is avoided by a Non-Contiguous OFDM (NC-OFDM) technique specifically designed for this system to support video transmission.

**Chapter Outline:** The rest of this chapter is organized as follows. In Sect. 2.2, a review on the related research in the literature is presented. In Sect. 2.3, the proposal and its mathematical framework are discussed. In Sect. 2.4, the performance results are presented, and finally, Sect. 2.5, summarizes the chapter.

## 2.2 Related Work

The first image transmission via acoustic waves occurred in Japan, where the system [139] demonstrated the transmission over a vertical path with a low frame rate. Low-bit-rate video compression is another solution investigated in the literature to combat the limitations of underwater acoustic channels. The authors in [87] presented an algorithm based on the use of data compression/coding implemented and tested over a $10\ \mathrm{m}$ vertical channel with $6090\ \mathrm{kHz}$ bandwidth. The feasibility of video transmission over short-length links was investigated in [118, 148], where MPEG-4 video compression and a wavelet-based transmission method were tested on coded OFDM. A joint optical/acoustic solution was presented in [44], which integrates high-data-rate and low-latency capabilities of optical communications in short transmission ranges with long-distance traveling of acoustics. Another acoustic/optical solution for video streaming was presented in [48], where the acoustic mode is used as backup in case of optical channel failure. Notice that, while optical-based techniques can support high data rates, all the optical solutions reported so far can only transmit at distances below $\sim 50\ \mathrm{m}$ due to scattering and laser-pointing-related issues. A software-defined underwater acoustic platform that supports higher data rates and provides flexibility and scalability for future underwater applications was discussed in [34].

However, despite all these works, the problem of robust video transmission is still unsolved, and achieving high video quality is still a challenge when we consider the limited available bandwidth along with the harsh characteristics of the underwater acoustic channel, which calls for novel high-spectral-efficiency methods. Recently, Non-Contiguous OFDM (NC-OFDM) has attracted the attention of researchers [83] due to its dynamic spectrum access and effective use of spectrum as a scares resource, which increases the spectral efficiency of conventional OFDM while avoiding interference with other users, especially in cognitive radios and frequency-selective channels. The authors in [16] have suggested Index Modulation (IM) as an effective technique for Fifth Generation (5G) wireless networks, in which the indices of the OFDM blocks convey additional information bits. IM can be applied to several modulation schemes such as Spatial Modulation (SM) [78] in order to achieve higher data rates.

Hydroflown sensor [33] is a MEMS-based hydrophone set that is able to measure particle velocity and Angle of Arrival (AoA) [17]. Several algorithms have been proposed for AoA estimation for acoustic vector sensors. Maximum likelihood (ML), as a conventional method, maximizes the likelihood of the received signal from a particular angle. MUltiple SIgnal Classification (MUSIC) [122] is an adaptive eigen-structure-based method that considers the noise subspace, while the signal subspace is considered in the Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT) [120], which assumes a displacement invariance for sensors. Matrix Pencil (MP) is similar to

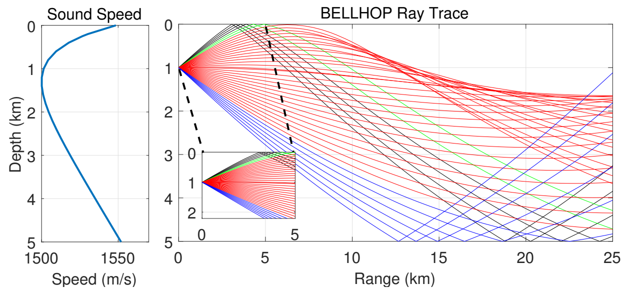



Figure 2.1: Bellhop ray tracing (right) for a standard Sound-Speed Profile (SSP) (left) indicating how acoustic beams travel through the channel [97]. Notice how the beams are almost straight for short/medium ranges (less than $5\ \mathrm{km}$).

ESPRIT but, instead of estimating the correlation matrix, it exploits the spatial samples of the data based on a snapshot-by-snapshot basis, and performs well in non-stationary environments [155].

Despite all these efforts, there are still open problems in signal processing and in the hardware needed to support real-time processing. Considering the characteristics of the underwater acoustic channel, direction-of-arrival estimation might be complicated since the position information of a vehicle is not accurate underwater. Although a vehicle may surface periodically to synchronize itself using Global Positioning System (GPS), which does not work underwater, self-inaccuracies in position estimation increases over time; also, the effect of drifting in ocean currents on the vehicle causes more uncertainties in the position, which leads to error in the angle-of-arrival estimation of the vehicle. These errors will lead to distortion in the video quality at the receiver, which translates to a low Quality of Service (QoS) for the user.

## 2.3 Proposed Architecture and Signaling Method

**System Assumptions:** Let us assume that the transmission occurs at *short/medium ranges*—up to a few kilometers—and that the direct beams are dominant over the reflected ones from the ocean surface/bottom, so that the receiver is not severely affected by multipath. For farther distances—above a few kilometers—and based on the Sound Speed Profile (SSP), the acoustic rays bend towards the region of lower acoustic speed (*"laziness law"*). This effect changes the Angles of Departure/Arrival (AoD/AoA) and their estimations. Using the Bellhop model [97] and considering a typical deep-water case, Fig. 2.1 illustrates the SSP (left) and the acoustic ray tracing (right) in the underwater channel for a sample source at a depth of $1\ \mathrm{km}$ and temperature of $39\ \mathrm{F}$. The bending

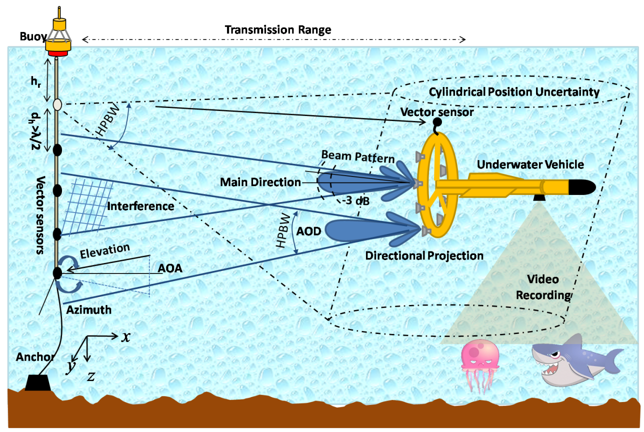



Figure 2.2: System architecture and geometry in which the vehicle is equipped with an Acoustic Vector Sensor (AVS) and a ring of beamforming transmitters, each one containing a Uniform Circular Array (UCA), while the buoy is armed with AVS hydrophones.

effect can be observed by going above a few kilometers; however, staying within a short/medium range, such bending is not notable, which explains the philosophy behind our signaling method in which the vehicle is steered via beamforming. Moreover, propagation delays in acoustic links are five orders of magnitude larger than in terrestrial RF links, so short/medium ranges are more appropriate for video transmission applications.

**Model Descriptions:** As in Fig. 2.2, the model consists of an anchored buoy and a vehicle traveling at a fairly smooth and constant horizontal speed ranging between $0.25$ to $0.5\ \mathrm{m/s}$ to capture data/video. Buoy and vehicle exchange control messages during the communication setup process. To avoid any interference with the data exchange process, we establish a separate control channel via Frequency Division Duplex (FDD); this will not have much impact on the overall data rate and bandwidth of the system as only a few bits per message are used. In this chapter, *downlink* $(BV)$ defines the direction of acoustic communications from the surface buoy, $B$, to the vehicle, $V$; whereas *uplink* $(VB)$ represents data transmission in the opposite direction.

Several transducers are installed on a vertical bar at the buoy's side starting from the depth of $h_r$ with antenna spacing of $d_h$—more than half of the wavelength to avoid spatial correlation. An *acoustic vector sensor array* is embedded on the bar, which gives us the measured scalar pressure and the beam's direction of arrival, in both the azimuth and elevation planes. The scalar response of the pressure sensor, which is omni-directional, is added to the responses of the vector sensors, which

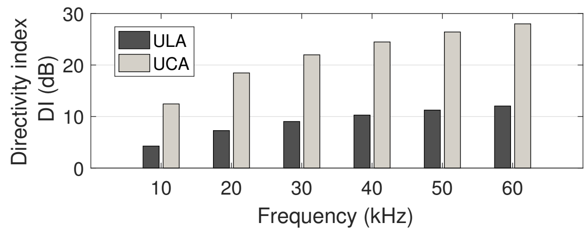



Figure 2.3: Directivity Index (DI) for both Uniform Circular Array (UCA) and Uniform Linear Array (ULA) at different frequencies.

measure particle velocity and output the information about the beam's direction of arrival. The sensitivity, the Directivity Factor ($D_f$), and the Directivity Index ($DI$) of these sensors depend on both the technology and the environment under study at a given range $r$, where $DI = 10 \log D_f / D_{ref}$, in which $D_{ref}$ is the omni-directional reference intensity and $D_f(\phi, \theta)$ can be defined as the ratio of the maximum acoustic intensity to the averaged intensity in all directions [74]. Plane wave is represented by $\phi \in [-\pi, \pi]$ and $\theta \in [-\pi/2, \pi/2]$ as the azimuth and elevation angles, respectively. The vehicle is equipped with *one* vector sensor for localization and control purposes in the downlink, and with *multiple* transmit antennae for sending data via its *circular arranged* acoustic transmitters capable of performing beamforming for the uplink communications. As a result, differently from the conventional linear arrangement, beamforming and direction-of-arrival estimation are performed in both azimuth and elevation planes.

As the array's elements can be placed according to different shapes (linear, circular, or rectangular), *Uniform Circular Array (UCA)* [57, 150] is exploited to leverage its higher performance compared with a Uniform Linear Array (ULA). Since the mobility of the vehicle might change the estimation of angles and, as a result, decrease the reliability of the reception, in this chapter, as an alternative configuration for the linear placement and linear array, a *circular arrangement* and UCA with $D > \lambda$ are proposed. $D$ and $\lambda$ represent the circle's diameter and the acoustic wavelength, respectively. Interestingly, the circular array can form the beam of $2\pi$ in azimuth plane and $\pi$ in elevation plane with little change in either the beamwidth or the sidelobe level [57].

In Fig. 2.3, DIs of UCA and ULA are compared, where DI accounts for the spatial gain in energy as a result of using directive antenna to the same antenna without directivity [74]. It confirms that UCA is a better choice to combat the channel attenuation, at higher acoustic frequencies.

**Proposed Signaling Method:** To take advantage of AVS, the *vehicle's exact location* should be

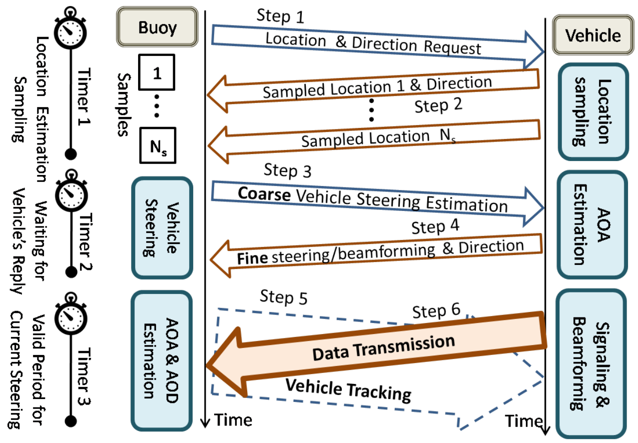



Figure 2.4: Protocol for vehicle steering and control message exchange. Interval estimation is used at the buoy for Angle of Departure (AoD) and coarse estimation of the uncertainty region (Fig. 2.2), while fine steering is performed via extra information, i.e., Angle of Arrival (AoA) estimation, offered by AVS.

determined. If the ocean currents are assumed unknown, the vehicle's drifting in the horizontal plane is identically and independently distributed (i.i.d.) and follows a normal distribution, which makes the horizontal projection of its confidence a circular region. Regarding the vehicle's movement along its trajectory, there is an *uncertainty* in the position of the vehicle. This uncertainty region is shown to be a cylinder [27, 108], as in Fig. 2.2. First, in Sect. 2.3, in order to determine the angles of departure/arrival, we study the uncertainty region and present the proposed protocol. We create spatial pipes towards the vehicle via beamforming based on the uncertainty region and its corresponding angles. Then, based on the estimated angles, the data will be transmitted via the proposed uplink signaling, i.e., SSFB. Afterwards, the receiver design and the discussion on the data rate compared with other methods are presented.

**Vehicle Steering Protocol:** The procedure is divided into six steps as shown in Fig. 2.4.

**Steps 1 & 2—Vehicle's Location Uncertainty Estimation:** We aim at estimating the location of vehicle given the inherent position uncertainty of objects underwater. The process starts by the buoy's request command and at the same time setting timer 1 until all the required location samples are gathered. Upon receiving this message and every $\Delta_t$ seconds, vehicle $V$ samples its current estimated location $loc_n^{(V)}$, $n = 1, ..., N_s$ via *Dead reckoning*, where $N_s$ is the total number of required samples. Timer 1 stops after $\tau_1 \approx N_s \ \Delta_t + t_p + t_t$, where $t_p$ is the time due to the propagation delays and $t_t$ stands for the distance-based transmission delays in short/medium ranges. There are internal- and external-uncertainties [27] in the trajectory and location estimation of the vehicle, which are considered in the next step.

**Step 3—Coarse Buoy's AoD Estimation:** Using $N_s$ vehicle's location samples gathered at the buoy, locations are converted to steering angles $\theta_n$ and $\phi_n$, $n = 1, ..., N_s$, in azimuth and elevation plains to estimate the *angular uncertainty region* of the vehicle via the method introduced in [108]. Let us perform the analysis for one of the planes, i.e., random variable $\theta$ with mean value $\mu$ and standard deviation $\sigma$. The *estimation* with mean value $\bar{\theta}'$ and standard deviation $\sigma'_\theta$ can be derived as, $\bar{\theta}' = \sum_{n=1}^{N_s} \theta_n / N_s$, and $\sigma'_\theta = \left[ 1/(N_{s-1}) \sum_{n=1}^{N_s} \left( \theta_n - \bar{\theta}' \right)^2 \right]^{\frac{1}{2}}$. The buoy's beamwidth is chosen in such a way that it is equal to the *confidence interval* [25] of $\bar{\theta}'$ [108], i.e.,

$$\Pr(\theta_L^{(B)} \leq \bar{\theta}' \leq \theta_U^{(B)}) \geq 1 - \alpha, \tag{2.1}$$

where $\Pr(.)$ represents the probability function, $\theta_L^{(B)}$ and $\theta_U^{(B)}$ are the lower and upper angular boundaries at buoy $B$ pointing at the bottom and top of the uncertainty region, $1 - \alpha$ is the confidence degree [25]. The angles can be calculated by the following equations, given $\mathcal{T}_{N_s-1,\alpha/2}$ is the student's t-distribution critical value with $N_s - 1$ degrees of freedom.

$$\theta_L^{(B)} = \bar{\theta}' - \mathcal{T}_{N_s-1,\alpha/2} \frac{\sigma'_\theta}{\sqrt{N_s}}, \tag{2.2a}$$

$$\theta_U^{(B)} = \bar{\theta}' + \mathcal{T}_{N_s-1,\alpha/2} \frac{\sigma'_\theta}{\sqrt{N_s}}. \tag{2.2b}$$

The buoy forms its estimated *Half Power Beam Width* $\widetilde{W}_\theta^{(BV)}$, in the interval of $\pm\mathcal{T}_{N_s-1,\alpha/2} \dfrac{\sigma'_\theta}{\sqrt{N_s}}$ around $\bar{\theta}'$, as $\widetilde{W}_\theta^{(BV)} = \theta_U^{(B)} - \theta_L^{(B)}$, while *AoD* from buoy towards vehicle is calculated as $\widetilde{\Gamma}_\theta^{(BV)} = \bar{\theta}'$. Similarly, it can be concluded that $\widetilde{W}_\phi^{(BV)} = \phi_U^{(B)} - \phi_L^{(B)}$ and $\widetilde{\Gamma}_\phi^{(BV)} = \bar{\phi}'$. If this calculation takes $t_2$ seconds, then timer 2 stops after $\tau_2 \approx t_2 + 2t_t + 2t_p$, when the estimation is sent back to the buoy.

**Step 4—Fine Steering Estimation Using AVS:** The angular estimation extracted via the vehicle's *vector sensor* is fed to vehicle's beamformers and is simultaneously reported to the buoy to be used as reference in its tracker. Tuning the antenna from vehicle to buoy (i.e., $W_\theta^{(VB)}$, $\Gamma_\theta^{(VB)}$, $W_\phi^{(VB)}$, and $\Gamma_\phi^{(VB)}$ as depicted in Fig. 2.5) is performed via the angles measured at vehicle's vector sensor, and is refined based on the prior coarse estimations, i.e., $\widetilde{W}_\theta^{(BV)}$, $\widetilde{\Gamma}_\theta^{(BV)}$, $\widetilde{W}_\phi^{(BV)}$, and

$\widetilde{\Gamma}_{\phi}^{(BV)}$, and the trajectory vector. Several methods were suggested for AoA estimation in the literature, ranging from correlation [155] and ML [132] to MUSIC [122] and ESPRIT [120], based on the assumptions and characteristics of the used elements. Generally, AoA estimation at the vehicle can be written as follows,

$$\mathbf{y}^{(V)}(t) = \mathbf{A}^{(V)}\left(\Psi^{(V)}\right) \mathbf{x'}^{(BV)}(t) + \mathbf{z}(t), \tag{2.3}$$

where $\mathbf{y}^{(V)}(t)$ is the signal at the vehicle's antenna, $\mathbf{x'}^{(BV)}(t)$ is the channel affected vector of signals from buoy to vehicle $\mathbf{x}^{(BV)}(t)$, $\mathbf{z}(t)$ is the underwater noise vector with the covariance matrix $\mathbf{Q}_z$, and $\mathbf{A}^{(V)}$ is the steering vector at the vehicle as a function of unknown parameter $\Psi^{(V)}$. In our case, $\Psi^{(V)} = [\theta^{(V)}, \phi^{(V)}]$, with unknown arriving angles and the estimated angles of $\widehat{\theta}^{(V)}$ and $\widehat{\phi}^{(V)}$ are geometrically proportional to $\widehat{\theta}_{op}^{(B)}$ and $\widehat{\phi}_{op}^{(B)}$ at the buoy. Regarding the coarse estimation, the optimum angles can be bounded as,

$$-\widetilde{W}_{\theta}^{(BV)}/2 < \widehat{\theta}_{op}^{(B)} < \widetilde{W}_{\theta}^{(BV)}/2, \tag{2.4a}$$

$$-\widetilde{W}_{\phi}^{(BV)}/2 < \widehat{\phi}_{op}^{(B)} < \widetilde{W}_{\phi}^{(BV)}/2. \tag{2.4b}$$

The scanning range in (2.4) is now narrower by the half power beamwidth instead of the whole angular range. Estimation variance is lower bounded by the *Cramer-Rao Bound (CRB)* and upper bounded by the variance of the coarse estimation. We describe the equation for azimuth plane as,

$$\mathbf{J}_{jj}^{-1} < var(\widehat{\theta}_{j}^{(V)}) < (\mathfrak{T}_{N_s-1,\alpha/2}\frac{\sigma'_{\theta_{op}}}{\sqrt{N_s}})^2, \tag{2.5}$$

where $var(.)$ shows the variance operation and $\mathbf{J}_{jj}^{-1}$ is the $jth$ diagonal entry of the inverse of the Fisher information matrix $\mathbf{J}$ [63]. CRB determines the lower bound of estimation of $\widehat{\theta}^{(V)}$ and $\widehat{\phi}^{(V)}$. The asymptotic error covariance of this estimation using ML is proposed in [71] in the presence of colored noise. Given the values of this estimation, vehicle and buoy are now ready to send and receive data while the buoy takes care of vehicle's tracking, as described in Step 5.

**Step 5—Tracking the Vehicle:** This step focuses on the vehicle's movement to ensure that it follows the planned trajectory in the estimated uncertainty region, as discussed in Step 2. Assume

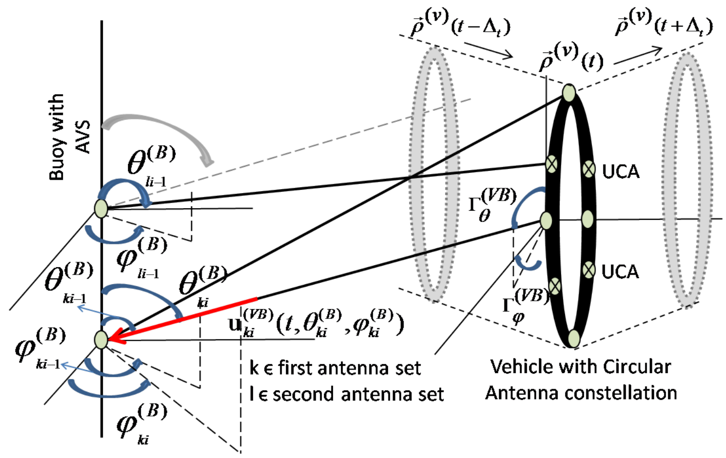



Figure 2.5: 3D spherical representation of angles in transmission from vehicle $V$ to buoy $B$. Both azimuth and elevation angles, $\phi$ and $\theta$, are used for detection. The proposed circular constellation on the vehicle contains two separate sets of $k$ and $l$.

the trajectory vector and the speed of the vehicle at the sampling time $\Delta_t(n)$, $n = 1, ..., N_s$, is defined as $\overrightarrow{\rho}^{(V)}\big(\Delta_t(n)\big)$ and $\overrightarrow{\nu}^{(V)}\big(\Delta_t(n)\big)$, respectively. The current trajectory vector, the previous and its next point are shown with $\overrightarrow{\rho}^{(V)}(t)$, $\overrightarrow{\rho}^{(V)}(t-\Delta_t)$, and $\overrightarrow{\rho}^{(V)}(t+\Delta_t)$ in Fig. 2.5. Every $\Delta_t$ seconds the vehicle broadcasts its velocity information, next turning points, and the destined location. If the initial position $loc^{(V)}_{n0}$ (e.g., initial deployment position on the surface) is known, then the next location coordinates can be written as $loc^{(V)}_{n0+\Delta_t} = loc^{(V)}_{n0} + \overrightarrow{\nu}^{(V)}\big(\Delta_t(n)\big)\Delta_t$. The average distance of horizontal traveling in the cylinder equals its radius and so the upcoming location of the vehicle should be inside the latest estimated cylinder [27]. While the vehicle is about to leave the region, timer 3 pauses the current estimation and requests a new sample, it estimates the new uncertainty region, and the process resumes its normal operation.

**Step 6—Uplink Data Transmission:** Assume the *unit vector* from transmitter set $k$ to sensor $i^{(B)}$ at time $t$ is shown by $\mathbf{u}^{(VB)}_{ki}\big(t, \theta^{(B)}_{ki}, \phi^{(B)}_{ki}\big)$ as,

$$[\sin\theta^{(B)}_{ki}\cos\phi^{(B)}_{ki}, \sin\theta^{(B)}_{ki}\sin\phi^{(B)}_{ki}, \cos\theta^{(B)}_{ki}], \tag{2.6}$$

in which $\phi^{(B)}_{ki} \in [-\pi, \pi]$ and $\theta^{(B)}_{ki} \in [-\pi/2, \pi/2]$ represent the azimuth and elevation angles, respectively. Figure 2.5 depicts the 3D spherical coordinate of the transmitted/received beams. As explained before, we *place transmitter's antennae on a circle* because this arrangement increases the degrees of freedom in both dimensions $\theta$ and $\phi$. Let us consider two separate sets of antennae,

$k$ and $l$, to provide simultaneous transmission via separate and independent channels.

Let us assume that AoA is the output of the AVS, as shown in Fig. 2.5. Every received beam can be distinguished with two parameters $\theta$ and $\phi$, while the received beam at the $i^{(B)}-th$ element of array can be expressed as,

$$\mathbf{y}^{(B)}\big(t, \theta_i^{(B)}, \phi_i^{(B)}\big) = \mathbf{r}'^{(VB)}_{ki}\big(t, \theta_{ki}^{(B)}, \phi_{ki}^{(B)}\big) + \mathbf{z}(t), \tag{2.7}$$

where $\mathbf{y}_B$ is the received signal as a function of AoA and $\mathbf{r}'$ stands for the effect of underwater acoustic channel on the signal $\mathbf{r}$. The transmitted signal can be written as follows,

$$\mathbf{r}_k^{(VB)}\big(t, \theta_k^{(VB)}, \phi_k^{(VB)}\big) = s_k(t)\mathbf{F}^{(VB)}\big(W_{\theta,\phi}^{(VB)}, \Gamma_{\theta,\phi}^{(VB)}\big). \tag{2.8}$$

In (2.8), $\mathbf{F}^{(VB)}$ stands for the beamforming vector determined by the array antenna. *We utilize UCA to steer the beam towards the desired antenna*, so the $\varpi - th$ element of the beamformer in UCA with the angle of $\gamma_\varpi$ w.r.t. the x-axis is calculated via (2.9), as in [150],

$$\exp\left(j\frac{2\pi}{\lambda}\left\{\frac{D_a}{2}\sin(\theta_k^{(VB)})\cos(\phi_k^{(VB)} - \gamma_\varpi)\right\}\right), \tag{2.9}$$

where $\lambda$ is the wavelength of the signal and $D_a$ is the diameter of UCA. $s_k(t)$ is the transmitted OFDM frame of the $k - th$ antenna-set as follows.

$$s_k(t) = \sum_{\xi=0}^{N_C-1} X_\xi^{(k)} exp(j2\pi f_\xi t), \quad X_\xi^{(k)} \in \left\{\delta_i e^{j\beta_i}, 0\right\}. \tag{2.10}$$

Total number of subcarriers is shown by $N_C$ while $f_\xi = \xi f_s$ represents the subcarrier frequency. Choosing $f_s = 1/(N_C T_s)$, where $T_s$ is the sampling interval, leads to orthogonality among different subcarriers in OFDM. $X_\xi$ can be either a data/pilot subcarrier or a null subcarrier. $\delta_i$ and $\beta_i$ stand for the amplitude and the phase of the desired constellation point, respectively. Note that the data/pilot subcarriers of each antenna are overlapped with the null subcarriers of other antennae, so all the antennae can transmit simultaneously without any inter-antenna-interference. The subcarrier assignment strategy is shown in the block diagram of Fig. 2.6(a), which explains how the transmitted signal $s_k(t)$ is created. Data stream is made by the conventional block-oriented standard H.264/MPEG-4, while Scalable Video Coding (SVC) [123] is used as the video compression technique. Since the system error rate is unpredictable and variable due to the vehicle's drifting, SVC encoder provides an adaptive quality video compression by partitioning the frames based on the amount of error fed back by the previous round of transmission.

We assume that each block of $N_c(\log_2 N + \log_2 M)$ bits consists of segments of $n+m$–bits; the first $n$=$\log_2 N$ bits define the transmitting antenna in every set containing $N$ antennae, while the last $m$=$\log_2 M$ bits define the appropriate transmitted signal regarding the chosen modulation scheme of order $M$. *Designation table* assigns the bits to the appropriate antenna and constellation point. *Frequency designation* is performed via transmitting signals of different antennae on the orthogonal subcarriers of $N$ *NC-OFDM* frames. The $\xi - th$ signal of the modulated stream is sent over the $\xi - th$ subcarrier of the *corresponding antenna*'s NC-OFDM frame and the other subcarriers in this frame are switched off.

$$\xi f_s \leftarrow \begin{cases} \Xi_i, & b_i, i \in [(\xi+1)n + \xi m + 1 \ldots (\xi+1)(n+m)], \\ 0, & otherwise. \end{cases} \tag{2.11}$$

Accordingly, every single antenna utilizes the unoccupied portions of the spectrum of other antennae. Note that unlike conventional Spatial Modulation (SM) [78], in which each symbol is sent via different antenna and so a very fast antenna switching is required, our proposed method first forms a complete frame of desired subcarriers for each antenna and then sends it at once. This feature is essential since fast switching signal transmission is not practical underwater because of the long propagation delay of the acoustic channel. In our method, all $N$ implemented antennae are active and the subcarriers are being used efficiently; however, for each transmission, only one antenna transmits in each subcarrier.

**AVS-based Receiver and Rate Comparison:** Fig. 2.6(b) describes the receiver of the system at the buoy, while AVS gives us an estimate to decide on the desired antenna $q$ from set $k$ as follows,

$$\widetilde{q} = \underset{q \in \{1,\ldots,N_k\}}{\arg\min} \left\| [\widetilde{\theta}_{ki}^{(B)}, \widetilde{\phi}_{ki}^{(B)}] - [\widehat{\theta}_{op}^{(B)}, \widehat{\phi}_{op}^{(B)}]_q \right\|^2, \tag{2.12}$$

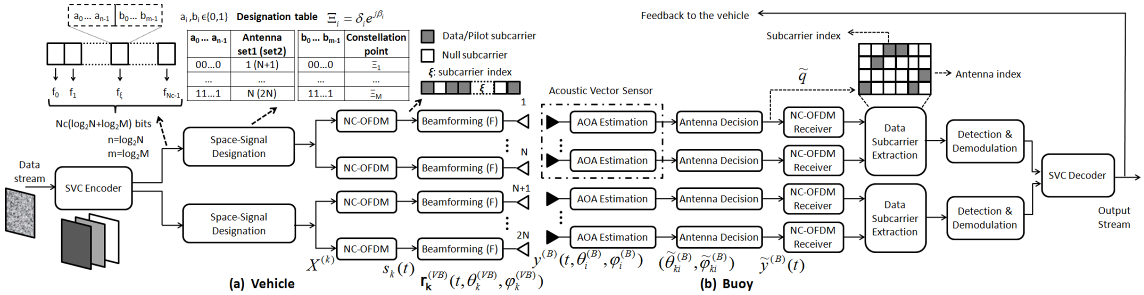



Figure 2.6: (a) Transmitter signaling blocks for the proposed SSFB method. The designation table shows how data bits are grouped. Each one of the $N$ implemented antennae in each set and in every transmission course exploits a portion of subcarriers in the NC-OFDM; (b) Receiver block diagram in which the received frames are treated.

where $[\widetilde{\theta}_{ki}^{(B)}, \widetilde{\phi}_{ki}^{(B)}]$ is the vector of estimated angles at the buoy and $[\widehat{\theta}_{op}^{(B)}, \widehat{\phi}_{op}^{(B)}]$ stands for the reference angles vector including all the antennae in every previously estimated set. To describe how AVS estimation error leads to the error in the detection, assume the angular decision on $q$ is made in the region of $\{[\theta^{(B)}, \phi^{(B)}] : -\pi/N_k < [\theta^{(B)}, \phi^{(B)}] < \pi/N_k\}$, with Probability Density Function (PDF) of $\mathcal{P}_{\theta,\phi}(\theta, \phi)$, the error probability, $P_e$, can be calculated as,

$$P_e = 1 - \iint_{-\pi/N_k}^{\pi/N_k} \mathcal{P}_{\theta,\phi}(\theta^{(B)}, \phi^{(B)}) d\theta d\phi, \tag{2.13}$$

where $N_k$ is the number of elements in antenna set $k$.

An OFDM frame consists of a preamble—for synchronization and Doppler estimation—and of data blocks. Let us define $\widetilde{y}^{(B)}(t)$ as the estimated received signal after antenna decision. NC-OFDM receiver performs Fast Fourier Transform (FFT) to detect data subcarriers.

$$\widetilde{Y}_\xi^{(k)} = \frac{1}{T_s} \int\limits_{T_s} \tilde{y}^{(B)}(t) e^{-j2\pi f_\xi t} dt, \tag{2.14}$$

where $\widetilde{Y}_\xi^{(k)}$ is an estimate of the transmitted signal $X_\xi^{(k)}$ at the receiver. Following NC-OFDM receiver block, data subcarrier extraction block makes an $N \times N_C$ matrix of estimated frames in which the $\widetilde{q} - th$ antenna frame is placed in the $q - th$ row. In the case of perfect frequency synchronization, only one element in each column contains data above the noise level and the others should be null. The non-zero data of row $q$ and column $\xi + 1$, (i.e., $\widetilde{\Xi}_{q,\xi+1}$) is demodulated in the

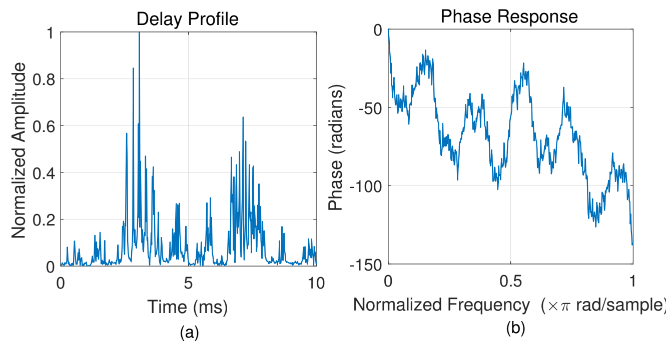



Figure 2.7: (a) Delay profile; (b) Phase response of emulated KAM08 channel.

next block and determines the output frame, using the designation table.

$$\begin{cases} \widetilde{\Xi}_{q,\xi+1} \to b_i, & i \in [(\xi+1)n + \xi m + 1 \ldots (\xi+1)(n+m)], \\ q \to a_i, & i \in [\xi(n+m)+1 \ldots (\xi+1)n + \xi m]. \end{cases} \tag{2.15}$$

**Rate Comparison:** While conventional SM transmits $\log_2 N + \log_2 M$ bits per transmission (bpt), Multiple Active SM (MA-SM), transmits $\left\lfloor \log_2 \binom{N}{N_A} + N_A \log_2 M \right\rfloor$ bpt from $N_A$ active antennae [16], where $\lfloor \cdot \rfloor$ rounds to the nearest lower integer.

OFDM-IM can transmit $\left\lfloor |\log_2 \binom{N_G}{K}| + K \log_2 M \right\rfloor G$ bpt, where $G \times N = N_G$. Here, $G$ stands for the number of groups, and $N_G$ shows the number of subcarriers in each group. The number of transmitted bits in each transmission course (for $I_B$ sets of antenna at the buoy as depicted in Fig. 2.5) for our proposed system is calculated as,

$$R = I_B N_C (\log_2 N + \log_2 M). \tag{2.16}$$

## 2.4 Performance Evaluation

**Simulation Settings:** We consider a short/medium range transmission (less than $2\ \mathrm{km}$), in which the anchored buoy has an attached bar with installed hydrophones and the vehicle can move with a smooth and almost constant horizontal speed on the other side, as depicted in Fig. 2.2. Two sets of circular antennae are on the vehicle, each one contains four UCAs while two separate vector sensors are on the buoy. To avoid spatial correlation between antennae, we need to keep the minimum

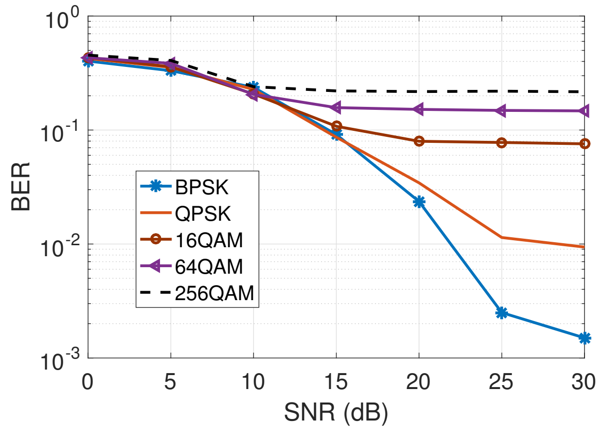



Figure 2.8: Bit Error Rate (BER) of the proposed system with different modulations in the presence of colored noise. Two sets of four transmitter antennae are considered.

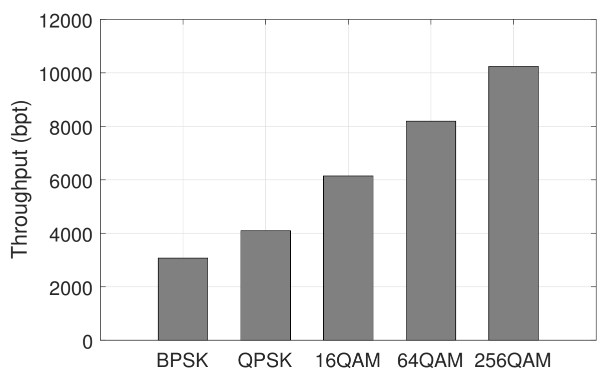



Figure 2.9: Throughput of the proposed system (SSFB NC-OFDM) with different modulations.

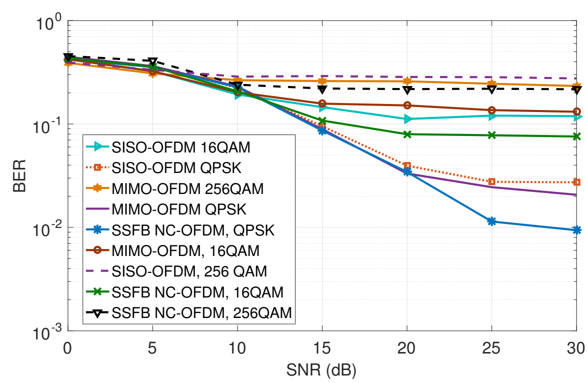



Figure 2.10: BER in SISO- and MIMO-OFDM for different signaling methods.

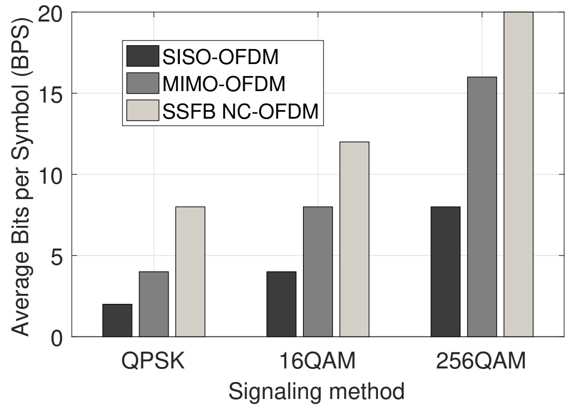



Figure 2.11: Average bits per symbol comparison between the proposed SSFB NC-OFDM and SISO- and MIMO-OFDM

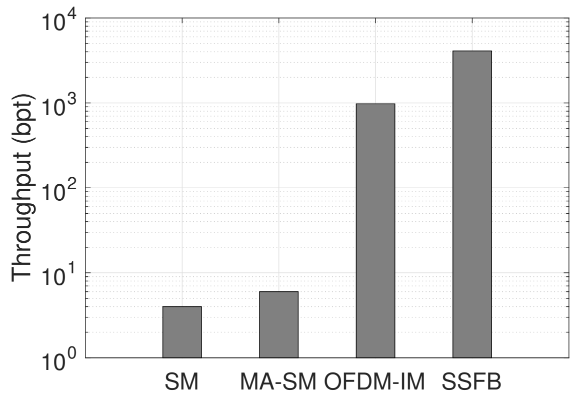



Figure 2.12: Comparing SSFB with conventional SM, MA-SM, and OFDM-IM in terms of number of bits transmitted in every transmitting course (bpt). Four transmitter antennae with QPSK modulation are considered.

distance between adjacent elements more than $\lambda/2$.

We assume a channel with a number of sparse and separable paths, in which the first path is the strongest part of signal. To extract the channel characteristics for our simulations, the channel was exploited from Kauai Acomms MURI (KAM08) experiment at the western coast of Kauai, HI, USA [54] where an anchored vertical hydrophone array with $3.75 \text{ m}$ inter–element spacing at the depth of $96 \text{ m}$ communicates with a source towed by a surface ship. The channel bandwidth is $25 \text{ kHz}$ and the sampling rate is $50 \text{ kHz}$. Figs. 2.7(a)-(b) show the normalized delay profile and the phase response of the channel, respectively. An Underwater colored ambient noise, leading to a $SNR \in [0, 30]$ dB, is considered as the background additive noise. Initially, we assume 512 subcarriers are present in every NC-OFDM frame. The proposed system contains an inherent zero

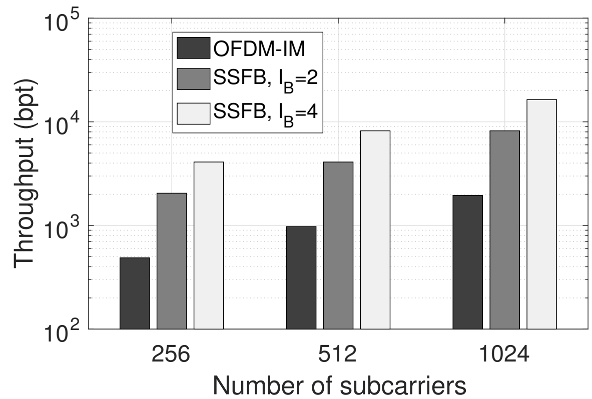



Figure 2.13: Throughput of the system with different antenna sets versus variable number of subcarriers compared with OFDM-IM.

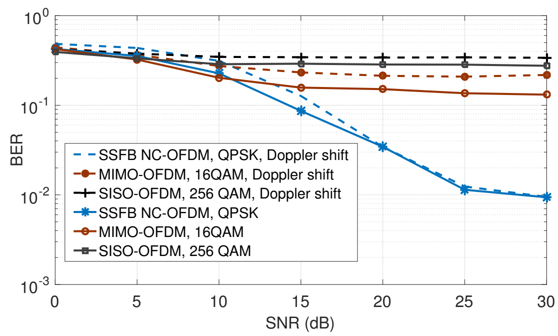



Figure 2.14: Effect of Doppler shift on the performance of proposed system compared to other conventional techniques.

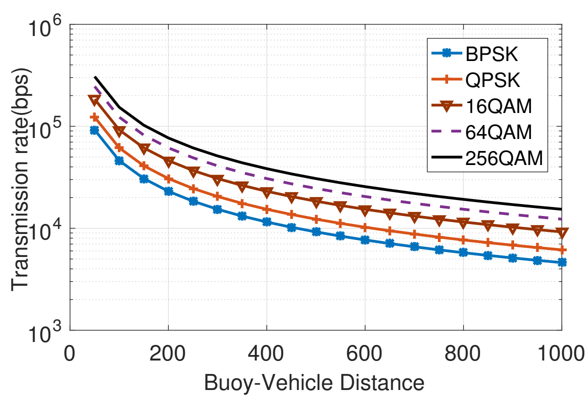



Figure 2.15: Transmission rate [bps] versus distance for SSFB with different modulations.

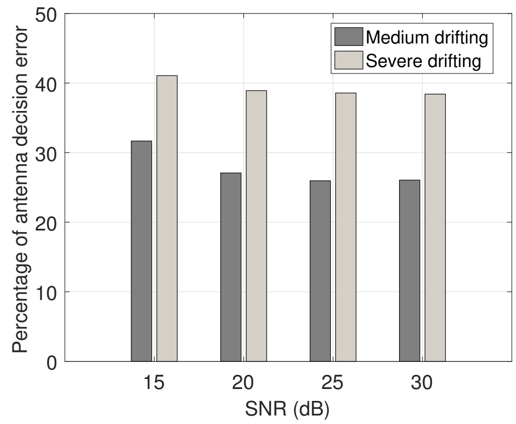



Figure 2.16: Percentage of error in antenna decision as the result of vehicle drifting.

padding by nulling the other interfering antennae. The recorded video which will be sent during the simulation is an MP4 RGB-24 video with the duration of $5.2$ s and frame resolution $240 \times 320$. Transmission starts with an initial delay of $(\tau_1 + \tau_2)$ seconds, as explained in Fig. 2.4.

**Results and Discussions:** Fig. 2.8 shows the performance of the proposed system with different modulation orders whereas its BER in the lower SNRs is not satisfactory, regardless of the modulation order. The reason is that null subcarriers are covered by the background noise and so their energy level is comparable with the data subcarriers; therefore, the receiver fails in the process of *data subcarrier extraction*. The throughput of the system is plotted in Fig. 2.9 in terms of number of bits per transmission course (bpt). Considering these two figures and due to the throughput-BER tradeoff, the most proper modulation scheme is chosen. Therefore, SVC adaptively changes the rate for the next round of transmission. Fig. 2.10 compares the proposed solution, i.e., SSFB, with the conventional OFDM-based methods. The proposed solution outperforms other techniques in terms of bit error rate, especially in higher SNRs.

Fig. 2.11 confirms that the average Bits Per Symbol (BPS) of SSFB is higher than of SISO- and MIMO-OFDM, when similar modulation order is used for all of them. Moreover, it is observed that SSFB can result in the same BPS values as SISO- and MIMO-OFDM, by using a lower modulation order with the lower error rate. Fig. 2.12 confirms that SSFB outperforms the conventional methods in terms of throughput. Regarding conventional SM, SSFB has a considerable data rate, while it covers the fast antenna switching problem of SM by forming the frames prior to transmission. SSFB rate is around double in comparison with OFDM-IM. In Fig. 2.13, SSFB throughput is compared

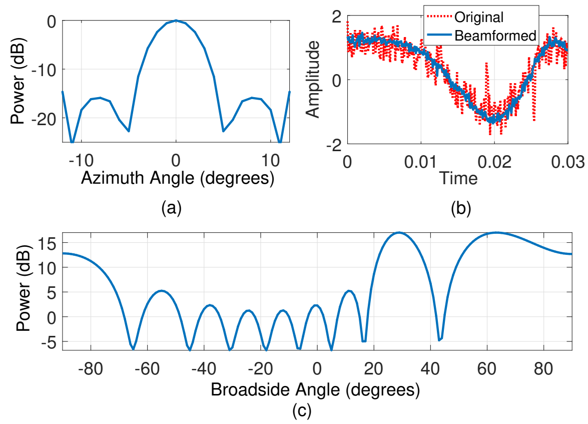



Figure 2.17: (a) An Azimuth cut of the array response of the UCA on the vehicle; (b) Output of the beamformer applied to the UCA; (c) Spatial spectrum of the AOA estimator output on the buoy.

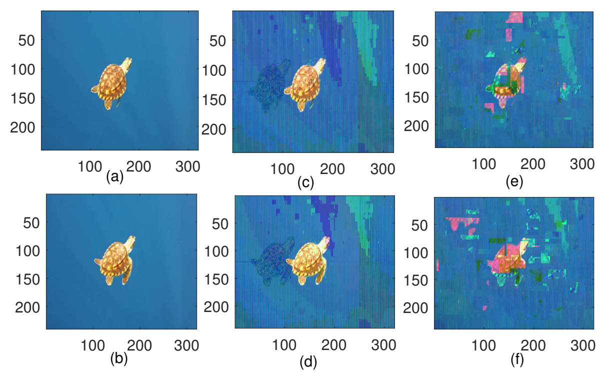



Figure 2.18: (a)-(b) Sample transmitted frames; (c)-(d) Received frames in the presence of noise; (e)-(f) Received frames when AoA estimation and antenna decision are erroneous.

with OFDM-IM for different number of subcarriers and antenna sets. The bpt of 4 antennae with 512 subcarriers equals using two antenna sets with 1024 subcarriers. Therefore, the later one is preferred regarding its lower probability of antenna beam interference.

Fig. 2.14 shows that SSFB is less vulnerable to Doppler shift than SISO- and MIMO-OFDM, especially in higher SNRs. Fig. 2.15 investigates the effect of horizontal buoy-vehicle distance on the transmission rate. In Fig. 2.16, we study the situation in which antenna decision making at the receiver is not successful due to the vehicle's drifts. Under severe drifts, the receiver suffers from error in AoA estimation and antenna decision, which leads to higher error rate. This error is fed back to the SVC encoder to be considered while making decision on the SVC layer for the next round of transmission. Meanwhile the angular calculation process is restarted, as explained in Fig. 2.4.

In Fig. 2.17(a), the array response of UCA is plotted. Fig. 2.17(b) shows the result of applying the designed beamformer to the UCA. It demonstrates that it can follow the original signal while the energy is focused on a specific direction. Fig. 2.17(c) confirms that AoA estimator is successful in separating two signals which are arrived at 30 and 65 degrees.

Simulated video transmissions through SSFB is shown for two frames in Figs. 2.18(a)-(b). In the first scenario, we assume that the vehicle has the least drift. AoA estimation and the antenna decision are performed successfully. This result is reflected in Figs. 2.18(c)-(d). Figs. 2.18(e)-(f) represent the second scenario, in which vehicle experiences AoA estimation and antenna decision error. Frames can not be recovered when the amount of AoA estimation error increases significantly. In this case, the error is fed back and the encoder tunes the parameters for the next round.

## 2.5 Summary

An Acoustic Vector Sensor (AVS)-based solution was proposed, called Signal-Space-Frequency Beamforming (SSFB), to transmit underwater videos at high data rates (in the order of hundreds of Kbps at the operating ranges of the application of interest, i.e., up to a few kilometers). To achieve this goal, the receiver (buoy) was equipped with AVS—hydrophones that measure acoustic particle velocity in addition to scalar pressure—in a multiple-antenna-array configuration, while the transmitter (vehicle) was equipped with a circular array of transducers. Data was modulated and transmitted via NC-OFDM, and detected via beam's angle of arrival at the receiver (buoy). In order to determine the angles of departure/arrival, the position uncertainty regions were studies and a protocol was presented. The vehicle utilized beamforming based on the uncertainty region and the estimated angles and the data was transmitted via the proposed uplink signaling, i.e., SSFB. Simulations showed that video transmission rates can enable applications such as coastal and tactical surveillance, which require multimedia acquisition and classification.

# Chapter 3

# Probabilistic Spatially-Divided Multiple Access in Underwater Acoustic Networks

Deploying Autonomous Underwater Vehicles (AUVs) is a necessity to enable a range of civilian/military underwater applications; yet, achieving a reliable coordination among the vehicles is a challenging issue due to the time- and space-varying characteristics of the acoustic communication channel. The design of a Medium Access Control (MAC) based on a probabilistic Space Division Multiple Access (SDMA) method for short/medium distances (less than $2\ \mathrm{km}$) is presented in this chapter. This method considers the inherent vehicle position uncertainty due to the inaccuracies in models and the drift of the vehicles. It minimizes the acoustic interference statistically by considering the angular position of neighboring vehicles via a two-step estimation and by keeping the transmitter antenna's beamwidth of each vehicle at an optimal value. Such value is chosen considering three contrasting goals, i.e.: $(i)$ spreading the signal beam towards the vehicle to combat position uncertainty using a coarse estimation; $(ii)$ focusing the beam to reduce acoustic energy dispersion through a fine estimation; and $(iii)$ minimizing interference to other vehicles. Simulation results show that this approach mitigates interference, reduces the probability of retransmission, and achieves higher data rates over conventional underwater MAC techniques.

## 3.1 Overview

Underwater wireless acoustic networks, which are composed of static sensors and mobile vehicles, underpin the underwater world and are instrumental to support next-generation ocean-observation systems, to enable both civilian and military applications, and to pave the way towards the futuristic Underwater Internet of Things (UW IoTs) paradigm. Oceanographic data collection, ocean pollution monitoring, offshore exploration, tsunami detection/disaster prevention, assisted navigation, and tactical surveillance are examples of some of such applications [51, 111]; for most if not all of these applications achieving communication and coordination is key, which in turn calls for the ability to transmit reliably signals underwater. This problem is still challenging in the harsh underwater environment in which Radio-Frequency (RF) waves are absorbed for distances above a few tens of meters, optical waves require narrow laser beams and suffer from scattering as well as ocean wave motion, communication through magnetic induction is only feasible up to a few meters, and acoustic waves lead to a communication channel that is very dynamic, prone to fading, spectrum limited with passband bandwidths of only a few tens of kHz due to high transmission loss at frequencies above $50$ kHz, and affected by non-Gaussian noise [136]. In addition, while the acoustic waves can propagate underwater up to several tens of kilometers—making the communication technologies based on them the only possible for distances above a hundred meters—their speed is not constant and depends on temperature, salinity, and pressure of the body of water traversed.

**Motivation:** One of the main challenges for the effective coordination of a team of autonomous vehicles underwater is how to design a *fair* and *efficient* Medium Access Control (MAC) protocol tailored to the harsh underwater acoustic environment. Due to the unique characteristics of the propagation of acoustic waves in the water, in fact, existing terrestrial MAC solutions are unsuitable for this environment [51]. One promising yet unexplored MAC technique for *sparse underwater networks* is Space Division Multiple Access (SDMA) which exploits the signal beam directivity and the spatial separation of the vehicles. It makes use of the fact that vehicles/mobile users can be served simultaneously when they are not located in the same area, so that the radiated energy for each user can be separated *in space* [117]. Among the possible channelized MAC techniques, SDMA is more efficient when the network is *sparse*—as in the underwater environment[1]—than Frequency Division Multiple Access (FDMA), which has limited acoustic bandwidth per channel, Time Division Multiple Access (TDMA), which requires long time guards and high signaling overhead (especially when some of the users are mobile and/or the body of water is large [77]), and Code Division Multiple Access (CDMA), which is affected by a low data rate [96, 109, 114].

Notwithstanding these considerations, there are strong arguments against a brute-force application of terrestrial SDMA to the underwater acoustic environment. In such MAC scheme, in fact,

[1]The network sparsity assumption here—leading to a high likelihood of spatial separation between vehicles—comes from the observation that the vehicle density in a body of water is often low due to cost and scalability.

the expected increase in the data rate is highly affected by the accuracy of the Channel State Information (CSI) at the transmitter; therefore, a considerable amount of effort should be dedicated to compensate for the partial information at the transmitter in a (possibly) rapidly changing underwater channel. Furthermore, a dramatic reduction in the throughput is observed if the feedback is delayed, since the CSI becomes outdated. Noticeably, in underwater acoustic channels, CSI is usually unknown to both transmitter and receiver. Moreover, in the terrestrial SDMA, every user can provide its own real-time position information, which is not feasible underwater as Global Positioning System (GPS) does not work and the position information is not readily available. Even if we had the users' locations, applying them would still not be feasible because achieving *perfect pointing* between transmitter and receiver is challenging due to the position uncertainty of the vehicles and the nature of acoustic wave propagation. Furthermore, considering the effect of ocean currents on the vehicle, inaccuracies in position estimation increases the position uncertainty [27, 108]; and this uncertainty becomes worse over time when the vehicle stays longer underwater due to error propagation, which leads to non-negligible drifts in the vehicle's position and thus making conventional concept of SDMA inapplicable for the underwater environment.

**Contributions:** Any channel access method that exploits a *deterministic* approach for interference mitigation/cancellation would not be an efficient solution underwater as it ignores the inherent position uncertainty of the vehicles caused by drifts, model errors, and unbounded errors, thus leading to performance degradation. For these reasons, we present a novel *probabilistic* and spatially-divided MAC to cancel/alleviate the interference while the inherent position uncertainty of vehicles is considered in a sparse underwater mobile network. An Angle of Departure (AOD)-based solution forms separate spatial beams via a probabilistic approach towards the target vehicles. Since the vehicles are mobile, a two-stage estimation scheme is required to calculate the beam parameters, i.e., $(i)$ a coarse interval estimation and $(ii)$ a fine estimation via unscented Kalman filtering to update the beam parameters for each antenna. An optimization problem mitigates the statistical interference between the expected overlapped vehicles by keeping the transmit beamwidth and direction within a desirable range. In the case the vehicles are entirely overlapped in space, we propose a *hybrid* probabilistic time and space MAC scheme, called T-SDMA, which takes *time* into account besides space, and outperforms conventional TDMA methods in terms of rate efficiency.

**Chapter Outline:** The remainder of this chapter is organized as follows. In Sect. 3.2, we provide a discussion on the related work and papers. In Sect. 3.3, we introduce the proposed system and provide solutions for probabilistic SDMA, where both spatially separable and non-separable scenarios are considered. In Sect. 3.4, we present our simulation results and discuss the benefits of our solution. Finally, in Sect. 3.5, we summarize the chapter.

## 3.2 Related Work

During the past few years, several MAC protocols have been proposed for underwater communications. Time sharing-based solutions are exploited in many real underwater scenarios. However, they will not be very efficient if the long propagation delay of the channel is not considered [73,158]. Authors in [140] proposed a distributed and energy-efficient MAC protocol called Tone Lohi (T-Lohi), as an energy efficient tone-based contention algorithm. This technique shows low channel utilization, when the number of nodes increases. A delay tolerant MAC protocol (DTMAC) was proposed in [69], in which the solution applies short-packets traffic to combat the effect of long propagation delay and mobility in sparse networks. A collision-free TDMA scheduling was discussed in [72] to improve the throughput by considering large propagation delays. Recently, the collected data at sea experiments confirms that there is no unique MAC solution for all the scenarios and configurations under various conditions [92]. Other random- and controlled-access MAC protocols such as Carrier-sense Multiple Access (CSMA) transmit multiple packets through the same underwater channel, which might lead to packet collisions at the receiver. The other method, which has both the carrier sensing and collision avoidance mechanisms, is Floor Acquisition Multiple Access (FAMA). The objective of this protocol is to ensure that a single sender reserves the channel via an RTS (Request to Send)/CTS (Clear to Send) handshake before transmitting a packet.

In this section, we briefly review conventional space sharing MAC algorithms and their related challenges, while keeping in mind that they cannot be directly used in underwater channels. Assuming spatial separation of the users, sectorized antenna can be a primitive application of terrestrial SDMA [117]. SDMA-based Smart antennae in mobile networks can improve the network capacity. A robust and self-organizing terrestrial SDMA is proposed for mobile ad-hoc networks in [15], where it is shown that the network bandwidth efficiency depends on the number of mobile users. In [55], an opportunistic terrestrial-based SDMA with threshold feedback algorithm for enforcing

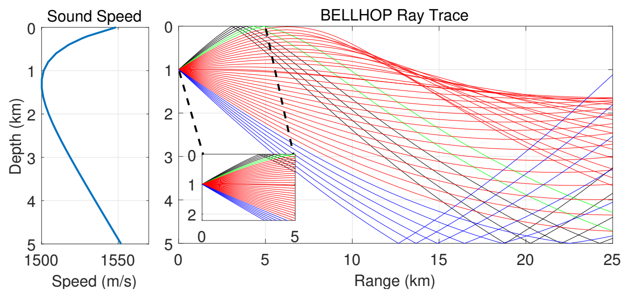



Figure 3.1: Bellhop ray tracing (right box) for a standard sound-speed profile (left box) indicating how acoustic beams travel through the underwater acoustic channel [97]. The beams are almost straight for short/medium ranges (less than $5\ \mathrm{km}$) when the transmitter is at a depth of $\sim 0.9\ \mathrm{km}$.

the sum feedback rate constraint was proposed. All these approaches still need the channel information to be fed back to the transmitter, which is not always feasible underwater due to the long propagation delays. Authors in [65] proposed a zero-forcing precoding with partial CSI, which is known at the transmitter in the terrestrial ad-hoc networks. In [101], beamforming was proposed for multiuser multi-antenna SDMA downlink systems. In [152], the authors presented a hybrid architecture for downlink beamforming with phased antenna arrays in indoor SDMA channels as a new generation of broadband terrestrial personal and local area networks. Recently, multi-beam smart antenna array system based on SDMA was presented, which is a promising candidate for next generation of wireless communications as it enables the antennas to adapt and to steer the energy towards a desired direction [6].

To implement any spatially-divided MAC protocol for the underwater channel, as we mentioned earlier, one of the main challenges to face is the inaccuracy and uncertainty in localization models of the underwater vehicles. Short Baseline (SBL) is one of the most common ways to localize vehicles underwater, in which the position estimate is performed via external transponder arrays. Long Baseline (LBL) system, similarly to SBL, also uses tethered external transponder arrays with fixed locations [64] in farther distances. Dead-reckoning estimation of position is based on accumulated measurement of the velocity compared to the surface. AUV Aided Localization (AAL) methods were also introduced in the literature, in which the distances to the AUV is estimated by each node, while the AUV is at different locations [42, 105]. In [27], an approach has been proposed to predict vehicles' position through statistical method. This method also estimates the position uncertainty of the vehicles and designs a routing protocol based on the vehicles' confidence region. Given the randomness of underwater channel and its long propagation delay, interference distribution under various signal propagation models was discussed in [73]. However, our proposed solution aims at mitigating the interference between the underwater vehicles caused by the position uncertainty in the water via spatial and time division techniques.

## 3.3 Proposed Probabilistic Solution in Sparse Networks

**Underwater Acoustics and the Requirements for Sparse Networks:** While sound waves travel through the underwater medium, part of the acoustic/elastic energy is absorbed; a well-known expression that models the medium absorption coefficient as a function of frequency is $a(f) = (0.11f^2)/(1+f^2) + (44f^2)/(4100+f^2) + 2.75 \times 10^{-4}f^2 + 0.003$ [134]. In this empirical formula, $10\log_{10} a(f)$ gives the channel attenuation in $\mathrm{dB/km}$. Propagation loss can be modeled via $P_a = \varsigma D^{\varpi} e^{a(f)D}$, in which $\varsigma$, $D$, and $\varpi$ stand for the scattering loss, distance, and spreading loss parameters, respectively [157].

Let us assume that the transmission occurs at *short/medium ranges* and that the direct beams are dominant over the reflected beams from the surface and the bottom of the sea so that the receiver is not severely affected by multipath. For farther distances (above a few kilometers) and based on the sound-speed profile, the acoustic rays bend towards the region of lower acoustic speed (the so-called *"laziness law"*). This effect changes the Angles Of Departure/Arrival (AOD/AOA) and their estimations. Using the Bellhop model [97] and considering a typical deep-water case, Fig. 3.1 illustrates the sound-speed profile (left) and the acoustic ray tracing (right) in the underwater channel for a sample source at a depth of $\sim 0.9$ km and a water temperature of $39°F$. The bending effect can be observed at distances of a few kilometers, however, staying within a short/medium range (less than $5$ km), such bending is not notable, which explains the philosophy behind our signaling method in which the vehicle is steered via beamforming.

To consider the time variability of an underwater acoustic channel, assume it varies after approximately $t_c$ seconds (channel coherence time). This parameter can be defined by Clarkes model [117] as $t_c = \sqrt{9/(16\pi {f_d}^2)} \approx 0.423/f_d = 0.423/(\alpha_d f_c)$, where $f_d$ is the Doppler shift, $f_c$ is the carrier frequency, and $\alpha_d$ represents the Doppler scaling factor. Let $D$ be the distance between the buoy and the vehicle, then the round trip delay time $t_D$ for the distance $2D$ should be less than $t_c$. It is easily shown that for an underwater channel with a specific sound profile and with $f_c = 20$ kHz,

$\alpha_d = 3 \times 10^{-5}$, and for distances greater than $D \approx 500$ m, the time for receiving feedback CSI signal could be larger than the coherence time of the channel [29]. It is worth mentioning that the proposed technique in Sect. 3.3 is not channel dependent, so the information sent back to the transmitter is usable for a longer time compared to channel variations. Furthermore, as discussed in [27], by assuming that ocean currents are unknown, the vehicle's drifting in the horizontal plane is identically and independently distributed (i.i.d.) and follows a normal distribution, which makes the horizontal projection of its confidence a circular region. Regarding the vehicle's movement along its trajectory, the uncertainty region is concluded to be a cylinder [27].

There are multiple methods for underwater vehicles to be localized that can be generally categorized as inertial dead reckoning, acoustic transponders and modems, and geophysical methods [85]. The selection of the localization method is dependent on the application, the environment and the desired accuracy. In some regions, depth sensors can be implemented to provide information about vertical position of the vehicles. In some applications, the receiver antenna arrays can be utilized to estimate the AUV's location. In [110], we utilized Acoustic Vector Sensors (AVS) to estimate the angle of arrival. In this chapter, we follow [27] and choose dead reckoning as the initial localization technique, although the localization technique does not directly impact our proposed solution. Dead reckoning method is subject to cumulative errors, but as a classic location estimator and because of its simplicity of implementation, it is still a widely-used solution in AUVs [85]. Each vehicle estimates its trajectory and position, using its own location estimates and considering the inherent position uncertainty of objects underwater. Every $\Delta t$ seconds, each vehicle estimates its current location by measuring its velocity and using the previous estimated locations. Vehicles adopt a polling model, as will be explained in the algorithm, to send back the measured position samples through a feedback channel. Notice that the type of vehicles that can benefit from this research depends on the application, but it does not directly impact our proposed solution. Buoyancy-propelled gliders—which follow a sawtooth-shaped glide path—move not as fast as conventional AUVs (a fraction of a meter per second), however, they are extremely efficient in terms of power consumption making them suitable for background monitoring missions, whereas propeller-driven AUVs are capable of operating at higher speeds.

**Proposed Probabilistic Solution:** In this section, we present our solution and provide more

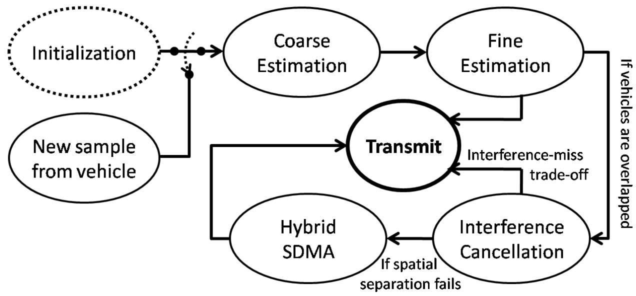



Figure 3.2: Framework depicting the interaction between different parts, called macro-states.

Table 3.1: Notations and Mathematical Terms: Coarse Estimation.

| Notation | Description |
|---|---|
| $a(f)$ | Medium absorption coefficient |
| $P_a$ | Propagation loss at distance $D$ |
| $\varsigma$, $\varpi$ | Scattering and spreading loss parameters |
| $t_D$ | Round trip delay |
| $t_c$ | Coherence time |
| $f_d$, $\alpha_d$ | Doppler shift and scaling factor |
| $n$, $N$ | Location sample index, total location samples |
| $\mathcal{J}$ | Total number of vehicles |
| $\Delta z$, $\Delta x$ | Vertical and horizontal location drift |
| $(i), (j)$ | Neighboring vehicles |
| $\Theta$ | Vector of parameters |
| $T_x$ | Buoy's antennae |
| $m$ | Antenna index |
| $\theta_n^{(j)}$ | Angle in elevation plane |
| $\phi_n^{(j)}$ | Angle in azimuth plane |
| $h_L$, $h_U$, $R$ | Cylinder's lower and upper height, and radius |
| $\theta'^{(j)}$ | Transferred angle to the center |
| $\bar{\theta}'^{(j)}$ | Centered AOD angle |
| $z_m^{(j)}$ | Vertical distance (depth) of $(j)$ to antenna $m$ |
| $\mathcal{T}$ | Student's t-distribution critical value |
| $\mathcal{N}$ | Normal distribution |
| $\bar{\theta}$ | Mean of $N$ samples |
| $\mu_\theta$ | Distribution mean |
| $\sigma_\theta^2$ | Standard deviation |
| $S_\theta$ | Standard deviation of samples |
| $1-\alpha$ | Confidence degree |
| $\theta_L^{(j)}$, $\theta_U^{(j)}$ | Lower and upper boundaries of vehicle $(j)$ |
| $W_\theta^{(j)}$ | HPBW towards vehicle $(j)$ |
| $\Gamma_\theta^{(j)}$ | AOD towards vehicle $(j)$ |

details for different parts of the system. Fig. 3.2 shows the system framework through the interaction between its *macro-states*. We define the notion of *macro-states* to distinguish them from the definition of states, which will be used later for beam's parameters. First, we form separate spatial beams probabilistically towards the target based on *coarse* and *fine* estimations, if we do not expect

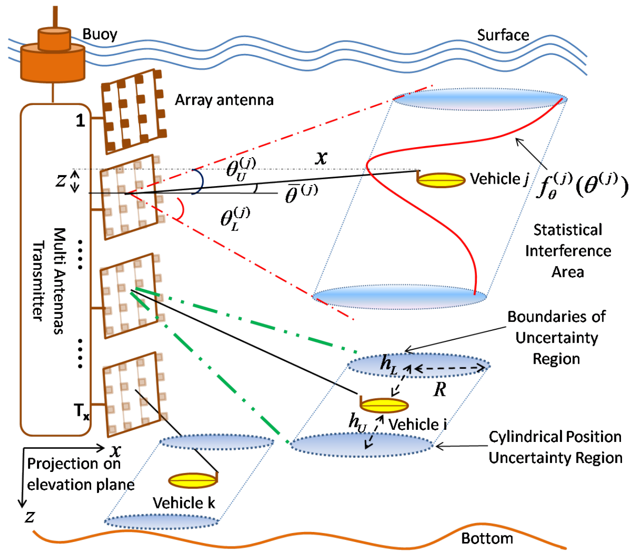



Figure 3.3: Geometric configuration representing surface buoy-to-vehicle communications. Position uncertainty regions of vehicles $j$, $i$, and $k$ are shown as cylinders. Angles $\theta_U^{(j)}$ and $\theta_L^{(j)}$ stand for the upper and lower angles of the beam.

any severe interference; otherwise, the system transits to the next macro-state, which consists in the *statistical interference cancellation*. If the probability of interference/miss trade-off is satisfactory, then we shift to the *transmit* macro-state. In the case that the vehicles are non-spatially separable, a *hybrid* probabilistic time and space MAC scheme, called T-SDMA, will be proposed. The details of each macro-state will be discussed as follows. First, in Sect. 3.3, we introduce the initialization procedure and the coarse estimation for our probabilistic SDMA via *interval estimation* of the antenna's parameters. Then, we present a fine estimation technique which is required for antenna AOD and beamwidth estimation via an *unscented Kalman filtering*. We propose a statistical interference cancellation via an optimization problem. All these discussions assume that the vehicles are separable in space. Finally, in Sect. 3.3, the estimated values are applied to spatially overlapped non-separable vehicles and a hybrid solution suitable for this scenario is presented. Table 3.1 summarizes the main mathematical terms used in the coarse estimation section.

**Coarse Confidence Interval Estimation:** In this step, we explore how the estimated position of the vehicles are used to achieve a coarse estimation of antenna parameters in a centralized fashion. Fig. 3.3 shows the general configuration of the system including buoy's antenna arrays, vehicles, and their cylindrical position uncertainty regions. A second set of arrays, which is not depicted in this figure, is implemented on the other side of buoy to cover the spherical space around it. Assume the neighboring vehicles $i$, $j$, and $k$ follow different trajectories. To calculate the uncertainty region, vehicle $j$ broadcasts its $n$th location sample at time $t$ as $loc_n(t) = [x_n^{(j)}, y_n^{(j)}, z_n^{(j)}]_{n=1}^N$. Assume the initial AODs towards two separate vehicles $i$ and $j$ are identified as $\theta^{(i)}$ and $\theta^{(j)}$. Since the variation of a vehicle's position inside the uncertainty region can be defined as a random variable [27], it is inferred that the variation of the angles at the buoy's transmitter is also a random variable with unknown mean and variance. We claim that each angle is a result of a drift in the $n$th location sample as $\theta_n^{(j)}(t) = \tan^{-1}(z_n^{(j)}(t-t_D/2)/x_n^{(j)}(t-t_D/2))$ where $t_D/2$ shows the delay in transmission, as discussed before. For simplicity of notation, let us drop the superscript $j$ in $z_n(t) = z_n(t-t_D/2) \pm \Delta z$ and $x_n(t) = x_n(t-t_D/2) \pm \Delta x$, where $\Delta z$ and $\Delta x$ stand for the vertical and horizontal drifts in the vehicle's location after $t_D/2$. For example, for a distance of $D = 1000$ m, the delay equals $t_D/2 \approx 0.66$ s and the vehicle moves $\Delta x \approx 0.16$ m, which is negligible compared to the transmission distance considering a vehicle's constant speed of 0.25–0.5 m/s [27]. Based on the Taylor's polynomial approximation, $\tan^{-1}(z_n/x_n) \approx z_n/x_n - 1/3(z_n/x_n)^3 + 1/5(z_n/x_n)^5 - ...$, where $z_n$ is the vehicle's vertical drift due to its position uncertainty. This value is very small in comparison with $x_n$, which is the horizontal distance between the vehicle and the buoy; consequently, $\theta_n \approx z_n/x_n$. Besides, [76] provides approximations to demonstrate in practice that many of the ratios of normal random variables are normally distributed if the denominator of $z/x$ is positive and its coefficient of variation is very small. Based on the numerical calculations in [49], we can conclude that $\theta_n$ follows a normal distribution.

*Definition* 1. Let $X = (X_1, ..., X_n)$ be a random sample observation from a distribution with parameter $\Theta$, then a random variable $\mathcal{V}(X_1, ..., X_n, \Theta)$ is called a pivotal quantity if its distribution is independent of all parameters $\Theta$ [25].

*Definition* 2. An interval estimate of a parameter $\Theta$, for any random sample observation $X$, is defined by the pair of $L_\Theta(X)$ and $U_\Theta(X)$, where $L_\Theta(X) \leq \Theta \leq U_\Theta(X)$. This interval, together with the probability $\Pr\big(\Theta \in [L_\Theta(X), U_\Theta(X)]\big)$, is called confidence interval [25].

*Statement* 1. Since $\theta_n$'s are random samples of a normal distribution with $\mathcal{N}(\mu_\theta, \sigma_\theta^2)$, from the statistical inference theorem [25] the mean $\bar{\theta}$ and the standard deviation of samples $S_\theta$ are also independent random variables and $\bar{\theta}$ has a normal distribution, $\mathcal{N}(\mu_\theta, \sigma_\theta^2/2)$. Besides, $\dfrac{\bar{\theta} - \mu_\theta}{S_\theta/\sqrt{N}}$ has a Student's t-distribution with $N - 1$ degrees of freedom.

*Statement* 2. Considering definitions (1)-(2) and for a pivotal quantity $\mathcal{V}(X, \Theta)$ and parameter $\Theta$, a confidence interval is defined as $\Pr\big(L_\Theta(X) \leq \mathcal{V}(X,\Theta) \leq U_\Theta(X)\big) \geq 1-\alpha$, where $1-\alpha$ shows its amount of confidence degree.

Given the above statements, after $\{\theta_n, \phi_n\}_{n=1}^N$ are derived from the vehicle's location, the *angular uncertainty region* of the vehicle will be calculated as follows. Let us perform the analysis for $\theta$ plane with mean value $\mu_\theta$ and standard deviation $\sigma_\theta$. It can be inferred that $\dfrac{\bar{\theta}-\mu_\theta}{S_\theta/\sqrt{N}}$ is a pivot since the Student's t-distribution does not depend on $\mu_\theta$ and $\sigma_\theta$. For vehicle $j$, it can be derived as,

$$\Pr(\theta_L^{(j)} \leq \mu_\theta^{(j)} \leq \theta_U^{(j)}) \geq 1-\alpha, \tag{3.1}$$

where $\theta_L^{(j)}$ and $\theta_U^{(j)}$ are the interval boundaries and can be calculated from the following equations.

$$\theta_L^{(j)} = \bar{\theta}^{(j)} - \mathcal{T}_{N-1,\alpha/2}\frac{S_\theta^{(j)}}{\sqrt{N}}, \tag{3.2a}$$

$$\theta_U^{(j)} = \bar{\theta}^{(j)} + \mathcal{T}_{N-1,\alpha/2}\frac{S_\theta^{(j)}}{\sqrt{N}}, \tag{3.2b}$$

where $\mathcal{T}_{N-1,\alpha/2}$ is the Student's t-distribution critical value with $N-1$ degrees of freedom. Furthermore, mean and standard deviation can be estimated as,

$$\bar{\theta}^{(j)} = \sum_{n=1}^{N}\frac{\theta_n^{(j)}}{N}, \tag{3.3}$$

$$S_\theta^{(j)} = \left[\frac{1}{N-1}\sum_{n=1}^{N}\left(\theta_n^{(j)} - \bar{\theta}^{(j)}\right)^2\right]^{\frac{1}{2}}. \tag{3.4}$$

In our solution, the transmitter adjusts its beam's direction towards each vehicle and modifies its beamwidth within a confident range to cover the desired user. Beam's direction in the elevation and azimuth planes are defined as $\bar{\theta}^{(j)}$ and $\bar{\phi}^{(j)}$, respectively, and the beamwidth is assumed homogeneous in both planes. For the elevation plane, the beamwidth is chosen in such a way that it is equal to the confidence interval of $\bar{\theta}^{(j)}$. In other words, the transmitter forms the beam in an interval of $\pm\mathcal{T}_{N-1,\alpha/2}S_\theta^{(j)}/\sqrt{N}$ around $\bar{\theta}^{(j)}$, i.e.,

$$\Gamma_\theta^{(j)} = \bar{\theta}^{(j)}, \quad W_\theta^{(j)} = \theta_U^{(j)} - \theta_L^{(j)}, \tag{3.5}$$

where $\Gamma_{\theta}^{(j)}$ and $W_{\theta}^{(j)}$ are the AOD towards vehicle $j$ and the Half Power Beam Width (HPBW), respectively. Similar calculations can be performed for $\Gamma_{\phi}^{(j)}$ and $W_{\phi}^{(j)}$. A reliable method to control the beam's direction is to use arrays of acoustic transducers. The required acoustic specifications, such as the maximum source level, maximum acoustic power, directivity index and AOD, number of steered beams, and beam widths define the number, geometrical arrangement, as well as the relative amplitudes and phases of the array elements [22, 82]. The antenna array contains multiple individual projectors to direct the acoustic power in the desired direction $\Gamma^{(j)}$ and beamwidth $W^{(j)}$ for vehicle $j$ in elevation or azimuth planes. Using more elements in the array leads to a higher gain and, therefore, a narrower beamwidth. The main beam always points at the desired direction, while the beamwidth depends on the effective aperture. Beamwidth can be controlled by changing the relative amplitudes and phases of the transducers. Novel array designs will provide the required narrow beams in the direction of interest [22]. A configurable software-defined platform, which utilizes a broadband phased array transducer, can achieve the required goals in a single unit on the fly, while the cost of having such multiple functions reduces [26].

**Fine Estimation:** As the vehicle moves, new measurements are acquired periodically and a fine estimation is required on top of the coarse estimation to reduce the energy dispersion and to support vehicle's mobility as explained below. We propose a two-stage solution to handle the estimation in a continuous manner when the vehicles move and take new location samples. The calculated parameters and predictions by interval estimation should be updated by a Kalman Filtering in order to track the vehicles. The following statements and theorems study the cases in details.

*Statement* 3. Let $\Omega$ be the space of infinite and countable states; the stochastic process $\{\gamma_n\}_{n\in N}$, whose components are in $\Omega$, is said to possess the *Markov property* if the probability follows $\Pr[\gamma_{n+1} = \mathcal{C}_{n+1} \mid \{\gamma_0 = \mathcal{C}_0, ..., \gamma_{n-1} = \mathcal{C}_{n-1}, \gamma_n = \mathcal{C}_n\}] = \Pr[\gamma_{n+1} = \mathcal{C}_{n+1} \mid \gamma_n = \mathcal{C}_n]$.

*Theorem* 1. Given $N$ samples $\{\theta_n\}_{n\in N}$, mean $\bar{\theta}_n$ and standard deviation $S_{\theta,n}$ are sequences of random variables and possess the Markov Property, when the new sample $\theta_{N+1}$ is accumulated in the sequence of $N+1$ samples.

*Proof* 1. While the accumulation is performed, we have $N$ samples at state $k$ and $N+1$ samples at state $k+1$. For these states we can write $\bar{\theta}_k = \dfrac{1}{N}\sum_{n=1}^{N}\theta_n$ and $\bar{\theta}_{k+1} = \dfrac{1}{N+1}\sum_{n=1}^{N+1}\theta_n = \dfrac{1}{N+1}(N\bar{\theta}_k + \theta_{N+1})$. Therefore, $\Pr[\bar{\theta}_{k+1} \mid \{\bar{\theta}_0, ..., \bar{\theta}_{k-1}, \bar{\theta}_k\}] = \Pr[\bar{\theta}_{k+1} \mid \bar{\theta}_k]$. Similarly, for the estimated standard deviation we can conclude that, $S_{\theta,k+1} = \left[\frac{1}{N+1}\sum_{n=1}^{N+1}\left(\theta_n - \bar{\theta}_{k+1}\right)^2\right]^{\frac{1}{2}} = \left[\frac{1}{N+1}\Big(\sum_{n=1}^{N}\left(\theta_n - \frac{N\bar{\theta}_k + \theta_{N+1}}{N+1}\right)^2 + \left(\theta_{N+1} - \bar{\theta}_{k+1}\right)^2\Big)\right]^{\frac{1}{2}} = \left[\frac{1}{N+1}\Big(\sum_{n=1}^{N}\left(\theta_n - \frac{(N+1)\bar{\theta}_k - \bar{\theta}_k + \theta_{N+1}}{N+1}\right)^2 + \left(\theta_{N+1} - \bar{\theta}_{k+1}\right)^2\Big)\right]^{\frac{1}{2}}$. If $\sum_{n=1}^{N}\left(\theta_n - \bar{\theta}_k\right)^2$ is replaced by $N{S_{\theta,k}}^2$, then $S_{\theta,k+1} = \left[\frac{1}{N+1}\left(N{S_{\theta,k}}^2 + N\left(\frac{\theta_{N+1} - \bar{\theta}_k}{N+1}\right)^2 + \left(\theta_{N+1} - \bar{\theta}_{k+1}\right)^2 - \frac{2(\theta_{N+1} - \bar{\theta}_k)}{N+1}\sum_{n=1}^{N}(\theta_n - \bar{\theta}_k)\right)\right]^{\frac{1}{2}}$. Since $\sum_{n=1}^{N}\theta_n = N\bar{\theta}_k$, the above equation results in $S_{\theta,k+1} = \left[\frac{1}{N+1}\left(N{S_{\theta,k}}^2 + N\left(\frac{\theta_{N+1} - \bar{\theta}_k}{N+1}\right)^2 + \left(\theta_{N+1} - \bar{\theta}_{k+1}\right)^2\right)\right]^{\frac{1}{2}}$.

Now, since $S_{\theta_{k+1}}$ is a function of two variables (i.e., $\bar{\theta}$ and $S_{(\theta)}$), where $\bar{\theta}_k$ and $\bar{\theta}_{k+1}$ are sequences of i.i.d. random variables, the Markov property holds by definition. Therefore, $\Pr[S_{\theta,k+1} \mid \{S_{\theta,0}, ..., S_{\theta,k-1}, S_{\theta,k}\}] = \Pr[S_{\theta,k+1} \mid S_{\theta,k}]$.

*Theorem* 2. Given $N$ samples $\{\theta_n\}_{n\in N}$, mean $\bar{\theta}_n$ and standard deviation $S_{\theta,n}$ are sequences of random variables and possess the Markov Property when the new sample $\theta_{N+1}$ is considered in the sequence of $N$ samples $n = 2, ..., N+1$.

*Proof* 2. There are $N$ samples at states $k$ and $k+1$, if we use a fixed-size sliding window. Therefore, $\bar{\theta}_{k+1} = \frac{1}{N}\sum_{n=2}^{N+1}\theta_n = \frac{1}{N}(\sum_{n=1}^{N}\theta_n + \theta_{N+1} - \theta_1) = \bar{\theta}_k + \frac{1}{N}(\theta_{N+1} - \theta_1)$ and the Markov property holds. Similar calculations can be written for state $k+1$ of standard deviation as, $S_{\theta,k+1} = \left[\frac{1}{N}\sum_{n=2}^{N+1}\left(\theta_n - \bar{\theta}_{k+1}\right)^2\right]^{\frac{1}{2}} = \left[\frac{1}{N}\Big(\sum_{n=1}^{N}\left(\theta_n - \bar{\theta}_k - \frac{\theta_{N+1} - \theta_1}{N}\right)^2 + \left(\theta_{N+1} - \bar{\theta}_{k+1}\right)^2 - \left(\theta_1 - \bar{\theta}_{k+1}\right)^2\Big)\right]^{\frac{1}{2}}$. By replacing $\sum_{n=1}^{N}\left(\theta_n - \bar{\theta}_k\right)^2$ by $N{S_{\theta_k}}^2$, we conclude $S_{\theta,k+1} = \left[\frac{1}{N}\left(N{S_{\theta_k}}^2 + \frac{\left(\theta_{N+1} - \bar{\theta}_k\right)^2}{N} + \left(\theta_{N+1} - \bar{\theta}_{k+1}\right)^2 - \left(\theta_1 - \bar{\theta}_{k+1}\right)^2 - \frac{2(\theta_{N+1} - \bar{\theta}_k)}{N}\sum_{n=1}^{N}(\theta_n - \bar{\theta}_k)\right)\right]^{\frac{1}{2}}$. Since the last term in this equation equals zero, we have $S_{\theta,k+1} = \left[\frac{1}{N}\Big(N{S_{\theta_k}}^2 + \frac{\left(\theta_{N+1} - \bar{\theta}_k\right)^2}{N} + \left(\theta_{N+1} - \bar{\theta}_{k+1}\right)^2 - \left(\theta_1 - \bar{\theta}_{k+1}\right)^2\Big)\right]^{\frac{1}{2}}$, which proves that $S_\theta$ also defines a Markov property.

Given the Markov property of beam parameters and the previous angle estimations for vehicle $j$ at state $k$, i.e., $\bar{\theta}_k^{(j)}$ and $S_{\theta,k}^{(j)}$, the new estimations are updated at state $k+1$ as $\bar{\theta}_{k+1}^{(j)}$ and $S_{\theta,k+1}^{(j)}$. To stabilize the variations, we consider both the uncertainties in new samples and in the latest estimation of the previous state until the current state. We assign a weight to them by using *Kalman Filtering*. We update the current state of parameter vector $\Theta$, which stands for the random variables

$[\Gamma_\theta(\bar{\theta}), W_\theta(S_\theta)]$ for vehicle $j$ as follows,

$$\Theta_{k+1}^{(j)} = \mathcal{F}(\Theta_k^{(j)}) + Z_k, \tag{3.6a}$$

$$Y_{k+1}^{(j)} = \mathcal{H}(\Theta_{k+1}^{(j)}) + V_k, \tag{3.6b}$$

where $Z_k$ and $V_k$ represent the process and measurement uncorrelated Gaussian noise (with covariance $Q_Z$ and $Q_V$), respectively. Functions $\mathcal{F}$ and $\mathcal{H}$ are the state-transition models, which are generally nonlinear. Based on the observed angle, the interval estimation calculates an estimation given $\mathcal{C}_{k+1}^{(j)}$—which is defined as a Markov state—with transition probability $\Pr[\widehat{\Theta}_{k+1}^{(j)} = \mathcal{C}_{k+1}^{(j)} | \Theta_k^{(j)} = \mathcal{C}_k^{(j)}]$. The probability of possible state for time instant $k+1$ equals to $\Pr[\widehat{\Theta}_{k+1}^{(j)} = \mathcal{C}_{k+1}^{(j)} | \Theta_k^{(j)} = \mathcal{C}_k^{(j)}] \Pr[\Theta_k^{(j)} = \mathcal{C}_k^{(j)}]$, where $\{\Theta_k^{(j)}\}_{k \in \mathcal{S}}$ takes its values from state space $\mathcal{S}$. For vehicle $j$, transition to the next state at time instant $k+1$ could be done via one of these three cases: *Case $I$ (**A**)*—sample accumulation; *Case $II$ (**S.W.**)*—fixed-size sliding window; or *Case $III$ (**C.E.**)*—staying in the current estimate.

$$\Pi_{k+1|k}^{(j)} = \begin{cases} \Pi_I^{(j)}, & \text{if} \{n\}_{k+1} = \{1, ..., N+1\}, \\ \Pi_{II}^{(j)}, & \text{if} \{n\}_{k+1} = \{2, ..., N+1\}, \\ \Pi_{III}^{(j)}, & \text{if} \{n\}_{k+1} = \{1, ..., N\}, \end{cases} \tag{3.7}$$

where $\Pi_{k+1|k}^{(j)} \in [0, 1]$ is the transition probability for vehicle $j$ and $\sum_{s=I}^{III} \Pi_s^{(j)} = 1$. Based on the following statement, two non-identical cases, Case $II$ and Case $III$, are defined in (3.7).

*Statement* 4. Considering every observation $\theta_{k+1}$ at time instant $k+1$, the behavior of the updated estimation ($Y_{k+1}^{(j)}$) in terms of mean and standard deviation is in general time dependent if the number of samples is not sufficiently large.

*Statement* 5. Let $\theta_{k+1}$ be the observed angle sample at time instant $k+1$. Our MAC protocol decides on the next case as follows. Case $I$ in (3.7) is superior when the updated estimations of the mean and the standard deviation need to be more accurate. Sample accumulation offers a better interval estimation for the beamwidth and AOD if the number of samples is bound to increase. This leads to a narrower beam, which possibly decreases the probability of interference but increases the probability of miss. On the other hand, keeping a constant number of samples, i.e., using a fixed-size sliding window, as in Case $II$, shows the time-dependent nature of the estimation as the

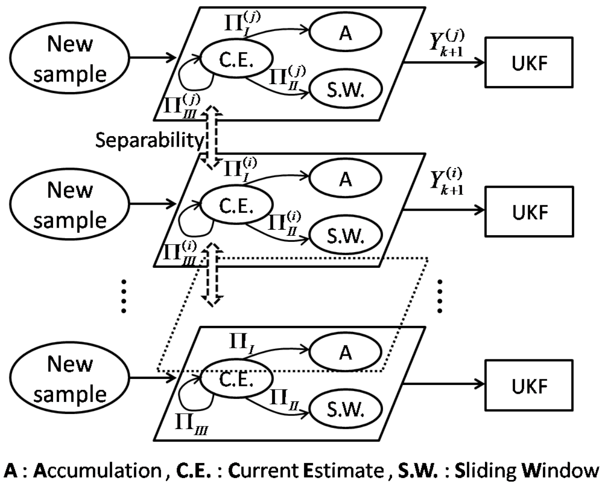



Figure 3.4: MAC protocol decides on the next estimate of each vehicle when new samples are acquired. Coarse estimation is updated via one of the cases, as in Statement 5.

vehicles change their positions. This fixed-size beamwidth does not increase the probability of miss. Case $III$ is preferred to reduce the probability of miss, when the probability of interference is very small, i.e., when the vehicles are separated in space.

Fig. 3.4 depicts how our MAC decides on the next case based on the new sample and also on the separability of the vehicles. Transition from one case to another in each vehicle, as part of coarse estimation, considers the given information of other neighbors so as to fulfill the separability of the users. The transition probabilities were defined in (3.7) and the output of this process is applied to the fine estimation, i.e., the Kalman Filtering.

*Unscented Kalman Filtering (UKF)* is applied to this problem as follows. Let $\widetilde{\Theta}_k$ and $\widetilde{Q}_{\Theta_k}$ be the mean and covariance of parameter $\Theta$ with the initial value calculated from the interval estimation output. To find the statistics of the random variable $Y$, we form $2\lambda + 1$ sigma points $\zeta$ for vehicle $j$—whose index is ignored in the following for the sake of notation simplicity—as follows,

$$\zeta_{i,k} = \begin{cases} \widetilde{\Theta}_k & i = 0, \\ \widetilde{\Theta}_k + (\beta\sqrt{\lambda \widetilde{Q}_{\Theta_k}})_i, & i = 1, ..., \lambda, \\ \widetilde{\Theta}_k - (\beta\sqrt{\lambda \widetilde{Q}_{\Theta_k}})_i, & i = \lambda + 1, ..., 2\lambda, \end{cases} \tag{3.8}$$

where $\beta$ is a scaling parameter and $(.)_i$ stands for the $i$th column of the square root matrix. Sigma

Table 3.2: Notations and Mathematical Terms: Fine Estimation.

| **Notation** | **Description** |
|---|---|
| $V_k$ | Measurement noise |
| $Z_k$ | Process noise |
| $Q_Z, Q_V$ | Covariance of noise $Z_k$ and $V_k$ |
| $\mathcal{F}, \mathcal{H}$ | State-transition models |
| $\Pi_{k+1\|k}^{(j)}$ | Transition probability for vehicle $(j)$ |
| $Y_{k+1}^{(j)}$ | Observation estimation |
| $\widetilde{\Theta}_k$ | Mean of the parameter $\Theta$ |
| $\widetilde{Q}_{\Theta_k}$ | Covariance of the parameter $\Theta$ |
| $\mathcal{K}_G$ | Kalman gain |
| $\zeta_{i,k}$ | Sigma points for vehicle $(j)$ |
| $\lambda$ | Number of the sigma points $\zeta$ |
| $\beta$ | Scaling parameter |
| $P_I^{(i,j)}$ | Interference of vehicles $(i)$, $(j)$ |
| $\mu_I$, $\sigma_I$ | Mean and standard deviation of interference |
| $\Theta'$ | Vector of centered variables |
| $\Theta'^*$ | Optimum parameters |
| $W_{\theta 0}^{(\jmath)}$ | Initial HPBW |
| $\Gamma_{\theta 0}^{(\jmath)}$ | Initial AOD |
| $S_{\theta' 0}^{(\jmath)}$ | Initial standard deviation |
| $\theta'^{(i)}_{0L}$, $\theta'^{(j)}_{0U}$ | Lower and upper boundaries |
| $\rho$ | Control parameter |
| $\mathcal{D}$ | Constraints domain |
| $(\bar{\theta}')^*$, $(S_{\theta'})^*$ | Optimum mean and standard deviation |
| $\mathcal{R}_I^{(j)}$ | Retransmission rate (interference) |
| $\mathcal{R}_m^{(j)}$ | Retransmission rate (miss) |
| $\mathcal{R}_{TS}$ | Time slot usage ratio for T-SDMA |
| $n_c$, $c \in \hat{c}$ | Total number of time slots, clusters |
| $M_1$ | Maximum number of vehicles in cluster |
| $M_2$ | Number of clusters in inter-clustered group |
| $M_3$ | Maximum vehicles per cluster in inter-clustered group |
| $\mathcal{R}_T$ | Time slot usage ratio |
| $M_T$ | Total number of time slots |
| $\eta$ | Rate efficiency |
| $\tilde{\mathcal{R}}_{TS}$ | Data rate per vehicle per transmitting frame |
| $T_s$ | Time duration of a time slot |
| $N_p$, $L_p$ | Number and length of packets |
| $SW$ | Number of attempts (retransmissions) |

vectors go through the nonlinear function $\mathcal{F}$, and the mean and covariance for $\chi_{i,k+1|k}$ are approximated via the following equations [149, 154],

$$\chi_{i,k+1|k} = \mathcal{F}(\zeta_{i,k}), \quad i = 1, ..., 2\lambda, \tag{3.9}$$

$$\widetilde{\Theta}_{k+1|k} = \sum_{i=0}^{2\lambda} \eta_i \chi_{i,k+1|k}, \tag{3.10}$$

$$\widetilde{Q}_{\Theta_{k+1|k}} = \sum_{i=0}^{2\lambda} \eta_i (\chi_{i,k+1|k} - \widetilde{\Theta}_{k+1|k})(\chi_{i,k+1|k} - \widetilde{\Theta}_{k+1|k})^T + Q_Z, \tag{3.11}$$

where the transpose operation is shown by $^T$ and the weights are defined as $\eta_0 = 1 - 1/\beta^2$ and $\eta_i = 1/(2\lambda\beta^2)$, for $i = 1, ..., 2\lambda$.

$$\mathcal{Y}_{i,k+1|k} = \mathcal{H}(\chi_{i,k+1|k}), i = 1, ..., 2\lambda, \tag{3.12a}$$

$$\widetilde{Y}_{k+1|k} = \sum_{i=0}^{2\lambda} \eta_i \mathcal{Y}_{i,k+1|k}, \tag{3.12b}$$

$$\widetilde{Q}_{Y_{k+1|k}} = \sum_{i=0}^{2\lambda} \eta_i (\mathcal{Y}_{i,k+1|k} - \widetilde{Y}_{k+1|k})(\mathcal{Y}_{i,k+1|k} - \widetilde{Y}_{k+1|k})^T + Q_V, \tag{3.13}$$

$$\widetilde{Q}_{\Theta Y_{k+1|k}} = \sum_{i=0}^{2\lambda} \eta_i (\chi_{i,k+1|k} - \widetilde{\Theta}_{k+1|k})(\mathcal{Y}_{i,k+1|k} - \widetilde{Y}_{k+1|k})^T. \tag{3.14}$$

Now, the UKF gain $\mathcal{K}_G$ is defined as $\mathcal{K}_G = \widetilde{Q}_{\Theta Y_{k+1|k}} \widetilde{Q}^{-1}_{Y_{k+1|k}}$.

The next state estimation and its covariance are applied to the antenna at state $k+1$. The corresponding estimations can be written as,

$$\widehat{\Theta}_{k+1} = \widetilde{\Theta}_{k+1|k} + \mathcal{K}_G (Y_{k+1} - \widetilde{Y}_{k+1|k}), \tag{3.15a}$$

$$Q_{\Theta_{k+1}} = \widetilde{Q}_{\Theta_{k+1|k}} - \mathcal{K}_G \widetilde{Q}_{\Theta Y_{k+1|k}} {\mathcal{K}_G}^T, \tag{3.15b}$$

while the error vectors can be written as $\epsilon_{k+1} = \Theta_{k+1} - \widehat{\Theta}_{k+1}$ and $\epsilon_{k+1|k} = \Theta_{k+1} - \widetilde{\Theta}_{k+1|k}$.

To quantify the real prediction error covariance matrix, the approximation method presented in [154] can be used. The residual of the observation can be computed as,

$$\upsilon_{k+1} = Y_{k+1} - \mathcal{H}(\widetilde{Y}_{k+1|k}), \tag{3.16}$$

while the residual covariance matrix is $Q_{\upsilon_{k+1|k}} = E[\upsilon_{k+1} {\upsilon_{k+1}}^T]$.

If the vehicles are still overlapped, the next macro-state will be the statistical interference cancellation; otherwise, the system shifts to the transmit state, as explained in Fig. 3.2. Table 3.2

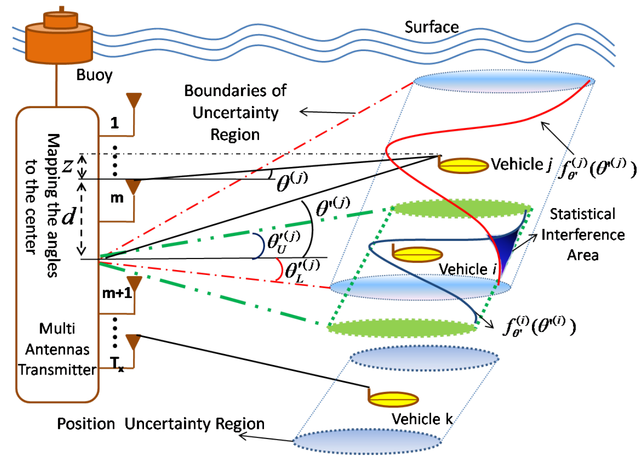



Figure 3.5: Configuration of the system while vehicles $j$ and $i$ overlap.

summarizes the main mathematical terms used in the fine estimation and the rest of this chapter.

**Statistical Interference Cancellation:** Let $\mathcal{J}$ be the number of underwater vehicles deployed in the body of water with $\mathcal{J} \leq T_x$, where $T_x$ is the number of buoy's antennae. Two vehicles $j$ and $i$ ($j, i \in \mathcal{J}$) probabilistically might overlap considering their uncertainty regions, as shown in Fig. 3.5. If the AODs for each pair of vehicles are $\Gamma_\theta^{(j)} = \bar{\theta}^{(j)}$ and $\Gamma_\theta^{(i)} = \bar{\theta}^{(i)}$, they can be mapped from the $mth$ and $(m+1)th$ antennae, $\forall m \in \{1, ..., T_x\}$ to their center point, where $\bar{\theta}'^{(j)}$ and $\bar{\theta}'^{(i)}$ are the transferred angles corresponding to vehicles $j$ and $i$ in the $\theta$'s plane. We can conclude that $\tan(\bar{\theta}'^{(j)}) = (1 + d_m/z_m^{(j)}) \tan(\bar{\theta}^{(j)})$, where $z_m^{(j)}$ is the vertical distance between the depth of vehicle $j$ and antenna $m$, and $d_m$ is the distance between the center point and antenna $m$. As $z_m^{(j)} \gg d_m$, it can be concluded that $\bar{\theta}'^{(j)} \approx \bar{\theta}^{(j)}$.

As shown in Fig. 3.5 and given the probability distribution of the vehicles' positions, interference might occur if the upper boundary in the uncertainty region of vehicle $i$ overpasses the lower boundary in the uncertainty region of vehicle $j$; we call this situation *statistical interference*.

*Statement* 6. For random variables $\theta'^{(j)}$ and $\theta'^{(i)}$, with means of $\bar{\theta}'^{(j)}$ and $\bar{\theta}'^{(i)}$ and standard deviations of $S_{\theta'}^{(j)}$ and $S_{\theta'}^{(i)}$, we define the *probabilistic separability* as $\Pr\left(\theta'^{(i)} < \theta'^{(j)}\right)$. If the interference occurs, the overlapping area of the distribution functions of $\theta'^{(j)}$ and $\theta'^{(i)}$, as specified in Fig. 3.5, represents the statistical interference of two vehicles as $P_I^{(i,j)} = 1 - \Pr\left(\theta'^{(i)} < \theta'^{(j)}\right)$.

*Lemma* 1. Interference is modeled by normal probability distribution function with the mean and the standard deviation of $\mu_I = \bar{\theta}'^{(i)} - \bar{\theta}'^{(j)}$ and $\sigma_I = \sqrt{(S_{\theta'}^{(j)})^2 + (S_{\theta'}^{(i)})^2}$, respectively.

*Proof* 3. $P_I^{(i,j)} = \Pr\left(\theta'^{(i)} > \theta'^{(j)}\right) = Pr\left(\theta'_I > 0\right)$, where $\theta'_I = \theta'^{(i)} - \theta'^{(j)}$. Since $\theta'_I$ is the weighted sum of two independent normal random variables, it has a normal distribution with mean and standard deviation values as specified in Lemma 1.

We present a method that minimizes the acoustic beam interference by considering the updated position of each vehicle. This method finds a focused beam for the overlapping vehicles so as to reduce the interference, however, this may cause the vehicles to fall out of the coverage. There is a trade-off between the probability of interference and the probability of miss; hence, an optimization problem is required to find the beam parameters to minimize the interference while satisfying coverage requirements. By keeping the transmitter antenna's beamwidth at an optimal value, the proposed method finds a desirable trade-off among three contrasting goals: $(i)$ spreading the signal beam towards the receiver to combat position uncertainty; $(ii)$ focusing such beam to reduce acoustic energy dispersion; and $(iii)$ minimizing interference to other vehicles in the surrounding. In the case that there are more than two interfering vehicles, the interfering vehicles will be sorted from high to low and interference cancellation will be performed for each pair of interfering vehicles.

*Lemma* 2. Minimizing the statistical interference is paramount to minimize $\dfrac{\sqrt{(S_{\theta'}^{(j)})^2 + (S_{\theta'}^{(i)})^2}}{\bar{\theta}'^{(j)} - \bar{\theta}'^{(i)}}$.

*Proof* 4. Starting from our definition of interference $P_I^{(i,j)} = Pr\big(\theta'_I > 0\big)$, we have,

$$P_I^{(i,j)} = \int_0^{\infty} f_I(\theta'_I) d\theta'_I = \frac{1}{\sigma_I\sqrt{2\pi}} \int_0^{\infty} \exp\left\{-\frac{1}{2}\left(\frac{\theta'_I - \mu_I}{\sigma_I}\right)^2\right\} d\theta'_I. \tag{3.17}$$

By defining the auxiliary variable $x = (\theta'_I - \mu_I)/\sigma_I$, we obtain,

$$P_I^{(i,j)} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(\mu_I/\sigma_I)} e^{-(x^2/2)} dx = \Phi(\frac{\mu_I}{\sigma_I}), \tag{3.18}$$

where $\Phi(.)$ is the Cumulative Distribution Function (CDF) of the standard normal distribution.

$$P_I^{(i,j)} = \Phi\left(-\frac{\bar{\theta}'^{(j)} - \bar{\theta}'^{(i)}}{\sqrt{(S_{\theta'}^{(j)})^2 + (S_{\theta'}^{(i)})^2}}\right) = 1 - \Phi\left(\frac{\bar{\theta}'^{(j)} - \bar{\theta}'^{(i)}}{\sqrt{(S_{\theta'}^{(j)})^2 + (S_{\theta'}^{(i)})^2}}\right), \tag{3.19}$$

where the equality is concluded via the rotational symmetry characteristic of $\Phi(.)$. From (3.19), it is inferred that,

$$\frac{\bar{\theta}'^{(j)} - \bar{\theta}'^{(i)}}{\sqrt{(S_{\theta'}^{(j)})^2 + (S_{\theta'}^{(i)})^2}} = \Phi^{-1}(1 - P_I^{(i,j)}). \tag{3.20}$$

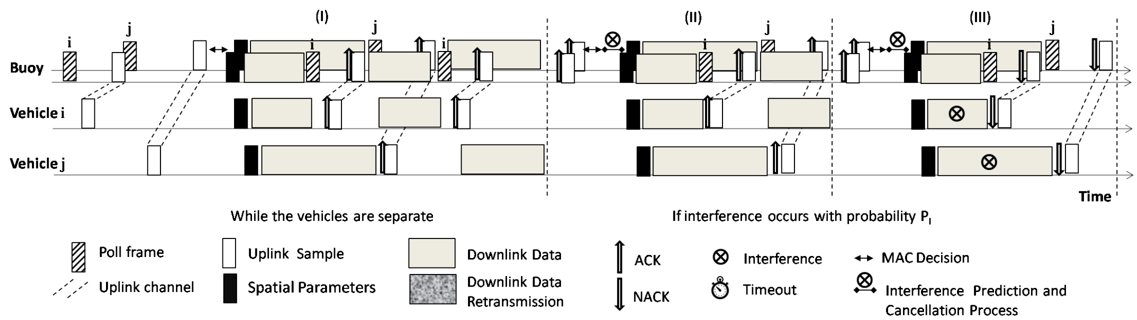



Figure 3.6: Timeline showing the interaction between the buoy and vehicles $i$ and $j$ in the transmission range. Comparison between different incidents is shown in different transmission rounds. In round $(I)$, reliable transmission is ensured since the vehicles are spatially separate; however, in rounds $(II)$ and $(III)$, the interference occurs.

Here, $\Phi^{-1}(.)$ is the Probit function and $\Phi^{-1}(1-P_I^{(i,j)}) \to \infty$ when $(1-P_I^{(i,j)}) \to 1$. Therefor, the left side of (3.20) approaches $\infty$ when $P_I^{(i,j)} \to 0$. It proves that we can minimize the statistical interference by minimizing $\frac{\sqrt{(S_{\theta'}^{(j)})^2 + (S_{\theta'}^{(i)})^2}}{\bar{\theta}'^{(j)} - \bar{\theta}'^{(i)}}$.

Let $\Theta' = [\bar{\theta}'^{(j)}, S_{\theta'}^{(j)}, \bar{\theta}'^{(i)}, S_{\theta'}^{(i)}]$ be the vector of variables. An optimization problem with the objective function introduced in Lemma 2 is cast as follows.

$$\text{Given}: \quad W_{\theta 0}^{(\jmath)}, \Gamma_{\theta 0}^{(\jmath)}, S_{\theta' 0}^{(\jmath)}, \theta_{0L}'^{(i)}, \theta_{0U}'^{(j)}, \rho, N, \alpha, \quad \jmath = i, j;$$

$$\text{Find}: \quad (\bar{\theta}'^{(\jmath)})^*, (S_{\theta'}^{(\jmath)})^*$$

$$\min_{\Theta'} \quad \Lambda(\Theta') = \frac{\sqrt{(S_{\theta'}^{(j)})^2 + (S_{\theta'}^{(i)})^2}}{\bar{\theta}'^{(j)} - \bar{\theta}'^{(i)}} \tag{3.21a}$$

$$\text{s.t.} \quad 2(\mathcal{T}_{N-1,\alpha/2}\frac{S_{\theta'}^{(\jmath)}}{\sqrt{N}}) \le W_{\theta_0}^{(\jmath)}, \ \jmath = i, j, \tag{3.21b}$$

$$\bar{\theta}'^{(j)} + (\mathcal{T}_{N-1,\alpha/2}\frac{S_{\theta'}^{(j)}}{\sqrt{N}}) \le \theta_{0U}'^{(j)}, \tag{3.21c}$$

$$\bar{\theta}'^{(i)} - (\mathcal{T}_{N-1,\alpha/2}\frac{S_{\theta'}^{(i)}}{\sqrt{N}}) \ge \theta_{0L}'^{(i)}, \tag{3.21d}$$

$$\bar{\theta}'^{(\jmath)} + (\mathcal{T}_{N-1,\alpha/2}\frac{S_{\theta'}^{(\jmath)}}{\sqrt{N}}) \ge \bar{\theta}_0'^{(\jmath)} + \rho\frac{S_{\theta' 0}^{(\jmath)}}{4}, \tag{3.21e}$$

$$\bar{\theta}'^{(\jmath)} - (\mathcal{T}_{N-1,\alpha/2}\frac{S_{\theta'}^{(\jmath)}}{\sqrt{N}}) \le \bar{\theta}_0'^{(\jmath)} - \rho\frac{S_{\theta' 0}^{(\jmath)}}{4}. \tag{3.21f}$$

This problem is a constrained multivariable fractional nonlinear programming as a ratio of chosen real-valued functions from a set $\mathcal{D}$ of space $\mathbb{R}^4$. Both the numerator and denominator are positive

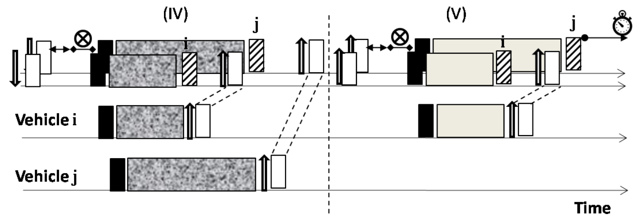



Figure 3.7: Timeline showing how MAC handles the transmission in the presence of possible interference when the vehicles are within the buoy's transmission range as presented in round $(IV)$. If the vehicle falls out of the angular coverage, the buoy will be notified by a timeout as presented in round $(V)$.

for all values of $\Theta'$ which are defined by the constraints domain set $\mathcal{D}$. The numerator is a *norm* function which is convex [19], while maximizing the denominator minimizes the objective function. Since $\bar{\theta}'^{(j)} > \bar{\theta}'^{(i)}$, the maximum value of the denominator would be $\max(\bar{\theta}'^{(j)}) - \min(\bar{\theta}'^{(i)})$ over the domain. However, the constraints optimize the mentioned value to satisfy the interference-miss trade-off as much as possible.

The objective function in (5.10a) minimizes the interference area by maximizing the distance between AODs of overlapping vehicles while their HPBWs in the numerator are minimized. Moreover, (3.21b) controls the beamwidth such that it does not exceed the initial estimated beamwidth. While constraints (6.6b)-(3.21d) keep the new beam boundaries, i.e., $\theta_L'^*$ and $\theta_U'^*$, inside the uncertainty regions, we prevent the beams to become too narrow via constraints (3.21e) and (3.21f). Through these constraints, the new boundaries are controlled so that the probability of miss does not exceed the MAC specified value. The tunable control parameter $0 \leq \rho < (\sqrt{2}\mathcal{T}_{N-1,\alpha/2})/\sqrt{N}$ in (3.21e)-(3.21f) is defined according to the target interference-miss trade-off; small values of $\rho$ mean MAC prefers handling larger probability of miss than probability of interference. In order to investigate the solution of the proposed fractional program, we use the *parametric* approach presented in [38, 121]. The set $\mathcal{D}$ is compact, and both numerator and denominator are continuous on it, so we cast the equivalent parametric program of (5.10a) as,

$$\mathcal{L}(\psi) = \min\{\sqrt{\big(S_{\theta'}^{(j)}\big)^2 + \big(S_{\theta'}^{(i)}\big)^2} - \psi\big(\bar{\theta}'^{(j)} - \bar{\theta}'^{(i)}\big)\}, \tag{3.22}$$

for $\Theta' \in \mathcal{D}$. The optimal solution of this problem is also the optimal solution of the fractional program. This iterative approach starts from an initial value for $\psi$ and solves (3.22) until convergence is achieved for a specific threshold through the following steps.

**Step 0**—Initialization: $\psi = 0$.

**Step 1**—Solve (3.22) and find $\Theta'^{*}$; if the value of $\mathcal{L}(0)$ does not satisfy the threshold, move to the next step.

**Step 2**—$\psi = \Lambda(\Theta'^{*})$.

**Step 3**—Solve (3.22) and update $\Theta'^{*}$ with the new result; if $\mathcal{L}(\psi)$ fulfills the threshold, stop; otherwise, repeat steps 2 and 3 until the threshold is satisfied. The optimum value is the one which is associated with the most recent $\Theta'^{*}$.

The solution of the presented program determines the optimum values of angles for minimum probability of interference at a tolerable miss rate. We define the *probabilistic retransmission rate* as a measure of probabilistic failure or disconnectivity, where the former is the result of interference and the latter is the consequence of missing the vehicle. This metric is calculated as in (3.23) if it is the result of interference, whereas it is computed using (3.24) in case of missing the coverage,

$$\mathcal{R}_I^{(j)} = \frac{P_I^{(i,j)}}{\int_{\theta'^{(j)}_L}^{\theta'^{(j)}_U} f_{\theta'^{(j)}}\left(\theta'^{(j)}\right) d\theta'^{(j)}}, \tag{3.23}$$

$$\mathcal{R}_m^{(j)} = \frac{\int_{\theta'^{(j)}_L}^{(\bar{\theta}'^{(j)})^* - 1/2(W^{(j)})^*} f_{\theta'^{(j)}}\left(\theta'^{(j)}\right) d\theta'^{(j)}}{\int_{\theta'^{(j)}_L}^{\theta'^{(j)}_U} f_{\theta'^{(j)}}\left(\theta'^{(j)}\right) d\theta'^{(j)}} + \frac{\int_{(\bar{\theta}'^{(j)})^* + 1/2(W^{(j)})^*}^{\theta'^{(j)}_U} f_{\theta'^{(j)}}\left(\theta'^{(j)}\right) d\theta'^{(j)}}{\int_{\theta'^{(j)}_L}^{\theta'^{(j)}_U} f_{\theta'^{(j)}}\left(\theta'^{(j)}\right) d\theta'^{(j)}}. \tag{3.24}$$

In both cases, the buoy retransmits the same data packets after re-tuning the transmit parameters.

Figs. 3.6 and 3.7 show the sequence of events that occur in multiple transmission rounds for a downlink/uplink communication between the buoy, as the initiator, and vehicles $i$ and $j$, when vehicles are inside the buoy's transmission range. We assume an asymmetric and separate spectrum utilization for downlink and uplink channels; i.e., a larger bandwidth in downlink for data transmission and a narrow-bandwidth uplink channel for acknowledgments and new vehicle's locations. Downlink exploits the proposed SDMA, while for the uplink, a polling technique, controlled by the buoy, is used to determine which vehicle is eligible to use this feedback channel at a given time. Each vehicle responds to the polling packet and sends back its new position information to the buoy, which is used in the next round of MAC decision process and spatial parameters calculation, as shown in Fig. 3.4. Returning ACK (acknowledgment) along with the position information to the

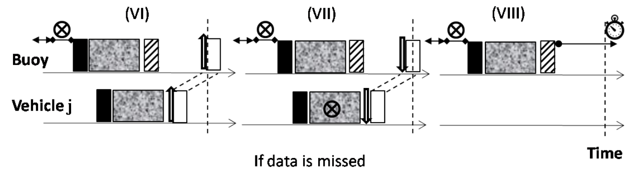



Figure 3.8: Possible situations that might happen when the data is missed, i.e., the vehicle is out of the buoy's antenna coverage but within the buoy's transmission range.

buoy confirms the reception, while a NACK signal (not acknowledgment) stands for an interference with probability of $P_I$. This probability is defined as $\Pr\left((\theta_L'^{(j)})^* < \theta' < (\theta_U'^{(i)})^*\right)$ and it means a retransmission is required, as depicted in rounds $(II)$-$(IV)$ of Figs. 3.6 - 3.7. This ACK/NACK message contains all the information on the received/discarded packets at the vehicle between two successive polling packets. As explained in round $(III)$, interference occurs when the vehicles are spatially overlapped, so data is not recoverable. Hence, the vehicles send back their NACKs in response to buoy's polling, one at a time. If the vehicle is out of coverage, then the data will be missed and other actions are required after a timeout. This incident is shown in transmission round $(V)$. Fig. 3.8 sketches the possible situations after the data is missed. In round $(VI)$, tuning the spatial parameters solves the coverage problem, so the ACK signal is received. In round $(VII)$, this tuning causes an interference between the vehicles, so the retransmission fails. Buoy is notified of missing the retransmission data by a timeout signal in round $(VIII)$. It is concluded that MAC has not accomplished a successful retransmission, so it switches to hybrid SDMA macro-state, which is discussed in the following section. Note that the proposed method makes two assumptions: $(i)$ for short/medium transmission range, in which the vehicle is within the transmission range $D$ of buoy, the above discussion works well; $(ii)$ for farther distances, i.e., in the zone $(D - R, D + R)$ and above, data may be missed and timeouts occur due to vehicles going out of transmission range. This scenario is covered in the algorithm.

**Spatially Non-separable Probabilistic SDMA:** Here, we discuss a solution for the situation in which the vehicles are fully overlapped and the separation is not possible in space. Vehicles at the same azimuth and elevation angles but different distances from buoy, are categorized as non-separable vehicles. We propose a hybrid TDMA-SDMA method, called *T-SDMA*, that uses time as the second domain and leads to an interference-free MAC solution for time-insensitive applications.

**Algorithm 1** Probabilistic SDMA

**Initialization:** Set $\Gamma_{\theta 0}, W_{\theta 0}, \forall\{n\}_1^N$: $\forall$ vehicles

1: Buoy polls each vehicle and waits for event $\big(loc_n\big)$ in uplink
2: $\theta_{k+1} \Leftarrow loc_n(k)$ % $\forall$ vehicles at $k$th time instant
3: Find $\Gamma_\theta,\ W_\theta,\ \theta_U,\ \theta_L$ % $\forall$ vehicles
4: Find neighbors $\mathcal{J}$ % $\mathcal{J} = (i,j) \mid \bar{\theta}'^{(j)} > \bar{\theta}'^{(i)}$
5: Perform interval estimation and compute $Y_{k+1}^{(\mathcal{J})}$
6: Perform UKF and find $[\Gamma_\theta^{(\mathcal{J})}(\bar{\theta}), W_\theta^{(\mathcal{J})}(S_\theta)]_{k+1|k}$
7: **if** the vehicles are spatially separate **then**
8:     Set spatial parameters; Transmit data
9: **else**
10:     $SW \leftarrow 0$; % number of attempts
11:     **while** vehicles $\mathcal{J}$ are overlapped: $\theta'^{(i)}_U < \theta'^{(j)}_L$ AND $SW \leqslant 10$ **do**
12:         Solve the optimization problem given $\rho_0$
13:         Find $\mathcal{R}_m^{(\mathcal{J})}$ and $\mathcal{R}_I^{(\mathcal{J})}$ % retransm. rate as in (3.23) and (3.24)
14:         **if** it returns optimized values **then**
15:             $\forall\mathcal{J}$ :Transmit with $P_I$ and poll % Prob. of interference
16:             **if** $ACK + loc_n$ received **then**
17:                 Stop
18:             **else if** $NACK + loc_n$ received **then**
19:                 Run steps 2-6 and 12-14 with $0 < \rho_{SW} < \rho_0$
20:                 SW← SW+1; Set spatial parameters; Retransmit
21:             **else if** $timeout$ **then**
22:                 **if** $the\ range < D(transmission\ range)$ **then**
23:                     Run 11-13 with $\rho_0 < \rho_{SW} < \sqrt{2/N}\mathcal{T}_{N-1,\alpha/2}$
24:                     SW← SW+1; Set spatial parameters; Retransmit
25:                 **else**
26:                     Drop the vehicle due to out of the range scenario
27:                 **end if**
28:             **end if**
29:         **end if**
30:     **end while**
31: **end if**
32: **if** $SW = 10$ **then**
33:     Reset the spatial parameters % Non-separable scenario
34:     Run Intra-cluster T-SDMA
35:     **if** $|\widehat{c}| > T_x$ **then**
36:         Run Inter-cluster T-SDMA
37:     **end if**
38: **end if**

Synchronization between the vehicles is not required since the solution is presented for downlink—buoy to vehicle—transmission. All data exchanges will be made through the buoy as the primary controller of the links. However, we assume a separate feedback channel for the uplink; therefore, when two or more vehicles are in the same area a polling packet is exchanged between the buoy and each vehicle separately to avoid any collision. Corresponding vehicle will be allowed to transmit through the feedback channel.

Firstly, clusters of non-separable vehicles are formed; then time sharing is applied inside each cluster. To compare this method with conventional TDMA, we define the *time slot usage ratio* as,

$$\mathcal{R}_{TS} = \frac{n_c}{M_2 \cdot \max(M_1, M_3)}, \tag{3.25}$$

where $n_c$ is the total number of time slots dedicated to each vehicle of cluster $c \in \widehat{c}$ in every $M_2$ frame. If $|\widehat{c}| > T_x$, an additional external clustering is required, which leads to an inter-cluster time sharing defined by parameter $M_2$. The denominator of (3.25) denotes the total number of time slots: $M_1$ is the maximum number of vehicles in the clusters that are communicating with separated antennae, while $M_3$ is the maximum number of vehicles per cluster in the group of inter-clusters. By selecting $\max(M_1, M_3)$ as the number of time slots of each frame, we guaranty that at least one time slot is dedicated to every vehicle within a cluster. The time slot usage ratio for conventional TDMA is calculated as $\mathcal{R}_T = 1/M_T$, where $M_T$ is the total number of time slots, i.e., it is equivalent to the number of vehicles. We define *rate efficiency* of the proposed T-SDMA technique over the conventional TDMA as $\eta = \mathcal{R}_{TS}/\mathcal{R}_T$.

Effective data rate per user over the total transmission time is defined [153] as $1/NB\log_2(1 + SNR)$, where $N$ is the number of time slots, $B$ is the bandwidth, and $SNR$ is the Signal-to-Noise Ratio. Effective data rate of a vehicle in *one* transmitting frame for T-SDMA can be written as,

$$\tilde{\mathcal{R}}_{TS} = \frac{N_p\, L_p\, n_c/M_2}{\max(M_1, M_3)T_s} = \frac{1}{\max(M_1, M_3)} B\log_2(1 + SINR_{TS}), \tag{3.26}$$

where $N_p$ and $L_p$ are the number and length of packets, respectively, and $T_s$ is the time duration of each time slot. The effective data rate of each vehicle in one transmitting frame for the conventional TDMA follows $\tilde{\mathcal{R}}_T = (N_p L_p)/(TM_T) = 1/M_T B\log_2(1 + SINR_T)$. Therefore, the ratio of the effective data rate in T-SDMA to TDMA can be formulated as,

$$\frac{\tilde{\mathcal{R}}_{TS}}{\tilde{\mathcal{R}}_T} = \frac{n_c M_T}{\max(M_1, M_3)M_2} = \frac{1/\big(\max(M_1, M_3)\big)B\log_2(1 + SINR_{TS})}{1/M_T B\log_2(1 + SINR_T)}. \tag{3.27}$$

From (3.27), the relation between SINR of T-SDMA and conventional TDMA is,

$$SINR_{TS} \approx (SINR_T)^{\frac{n_c}{M_2}}. \tag{3.28}$$

Algorithm 1 reports the set of rules in pseudo-code under different circumstances. When the new location sample is observed, the interval estimation is updated given the probabilities of possible cases, i.e., keeping the current AOD and beamwidth if the vehicles are far apart, updating the AOD and beamwidth by sliding the estimation area of the samples, or switching to a new uncertainty region by accumulating the samples. It will be applied to a Kalman estimator to decide on the next upcoming state. While the neighboring vehicles overlap, the optimization problem is solved. If it returns the optimized values, the vehicles are assumed separable and a similar procedure as described for separate vehicles will be applied to them, however, MAC will decide on the interference-miss trade-off. Since it is a probabilistic MAC, transmission is performed by taking the risk of miss. If the ACK message gets back, then it will continue the transmission without any interruption; otherwise, retransmission will be performed in the next round. Parameter $SW$ is used to count the attempts in order to get the optimal values. If it fails, then it triggers the algorithm to apply the hybrid MAC solution, i.e., T-SDMA, to cover the problem of non-separable vehicles.

## 3.4 Performance Evaluation

We consider a short/medium range transmission (less than $2$ $\mathrm{km}$) with an anchored buoy and several moving vehicles. Assume each vehicle can communicate with one of the buoy's directional hydrophone arrays. An underwater acoustic channel is simulated with a limited number of separable paths in which the first path is the strongest one (lowest transmission loss) and with the smallest delay. The required specifications of the underwater acoustic channel for the simulation are extracted from the Kauai Acomms MURI (KAM08) experiment at the western coast of Kauai, HI, USA [54]. The normalized delay profile and the phase response of the sample emulated channel is depicted in Fig. 3.9, where an anchored vertical hydrophone array with $3.75$ $\mathrm{m}$ inter-element spacing at the depth of $96$ $\mathrm{m}$ communicates with a source towed by a surface ship.

The other simulation parameters are listed as follows. The channel bandwidth is $5 - 25$ $\mathrm{kHz}$, the sampling rate is $50$ $\mathrm{kHz}$, and an underwater ambient noise leads to an $SNR \in [0, 20]$ $\mathrm{dB}$.

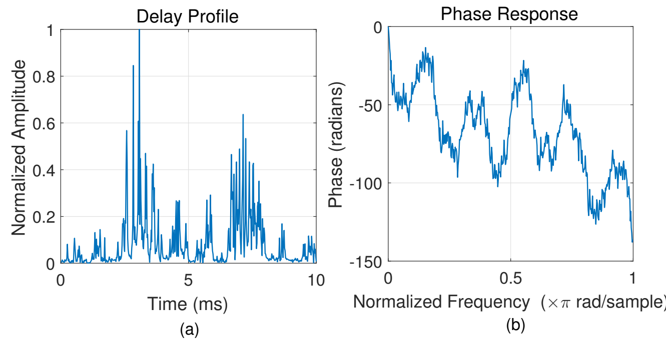



Figure 3.9: (a) Delay profile; (b) Phase response of emulated KAM08 channel.

Table 3.3: Initial Angular Specifications of Vehicles.

| | **Vehicle** 1 | **Vehicle** 2 | **Vehicle** 3 | **Vehicle** 4 |
|---|---|---|---|---|
| AOD ($\Gamma_\theta$) | -65.0325 | -31.9734 | 17.9087 | 42.9319 |
| HPBW ($W_\theta$) | 32.4749 | 20.6128 | 20.5878 | 32.4613 |
| AOD ($\Gamma_\phi$) | 45.3425 | 10.0598 | -20.0776 | -60.0550 |
| HPBW ($W_\phi$) | 53.9771 | 34.2830 | 34.2941 | 53.9718 |

We use the Gaussian noise model, introduced in [134], in which the overall power spectral density of the ambient noise is assumed as $N(f) = N_t(f) + N_s(f) + N_w(f) + N_{th}(f)$. Here, $N_t$, $N_s$, $N_w$, and $N_{th}$ stand for the turbulence noise, shipping activity, wind-driven noise, and thermal noise, respectively. The vehicles are randomly deployed at different depths from the surface $[0, 500]$ m and horizontal distances $[100, 2000]$ m. After converting these initial locations to the required antenna parameters at the buoy, Table 3.3 reports the initial state via an initial interval estimation. Vehicles move with a constant speed of $0.5$ m/s and transmit a new position information when the buoy polls them, according to the explained procedure in the algorithm, every $10$ s for the sampling interval of $50$ minutes. Finally, we set the percentage of confidence, i.e., $(1-\alpha)$ in (3.1), to the common value of $95\%$, which means that the true data is in the confidence interval with $95\%$ confidence.

We explore a scenario where two vehicles start their mission from different locations with a common target; therefore, their trajectories tend to converge to the same area. Fig. 3.10 shows the simulation assumption on the trajectory of these vehicles which is used in performing the system simulation. However, there are uncertainties in the estimation of the location. To handle these uncertainties, we adopt the statistical approach to estimate the position of each vehicle. Based on the procedure, explained in Fig. 3.4, MAC decides on the next estimate of each vehicle's AOD and beamwidth. In Fig. 3.11, the resultant AODs and beam boundaries are plotted for two vehicles

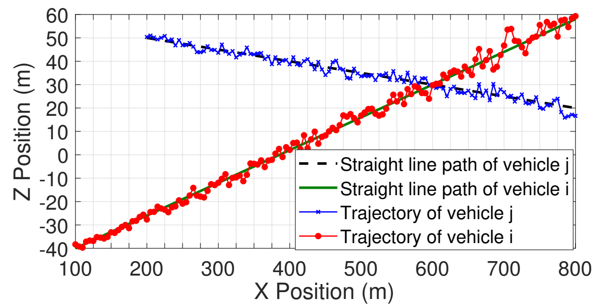



Figure 3.10: Trajectory of two vehicles in XZ plane for the simulated scenario.

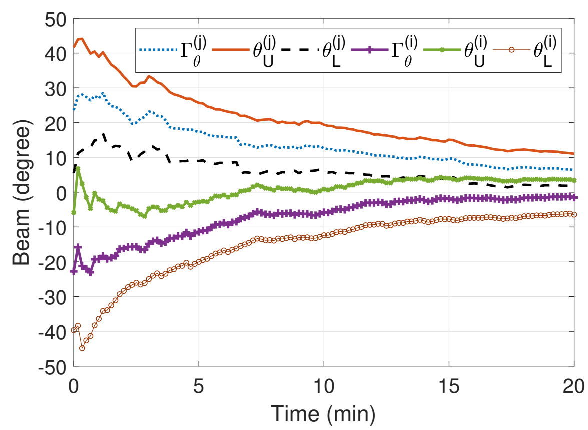



Figure 3.11: AOD and beam boundaries for two moving vehicles.

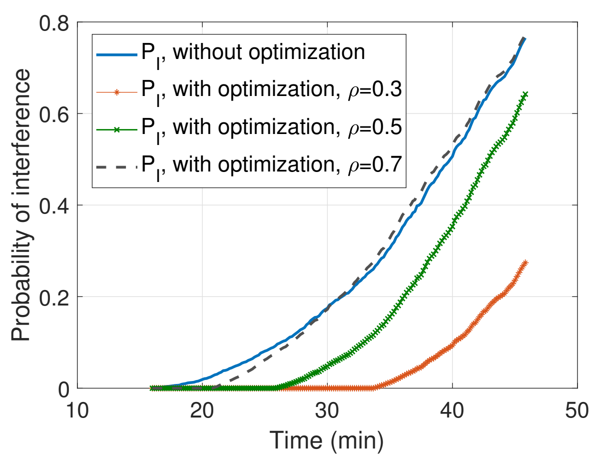



Figure 3.12: Probability of interference for moving vehicles, with and without optimization, for different values of $\rho$.

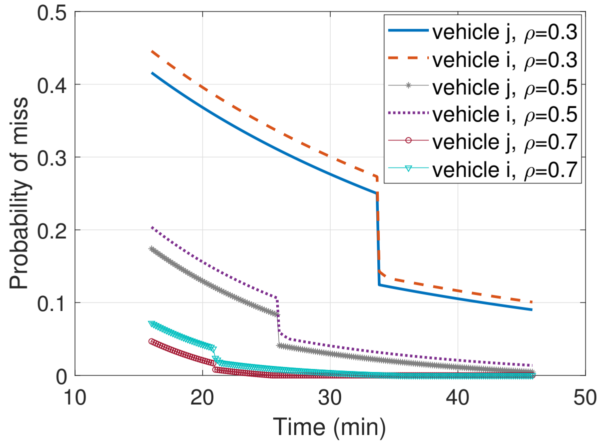



Figure 3.13: Probability of miss for moving vehicles, after optimization, for different values of $\rho$.

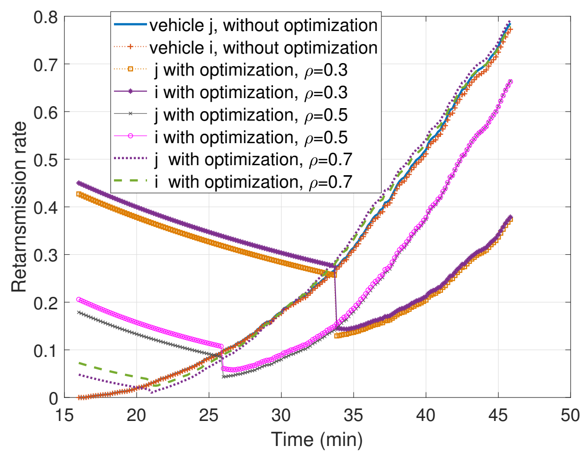



Figure 3.14: Retransmission rate for moving vehicles, with and without optimization and for different values of $\rho$.

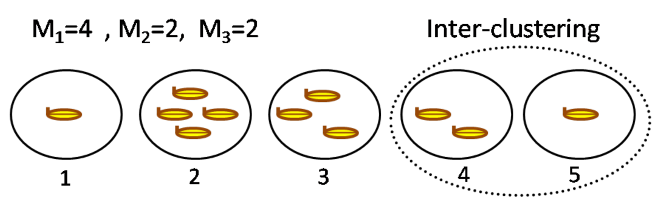



Figure 3.15: Clustering scenario in T-SDMA including 5 clusters and an additional inter-clustering.

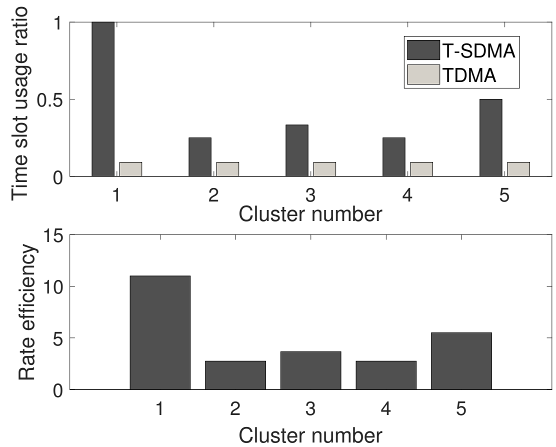



Figure 3.16: Comparison between probabilistic hybrid SDMA-TDMA and conventional TDMA in terms of time slot usage ratio and efficiency.

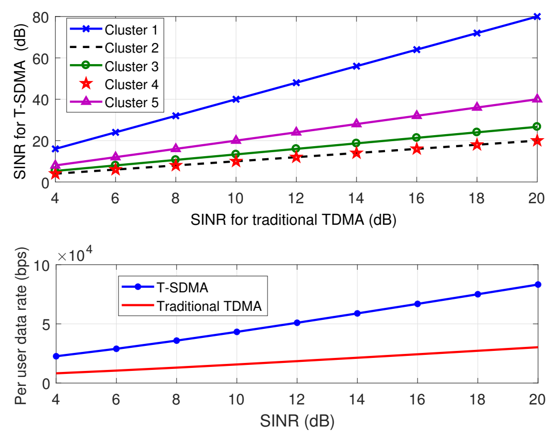



Figure 3.17: T-SDMA and TDMA comparison in terms of SINR and data rate per user.

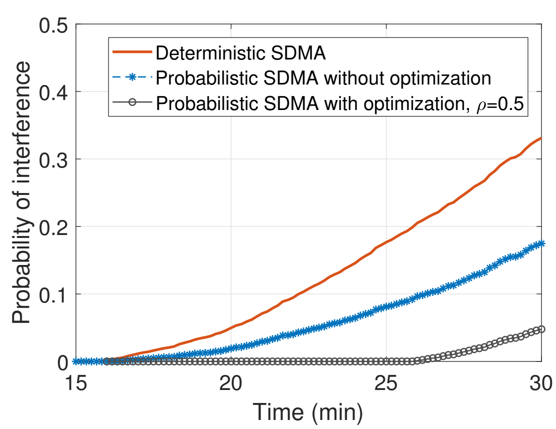



Figure 3.18: Probability of interference for deterministic SDMA in comparison with the proposed probabilistic SDMA.

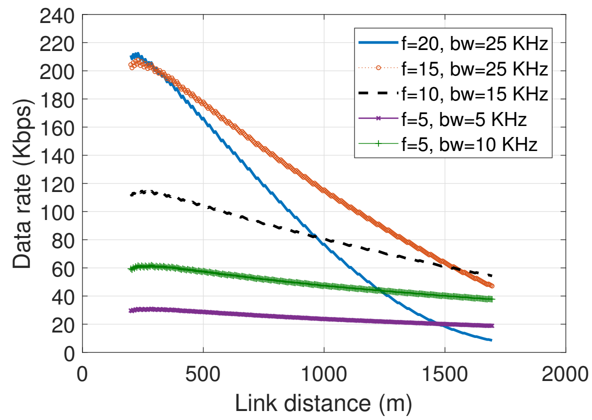



Figure 3.19: Maximum achievable rate in downlink for probabilistic SDMA, in terms of link distance, frequency and bandwidth.

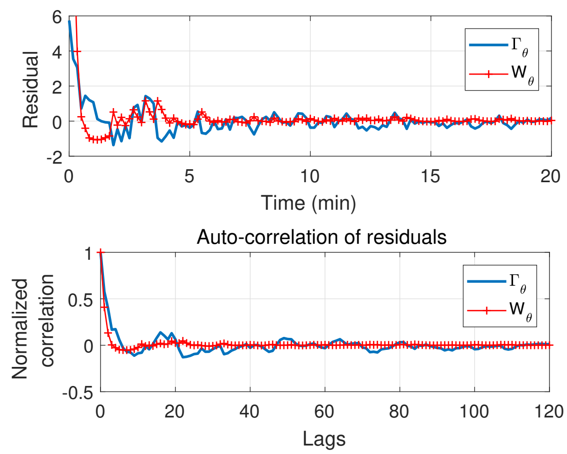



Figure 3.20: Residual and its auto-correlation of UKF estimation.

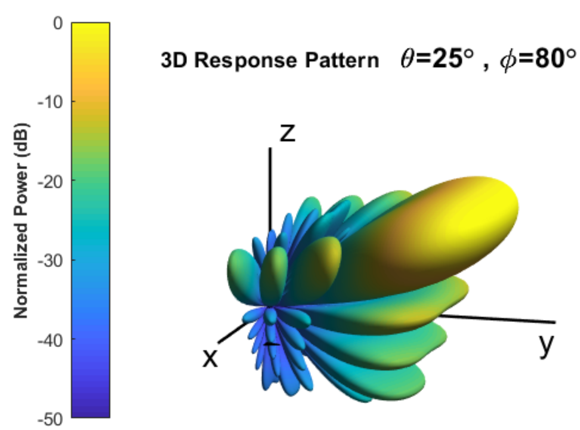



Figure 3.21: 3D pattern of one of the buoy's antenna arrays when steered towards a vehicle in a sample location at maximum frequency 45 kHz.

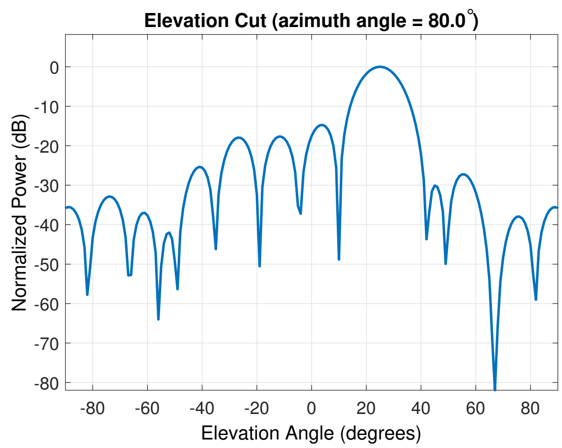



Figure 3.22: Elevation cut beampattern of one of the buoy's antenna arrays when steered towards a vehicle in a sample location at maximum frequency $45\ \mathrm{kHz}$.

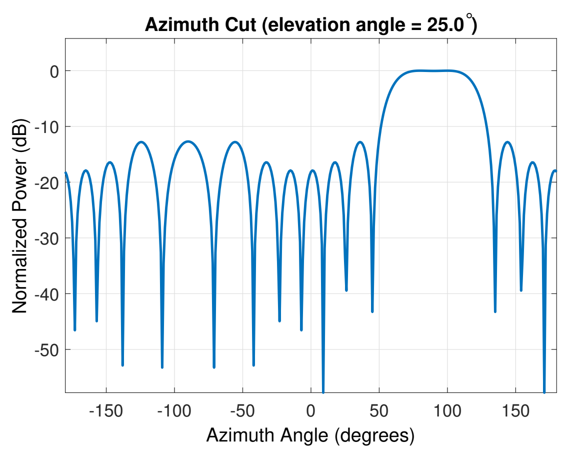



Figure 3.23: Azimuth cut beampattern of one of the buoy's antenna arrays when steered towards a vehicle in a sample location at maximum frequency $45\ \mathrm{kHz}$.

versus time. The figure shows that after 15 minutes the corresponding beams penetrate each other, and hence interference occurs.

Figs. 3.12, 3.13, and 3.14 evaluate the corresponding parameters after the interference occurs. In particular, Fig. 3.12 shows the probability of interference as time passes and vehicles get closer to each other; by updating the antennae' beam direction and beamwidth via the optimization problem, the probability of interference decreases. However, this outcome depends on the value of $\rho$, which is defined by MAC to control the optimized beamwidth and to regulate the interference-miss trade-off. Fig. 3.13 describes the decrease in the corresponding probability of miss by time and by increasing $\rho$, which contradicts the interference trend. When the objective function of the optimization problem does not change in a feasible direction, the solver finds a local minimum that satisfies the constrains, which may lead to a sudden change in the curves. In Fig. 3.14, the variation of retransmission rate is studied as time passes and for different values of $\rho$. Since this parameter reflects the effect of both interference and miss probabilities, it is an appropriate measure for MAC to choose a proper value for $\rho$. As an example, $\rho = 0.5$ keeps the retransmission rate below $20\%$ for almost $2/3$ of the transmission window, i.e., $20$ min.

Based on the evaluation of retransmission rate, when its projected value surpasses a tolerable quantity, MAC switches to T-SDMA mode and clusters the vehicles. As an example, in Fig. 3.14, when retransmission rate crosses $50\%$, it seems more reliable to switch to T-SDMA. In order to provide an example of the proposed hybrid T-SDMA, we investigate the scenario that is displayed in Fig. 3.15. We assume $11$ vehicles are categorized in $5$ clusters, which communicate with $4$ antennae at the buoy. MAC performs an additional inter-cluster time sharing for the last two clusters. Figs. 3.16 and 3.17 compare the performance of the proposed hybrid T-SDMA method with the conventional TDMA. Fig. 3.16 shows the time slot usage ratio and the rate efficiency and confirms that T-SDMA outperforms the conventional TDMA. Fig. 3.17 compares these methods in terms of SINR and effective data rate per vehicle. Appropriate clustering along with the beamforming ensures we achieve a gain for maximum achievable data rate in T-SDMA compared to the traditional TDMA, while there is no interference between the vehicles. In other words, we exploit space in order to increase the rate compared to the traditional TDMA, which is directly related to the maximum required SINR.

In Fig. 3.18, the probability of interference of the proposed probabilistic SDMA is compared with the non-probabilistic SDMA. In the latter, we assume that the transmitter estimates the vehicle's location and steers the beam towards it with a constant beamwidth. In order to provide a fair comparison, we suppose its beamwidth is equal to that of the probabilistic SDMA at the 15th minute when interference begins. The figure confirms that the probability of interference of probabilistic SDMA is less than the deterministic method, since our method updates the beam specifications by statistical calculations. This superiority increases when probabilistic SDMA applies the optimization to it. In Fig. 3.19, we investigate the effect of link distance, frequency, and bandwidth on the maximum achievable data rate in downlink when all vehicles are spatially separable from the other vehicles in the surrounding. As each vehicle moves and more samples are accumulated to make a better focused beam, the antenna gain and the directivity change. These variations lead to a change in SINR which is reflected in the data rate variation. In Fig. 3.20, the validation of UKF state estimation is verified by plotting the residual signals. The corresponding curves show that the residual magnitudes are small, their mean values are zero, and the autocorrelation functions are zero except at zero lag. Finally, in Figs. 3.21, 3.22, and 3.23, we provide the beampattern of one of the buoy's arrays while it is steered towards a vehicle at sample elevation and azimuth angels of $25°$ and $80°$, respectively. The figure shows the 3D pattern, elevation, and azimuth cut, when the frequency equals to the maximum value of $45$ kHz.

## 3.5 Summary

Achieving reliable communications and interference mitigation for an efficient MAC protocol is a necessity for a team of AUVs in the time- and space-varying underwater environment. In this chapter, a novel probabilistic MAC based on Space Division Multiple Access (SDMA) for short/medium distances was proposed to leverage inherent position uncertainty of the moving vehicles. A two-stage estimation technique was presented based on interval estimation and Unscented Kalman Filtering (UKF) to estimate the position of the vehicle and focus the beam, respectively. Spatially separable and non-separable scenarios were studied and an optimization problem was solved to minimize the statistical interference. The method was extended to the scenario of non-separable vehicles via a hybrid T-SDMA solution. Simulation results demonstrated that the proposed approach could handle the interference, while the vehicles were moving, and so it could achieve a high data rate and reliability. Note that the proposed method may be used in the uplink too if there is the possibility of mounting a beamformer and an array of hydrophones at each vehicle; also, since currently the coordination is performed at the buoy in a centralized fashion, a distributed in-network coordination among the vehicles in the uplink would be needed.

# Chapter 4

# In-Network Collaboration for Reliable Underwater Acoustic Communications

Achieving high throughput and reliability in underwater acoustic networks for transmitting distributed and large volume of data is a challenging task due to the bandwidth-limited and unpredictable nature of the acoustic channel. In a multi-node network, such as in the Internet of Underwater Things (IoUT), communication link efficiency varies dynamically: if the channel is not in good condition, e.g., when in deep fade, channel coding techniques may fail to deliver the information even with multiple retransmissions. Hence, an efficient and agile collaborative strategy is required to allocate appropriate resources to the communication links based on their status. The proposed solution in this chapter adjusts the physical- and link-layer parameters collaboratively for a Code Division Multiple Access (CDMA)-based underwater network. An adaptive Hybrid Automatic Repeat Request (HARQ) solution is employed to guarantee reliable communications against errors in poor links. Results were validated using data collected from the LOON testbed—hosted at the NATO STO Centre for Maritime Research and Experimentation (CMRE) in La Spezia, Italy—and from the REP18-Atlantic sea trial conducted in Sept'18 in Portuguese water.

## 4.1 Overview

Over the past decade, Underwater Acoustic Networks (UANs) have attracted the attention of researchers, engineers, and practitioners, as they enable a wide range of applications such as oceanographic data collection, offshore exploration, tactical surveillance, pollution and noise monitoring, disaster prevention, and assisted navigation [111]. These networks face various challenges due to the unique and harsh characteristics of the propagation of underwater acoustic waves [51, 94, 136]. In applications as the Internet of Underwater Things (IoUTs), data is usually distributed across a high number of nodes, while a single node (sink) is used for data collection, fusion, and processing [89]. The efficiency of an IoUTs system in mission-critical applications relies on the robustness of the communication algorithms and protocols that control the components of such a system. To maximize the achievable throughput of such a network, one of the major challenges is the design of a secure, robust, and scalable Medium Access Control (MAC) and an Error Control (EC) strategy. These solutions need in fact to guarantee low channel access delay, low energy consumption, and fairness among competing and/or collaborating nodes in the face of the harsh characteristics of the underwater acoustic propagation medium [90, 92, 108, 124].

**Motivation:** Terrestrial and conventional MAC/link-layer communication techniques fail to provide the required robustness and reliability for futuristic applications due to the characteristics of the underwater acoustic channel [94, 113]. Direct-Sequence Spread Spectrum Code Division Multiple Access (DSSS-CDMA) is a promising physical-layer and multiple-access techniques for UANs since i) it is inherently robust to frequency-selective fading, ii) it compensates for the effect of multipath at the receiver by using filters that can collect the transmitted energy spread over multiple paths, and iii) it allows receivers to distinguish among signals simultaneously transmitted in the same frequency band by multiple devices [96, 135, 146]. The use of an efficient CDMA scheme, supporting an adaptive EC strategy such as Hybrid Automatic Repeat Request (HARQ), has the potential to increase channel reuse and to reduce the number of packet retransmissions, thus increasing network reliability and achievable throughput, while decreasing the network energy consumption. However, since the number of retransmissions is limited in the practical truncated ARQ/HARQ error coding strategies, the receiver might start dropping packets, thus significantly limiting the capability of delivering data in the network.

**Contribution:** In this work, we extend the concept of point-to-point HARQ to an implicitly collaborative scenario in combination with a DSSS-CDMA approach. A transmitting node with low-quality communication links piggybacks on its neighboring nodes' transmissions when protecting its data against errors, in order to increase the system throughput. We propose a solution to achieve the following objectives: i) high network reliability and throughput by allocating an appropriate share of system resources to different nodes; ii) latency problem alleviation caused by the conventional HARQ retransmission strategy; iii) simultaneous transmission on the available bandwidth via easily- and locally-generated CDMA chaotic codes using a secret seed with a flexible and large family size; and iv) low energy consumption via efficient output power allocation. Data collected using the CMRE LOON testbed [11] were investigated to validate the proposed method. This testbed is hosted in the Gulf of La Spezia, Italy, close to the CMRE premises and it is characterized by *shallow-water* communications (occurring at a maximum depth of $15\ \mathrm{m}$), which may be heavily affected by multipath. Our solution was able to find dynamically the optimal trade-off among these four objectives according to the application requirements. To evaluate the protocol performance under different conditions, additional data was then collected in a *deep-water* scenario during the REP18-Atlantic sea-trial. This trial was organized by CMRE, the Portuguese Navy (PRT-N), and the Faculty of Engineering of the University of Porto (FEUP) in Portuguese water, between Sines and Sesimbra, in Sept'18.

**Chapter Outline:** The remainder of this chapter is organized as follows. In Sect. 4.2, we present a summary of prior work on different acoustic data transmission techniques in the underwater environment, and CDMA is discussed as a candidate for the UANs. In Sect. 4.3, we define the required parameters for the proposed solution, and then present the proposed collaborative hybrid ARQ technique to achieve reliable communications in underwater CDMA networks. Then, in Sect. 4.4, we provide experimental and simulation results along with observations based on data collected via the CMRE LOON testbed and during the REP18-Atlantic sea trial. The results confirm the efficiency of the proposal. Finally, in Sect. 4.5, we summarize this chapter.

## 4.2 Related Work

Conventional ARQ, as a feedback-assistant EC technique, requests a retransmission for the erroneously received data packet. When the error is detectable, the packet is discarded until the same packet is successfully received in the next round. Retransmission is an appropriate solution to achieve a certain level of reliability in the underwater channel, specially when the Forward Error Correction (FEC) schemes are not able to correct the burst errors alone. On the other hand, because of the long propagation delay in underwater channels, the performance drops significantly since a technique such as stop&wait and other similar ARQ techniques fail to provide a reasonable throughput. Furthermore, having a feedback link might not be feasible in some practical systems or might be erroneous if it is available. Therefore, to reduce the number of retransmissions and to increase the system reliability under poor channel conditions, a powerful FEC code should be used, which makes the decoding hard to implement [31, ch.22].

In practical experiments—when usually the channel is error prone and therefore unreliable—multiple rounds of retransmissions should be performed to deliver the intended data; consequently, a huge amount of time is wasted given the long propagation delay in the underwater channel. Therefore, a proper combination of the ARQ and FEC is required in an efficient scheme to overcome the mentioned problems. This combination of ARQ and FEC leads to a Hybrid approach, i.e., HARQ, which reduces the number of packet retransmissions and increases the system reliability, specially under poor channel conditions. If the data is not decodable, the receiver sends back a Negative Acknowledgement (NACK) to the transmitter and asks for additional/duplicated FEC, which eventually increases the probability of successful transmission [130]. However, if the channel is very noisy, even using multiple retransmissions may not work. In the truncated ARQ/HARQ, the number of retransmissions is limited. Therefore, the receiver might drop the data, which detrimentally affects the throughput of the network.

A type-I HARQ discards the erroneous received packet after a failed attempt to correct it, then the transmitter repeats the same packet until the error is corrected. This method might be inefficient in time-varying underwater acoustic channel. When the channel is in good condition, i.e., retransmission is not required, FEC information is more than it requires and so the throughput drops. On the other hand, if the channel is not in good condition, e.g., when in deep fade, the pre-defined FEC might not be adequate and the throughput drops again because of multiple retransmissions [133].

A type-II HARQ requires a larger buffer size and has a higher complexity and efficiency compared to type-I. It adapts itself with the channel in such a way that it first transmits the packet along with the error detection bits—similar to one of the ARQ schemes—when the channel is good. While the channel becomes worse and after detecting the erroneous packet, a NACK message is sent back and—rather than retransmitting the same packet as type-I does—FEC is transmitted to help decode the stored packet in the receiver's buffer. If the error persists, the second NACK is issued and the same FEC might be retransmitted or extra FEC might be added depending on the coding strategy. Incremental Redundancy (IR) HARQ, which shows a higher throughput efficiency in terrestrial time-varying channels, adds extra redundant information in each round of retransmission after receiving the NACK message [130]. Terrestrial standards such as in High Speed Packet

Access (HSPA) and Long Term Evolution (LTE) have exploited HARQ synchronously for the uplink, and asynchronously in the downlink direction. Authors in [86] discuss the requirements for designing a user-centric and network-optimized HARQ for the fifth generation (5G) of mobile communications. Given the necessity of supporting futuristic applications such as in IoUT, we believe that a new design for HARQ is essential.

Using numerical simulations, authors in [8] used the random linear packet coding to control the packet loss in a hierarchical definition of packets in the stop&wait ARQ protocol for the channels with a long propagation delay. In [24], the authors applied fountain codes to HARQ in underwater networks to reduce retransmissions and achieve optimal broadcasting policies. An adaptive coding approach based on the IR-HARQ was proposed in [36] to improve the packet error rate in a time-slotted underwater acoustic network. In [109], we proposed a scheme based on HARQ that exploits the diversity gain offered by independent links of an underwater acoustic Multiple Input Multiple Output (MIMO) channel. A large number of papers can be found in the literature that investigate the efficiency of point-to-point HARQ, especially in the terrestrial environment.

Various works have been proposed addressing separately CDMA and HARQ for UANs. Authors in [135] discuss DSSS CDMA as a candidate MAC for mobile UANs in which multiple nodes connect to a central receiver. A distributed single-carrier CDMA underwater MAC was proposed in [96], which aims at achieving high network throughput, low channel access delay, and low energy consumption. Although Pseudo-Noise (PN) codes have been extensively employed in spread spectrum communication systems, considering their limitation in the number of different PN sequences and their cross-correlation properties, [50] proposed using chaotic sequences in DSSS communications. These sequences can be generated through an uncomplicated deterministic map. Moreover, since chaotic systems are extremely dependent on the initial conditions, they can produce an infinite set of orthogonal uncorrelated sequences. One widely studied chaotic set that has been employed for underwater communications [13] is generated based on the Logistic map. This map is able to produce a variety of distinct sequences for different users, by just changing the initial states and/or its bifurcation parameter. The proposed algorithm in [96] uses locally-generated chaotic codes to spread transmitted signals on the available bandwidth, which guarantees secure protection against eavesdropping (as packets can only be decoded with the proper chaotic code, which depends on the

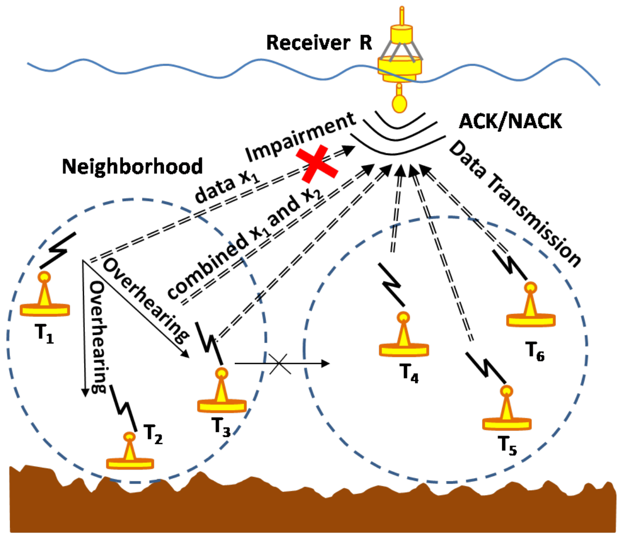



Figure 4.1: Architecture showing transmitting nodes, $T_i$, in neighborhood $m$, when the channel quality varies from one link to another. As an example, data from node $T_1 \in N_1$ fails to reach the receiver. CDMA is exploited and the nodes overhear and collaborate in the HARQ procedure based on their communication links quality.

secret initial conditions/seed), transmitter-receiver self-synchronization, and good auto- and cross-correlation properties [21].

## 4.3 Proposed In-Network HARQ Solution

While HARQ is a reliable EC solution based on packet retransmission, in practical scenarios and in the truncated HARQ, the number of retransmissions is limited. Therefore, if the underwater channel quality cannot be guaranteed, then multiple retransmissions may be not sufficient to correctly deliver the intended data, thus detrimentally affecting the reliability and the total throughput of the network. The problem gets worse when the interference from other users is involved in the performance. Therefore, a solution should be provided especially for multiuser UANs. CDMA is a promising technique for UANs since it can provide the required robustness and security of low-data-rate communications in multiuser scenarios in which all the nodes can overlap at the same time and at the same frequency band—despite the limited bandwidth—without any interference. Although M-sequences are popular in many CDMA systems, their cross-correlation shows some partial correlation for larger length of sequences in the multipath channels [50]. Moreover, the number of users that can be supported is limited by a sequence, which is a serious restriction for IoUT scenarios.

Fig. 4.1 shows the case in which $N$ transmitting nodes, $T_i,\ i = 1, 2, ..., N_m$, in neighborhood areas $m = 1, ..., M$ combine independently-sensed information, $x_i(r)$ (from independent nodes) for data fusion at the receiver $R$ (the sink) at different transmission rounds $r = 1, 2, ..., r_T$. Nodes in the same neighborhood can overhear other transmissions and collaborate to deliver the intended data, similarly to what $T_2$ is doing for $T_1$ in Fig. 4.1.

We use chaotic sequences in CDMA to increase the security in the communications. Albeit deterministic, chaotic codes look like noise, similarly to PN sequences; however, they are different for every bit of transmitted data. Hence, it is much harder for an eavesdropper, i.e., an unauthorized node outside of the neighborhood without the knowledge of the used codes (seed plus generating map), to regenerate the sequences and extract the data. This property allows us to have authorized nodes collaborating in a more secure way in the defined scenario and to also guarantee the service to a large number of users. Nodes that need to exchange data have to share the same family/seed of the code, with the chaotic codes generated in a deterministic way once the family/seed is known. This concept of sharing the same family/seed of the code is similar to the sharing of an encryption key in the context of symmetric cryptography. Before deploying the network, nodes are assigned a key or set of keys to talk to the other nodes. Nodes sharing the same key can read each others data. The use of more keys enables to create different cluster of nodes capable of sharing data. Different keys can be also used for different nodes or type of messages, e.g., some message may have higher priority or classification level and need to be shared with (or handled by) a reduced number of nodes so a specific key has to be used. Not only chaotic sequences provide the security in the channel, but they also show a considerable robustness against the multipath effect due to their good auto-and cross-correlation functions. Yet, CDMA requires to optimize the transmit power and spreading code length to limit the near-far problem and to maximize system throughout.

**CDMA-based Collaborative HARQ:** The proposed collaborative HARQ for data protection, combined with a CDMA (using chaotic codes) for secure and interference-free transmissions, relies on a closed-loop strategy based on measurements sent back by the receivers. This is to avoid relying on the unrealistic symmetric-link assumption, which does not usually hold in the underwater environment. Each receiver periodically collects information on the channel state. This information is then provided to the neighbors by transmitting short ACK/NACK messages.

For each neighborhood, the error probability on the decoded sequence $\tilde{x}_i$ for the transmitted codeword $x_i$ from node $T_i$ can be upper-bounded using the *Bhattacharyya bound* [100, 130] as, $P_e\big(x_i, \tilde{x}_i\big) \leqslant B_i^h$, where $h$ is the Hamming distance and $B_i$ is the Bhattacharyya Parameter (BP) in a noisy channel. Here we assume that channel does not change during one transmission round. This parameter is defined for every transmitted bit $x$ and received bit $y$ as $B = \sum_{y\in\Omega} \sqrt{\Pr(y|x=0)\Pr(y|x=1)}$. Here, $\Omega$ stands for the output alphabet and $\Pr(y|x=0)$ and $\Pr(y|x=1)$ are transition probabilities, $\forall y \in \Omega$. This parameter is considered as a channel reliability metric as it is an upper-bound on the probability of error in a typical Maximum-Likelihood (ML) detection problem, where larger BP values suggest channel unreliability and viceversa. The union-Bhattacharyya bound [130] can be calculated for each channel for node $T_i$ as,

$$P_{e\,i}^{c} \leqslant \sum_{h'=1}^{n} A_{h'} {B_i}^{h'}, \tag{4.1}$$

where $P_{e\,i}^{c}$ denotes the codeword error probability of code $c$ from family code $\mathbb{C}$ and $A_{h'}$ represents the codewords with weight $h'$. Data is encoded using a pre-defined mother code. Data and the first portion of parity bits are transmitted in the first round. If the receiver cannot decode the data, then a NACK will trigger the transmitter to send the second portion of parity bits in the next round, so that the receiver possibly can decode with the help of both portions. We transmit the coded data in $(r_T)$ transmission rounds from the selected nodes based on Algo. 2, which will be discussed later.

For each round of transmission ($r$), the error probability of the transmitted packet from node $T_i$, in a DSSS-CDMA system, is upper-bounded as [100],

$$P_{ei}(r) \leqslant (2^k - 1) \cdot Q\left(2\sqrt{\frac{P_i}{J_i} SL_i \; R_{ci} \; d_i}\right), \tag{4.2}$$

where $P_i$ is the transmitting power of $T_i$, $Q\big(.\big)$ is the Q-function, $J_i$ is the total interference and noise experienced by $T_i$, $R_{ci} = k/n$ is the coding rate of a code $c(n, k)$, $d_i$ is the minimum hamming distance ($h_i$), and $SL_i$ is the length of CDMA spreading code. Note that (4.2) implies that the transmitted power, amount of interference, rate, strength of channel coding, and the processing gain of CDMA system, all affect the probability of error.

The maximum achievable spectral efficiency achieved by each node in a neighboring area $m$, at round $r$ can be expressed as,

$$R_i(r) = \log\left(1 + \frac{\alpha_i P_i g_i}{N_0 W + \sum_{j=1, j\neq i}^{N_m} \alpha_j P_j g_j + \mathbb{J}_m}\right), \tag{4.3}$$

where $W$ is the channel bandwidth, $N_0$ is the noise Power Spectral Density (PSD), and $g_i$ is the channel gain. The transmit power is controlled by $P_i$ and $\alpha_i$ in each round of transmission. $\mathbb{J}_m$ is the interference from other neighborhoods of $m$ defined as,

$$\mathbb{J}_m = \sum_{k=1, k\neq m}^{M} \sum_{j=1}^{N_k} \alpha_j P_j g_j. \tag{4.4}$$

Note that $\sum_{j=1, j\neq i}^{N_m} \alpha_j P_j g_j + \mathbb{J}$ is the total interference that is undesirable from the CDMA perspective; however, we leverage the first term in our proposed HARQ to engage other nodes in the same neighborhood to collaborate in the process.

The other promising metric is the long-term throughput, which is defined based on the renewal reward theorem [159] as, $\eta = \mathbb{E}[\tilde{X}]/\mathbb{E}[\tilde{T}]$. Hence, $\mathbb{E}[.]$ defines the expectation of a random variable, $\mathbb{E}[\tilde{X}] = X\big(1 - \overline{\mathrm{Pr}_{out}}\big)$ is the number of decoded information nats, i.e, natural information unit. $\overline{\mathrm{Pr}_{out}}$ can be defined as the probability that the data has not been decoded after $(r)$ rounds. $\mathbb{E}[\tilde{T}]$ is the number of attempts for channel use during a packet transmission period. We decrease $\overline{\mathrm{Pr}_{out}}$ via node collaboration and by adjusting the corresponding parameters so as to improve the network long-term throughput. The probability of decoding in round $(r)$ given that the data has not been decoded in the previous $(r-1)$ rounds is equivalent to $\Pr\big(\mathrm{NACK_1}, ..., \mathrm{NACK_{r-1}}, \mathrm{ACK_r}\big)$. In our proposed method, when an impaired node gets its first NACK, we prevent getting more NACKs via the help of the collaborating nodes. Therefore, the average number of transmissions after $(r_T)$ rounds for an impaired node $i$ can be calculated as,

$$\big(1 - P_{ei}(r=1)\big) + \sum_{r=2}^{r_T}\Big[r(1-q)P_{ei}^{r-1}(r)\big(1 - P_{ei}(r)\big) + rqP_{e\kappa}^{r-1}(r)\big(1 - P_{e\kappa}(r)\big)\Big], \tag{4.5}$$

where $q$ is the probability of collaboration and $\kappa$ is the collaborating node.

**Total Rate Maximization:** In a multi-node system, in which nodes are influenced by each others' activities and are affected by the collaborators' coding scheme, those—which are involved in throughput and rate—should be selected in an optimal way. These parameters can be discussed under two major constraints, as follows.

*Signal-to-Interference-Noise-Ratio (SINR) Constraint:* To find the constraint for the multiuser interference in a CDMA system, we should reassure that a minimum required SINR and so the minimum error rate is satisfied at the receiver. This parameter—as a popular metric for the Quality of Service (QoS)—is a factor of processing gain, coding gain, and the signal power to interference ratio. Processing gain in CDMA represents the gain that is obtained by expanding the bandwidth of the signal and is shown by the spreading length [100]. Performance of the channel coding is decided by its coding gain and the minimum hamming distance.

*Power Constraint:* When we amplify the transmit power, the received SNR will be improved; however, it causes more interference to the other nodes. We try to regulate the transmit power and to reduce the interference to the other neighborhoods by a power control strategy. The peak transmitting power of each node in every neighborhood should be bounded to a pre-defined maximum power $P_{max}$, i.e., $P_i \in [0, P_{max}]$. As a result of the interference, SINR, and channel impairment, a power-control coefficient $\alpha_i$ is decided in each round that matches the HARQ procedure.

*General Optimization Problem:* To maximize the total rate and to satisfy the performance and power constraints, we cast an optimization problem to find the optimum parameter vector $\Theta = [\theta_1, ..., \theta_{N_m}]$, where $\theta_i = [P_i, R_{ci}, d_i, SL_i, \alpha_i]$ and $i = 1, ..., N_m$.

$$\max_{\Theta} \ \mathcal{F}(r) = \sum_{i=1}^{N_m} R_{ci} R_i(r) \tag{4.6a}$$

$$\text{s.t.} \quad \text{SINR constraint}: \gamma_\mathrm{i}(\mathrm{r}) = \big(2\mathrm{SL_i}\big)_{\mathrm{dB}} + \big(\mathrm{R_{ci}d_i}\big)_{\mathrm{dB}} + \big(\frac{\alpha_\mathrm{i}\mathrm{P_i g_i}}{\mathrm{J_i} + \mathbb{J}_\mathrm{m}}\big)_{\mathrm{dB}} \geq \gamma_{\min}, \tag{4.6b}$$

$$\text{Power constraints}: \sum_{\mathrm{i}=1}^{\mathrm{N_m}} \alpha_\mathrm{i}\mathrm{P_i g_i} \leq \mathrm{P_{th}}, \tag{4.6c}$$

$$P_i - P_{max} \leq 0, \quad i = 1, 2, ..., N_m \tag{4.6d}$$

$$\sum_{i=1}^{N_m} \alpha_i \leq N_m, \quad \alpha_i \in \{0, 1\}, \tag{4.6e}$$

where $\gamma_i(r)$, in $dB$, is the received Signal-to-Interference-Noise-Ratio (SINR) from $T_i$ at round $r$, $J_i = N_0 W + \sum_{j=1, j\neq i}^{N_m} \alpha_j P_j g_j$, and $\gamma_{min}$ is the minimum SINR, which is proportional to the probability of error in HARQ and determines the level of performance. $P_{th}$ guarantees that the total received power in $m$ does not affect other neighborhoods.

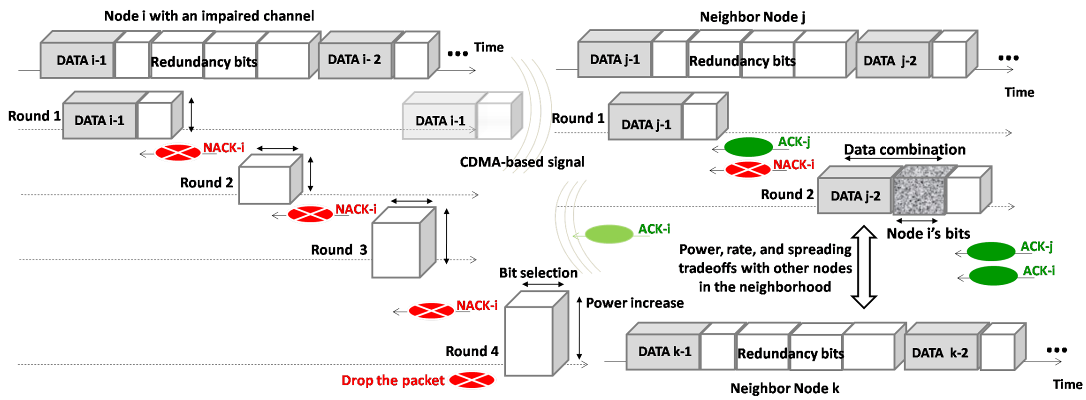



Figure 4.2: Proposed protocol for the interaction between the nodes. Without using this protocol (left side), node $i$ would keep sending incremental redundancy and the packet would drop after 4 rounds, while using our collaborative CDMA-based method (right side), 2 rounds are sufficient for delivering the data.

Let $\Psi = [\psi_1, ..., \psi_L]$ be the vector of Lagrange multipliers and $L = 2N_m + 2$ be the number of constraints. We form the Lagrangian function as $\mathcal{L}(\Theta, \Psi) = \mathfrak{F} - \sum_{l=1}^{L} \psi_l(g_l(\Theta) - b_l)$, where each $g_l(\Theta)$ and $b_l$ are determined by each constraint such that $g_l(\Theta) \leq b_l$.

$$\begin{aligned} \mathcal{L}(\Theta, \Psi) = \mathfrak{F} - \sum_{l=1}^{N_m} \psi_l \left(P_l - P_{max}\right) - \sum_{l=1}^{N_m} \psi_{l+N_m} \left(-\gamma_l + \gamma_{min}\right) \\ -\psi_{2N_m+1} \left(\sum_{l=i}^{N_m} \alpha_i P_i g_i - P_{th}\right) - \psi_{2N_m+2} \left(\sum_{i=1}^{N_m} \alpha_i - N_m\right). \end{aligned} \tag{4.7}$$

To find the optimum values, using Kuhn-Tucker condition, $\bigtriangledown_\Theta \mathcal{L}(\Theta, \Psi) = 0$ should be solved. There are $L$ complementary equations that should be held as $\psi_l\big(g_l(\Theta) - b_l\big) = 0,\ l = 1, ..., L$, such that $\psi_l \geq 0$. The feasible results of these equations determine the optimum parameter that results in maximum spectral efficiency. A numerical solution is presented in Sect. 4.4 for this problem based on the experimental data collection.

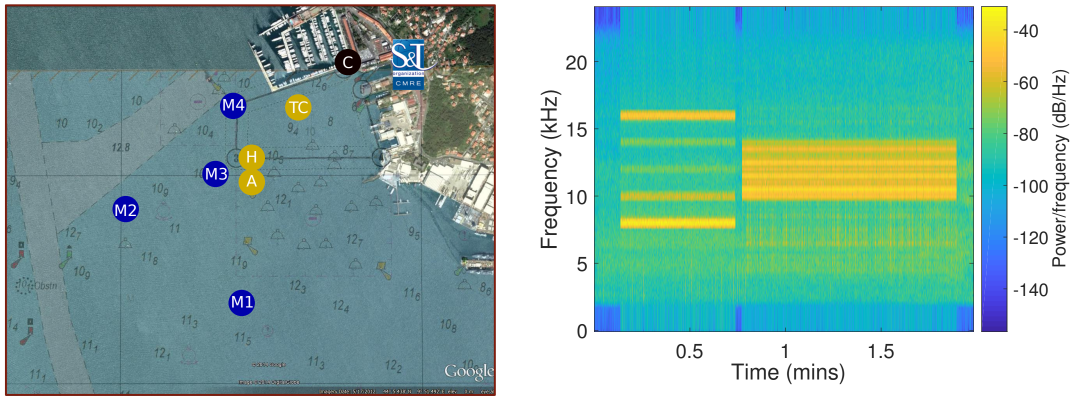
Fig. 4.2 visualizes the procedure for two sender nodes, one with an impaired channel (node $i$) and the other with a good channel (node $j$), as an example. The HARQ protocol at node $i$ decides on the bits to send in each round. If a NACK is received, the next portion of redundant bits will be transmitted. However, in our proposed algorithm, the impaired node is switched off after the first NACK arrives in round 1. The collaborating neighbor $j$ overhears the impaired node, stores its data, and transmits it in round 2. Algorithm 2 reports the pseudo-code executed by sender nodes in

**Algorithm 2** Collaborative HARQ for nodes $T_i \in \mathcal{N}_m$ in the same neighborhood

1: $\forall i = 1, ..., N_m :\ r_i \leftarrow 1$ as the index of transmission round; define $r_T$
2: choose collaborators $\hat{c}$ based on Eq. (4.1) and the neighbor discovery algorithm; share the chaotic map with $\hat{c}$
3: **if** event $\big($request to send for $T_i \in \mathcal{N}_m$ $\big)$ **then**
4:   **while** time-out codeword( ) **do**
5:     $\forall T_i$: Make the packets considering Eq. (4.1); HARQ block formation
6:     generate the chaotic code; solve problem in Eq. (6.6) considering Eq. (4.2)
7:     puncture the codeword; HARQ transmission procedure ( )
8:     event (wait for $\mathrm{ACK_i/NACK_i}$)
9:     **while** $\mathrm{NACK_i}$ AND $r_i \leqslant r_T$ **do**
10:       $r_i \leftarrow r_i + 1$; construct the codeword with additional redundancy; rate and power selection
11:       HARQ Retransmission procedure( )
12:     **end while**
13:     **if** $\mathrm{ACK_i}$ **then**
14:       goto end
15:     **else**
16:       **for** $\forall\ T_{\hat{i}}(r_T)$ (impaired nodes) **do**
17:         $\alpha_{\hat{i}} \leftarrow 0$; revise the collaborators $j \in \hat{c}$ based on Eq. (4.1) AND the received $\mathrm{NACK_{\hat{i}}}$ AND $\mathrm{ACK_{j \neq \hat{i}}}$
18:         **for** $j = 1 : \hat{c}$ **do**
19:           solve Eq. (6.6) for new $\Theta_j^*$ for all $\hat{c}$
20:           **if** $\arg\max\ \{\mathrm{R_j}\}_{\mathrm{j \in \hat{c}\ \&\ j \neq \hat{i}}} > \mathrm{R_{\hat{i}}}$ **then**
21:             **repeat** steps 3-12 for *combined* data of $j$ and $\hat{i}$ until $\mathrm{ACK_j}$ OR $r_T$   *% the new collaboration*
22:           **else if** $R_{\hat{i}} < \sum R_{\hat{c}}$ **then**
23:             update $\sum R_{\hat{c}}$ with the next collaborator $\arg\max\ \{\mathrm{R_j}\}_{\mathrm{j \in \hat{c}\ \&\ j \neq \hat{c}(max)}}$
24:             **repeat** steps 3-12 for *combined* data of $j$ and $\hat{i}$ until $\mathrm{ACK_j}$ OR $r_T$
25:           **else**
26:             packet drop
27:           **end if**
28:         **end for**
29:       **end for**
30:     **end if**
31:   **end while**
32: **end if**

a neighborhood $\mathcal{N}_m$. For the neighbor discovery algorithm, an approach similar to the one used by the DIVE protocol [91] can be employed. DIVE has a built-in mechanism to cope with unreliable channels. This approach can also be extended to share relevant link quality information when the network is deployed, thus supporting the cooperative strategy. This information can then be updated over time by piggybacking on regular data packets.

Data packets in the forward channel should be acknowledged successfully without error in the feedback transmission. The assumption of error-free feedback reception is not unreasonable since the length of this message is very short and therefore it can be protected by a strong channel coding technique. However, in a situation in which the ACK/NACK is lost, the timer expires to setup




(I) (II)

Figure 4.3: (I) Geographic configuration of the CMRE LOON testbed, in the Gulf of La Spezia, Italy. $M1$-$M4$ are the modem tripods, $C$ is the shore side container lab (control station), $TC$ is the thermistor chain, $H$ is the hydrophone array, and $A$ is an ADCP. (II) Spectrum of a sample received signal from the LOON in two successive transmission time slots while the spreading length and coding rate have changed.

the retransmission process. In the conventional scheme, if $T_i$ does not receive the ACK before a timeout expires, it will keep transmitting extra information in the next packets under the HARQ policy considering the previous channel state. However, in the proposed scheme, because of the collaboration among the nodes (i.e., $T_{\hat{c}}$), the probability of reception is increased by leveraging the statistical independency of the channels (i.e., channel diversity).

## 4.4 Performance Evaluation

In this section, we provide the performance results when using data collected from the CMRE LOON testbed and during the REP18-Atlantic sea-trial.

**CMRE LOON Testbed:** The geographic configuration of the CMRE LOON testbed is depicted in Fig. 4.3(I) for underwater communications and networking. It consists of four bottom-mounted tripods (M1-M4) installed at a depth of about $10$ m. Each tripod is equipped with heterogeneous communications technologies and sensors, and it is cabled to a shore control station (C) providing data connection and power supply. The LOON tripods also support arbitrary waveform transmission/recording. Additionally, the LOON includes a high-definition acoustic data acquisition system (at frequencies above $1$ kHz) from an array of hydrophones (H), a thermistor chain (TC), sound velocity sensors, an Acoustic Doppler Current Profiler (ADCP) with waves measurement (A), and a meteorological station. These sensors are used to correlate the characteristics of the acoustic channel with the performance of the investigated protocols. The LOON provides therefore a comprehensive data set of environmental, acoustic, and packet measurements to study the communication processes at different communication layers.

**Experiment Settings:** A variety of scenarios can be considered in the shallow-water environment where the LOON is deployed to capture the system outputs. We considered a point-to-point transmission from node $M4$ to $H$ for modeling the link of several rounds of transmissions. Packets are transmitted using baseband Binary Phase Shift Keying (BPSK) and Quadrature Phase Shift Keying (QPSK) modulations over the passband channel $4 - 19$ kHz by exploiting Reed-Solomon channel coding, $(7, 3)$ or $(15, 9)$. A logistic map is used to generate a chaotic spreading code with various lengths, i.e., $SL = [10, 40]$. As an example, we have measured the average Signal-to-Noise Ratio (SNR) equal to $32.68$ dB for a QPSK transmitted signal with $SL = 22$, and an average Bit Error Rate (BER) of approximately $4.6365 \times 10^{-4}$ is achieved.

**Simulation Settings:** We focus on the collaboration among transmitting nodes and assume that ACK/NACK feedback links are free of errors. Algorithm 2 has however a mechanism with a timer to retransmit the data if a feedback is not received within the expected time. Nodes in the same neighborhood can overhear each other, while nodes in adjacent areas do not receive the data since the chaotic CDMA sequence protects from unauthorized overhearing. Simulation are conducted for a neighborhood of 3 nodes. We use the data collected using the LOON testbed to model a multiuser scenario, then we optimize the parameters, as described in Sect. 4.3. The computed values are passed through the channels extracted from the LOON in a close-loop manner. We evaluate the system performance in MATLAB by considering the following metrics: SINR, long-term throughput ($\eta$), neighborhood efficiency rate, and effective rate per node.

**Results:** Fig. 4.3(II) shows the frequency spectrum of a sample received signal from the LOON while the spreading length and coding rate changes for two successive transmitted signals. In Figs. 4.4(I) and 4.5(II), two experiments with different settings are shown (for BPSK and QPSK scenarios, respectively). The PSD of the transmitted and received signals in passband and decoded baseband are plotted for comparison. Received SNR versus bandwidth, channel profile for the duration of the transmission, and scatter plot of the estimated symbols are provided. The transmitted signal parameters, BER, and SNR are also included in the figure. In Fig. 4.6(I), the received SINR

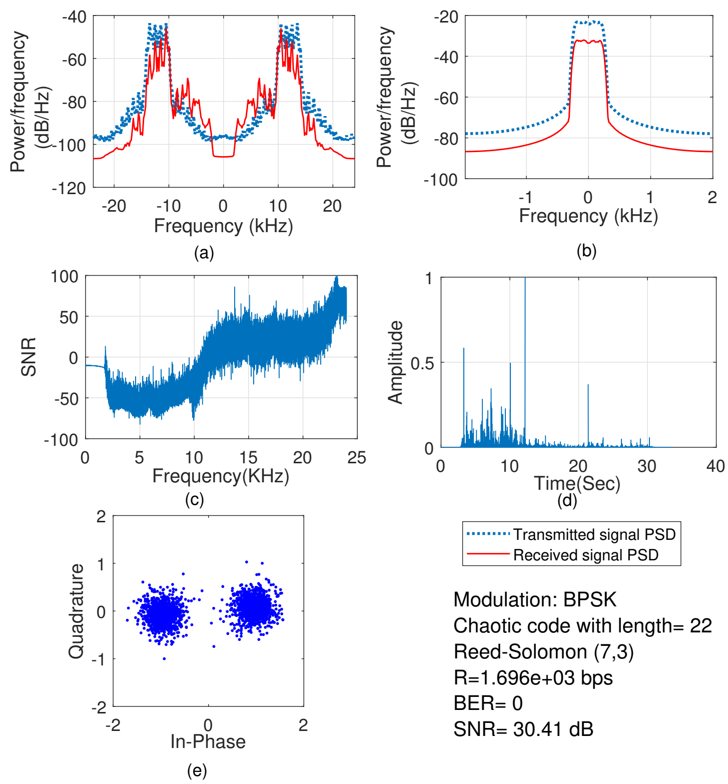



Figure 4.4: Experiment results; (a) PSD of the transmitted and received passband signals; (b) PSD of the baseband decoded received signal in comparison with transmitted signal; (c) SNR of the received signal per frequency; (d) Experienced channel profile; (e) Constellation of the equalized baseband received signal.

in a neighborhood of three nodes is presented to investigate the effect of multiuser interference. Fig. 4.6(I-a) presents the case where only one node in the area is transmitting. In this case, without interference, the received signal has a considerably better SINR. In Fig. 4.6(I-b), the data transmission in the area is performed by three nodes, so there is a multi-user interference. Fig. 4.7(II-a) depicts the total efficient rate in a neighborhood of three nodes. The plot shows how the collaboration strategy handles channel impairments and distributes the traffic load in the neighborhood. As the result of collaboration, when there are fewer nodes to perform data transmission, multiuser interference drops and spectral efficiency improves. Fig. 4.7(II-b) presents the effective received rate per node. The plot shows that in the case where an impairment occurs in $T_1$ link, $T_2$ collaborates in data transmission. In Fig. 4.8(III), long-term throughput for the proposed collaborative method is investigated. In Fig. 4.8(III-a), $\eta$ is plotted for different values of $r_T$ when none of the nodes experiences channel impairment. The plot confirms that power control can improve the long-term throughput.

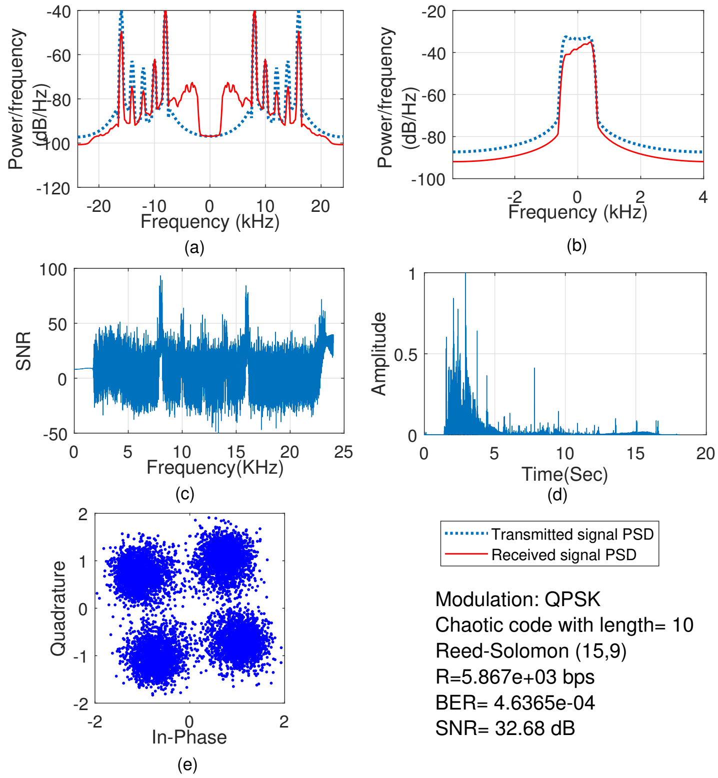



Figure 4.5: Experiment results; (a) PSD of the transmitted and received passband signals; (b) PSD of the baseband decoded received signal in comparison with transmitted signal; (c) SNR of the received signal per frequency; (d) Experienced channel profile; (e) Constellation of the equalized baseband received signal.

Finally, in Fig. 4.8(III-b), collaborative HARQ is compared with the conventional method to confirm that collaboration improves long-term throughput under channel impairment. The figure also shows the positive effect of power control.

**Sea Experiment Setup:** Sea experiments were conducted during the REP18-Atlantic (Recognize Environmental Picture) trial organized by CMRE, the Portuguese Navy (PRT-N), and the Faculty of Engineering of the University of Porto (FEUP). The trial took place place from the $1^{st}$ to the $20^{th}$ of September 2018, in the Atlantic Ocean off the coast of Portugal between Sines and Sesimbra. The area of operations is depicted in Fig. 4.9(I). The scope of this exercise was to investigate, evaluate, and demonstrate novel technologies and solutions in the domain of underwater communications and networking; as well as aerial, surface, and underwater robotic solutions and autonomous strategies.

During the REP18 sea trial, dedicated tests were conducted to investigate the use CDMA signals

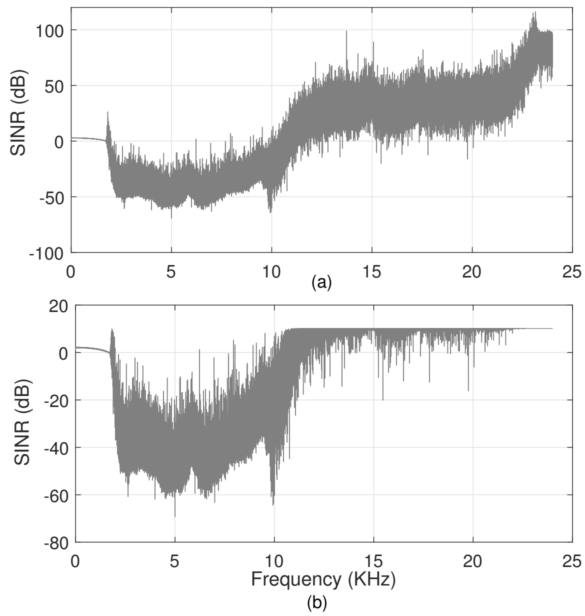



Figure 4.6: SINR (simulation) when the number of transmitting nodes in the same neighborhood is (a) One node; (b) Three nodes.

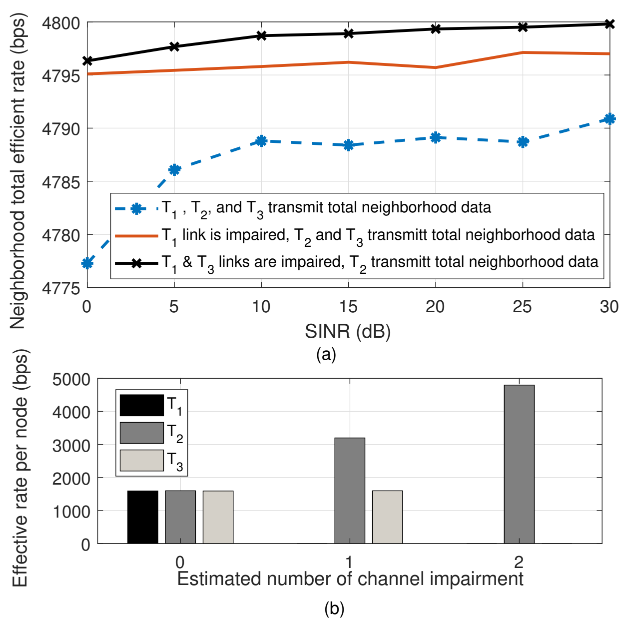



Figure 4.7: (a) Efficient rate of the neighborhood; (b) Effective rate per transmitting node when channel impairment is detected.

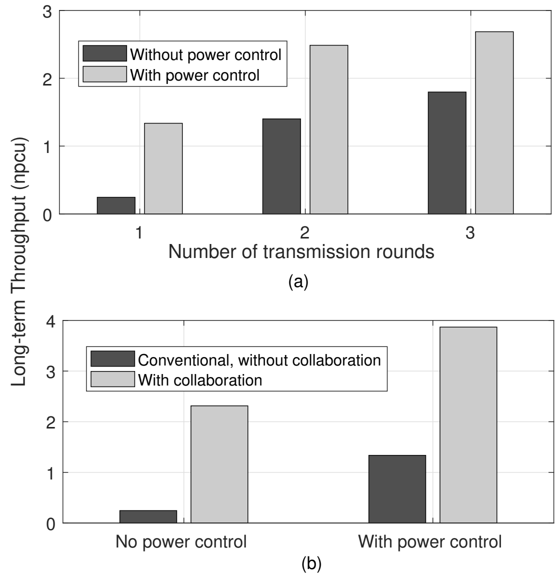
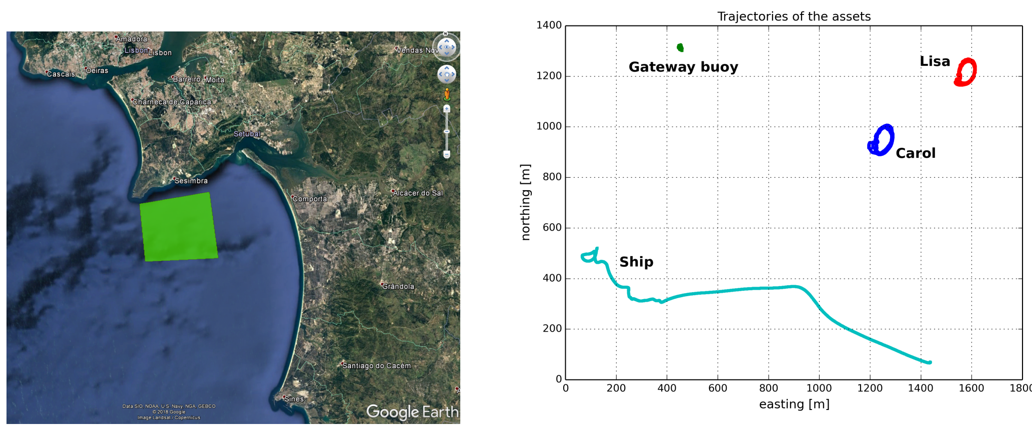



Figure 4.8: Long-term throughput in nats-per-channel-use (npcu) (a) For different number of transmission rounds when no channel impairment occurs, with/without power control; (b) Comparing traditional HARQ with our collaborative method.



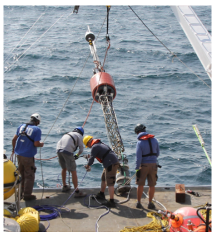

(I) (II)

Figure 4.9: (I) REP18 area of operations; (II) Trajectories of the nodes.

for underwater acoustic networking. These tests were scheduled in the night between the $8^{th}$ and the $9^{th}$ of September making use of four CMRE assets. In the area of the experiment a maximum depth of $\sim 130$ m was experienced. Table 4.1 details about the deployed assets, while Fig. 4.9(II) displays about the nodes trajectories and distances during the conducted experiment. A source level of 184 $dB\ re\ \mu$Pa@1m was considered at the transmitter, which is in line with that of many commercial acoustic modems currently available on the market.

Table 4.1: Deployed CMRE Assets.

**WaveGlider SV3** (x2): the wave glider is a self-propelled unmanned surface vehicle which uses wave motion to navigate. The SV3 version is also equipped with an auxiliary propeller. The wave glider enables long duration exploration and monitoring operations. Two wave gliders (named Lisa and Carol) were deployed, both equipped with an embedded board to run locally the required software and the capability to record acoustic signals using the iclisten smart hydrophone [81]. Lisa was also equipped with the capability of transmitting arbitrary waveforms using the Neptun T313 transducer [131]; On Lisa the acoustic payload was deployed at a depth of $\sim 40$m, while for Carol the depth was $\sim 20$m.



**Moored gateway buoy** (x1): Moored buoy equipped with dual radio connectivity (Wi-Fi $2.4$ GHz and Freewave $900$ MHz), an embedded board to run locally the required software and the capability to transmit/receive arbitrary waveforms using the Neptun T313 transducer and the iclisten smart hydrophone, respectively. The acoustic payload was deployed at a depth of $\sim 80$m.

**Manta portable node** (x1): this is a portable node including radio connectivity (Wi-Fi $2.4$ GHz), an embedded board to run locally the required software, and the capability to transmit/receive arbitrary waveforms using the ITC3013 [58] transducer and the iclisten smart hydrophone, respectively. It was deployed from the NRP Almirante Gago Coutinho ship during the conducted activities, with the acoustic payload at a depth of $\sim 20$m. The ship was left drifting during the experiment to avoid impacting the data collection with the noise produced by the propellers.

**Experiment Description:** Two main scenarios were considered, i.e., single transmitter and simultaneous synchronized transmitters. Four nodes were deployed, as depicted in Fig. 4.9(II). One node, named Carol, was equipped with reception capability. The remaining three nodes (i.e., Gateway, Lisa and Manta/Ship) were instead provided with both transmission and reception capability. Multiple signals were transmitted and received during the conducted experiments. The usable bandwidth was between $7-16$ kHz and the receivers were able to record the sound with the rate of $64$ kHz. The modulation methods were BPSK and QPSK. A chaotic DS-CDMA sequence with Logistic and Bernoulli maps were created with the spreading lengths of $10$, $30$, $40$. A Reed-Solomon channel coding, generally shown as RS(n,k) with s-bit symbols, was used with different coding strengths as $(7,3)$ $(15,9)$, $(31,19)$, $(31,21)$, $(31,23)$, $(31,25)$.

**Results:** Fig. 4.10 shows the channel response at two sample receivers, i.e., Gateway at time 20:00 and 21:00 and Carol at time 20:00, while the Manta/Ship node was used as transmitter. In

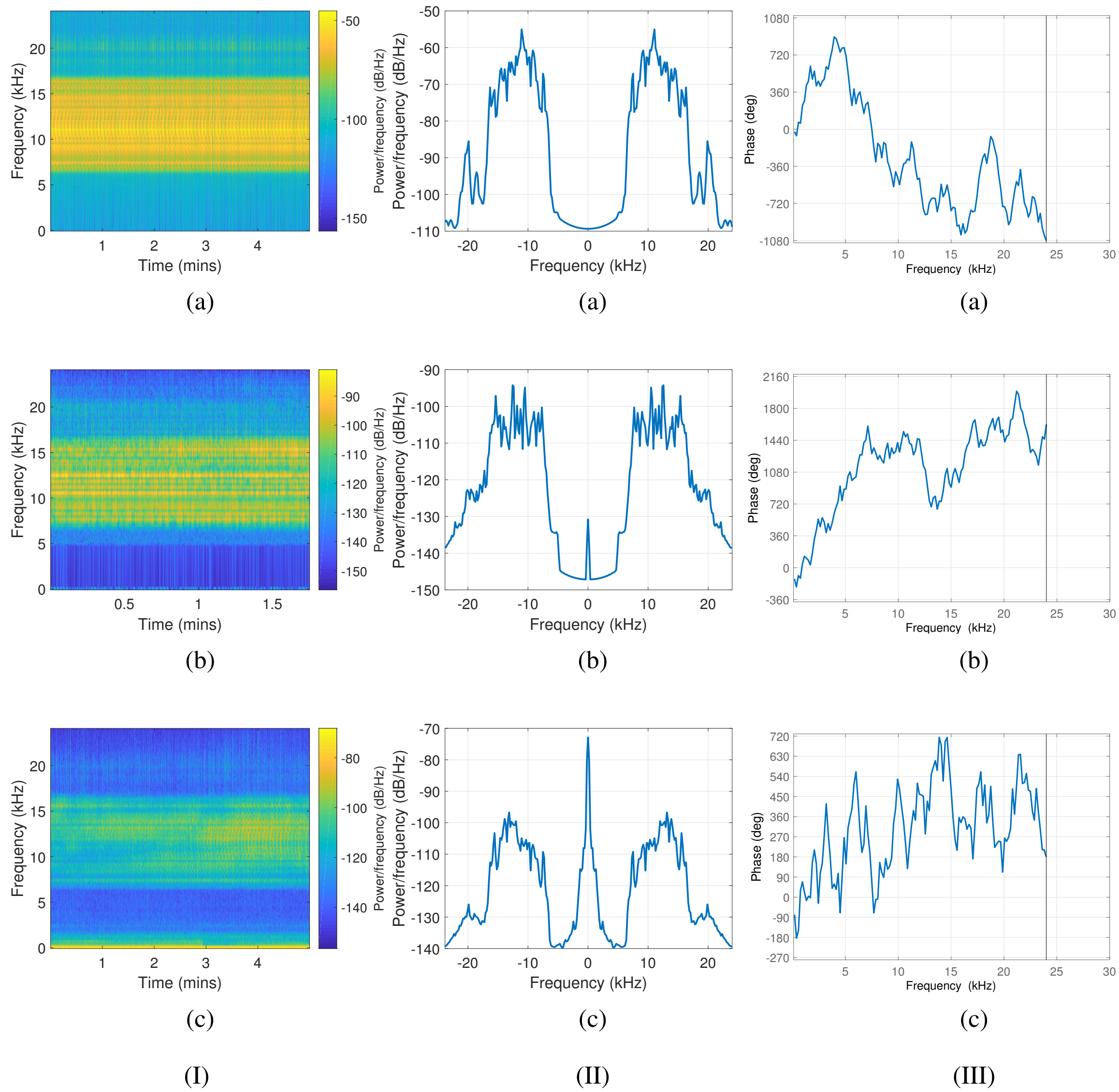



Figure 4.10: Channel response while receiving the signal at (a) Gateway at time 20:00; (b) Gateway at time 21:00; (c) Carol at time 20:00. Manta node deployed from the Ship was used as transmitter. The columns show (I) Power versus frequency and time; (II) PSD of the received signals; (III) Phase variations versus frequency.

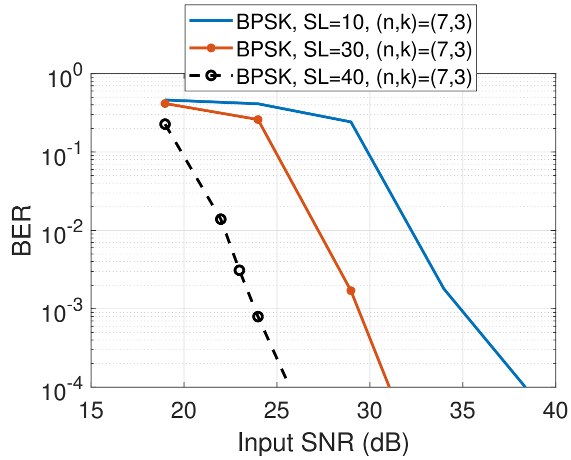



Figure 4.11: Single node transmission scenario; The Bit Error Rate (BER) for different spreading lengths of the BPSK chaotic DS-CDMA signal.

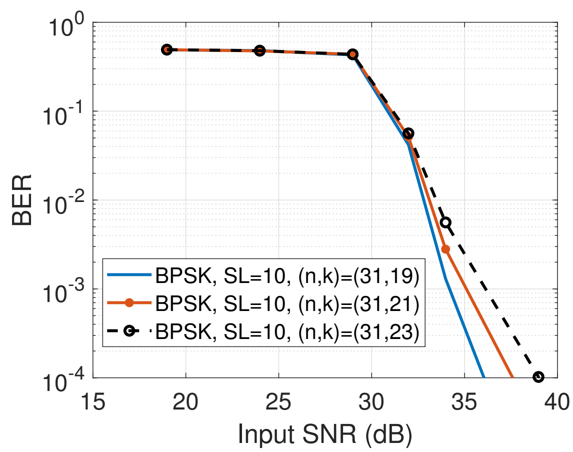



Figure 4.12: Single node transmission scenario; BER versus SNR shows the effect of coding strength in the BPSK chaotic DS-CDMA signal.

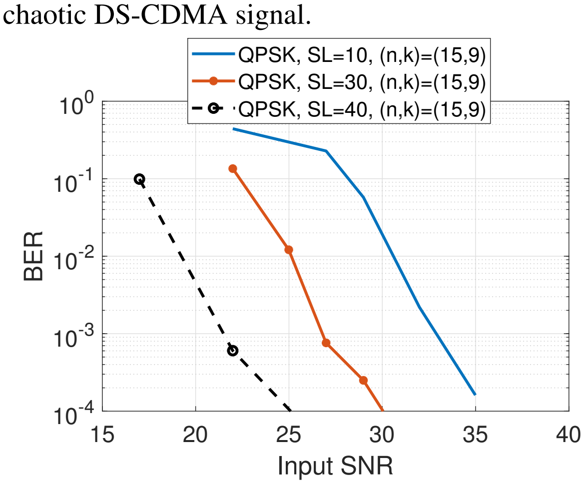



Figure 4.13: Single node transmission scenario; BER for a RS $(15, 9)$ coded QPSK with different spreading lengths.

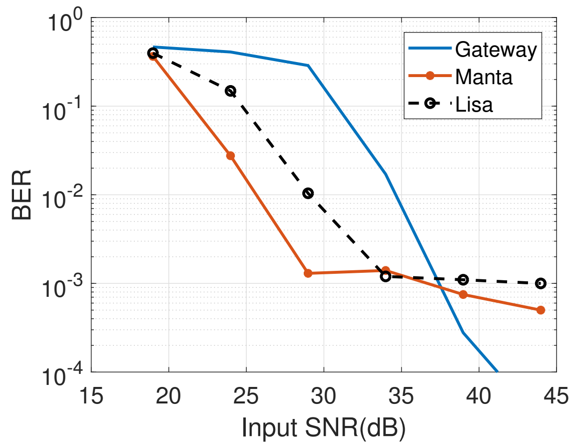



Figure 4.14: BER in multiple-node transmission, scenario $(a)$ in which all the channels have a satisfactory quality.

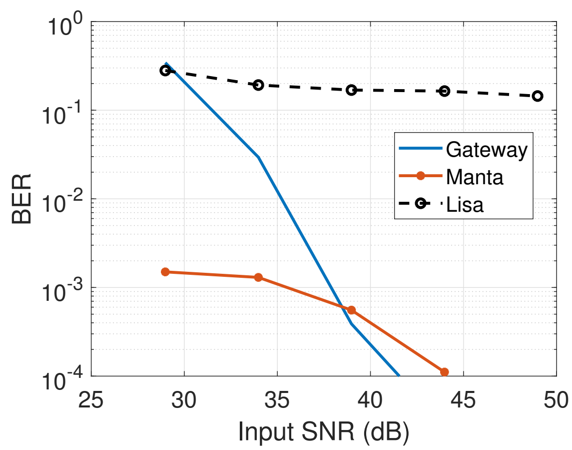



Figure 4.15: BER in multiple-node transmission, scenario $(b)$ in which the signal which is transmitted from Lisa does not experience a good channel

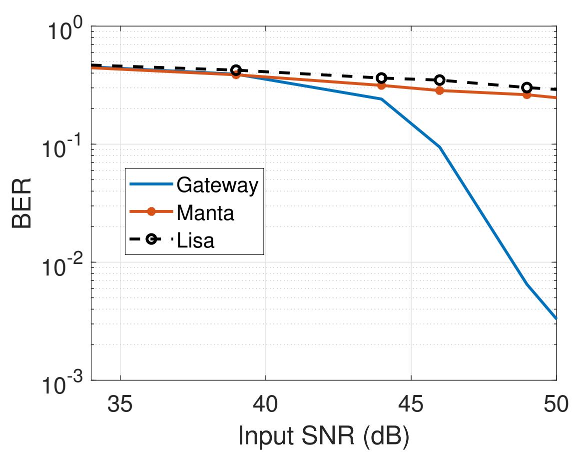



Figure 4.16: BER in multiple-node transmission, scenario $(c)$ in which the signals coming from both Lisa and Manta are impaired, but Gateway has an acceptable channel.

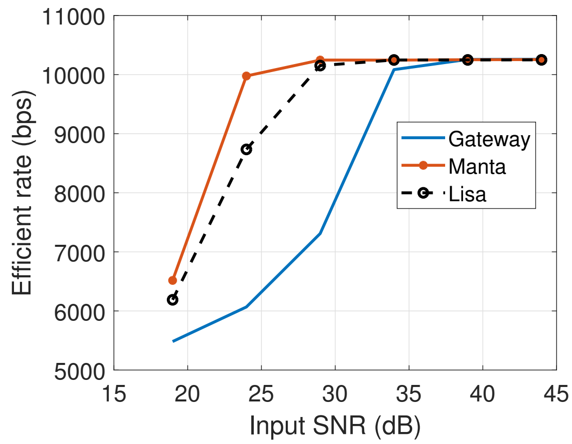



Figure 4.17: Efficient rate when all three channels are reliable as described in scenario (a).

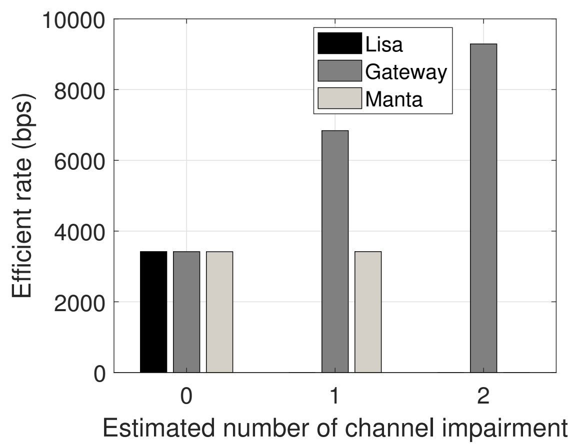



Figure 4.18: Efficient rate comparison and collaboration in different scenarios (a) to (c).

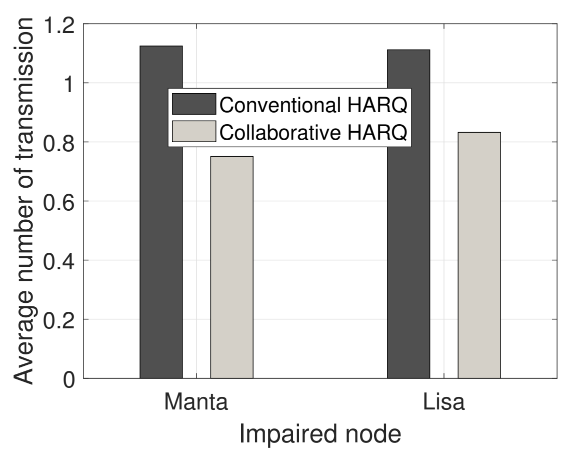



Figure 4.19: Average number of transmissions in the proposed solution versus the conventional HARQ. Two cases are compared; when Manta is impaired and Gateway collaborates and when Lisa is impaired and Manta is involved in the collaboration.

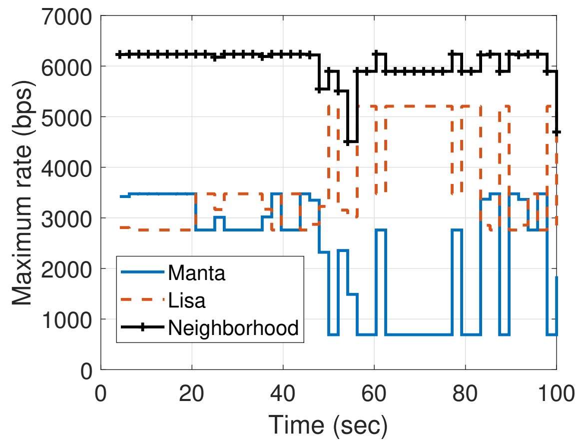



Figure 4.20: Maximum rate, for 100s, when Manta and Lisa are the transmitters.

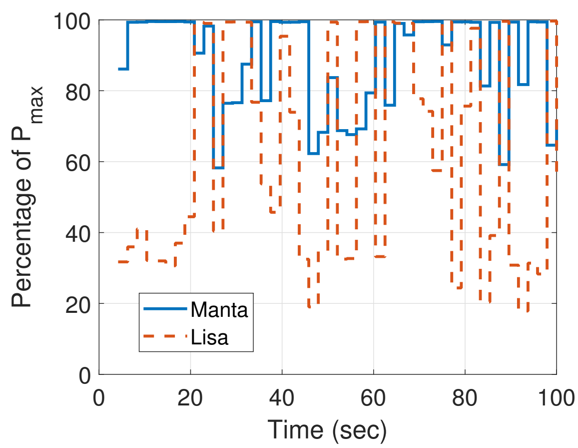



Figure 4.21: Transmit power percentage, for 100s, when Manta and Lisa are two transmitters.

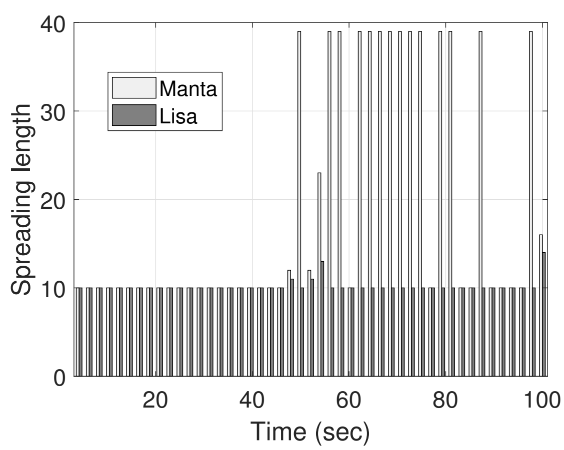



Figure 4.22: Optimal CDMA spreading length, for 100s, when Manta and Lisa are two transmitters.

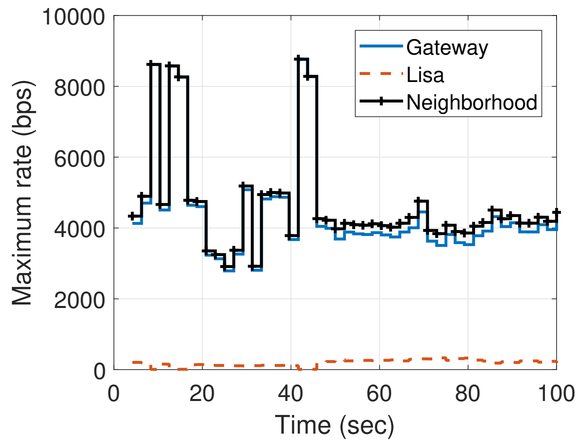



Figure 4.23: Maximum rate, for 100s, when Gateway and Lisa are two transmitters.

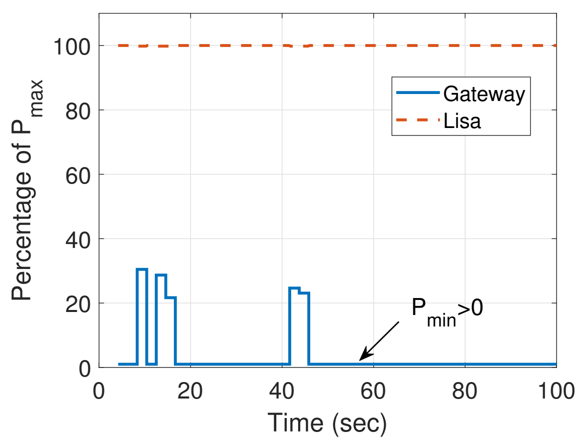



Figure 4.24: Transmit power percentage, for 100s, when Gateway and Lisa are two transmitters.

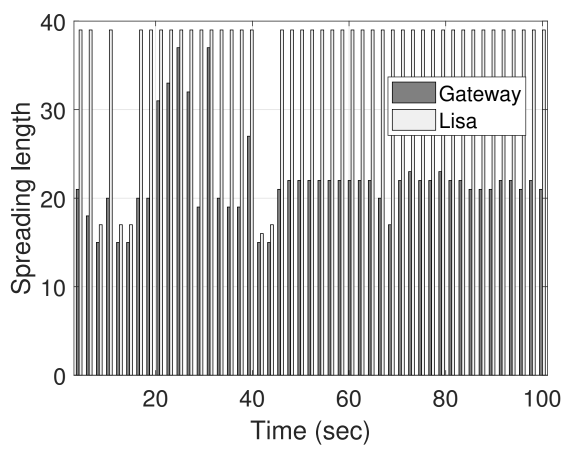



Figure 4.25: Optimal CDMA spreading length, for 100s, when Gateway and Lisa transmit.

these figures, column (I) shows the frequency spectrum of a sample received signals through different receivers for a specific duration. As shown in these figures, the received signals, which are recorded at different receivers (and also at different times), experience various channels. The effect of the amplitude response of the channel is reflected in the power spectral density of the received signals, as shown in Figs. 4.10(II). Figures 4.10(III) present the phase variations with respect to the frequency for these channels.

The experiments were performed for each setting with specific SNR. We scaled the data, applied power control, and added extra ambient noise at different noise levels to be able to present the performance in different SNRs for each experimented channel. Figs. 4.11, 4.12 and 4.13 refer to the single node transmission scenario from the Gateway transmitter. In Fig. 4.11, the BER for different spreading lengths for a BPSK signal is shown and confirmed that a higher spreading length leads to a lower BER. In Fig. 4.12, the effect of coding rate on the performance was investigated. Changing the coding strength can improve the performance in high SNRs as shown in this figure. Fig. 4.13 shows the BER for a QPSK signal with a $(15, 9)$ coding and different spreading lengths.

Figs. 4.14, 4.15, and 4.16 show the BER in multiple-node transmission scenario, where all the nodes (i.e., Gateway, Manta, and Lisa) were transmitting simultaneously different signals. Fig. 4.14 shows scenario $(a)$ in which all three channels have a good quality and so there is no channel impairment, while Fig. 4.15 represents a scenario $(b)$ in which the signal coming from Lisa does not experience a good channel; therefore, the network has one channel impairment. In this case, conventional HARQ will fail to deliver the data from this channel even with multiple retransmissions. The proposed collaborative solution will solve this problem, as discussed in Figs. 4.17, 4.18, and 4.19. The other scenario, called $(c)$, is reported in Fig. 4.16. This time two channel impairments are considered, both Lisa and Manta fail to deliver the data in the presence of a good channel from Gateway. Gateway will collaborate in our proposed solution to improve the total network efficiency, as reported in Figs. 4.17, 4.18, and 4.19. Efficient rate is shown in Fig. 4.17 when all three channels are reliable as described in the scenario $(a)$. Fig. 4.18 compares the efficient rate and the collaboration in the aforementioned scenarios. As an example, Gateway handles the Lisa's data in scenario $(b)$ and the whole network's data in scenario $(c)$. Fig. 4.19 compares the average number of transmissions in the collaborative HARQ with the conventional one for two cases: $(i)$ the receiver returns a NACK to Manta. Gateway then collaborates with Manta in transmitting the extra

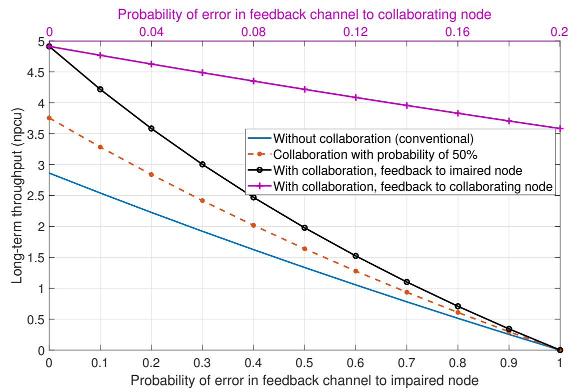



Figure 4.26: Long-term throughput simulation for the feedback ACK/NACK under different assumptions. The case in which the impaired node does not collaborate, i.e., the conventional approach, is compared against the case in which collaboration occurs.

redundancy which leads to a reduction in the average number of retransmissions as a result of this collaboration; $(ii)$ Lisa is and Manta is involved in the collaboration. By comparing the cases, we conclude that Gateway was a better collaborator in a comparable situation.

To verify the energy efficiency of the proposed solution, we investigate the trade-offs between the maximum rate, power, and CDMA spreading length in Figs. 4.20- 4.25. We considered the experimental data from the channel of Lisa and Manta transmitters for a period of $100$ s and plotted the offline results of the optimization for maximum rate, transmit power, and the optimal CDMA spreading length, since the experiment was open loop. The goal is to maximize the neighborhood data rate which is shown in Fig. 4.20. In Fig. 4.21, the required power for each transmitter to reach this rate is plotted. This figure confirms the efficiency of power allocation since the transmitters transmit with a fraction of the maximum power. Fig. 4.22 shows how the spreading length of these two transmitters adapts with the channel situation.

Figs. 4.23, 4.24 and 4.25 show the case in which Lisa and Gateway are the transmitters. Here, Gateway's signal is dominant and the signal that comes from Lisa experiences a poor channel. Lisa has to use its maximum power to defeat the interference coming from the strong signal of Gateway; however, it conveys a very low data rate. On the other hand, Gateway keeps its transmitter at the minimum power, while transmitting a great portion of data, as in Figs. 4.23 and 4.24. Our solution here is to switch off the node related to the weaker link, as described in the previous section. The other dominant node can handle the procedure with lower interference and better performance.

To evaluate the effect of errors in the feedback channel, a two-round HARQ simulation was investigated. Fig. 4.26 depicts the results in which two different assumptions were considered. Firstly, we assumed that the quality of feedback channel to both the impaired and collaborating nodes are similar in terms of probability of error. Secondly, we consider different probabilities of errors for those two feedback channels and simulated the long-term throughput for both cases. It is shown that the long-term throughput increases when the proposed algorithm is used.

## 4.5 Summary

In this chapter, a collaborative strategy for a CDMA-based underwater Hybrid ARQ was introduced to increase the overall throughput of the network. The solution leveraged both chaotic CDMA and HARQ properties to adjust the physical- and link-layer parameters and to compensate for the poor underwater acoustic communication links. System performance improvement and power control were considered, while the total throughput of the system was optimized. Experimental data was first collected in a shallow-water configuration using the CMRE LOON testbed and processed to extend the results to other nodes via simulation. Additional data was then collected in a deeper-water scenario during the REP18-Atlantic sea-trial to achieve a meaningful comparison under different conditions. This research has the potential to be implemented in larger underwater networks with heterogeneous nodes with higher volume of data and to analyze the scalability of the solution.

# Chapter 5

# Software Defined MIMO-Based Underwater Adaptive Video Transmission

Achieving reliable acoustic wireless video transmission in the extreme and uncertain underwater environment is a challenge due to the limited bandwidth and the error-prone nature of the channel. Aiming at optimizing the video quality and the user's experience, an adaptive solution for underwater video transmission is proposed that is specifically designed for Multi-Input Multi-Output (MIMO)-based Software-Defined Acoustic Modems (SDAMs). To keep the video distortion under an acceptable threshold and to increase the physical-layer throughput, cross-layer techniques utilizing diversity-spatial multiplexing and Unequal Error Protection (UEP) are presented along with the scalable video compression at the application layer. Specifically, the scalability of the utilized SDAM with high processing capabilities is exploited in the proposed structure along with the temporal, spatial, and quality scalabilities of the Scalable Video Coding (SVC) H.264/MPEG-4 AVC compression standard. Experiment results at the Sonny Werblin Recreation Center, Rutgers University-NJ, are presented and several scenarios are experimentally considered for both the known and unknown channel information at the transmitter and for both the video broadcasting and multicasting. Hydrophones are placed in different locations in the pool to achieve the required SVC-based video Quality of Service (QoS) and Quality of Experience (QoE) given the channel state information and the robustness of different SVC scalability. The video quality level is determined by the best communication link while the transmission scheme is decided based on the worst communication link, which guarantees that each user is able to receive the video with an appropriate data rate and quality.

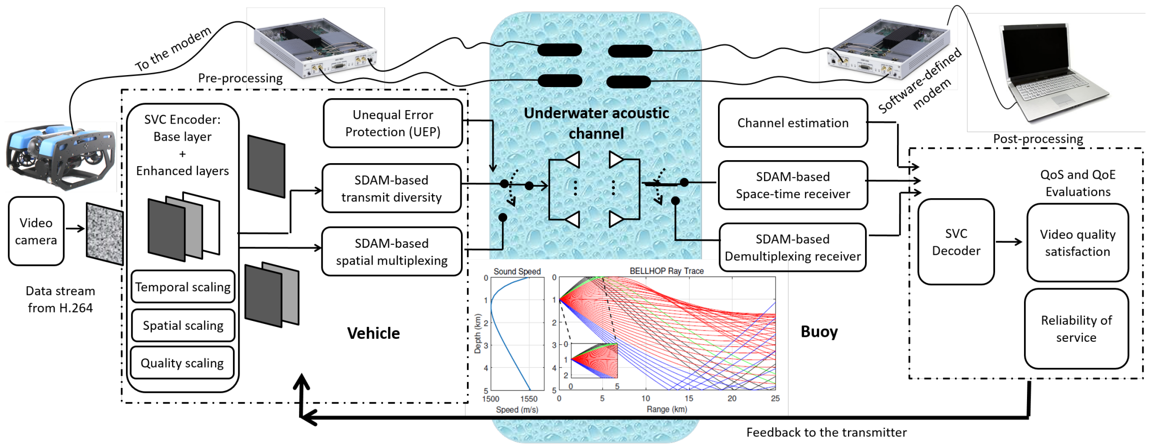



Figure 5.1: Proposed system for the MIMO-based software-defined acoustic transmission. Transmission techniques that utilize diversity and spatial multiplexing are the modalities.

## 5.1 Overview

Video transmission enables a wide range of applications in the underwater environment such as coastal and tactical multimedia surveillance, marine debris detection and monitoring, undersea and offshore exploration, oil pipe/bridge inspection, monitoring of geological/biological processes from the seafloor to the air-sea interface. In order to enable these applications, which all require real-time or near-real-time video acquisition, processing, and transmission [107], and to pave the way towards the futuristic Internet of Underwater Things (IoUTs) paradigm [116], achieving *reliable multimedia transmission* is a necessity, especially from places where humans cannot easily/safely go. Moreover, any communication solution aiming at enabling these applications should support different Quality of Service (QoS) requirements ranging from delay sensitive to delay tolerant and from loss sensitive to loss tolerant [95].

In practical scenarios, underwater Remotely Operated Vehicles (ROVs) are usually used, which are tethered to the supporting ship by a high-speed cable. This constrains the mission as well as the number of ROVs that can operate simultaneously in the same body of water. This is a serious limitation in the $(i)$ development of underwater systems for future applications; $(ii)$ maneuverability and range of the vehicles engaged in the mission; and $(iii)$ coordination of multiple vehicles in the mission. In other cases, when not tethered, the vehicles have to rise periodically to the surface to communicate with a remote station via Radio-Frequency (RF) signals. Resurfacing periodically does not guarantee interactivity as well and leads to considerable energy/time inefficiencies.

**Motivation:** Having a reliable and high-speed wireless transmission underwater is a challenge in such an environment in which RF waves are absorbed for distances above a few tens of meters, optical waves suffer from scattering and ocean wave motion, and acoustic waves—while being able to propagate up to several tens of kilometers—lead to a communication channel that is dynamic, prone to fading, spectrum limited with the bandwidth of only a few tens of $\mathrm{kHz}$ due to high transmission loss at frequencies above $50\ \mathrm{kHz}$, and affected by the ambient non-white noise [111]. While conventional underwater acoustic modems with their fixed-hardware designs [27] hardly meet the required data-rate and flexibility to support video requirements for futuristic applications, recently other solutions based on open-source and reconfigurable architectures employing software-defined modems have been proposed. Using software-defined modems helps scientists explore different protocols and techniques on a single hardware, perform in-network analysis, and transmit a high volume of data to a remote node depending on environment and system specifications.

**Approach:** To adapt to the underwater channel with variable video qualities and also leverage the benefits of using a software-defined modem, Scalable Video Coding (SVC) is proposed [123], which provides scalability in the processing of video and adaptation to the preferences of end-users as well as to the varying characteristics of the acoustic channel. Common types of scalability include temporal (frame rate), spatial (frame size), and quality (fidelity), which can all be adaptively chosen according to the channel conditions. An SVC video can be decoded with a high flexibility based on the knowledge of the receiver's channel. Also, thansk to the layering technique, an SVC video can reach high error robustness and video quality even with limited bandwidth.

The limited capacity of the underwater acoustic channel leads to a low data-rate and a restrained utilization of SVC video standard. To make full use of this channel, our approach consists in exploiting spatial diversity and multiplexing in a Multiple-Input Multiple-Output (MIMO) structure in cooperation (i.e., in a cross-layer manner) with the SVC and Unequal Error Protection (UEP). Given these limitations, the video should be reconstructed without much distortion notwithstanding a limited data rate. An optimization is thus required to select the optimal video transmission scheme in an adaptive manner. Our cross-layer (application, physical, and error control) algorithm is evaluated via field experiments to verify the efficiency of video transmission, in terms of quality and data rate, when the Signal-to-Noise Ratio (SNR) varies.

**Contributions:** We propose an adaptive cross-layer solution for underwater video transmission using a MIMO-based reconfigurable Software-Defined Acoustic Modem (SDAM) given the latest Universal Software Radio Peripheral (USRP) family product designed by the National Instrument (NI) [5]. For the application layer, we apply videos with different types of SVC scalability, which show different error robustness with varying levels of environment SNR. For the physical layer, given the underwater channel-compatible scalable coded video with a user-defined tolerable level of distortion, we navigate the *multiplexing-diversity tradeoff* with the MIMO structure to balance transmission data rate and reliability. Experiment results show that "multiplexing" improves the data rate significantly at high SNRs, while "spatial diversity" enhances the video quality at low SNRs. For the error-control layer, we apply Unequal Error Protection (UEP) to improve the system robustness without scarifying the physical-layer throughput by encoding more important parts of the data packet to achieve higher reliability so as to avoid error propagating to less important parts. With this cross-layer solution, the channel capacity can be improved by joint work of MIMO and UEP within the limited underwater acoustic channel, the video can reach the optimal data rate within the channel capacity and video scalability, and video distortion can be reduced. While optimizing the video quality, results show that the optimal QoS cannot stand for QoE completely, so we consider both the objective and subjective metrics in our algorithm to make the optimization results closer to the human experience. Several experiments have been conducted at the Sonny Werblin Recreation Center at Rutgers University on a camera-equipped SDAM-based underwater vehicle, and the results are presented in this chapter. The adaptivity of our system is discussed based on the experimental results under various scenarios.

**Chapter Outline:** The remainder of the chapter is organized as follows. Sect. 5.2 presents the relevant publications and related work. Sect. 5.3 discusses the framework and proposes the solution. Sect. 5.4 presents the performance results based on the conducted experiments; finally, Sect. 5.5 summarizes the chapter.

## 5.2 Related Work

Conventional video coding does not meet the underwater video transmission requirements for the futuristic applications. This goal is even harder to achieve in distances above hundred meters through the acoustic channel, as acoustic waves suffer from attenuation, Doppler spreading, high propagation delay, and time-varying propagation characteristics [95]. To achieve higher data rates in the bandwidth-limited underwater acoustic channel, several techniques should be combined holistically. In [109], a Hybrid Automatic Repeat Request (HARQ)-based solution is proposed that exploits the diversity gain offered by independent links in an underwater acoustic MIMO system. An Orthogonal Frequency-Division Multiplexing (OFDM) modulated dynamic coded cooperation scheme is proposed for the underwater relay network in [28]. Authors in [35] discuss the relationship between underwater acoustics and optics for long-range and short-range distances, respectively, to determine the correlation between the properties and the reliability of the acoustic/optical links. In [110], a signaling method for video transmission is proposed that makes use of multiple domains to leverage the benefits of Acoustic Vector Sensors (AVS). Scalable Video Coding (SVC), as an extension of H.264/MPEG-4 AVC, can be the solution for video delivery in harsh environments, such as in underwater, by offering more flexibility via different modalities—temporal (frame rate), spatial (frame size), and quality (fidelity or SNR)—to match the lossy video compression and erroneous transmission environments. It can also support the scalability in the complexity and in the Region Of Interest (ROI) [123]. An adaptive mechanism based on Scalable High Efficiency Video Coding (SHVC) is proposed for surveillance video coding [66], which achieves an enhancement of bitrate compared with the traditional SHVC video coding benchmark. The effect of scalability in SVC with the goal of providing guidelines for an adaptive strategy to select the optimal suggestion for a given bandwidth is discussed in [68]. An automatic tool for measuring the subjective metric—Mean Opinion Score (MOS)—of SVC video and improving the Quality of Experience (QoE) by using a random neural network is introduced in [129]. Authors in [14] propose an algorithm to estimate the SVC video distortion by assessing an objective metric, the Structural Similarity (SSIM). A public database for image and video quality evaluation with both the subjective and objective metrics is introduced in [2].

The Centre for Maritime Research and Experimentation (CMRE) proposed a structure [98] towards designing a Software-Defined Open-Architecture Modems (SDOAM) that is compatible with the JANUS standard [99]. Authors in [34] proposed a networking platform for short-range acoustic SDAMs, called SEANet. A survey on the past and current SDAMs was presented in [39], where the joint project between The Netherlands Organization for Applied Scientific Research (TNO) and the

Norwegian Defense Research Establishment (FFI) focusing on building a programmable modem, called NILUS softmodem, was discussed. The other platform, presented in [144], adapts some of the terrestrial radio and network development with the underwater acoustic environment. Based on the investigation of Multi-Stream Frequency-Repetition Spread-Spectrum (MSFRSS) modulation, authors in [18] examined the feasibility of underwater acoustic streaming of camera and sonar data on the TNO's testbed.

## 5.3 Proposed Solution for Video Transmission

In this section, we describe our system model, followed by the proposed cross-layer multimedia communication approach that leverages the MIMO structure and scalability characteristic of the compressed video to mitigate the overall distortion.

**System Model:** As illustrated in Fig. 5.1, a camera-equipped underwater robot initially records and encodes the video in the pre-processing block using an SVC encoder. Data is protected against the noisy channel with a proper channel coding technique, i.e., UEP, as well as an appropriate MIMO scheme using either spatial diversity or spatial multiplexing. At the receiver side, post-processing will be performed, and the human will participate dynamically in a closed-loop manner to tune the system based on the video quality satisfaction and the reliability of service in the received video stream. The decision is optimized, and the transmitter is notified accordingly.

We consider an SVC-based video bitstream, divided into chunks/segments, consisting of a *base layer* plus $L$ *enhancement layers* adopting different communication modalities. The chunk size is determined by the base and enhancement layer Group of Pictures (GoP) of the SVC file. The modality is being decided at the pre- and post-processing blocks, based on the Rate-Distortion (RD) requirements of the system. For a compressed video [137],

$$D_e(R_e) = \frac{\theta}{R_e - R_0} + D_0, \tag{5.1}$$

where $D_e$ represents the distortion of the encoded video and $R_e$ is the output rate of the encoder; the other remaining variables, $\theta$, $R_0$, and $D_0$, depend on the encoded video and on the model, and are estimated empirically. To quantify and measure the video distortion over the underwater acoustic channel, the Peak Signal-to-Noise Ratio (PSNR) is used as a metric for measuring the distortion $D_e$ based on the overall Mean Square Error (MSE). Other metrics, such as Physical-Layer Throughput (PLT), Structural Similarity (SSIM), and Mean Opinion Score (MOS) are also used to predict the perceived quality of the video. To reduce the amount of distortion, SVC provides hierarchical prediction structures for temporal scalability, inter-layer prediction of motion for spatial and quality scalability, and key pictures definition for drift control in packet-based quality scalable coding with hierarchical prediction structures [123]. Note that the total amount of distortion is composed of the errors caused by the lossy compression $D_e$ and the errors caused by the underwater acoustic channel, which can be alleviated by choosing an appropriate scheme.

While sound travels through the underwater medium, part of the acoustic energy is absorbed. An expression that models the medium absorption coefficient as a function of frequency $f$ is, $a(f) = (0.11f^2)/(1+f^2) + (44f^2)/(4100+f^2) + 2.75 \times 10^{-4} f^2 + 0.003$ [134]. In this empirical formula, $10 \log_{10} a(f)$ represents the channel attenuation. Propagation loss can be modeled via $P_a = \varsigma \Delta^{\varpi} e^{a(f)\Delta}$, in which $\varsigma$, $\Delta$, and $\varpi$ stand for the scattering loss, distance, and spreading loss, respectively [157]. When considering multiple propagation, in which the signal at the receiver is the outcome of several delayed signals of the original signal, the Channel Transfer Function (CTF) of each path $p$ is $H_p(f) = \Lambda_p / \sqrt{P_a}$, where $\Lambda_p$ is the cumulative reflection coefficient of surface and bottom reflections for each path. The overall CTF is calculated as $H(f) = \sum_p H_p(f) e^{j\theta_p(f)}$, in which $\theta_p(f)$ is the phase response characteristic for path $p$. Delay characteristic can be defined as $\tau_p = -\frac{1}{2\pi} \frac{d\theta_p(f)}{df}$, and it represents the propagation delay associated with path $p$. This delay is highly related to the sound speed profile, which is a function that increases with the increase of water pressure (i.e., depth), salinity, and temperature [111].

**Diversity and Multiplexing Modalities:** For an underwater acoustic MIMO system with $m$ transmit and $n$ receive hydrophones, the received signal in a flat-fading channel can be represented by $y = \mathbf{H}x + \mathcal{N}$, where $\mathbf{H}$ is the $n \times m$ channel matrix, $x$ is the transmit signal, $\mathcal{N}$ is a zero-mean Gaussian noise. We utilize Space-Time Coding (STC) and Spatial Multiplexing (SM) to achieve spatial diversity and multiplexing gains, respectively, in order to adapt to the acoustic channel's conditions. Using SM, multiple data streams are transmitted simultaneously and the data rate is improved without extra bandwidth occupation [62]. However, for a MIMO system with $m$ transmitters, each data stream interferes with the other $m-1$ streams; hence, the receiver should be capable of eliminating this interference. Using spatial diversity, one single data stream is space-time coded over multiple hydrophones. Thus, communication channels with different fading and interference characteristics can be utilized to collect different versions of the received data so as to improve the system's reliability [7]. Given this fundamental tradeoff, the achievable diversity-multiplexing equation can be written as, $q(r) = (m - r)(n - r)$, where $q(r)$ shows the diversity gain and $r \in \mathbb{Z}$ represents the multiplexing gain, which can be defined as, $r = 0, 1, ..., \min(m, n)$. As two special cases, we have $q_{max} = mn$ and $r_{max} = \min\{m, n\}$. The tradeoff curve confirms that while the rate increases by $r$ bps/Hz over an increase of 3 dB in SNR, the error rate is reduced by order of $2^{-q(r)}$.

This tradeoff is achieved only under *ideal conditions*, i.e., assuming that the SNR approaches infinity for i.i.d. Rayleigh-fading channels. This asymptotic definition breaks if the SNR is limited, as is the case in real scenarios [80]. The realistic diversity and multiplexing gains for a low SNR $\gamma$, array gain $g$, spectral efficiency $R$, and outage probability $P_{out}(r, \gamma)$, are calculated as follows,

$$r = \frac{R}{\log_2(1 + g\gamma)}, \quad q(r, \gamma) = -\frac{\partial \ln P_{out}(r, \gamma)}{\partial \ln \gamma}, \tag{5.2}$$

$$P_{out}(r, \gamma) = \Pr\left[\log_2 \det(\mathbf{I}_n + \frac{\gamma}{m}\mathbf{H}^*\mathbf{H}) < R\right], \tag{5.3}$$

where $\mathbf{I}_n$ represents the $n \times n$ identity matrix and superscript $^*$ stands for the conjugate-transpose operation. When STC is exploited to achieve diversity, the outage probability can be approximated given the fading distribution of the channel. It was shown in [88] that for uncorrelated MIMO channels, $\mathbf{H}$ can be represented by the variances of the power gains of channel as $\mathrm{var}\,\|\mathbf{H}\|_{\mathrm{F}}^2 = \sum_{\mathrm{i=1}}^{\mathrm{n}} \sum_{\mathrm{j=1}}^{\mathrm{m}} \mathrm{var}\,|\mathrm{h_{ij}}|^2$. These values can be obtained by estimating the mean powers of the channel matrix in the experiment.

At the receiver, Zero Forcing (ZF) is utilized, where the demultiplexed signal is expressed as,

$$\hat{x} = (\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^* y, \tag{5.4}$$

where $(.)^{-1}$ represents the inverse transform, since $(\mathbf{H}^*\mathbf{H})$ is a non-singular square matrix. To estimate the channel, pilot symbols $X_p$ are inserted after every two data symbols and channel state information is calculated by analyzing the received pilot $\hat{X}_p$. Therefore, the estimated channel $\mathbf{H}$

can be calculated as follows

$$\hat{\mathbf{H}} = (X_p^* X_p)^{-1} X_p^* \hat{X}_p. \tag{5.5}$$

In practical scenarios, in which the underwater channel is not known at the transmitter, we estimate a lower bound for the outage probability given only the statistics of a statistically-equivalent channel with the same distribution and with the eigenvalue set $\{\zeta_i\}_1^m$ to initiate the process as,

$$P_{out}(r,\gamma) \sim \Pr\,[\log_2 \prod_{i=1}^{m} \big(1 + \frac{\gamma}{m}\zeta_i\big) < R]. \tag{5.6}$$

Some underwater acoustic channels show the behavior of a Rayleigh fading [136] or a Rician channel, particularly in short distances (saturation condition due to heavy multipath). Therefore, a lower bound on the outage probability for finite SNRs in a Rician fading channel with the equivalent channel matrix $\mathbf{H}_{eq} = (K/(K+1))^{-0.5}\mathbf{H}_{LOS} + (K+1)^{-0.5}\mathbf{H}_w$ with line of sight ($\mathbf{H}_{LOS}$) and non-line of sight ($\mathbf{H}_w$) components, and with parameter $K$, can be estimated as in [80],

$$P_{out}(r,\gamma) > \prod_{i=1}^{m} F_i(\epsilon), \tag{5.7}$$

where $F_i(x)$ is a Cumulative Distribution Function (CDF) with the following description while no full Channel State Information (CSI) is assumed at the transmitter,

$$F_i(x) = \begin{cases} \Phi_i(x) & i = 1, ..., m-1 \\ e^{-Knm} \sum_{j=0}^{\infty} \frac{(Knm)^j}{j!} \Phi_{m+j/2}(x) & i = m. \end{cases} \tag{5.8}$$

Here, $\Phi_i(x) = \dfrac{\hat{\Gamma}\big(n-m+2i-1, (K+1)x\big)}{\Gamma(n-m+2i-1)}$, where $\epsilon \propto (m, \gamma, g, \vartheta)$. Furthermore, $\Gamma(.)$ and $\hat{\Gamma}(.)$ are gamma and incomplete gamma functions, respectively. In order to start the process, diversity gain can be initially estimated as,

$$q(r,\gamma) = \sum_{i=1}^{m} \frac{F_i'(\epsilon)}{F_i(\epsilon)} \big[\epsilon - \frac{mg}{1+g\gamma} (\vartheta_i^*(1+g\gamma)^{\vartheta_i^*} - \vartheta_{i-1}^*(1+g\gamma)^{\vartheta_{i-1}^*}\big]. \tag{5.9}$$

Here $(.)'$ stands for the derivative operation, $\vartheta_i^*$ is the value that maximizes the lower bound of the outage probability in (5.7). Note that when the SNR is high, the diversity gain follows the asymptotic diversity in (5.2) for both Rayleigh fading with a full-rank transmit covariance matrix and Rician fading channels [80]. The low SNR analysis is essential in MIMO systems in realistic propagation conditions. With SNR and diversity gain known, the estimated bit error rate at the physical-layer and the corresponding video distortion with different scalable coded video layer reconstructions can be calculated.

**Pre-processing and Optimization:** Let $R_e(c,l)$, with SVC layers $\{l = 1, ..., L+1\}$, denote the rate for layer $l$ of video chunk $c$. An appropriate data rate $R_i \geq R_e(c,l)$ for reliable communication should be assigned to layer $l$ in order to maximize the total transmission rate, i.e., transmitting as many video layers as possible without getting an outage or erroneous reception, given the bandwidth limitations and the quality of the underwater channel as well as the maximum allowed distortion. The following optimization problem justifies the aforementioned discussion,

$$\max_{\alpha_l} \ \mathcal{F}_{\mathcal{R}} = R_e(c,1) + \sum_{l=2}^{L+1} \alpha_l \alpha_{l-1} R_e(c,l) \tag{5.10a}$$

$$\text{s.t.} \quad \sum_{i=1}^{L+1} \alpha_i \alpha_{i-1} R_i \leq R_{max}, \tag{5.10b}$$

$$\alpha_1 = 1, \ \alpha_i \in \{0,1\}, \forall i \in \{2, ..., L+1\}. \tag{5.10c}$$

The first problem is a *knapsack program*, which defines the enhancement layers of rate $R_e(c,l)$ that could be transmitted over the underwater channel with maximum achievable communication data rate $R_{max}$. Coefficients $\{\alpha_i\}$ determine the set of enhancement layers that can be passed through the channel given the mentioned constraints. Selecting each layer depends on the presence of the preceding layer. This optimization guarantees that the base and enhancement layers are correctly transmitted (and received) given the limited capacity of the underwater acoustic channel. As stated in (5.6) and (5.3), the MIMO transmission scheme takes full advantage of the channel, leading to an improvement of $R_{max}$.

**Post-processing and Video Quality Decision:** To ensure that the desired Quality is achieved, an optimization problem finds the required parameters for the minimum possible distortion. Video header packets hold the general information of the H.264/SVC file, parameter set packets define the syntax structure of video, and slice data packets contain the detailed messages in the video. We define the distortion vector as **d** $= [d_e \ d_c]^T$, where $d_e$ is the distortion imposed by the codec as presented in (5.1) and $d_c$ is the distortion imposed by the channel (and is determined via experiments as it is related to the channel effective loss rate ($\lambda$)). If we assume that $\mathbf{d_h} = [\mathbf{d_e^h}, \mathbf{d_c^h}]^\mathbf{T}$ is the distortion at the stream header and $\mathbf{d_b} = [\mathbf{d_e^b}, \mathbf{d_c^b}]^\mathbf{T}$ is the distortion at the stream body of the transmitted video, then the total distortion can be written as,

$$d_e = D_e\left(\sum_{l=1}^{L+1} \alpha_l R_e(c, l)\right). \tag{5.11}$$

Hence, the total distortion is modeled as $D = \boldsymbol{\mu}^T[\mathbf{d_h}\ \mathbf{d_b}]\boldsymbol{\nu}$, where $\boldsymbol{\mu} = [\mu_e\ \mu_c]^T$ is a weighting vector specifying the influence of the encoder distortion and channel distortion; and $\boldsymbol{\nu} = [\nu_h\ \nu_b]^T$ is a weighting vector specifying the influence of the distortion at the header and the body of video stream. We can cast the following optimization problem,

$$\min_{R_i} \quad \boldsymbol{\mu}^T[\mathbf{d_h}\ \mathbf{d_b}]\boldsymbol{\nu} \tag{5.12a}$$

$$\text{s.t.} \quad \sum_{i=1}^{L+1} \alpha_i R_i \geq R_{min}, \tag{5.12b}$$

$$R_i \geq R_e(c, l), \tag{5.12c}$$

$$D \leq D_T. \tag{5.12d}$$

In the constraints, $R_{min}$ stands for the minimum required rate to avoid $P_{out}$, and $D_T$ represents the acceptable end-user distortion threshold. The problem can be optimized through a piecewise linear approximation method, which leads to a convex approximation function for (5.12a). The encoder distortion $d_e$ is determined by the video codec and the channel distortion $d_c$ will be alleviated by selecting an appropriate UEP scheme based on the weighting vector $\boldsymbol{\nu}$.

**Unequal Error Protection (UEP):** Given the structure of the video, if an error occurs in the stream header packets, the video cannot be decoded. Similarly, if the error occurs in the parameter set packets, the structure of the video will be damaged, which will lead to an extremely low-quality video. However, if the error occurs in the slice data packets, the video can be decoded successfully with a good quality. To achieve a high-quality video transmission, the received stream header with a negligible bit error rate is required, or the transmission of the whole video stream will fail. Since the video is much more sensitive to errors in the header and parameter set than those in the stream body, i.e., $\nu_h \gg \nu_b$, we utilize the UEP scheme to improve the received video quality. Therefore,

**Algorithm 3** Adaptive video transmission.

```
Layers = scalableVideoCoder( );        % Decide video layers
Transmit(baseLayer); s ← 1        % s is the number of trials
while t < Chunk Time do
    Receive(feedback)
    if channelState.rollingAverage > threshold.distortion then
        transmitter.switchTo('Multiplexing')
    else
        transmitter.switchTo('Space-Time Coding')
    end if
    Estimate(diversityGain, outageProbability)
    if MeanOpinionScore < threshold.opinionScore then
        Decide(channelCoding)
        Reconstruct(Layers); s ← s + 1
    end if
    Transmit(Layers)
    channelState.update( )
    if s = ∑ iα_i % α_i stands for the layer coefficient then
        Goto 1        % Done Transmitting this chunk
    end if
end while
```

by adding more redundancy in the header and parameter set, the receiver will have the capability to recover the header and parameter set more accurately than the body [125].

**Cross-layer Optimization:** Algorithm 4 describes the procedure for transmitting the underwater video adaptively, where the transmitter decides on the channel coding scheme, MIMO scheme, and type of video scalability. Given the objective and subjective metrics, our SDAM adaptively self-reconfigures by solving the optimization problem so as to be able to switch between the two MIMO transmission modes, i.e., diversity-based and multiplexing-based, and decides on the number of video layers to achieve the required goals. The base-layer stream, which contains the highest priority information of the video, requires the highest reliability, while the enhancement layers require a higher data rate, $R_{max}$. This fact, on the other hand, might result in more communication errors if the channel condition is not good. In our algorithm, we consider objective and subjective metrics jointly, given the fact that the QoE is more related to the user's experience. Given similar QoS but different scalability, the QoE might be different.

**SVC-based Multicasting:** With the SVC standardization, the low-quality video subset bitstreams can be derived and decoded from a high-quality SVC video bitstream by dropping some packets. Therefore, video bitstreams with different quality levels can be received by different users which can decode video bitstreams adaptively according to their experienced acoustic channel.

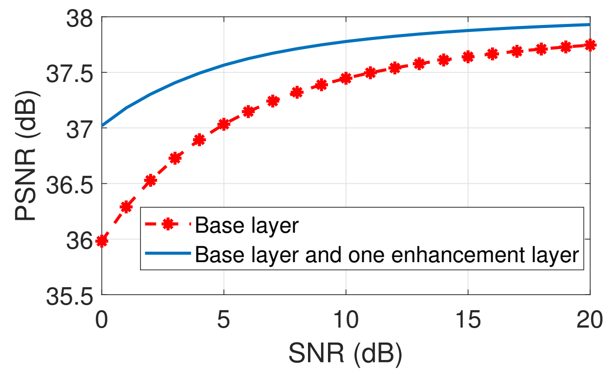



Figure 5.2: Simulated distortion (PSNR) versus SNR.

When the bit error rate is high, the low-quality video stream will be decoded; whereas when the bit error rate is low, the high-quality video stream can be decoded. The video quality level will be determined by the feedback from the best communication link, while the transmission scheme will be determined by the feedback from the worst communication link, which guarantees that each user is able to receive the video stream with an appropriate data rate.

**Objective and Subjective Metrics:** Some objective metrics are efficient to assess automatically and are of low computational cost, including PSNR, Physical-Layer Throughput (PLT), and Structural Similarity (SSIM). PLT is a physical-layer performance metric that shows the actual amount of transmitted data per second and is calculated as,

$$\text{PLT} = \frac{MK_{FFT}R_{chc}R_{st}}{2L_T L_F T_{OFDM}}(1 - p_c), \tag{5.13}$$

where $K_{FFT}$ stands for the FFT-size of the Orthogonal Frequency Division Modulation (OFDM), $M$ represents the order of baseband modulation ($M = 1$ for BPSK), $R_{chc}$ is the channel coding rate, $R_{st}$ is the number of streams transmitted simultaneously ($R_{st} = 1$ for SISO and 2-by-2 STBC, $R_{st} = 2$ for 2-by-2 V-BLAST [30]), $p_c$ is the bit error rate of the received data stream, and $T_{OFDM}$ represents the period of one OFDM symbol.

On the other hand, subjective metrics will correlate better with the human perception. The SSIM measures the fidelity of the video and is calculated based on the similarity of the local area luminance, local area contrast, and local patch structure. We apply Mean Opinion Score (MOS), as an objective metric, which has a scale from $0$, i.e., cannot play, to $100$, i.e., fully satisfied.

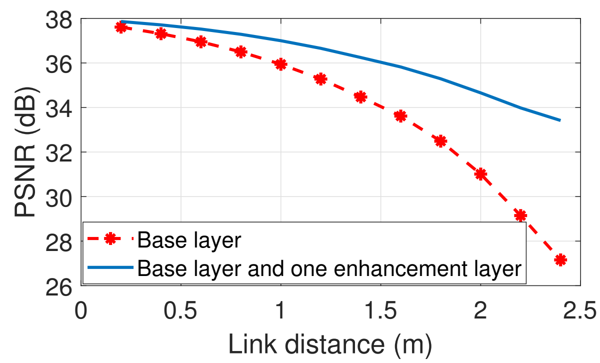



Figure 5.3: Simulated distortion (PSNR) versus link distance.

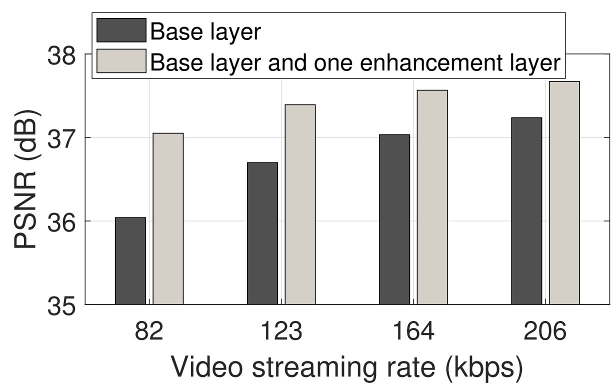



Figure 5.4: Simulated distortion (PSNR) for different video streaming rates when $SNR = 5\ dB$.

## 5.4 Performance Evaluation

Several rounds of experiments are conducted in the swimming pool as well as computer simulations to validate the proposal. Both the objective and subjective assessments of the received video on the application and physical-layer design are presented and the adaptivity of this solution is discussed to balance MIMO transmission and channel coding as well as SVC video scalability.

**Testbed Description:** We modified an existing tethered Remotely Operated Vehicle (ROV), called BlueRov2 [1], as shown in Fig. 5.1, to operate in the autonomous mode while capturing the video with its 1080p camera. The video is passed through the acoustic modem and transducer to be sent to the buoy on the other side of the link, as shown in Fig. 5.1. A high-performance and scalable platform using a programmable Kintex-7 Field-Programmable Gate Array (FPGA), called X-300, designed by Ettus Research Group with the National Instruments Corporation (NI) [5], is utilized in

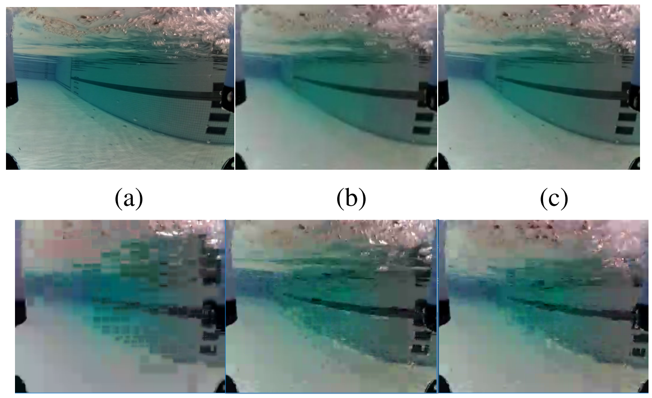
Table 5.1: Hardware Specifications.

| Part | Parameter | Value |
|---|---|---|
| Transducer | Frequency range | 1–180 kHz (Omnidirectional) |
| | Receiving sensitivity | $-211$ dB $\pm$ 3 dB re 1 V/$\mu$Pa |
| | Transmit sensitivity | 130 dB $\pm$ 3 dB re 1 V/$\mu$Pa |
| Preamp. | Frequency (Gain) | 0.5–500 kHz (0–50 dB) |
| | HP/LP filters | 1 Hz–250 kHz/1 kHz–1 MHz |
| Power Amp. | HP filters (Gain) | 1 Hz–20 kHz (0–36 dB) |
| Modem | Mainboard | Kintex-7 FPGA |
| | Frequency (Clock) | 0–30 MHz (10 MHz/1 PPS) |
| | ADC sample rate | 2 yuchannels, 200 MS/s (14 bits) |
| | DAC sample rate | 2 channels, 800 MS/s (16 bits) |
| MIMO | Uplink Structure | Upto 2x2 MIMO |
| | Feedback Structure | 1X1 SISO (FDD Duplexing) |
| Camera | Standard | H.264 1080p ($1X1.7$ mm lens) |
| | Tilt range & H. FOV | $\pm 90°$ & $110°$ |

this research. This platform contains a main-board that provides the basic functionalities of the modem and daughter-boards that take care of signal up/down conversions and other required bandpass signal processing procedures. Teledyne Marine RESON TC4013 omnidirectional transducers [3] with a frequency range of 170 kHz are used in our testbed. The specifications of the system are summarized in Table 5.1.

Joint Scalable Video Model (JSVM) software is used as the reference package for implementing SVC. Using the FixedQPEncoder program, test videos were down-sampled and encoded into multiple layers of different qualities. Each layer has a target bit rate, and the Quantization Parameter (QP) can be varied in order to optimize the PSNR while staying under the target rate. As the second decoder, OpenSVC is used for decoding due to its implementation of error concealment and its integration with Mplayer for video streaming.

**Simulations Results:** Figs. 5.2-5.4 show the result of solving optimization problems (5.10a) and (5.12a) in the pre- and post-processing blocks of the system. In these figures, the PSNR is plotted as a measure of distortion given the decision of the transmitter on the structure of video, i.e., the base layer only, or base and enhancement layers. In Fig. 5.2, the effect of SNR—as a metric of communication channel quality—on the distortion is investigated. Fig. 5.3 shows the decreasing trend of PSNR when the link distance increases. Fig. 5.4 shows the distortion metric for different video streaming rates. Given the promising results of these simulations, we conduct experiments in



(a) (b) (c)

(d) (e) (f)

Figure 5.5: (a) A frame from original video; (b) A frame of reconstructed video from base layer; (c) A frame of reconstructed video from base and enhancement layers; (d)-(f) Show the reconstructed frames of all the layers after experiencing a harsh error.

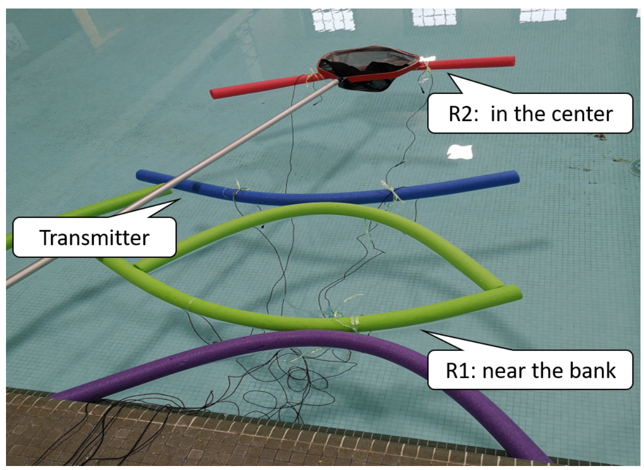



Figure 5.6: Testbed in the pool experiments. The receivers near the bank and in the center are named as R1 and R2, respectively.

the pool, as explained in the following section.

**Pool Experiments:** For our extensive experiments, hydrophones are placed in a large pool. The tests are repeated for variable distances and depths. The transmission is performed with the maximum data rate of $100\ \mathrm{kBd}$ and with H264/AVC codec JSVM signals, as shown in Fig. 5.5. We considered placing the hydrophones near the wall and also in the center of the swimming pool (Fig. 5.6), which changes the results due to the multipath effect. The stream bits are modulated with Binary Phase Shift Keying (BPSK) with different MIMO schemes. For the frame structure, the pilot symbols are inserted after every two data symbols for channel estimation. Assume the coherent time is $3\ \mathrm{ms}$ with a transmission rate of $100\ \mathrm{kbps}$; therefore, the interval time between two pilot symbols

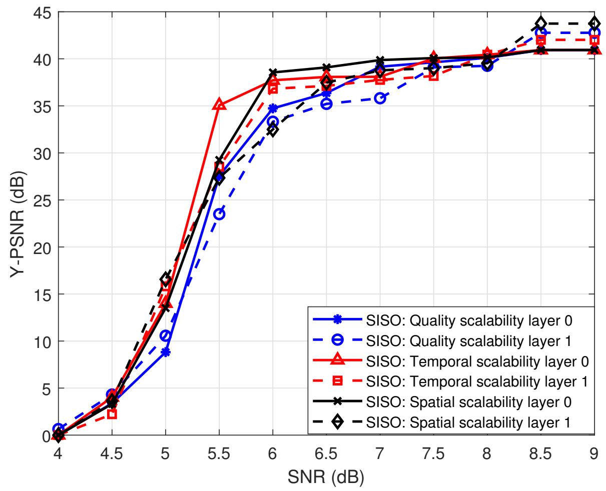



Figure 5.7: Applying SISO scheme, PSNR of the video received near the bank of the pool.

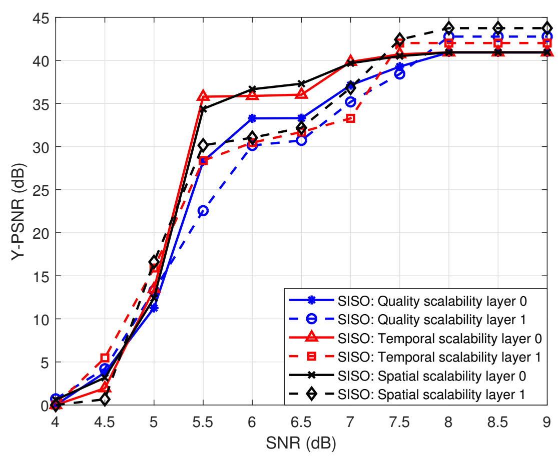



Figure 5.8: Applying SISO scheme, PSNR of the video received in the center of the pool.

is 20 $\mu$s, which is far less than the coherence time of the channel. To mitigate the multipath effect as well as to enhance the spectrum efficiency, the OFDM modulation is applied in the underwater transmission. The OFDM FFT size is chosen to be $6144$ with FFT duration of $61.44$ ms. We choose the cyclic prefix length to be $10.24$ ms. With the OFDM system bandwidth to be $100$ kHz, overall the OFDM symbol length is $71.68$ ms, and the subcarrier spacing is $16.28$ Hz. The specifications of the SVC encoder are summarized in Table 5.2.

Figs. 5.7-5.9 show the PSNR of different reconstructed videos with the SISO scheme with spatial scalability, and also the PLT in the center of the pool with spatial scalability layer $1$. We observe that when the SNR is low, the video with layer $0$ has a higher PSNR than that with layer $1$. When

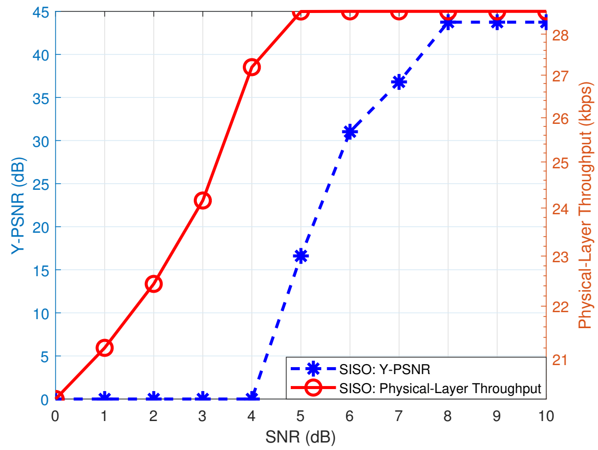



Figure 5.9: Applying SISO scheme, PLT for the video with spatial scalability layer 1, which is received in the center of the pool.

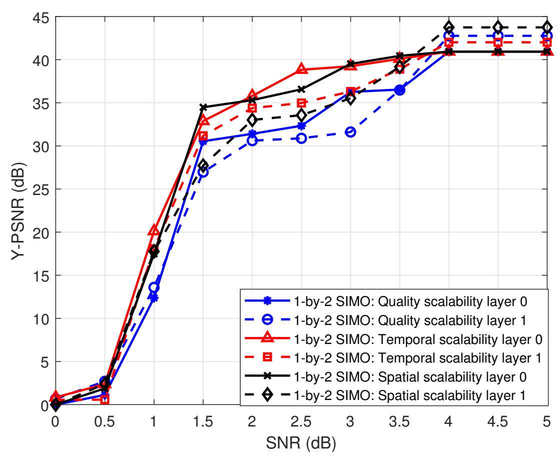



Figure 5.10: PSNR of the video received near the bank of the pool with 1-by-2 SIMO

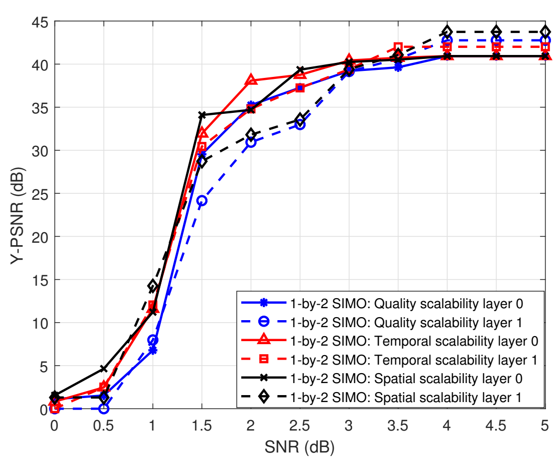



Figure 5.11: PSNR of the video received in the center of the pool with 1-by-2 SIMO

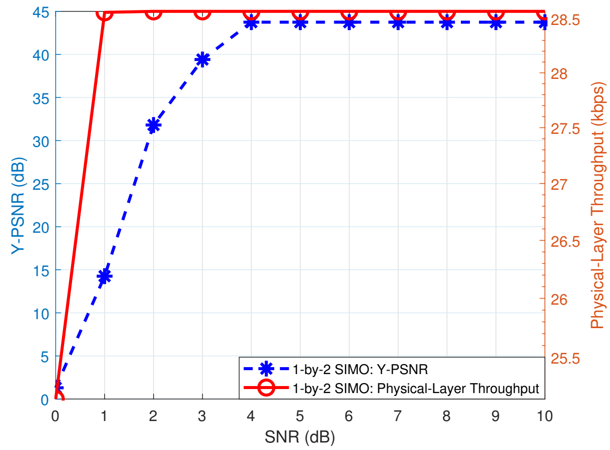



Figure 5.12: PLT for the video with spatial scalability layer 1 received in the center of the pool with 1-by-2 SIMO.

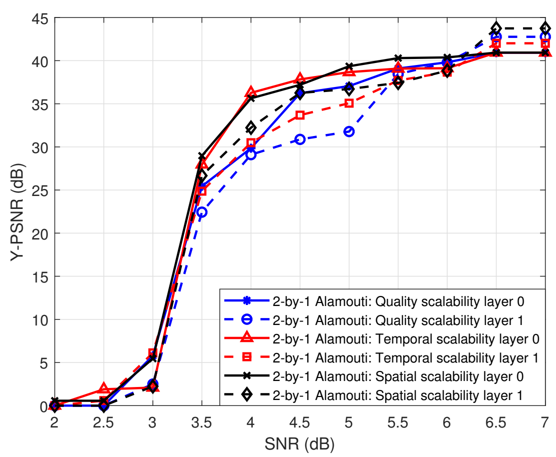



Figure 5.13: PSNR of the video received near the bank of the pool with 2-by-1 Alamouti.

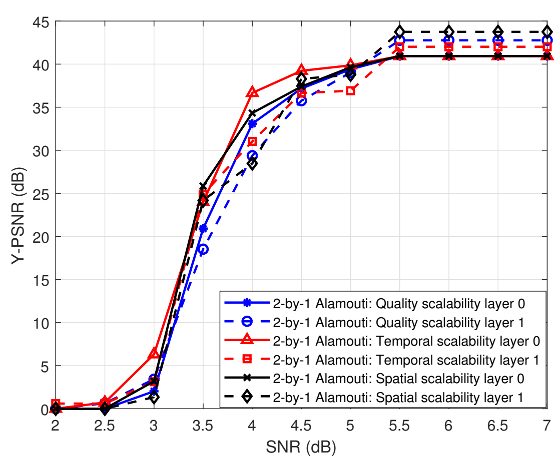



Figure 5.14: PSNR of the video received in the center of the pool with 2-by-1 Alamouti.

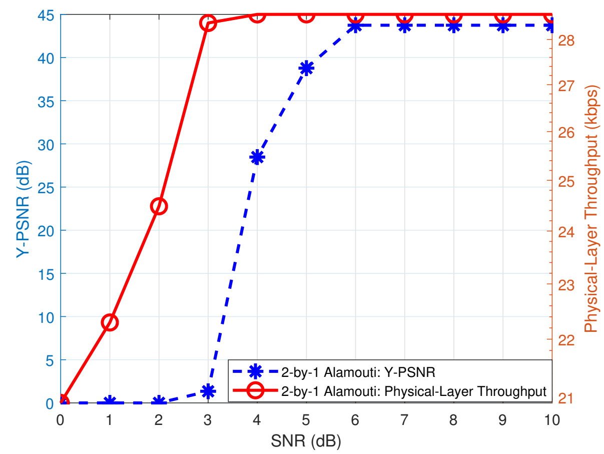



Figure 5.15: PLT for the video with spatial scalability layer 1 received in the center of the pool with 2-by-1 Alamouti.

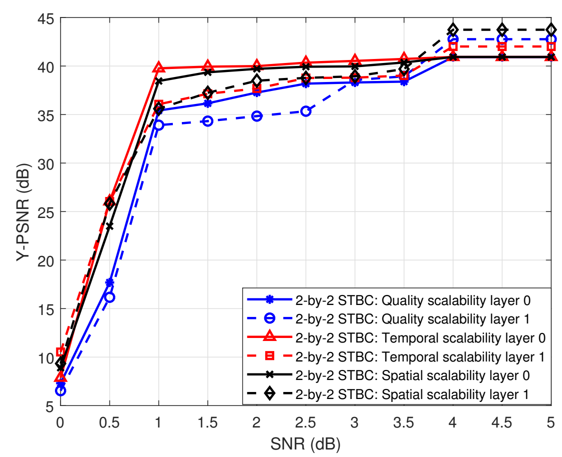



Figure 5.16: PSNR of the video received near the bank of the pool with 2-by-2 STBC.

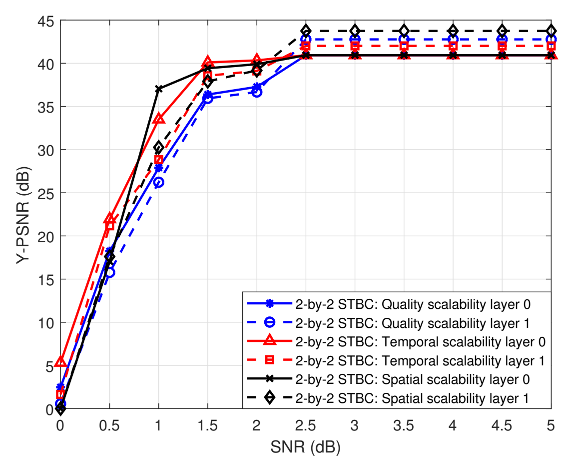



Figure 5.17: PSNR of the video received in the center of the pool with 2-by-2 STBC.

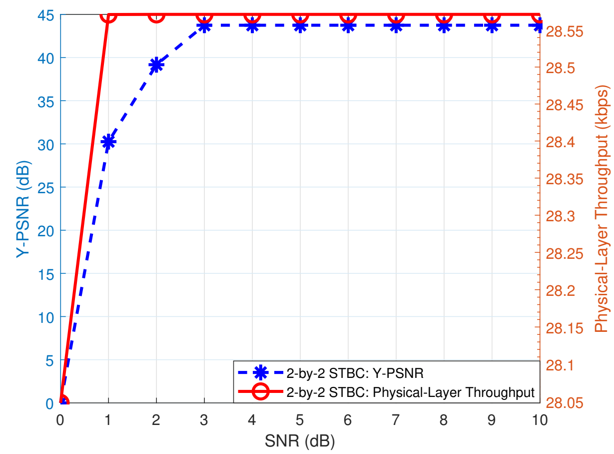



Figure 5.18: PLT for the video with spatial scalability layer 1 received in the center of the pool with 2-by-2 STBC

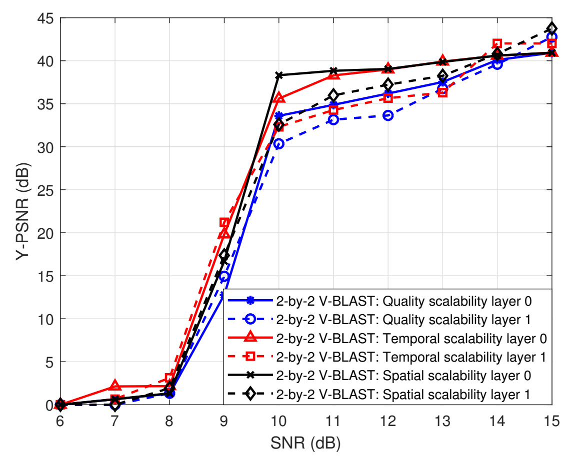



Figure 5.19: PSNR of the video received near the bank of the pool with 2-by-2 V-BLAST.

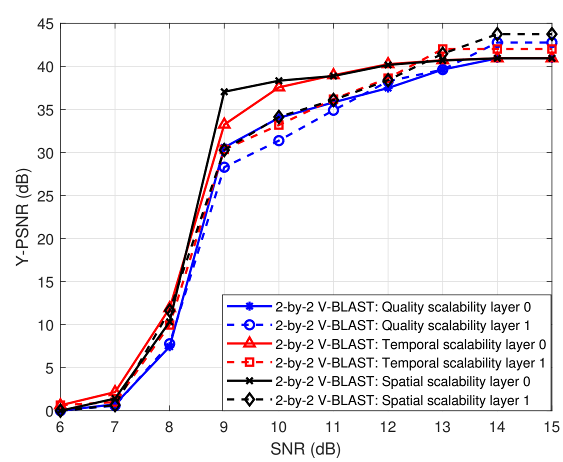



Figure 5.20: PSNR of the video received in the center of the pool with 2-by-2 V-BLAST.

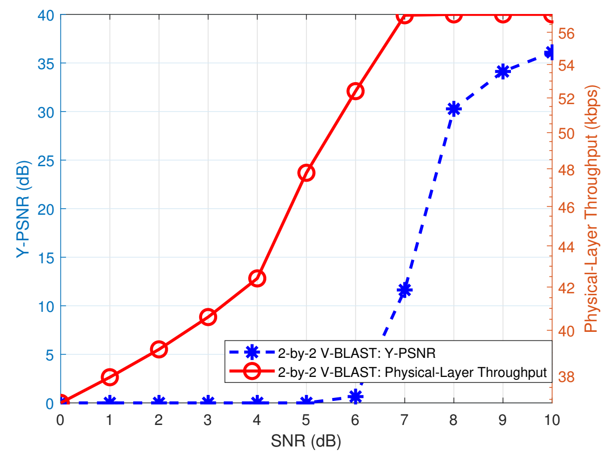



Figure 5.21: PLT for the video with spatial scalability layer 1 received in the center of the pool with 2-by-2 V-BLAST.

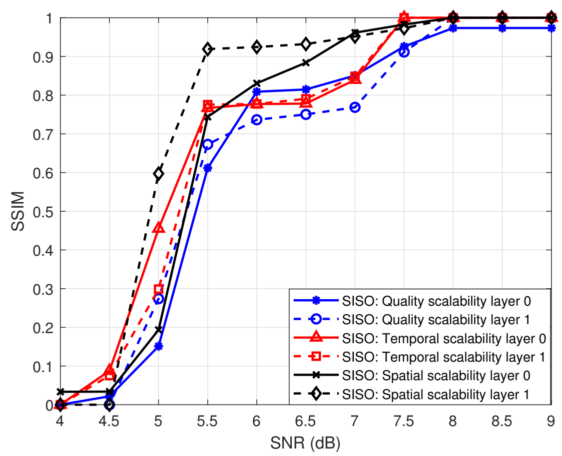



Figure 5.22: SSIM of the video with different transmission scheme for SISO.

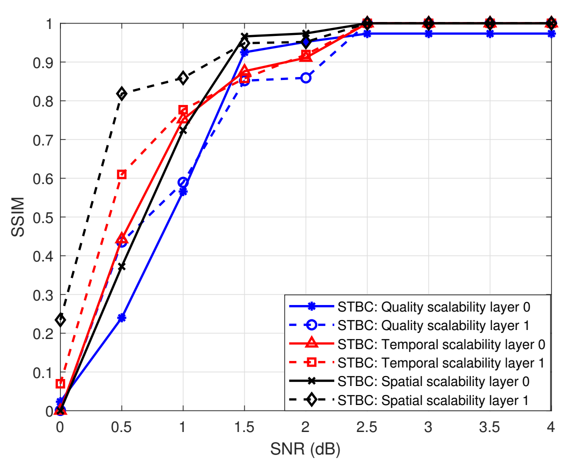



Figure 5.23: SSIM of the video with different transmission scheme for a 2-by-2 STBC.

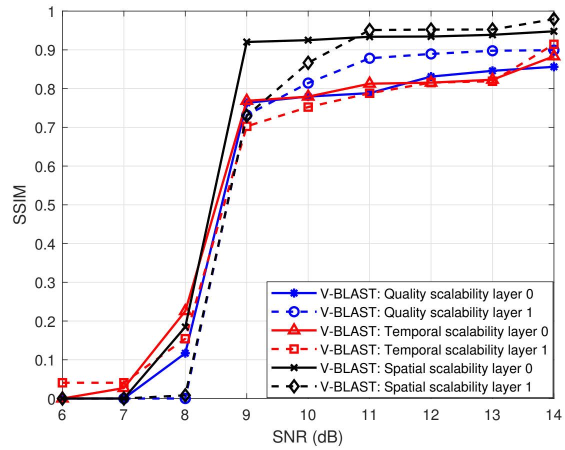



Figure 5.24: SSIM of the video with different transmission scheme for a 2-by-2 V-BLAST.

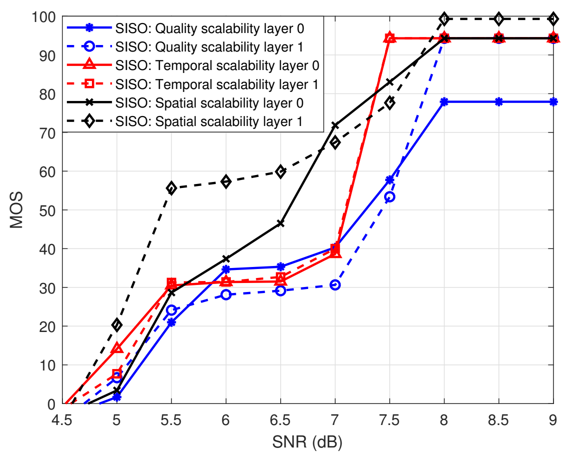



Figure 5.25: MOS of the video with different transmission scheme for SISO.

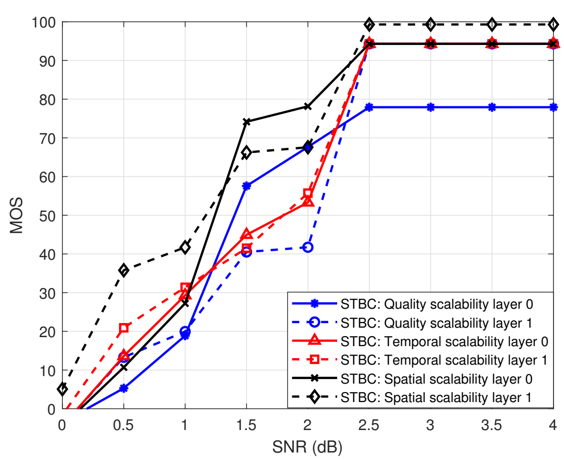



Figure 5.26: MOS of the video with different transmission scheme for a 2-by-2 STBC.

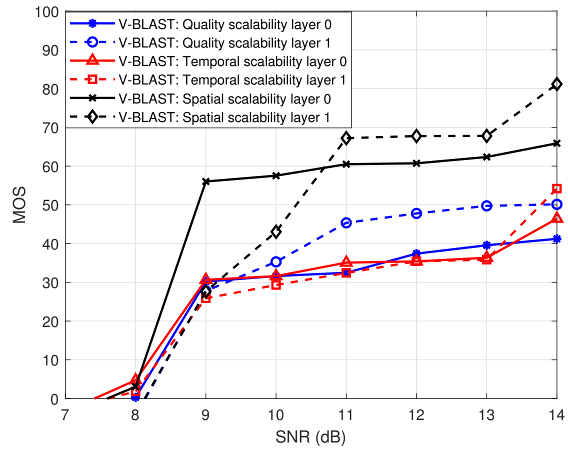



Figure 5.27: MOS of the video with different transmission scheme for a 2-by-2 V-BLAST.

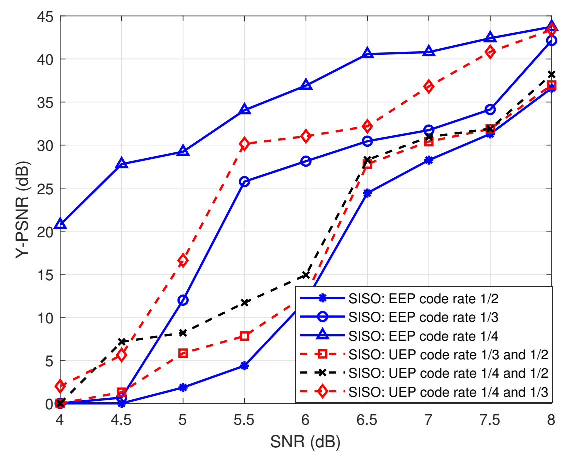



Figure 5.28: PSNR of the video with different error protection scheme for SISO.

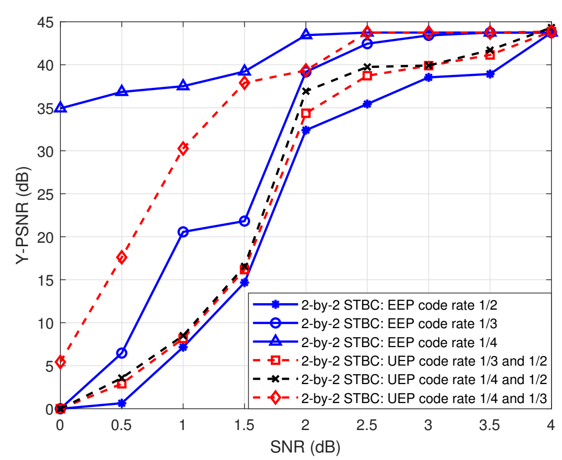



Figure 5.29: PSNR of the video with different error protection scheme for 2-by-2 STBC.

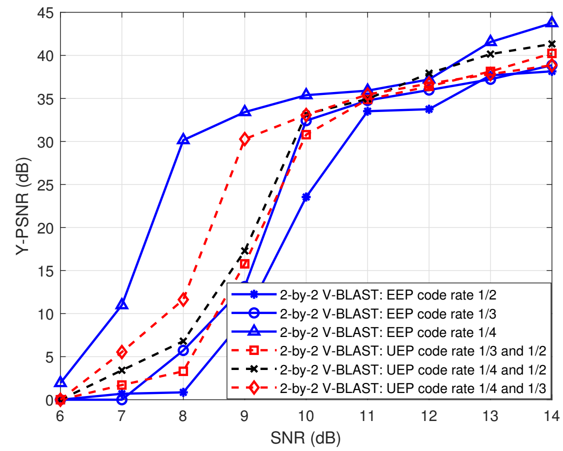



Figure 5.30: PSNR of the video with different error protection scheme for 2-by-2 V-BLAST.

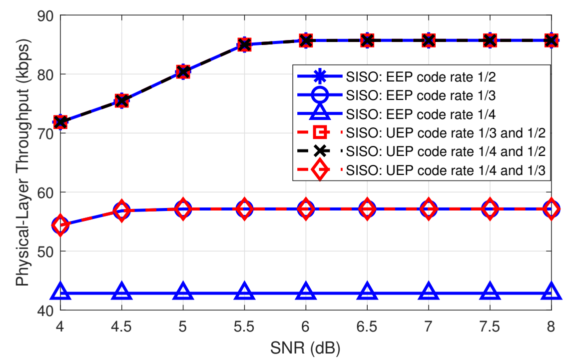



Figure 5.31: PLT of the video with different error protection scheme: SISO

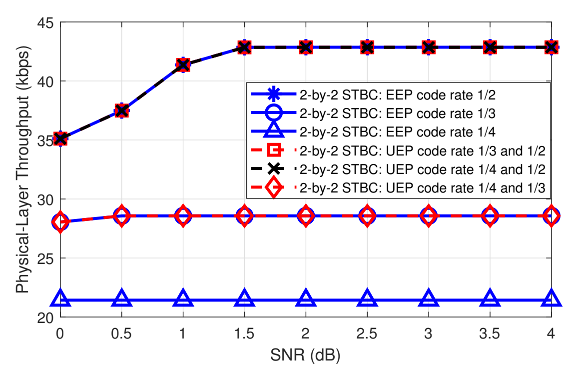



Figure 5.32: PLT of the video with different error protection scheme: 2-by-2 STBC.

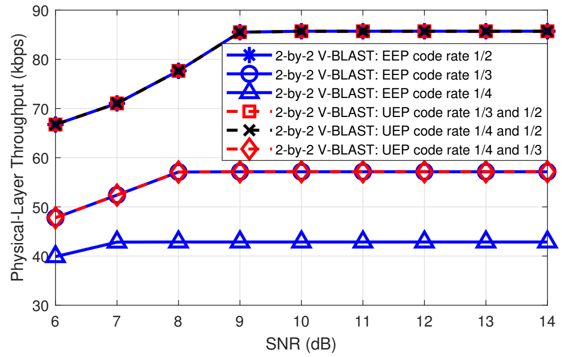



Figure 5.33: PLT of the video with different error protection scheme: 2-by-2 V-BLAST.

Table 5.2: SVC Encoder Specifications.

| **Part** | **Parameter** | **Value** |
|---|---|---|
| Base layer | Frame rate | 15 fps |
| (QP = 32) | Spatial Resolution | $640 \times 368$ |
| Quality enhancement layer | Frame rate | 15 fps |
| (QP = 30) | Spatial Resolution | $640 \times 368$ |
| Temporal enhancement layer | Frame rate | 30 fps |
| (QP = 32) | Spatial Resolution | $640 \times 368$ |
| Spatial enhancement layer | Frame rate | 15 fps |
| (QP = 32) | Spatial Resolution | $1280 \times 720$ |

the SNR is high, the video with layer $1$ has a higher PSNR than that with layer $0$, so we apply lower-quality video in bad channels and higher-quality videos in good channels. Specifically, the effect of error correction varies with different enhancement layers. For example, when the SNR is $5$ dB, the video stream with spatial scalability performs better than that with quality scalability. When the SNR is $5.5$ dB, the video stream with temporal scalability layer $0$ performs best. When the SNR ranges in $6-7.5$ dB, the stream with spatial scalability layer $0$ performs best. When the SNR is higher than $8.5$ dB, the Bit Error Rate (BER) is $0$, so the decoded streams with layer $0$ reach the equal PSNR. The stream with spatial scalability layer $1$ reaches the highest PSNR.

Figs. 5.10-5.15 show the PSNR and PLT of the received video with 1-by-2 SIMO and 2-by-1 multi-hydrophone Alamouti schemes [12]. Compared with SISO in Figs. 5.7-5.9, both SIMO and Alamouti improve the robustness of the system for each receiver that gets the transmitted stream with diversity order of 2. Compared with Alamouti, SIMO has an SNR gain of $2.5$ dB for R1 (the hydrophone in the center) and $1.5$ dB for R2 (the hydrophone in the side). While the Alamouti scheme is able to transmit two streams simultaneously, it suffers more distortion than SIMO.

The redundancy we add into the video streams will improve the system robustness, but will reduce the PLT. Figs. 5.16-5.21 show the PSNR and the PLT of the video received with 2-by-2 STBC and 2-by-2 V-BLAST [30]. Compared with SISO in Figs. 5.7-5.9, we can observe that the STBC improves the robustness significantly with an SNR gain of $4.5$ dB for R1 and $5.5$ dB for R2. While the V-BLAST suffers more distortion than SISO, but it almost doubles the PLT. Hence, we require the multiplexing-diversity tradeoff. We note that STBC is efficient when the SNR is low, whereas V-BLAST is efficient when the SNR is high.

Figs. 5.22-5.27 present the SSIM and MOS of the received video with different transmission schemes, where we can observe that the proposed cross-layer design improves the objective and subjective metrics. It is shown that when the SNR is high, the SSIM with different SVC scalability is almost the same, while the MOS performances are quite different. Moreover, the spatial scalability with layer $1$ has the highest MOS while the quality scalability with layer $0$ has the lowest MOS, even though the PSNR (Figs. 5.16-5.21) and SSIM (Figs. 5.22-5.27) performances are close to each other. Similar to Figs. 5.16-5.21, STBC improves the video quality and system robustness in the low-SNR environment, while V-BLAST can only work in the high-SNR environment but offers a higher transmission data rate.

To enhance the quality of the video, performances of UEP and Equal Error Protection (EEP) are compared. Figs. 5.28-5.30 show the PSNR of the received video in the center of the pool, using all layers with spatial scalability. We observe that the EEP with $1/4$ code rate performs the best; however, the PLT is reduced significantly. In contrast, the UEP with $1/4$ code rate for the header and $1/3$ code rate for the body also performs well. In Figs. 5.31-5.33, the UEP with $1/2$ code rate only for the body are almost overlapped with the EEP with $1/2$ code rate. The UEP with $1/3$ code rate only for the body are almost overlapped with the EEP with $1/3$ code rate. Even though we add some redundancy in the header, the length of the header is far shorter than that of the body, so the UEP reduces the PLT slightly but improves the PSNR greatly. In Fig. 5.28, we find that when the SNR is low, the UEP with $1/4$ code rate for the header and $1/2$ code rate for the body performs better than the EEP with $1/3$ code rate, for its header is protected with higher robustness by $1/4$ code rate. Figs. 5.34-5.37 represent the channel response with different transmission scheme experienced in this testbed, containing the phase of the channel in Figs. 5.34-5.35 and its power spectrum in Figs. 5.36-5.37. Due to the high transmission loss in high frequency band, the spectrum limited with passband bandwidths of only a few tens of kHz.

**Adaptivity of Our Solution:** With the cross-layer optimization algorithm described in Sect. 5.3, we can jointly improve the data rate, the system robustness, and the video quality. Based on the optimization process in Figs. 5.7-5.27, we can select the optimal video transmission scheme. As shown in these figures, R2 performs better than R1 since it suffers from less multipath delay due to the reflected signals from the bank. Given the PSNR threshold of $30$ dB and the EEP initial channel coding with Turbo coding rate of $1/3$, when the SNR is $5.5$ dB the optimal video transmission scheme is SISO with temporal scalability layer $0$ for R1 and R2, so the transmitter only needs to transmit the video stream with temporal scalability layer $0$. When the SNR is $1$ dB, the optimal scheme is 2-by-2 STBC with temporal scalability layer $0$ for R1 and spatial scalability layer $0$ for R2. With the STBC scheme, each receiver gets up to $4$ versions of received signals, which improves the reliability after gain combing. The transmit stream needs to contain both the temporal and spatial enhancement layers. When the SNR is $10$ dB, the optimal video transmission scheme is 2-by-2 V-BLAST with spatial scalability layer $0$ for R1 and R2, as the V-BLAST transmission scheme enables the transmitter to transmit two different streams simultaneously and achieves the highest transmission data rate. When the SNR is $9$ dB, the channel coding will be switched to UEP for the V-BLAST scheme based on Fig. 5.30. The UEP puts more redundancy in the stream header, which sacrifices slightly the transmission rate but almost doubles the PLT compared with the STBC. The detailed composition of different possible transmission schemes and the corresponding channel coding schemes are reported in Tables 5.3 and 5.4.

## 5.5 Summary

In this chapter, a novel scheme was proposed to layerize and transmit a video underwater using a Multi-Input Multi-Output (MIMO)-based Software-Defined Acoustic Modem (SDAM). The balance between data rate and reliability, i.e., the *multiplexing-diversity tradeoff*, as well as Scalable Video Coding (SVC) were achieved to transmit a video with a defined level of distortion, which was a result of the encoder and the error-prone underwater acoustic channel. The proposed optimization provided the scalability in the video bitstream processing and Unequal Error Protection (UEP) to adapt to the preference of end-users as well as to the varying characteristics of the network.

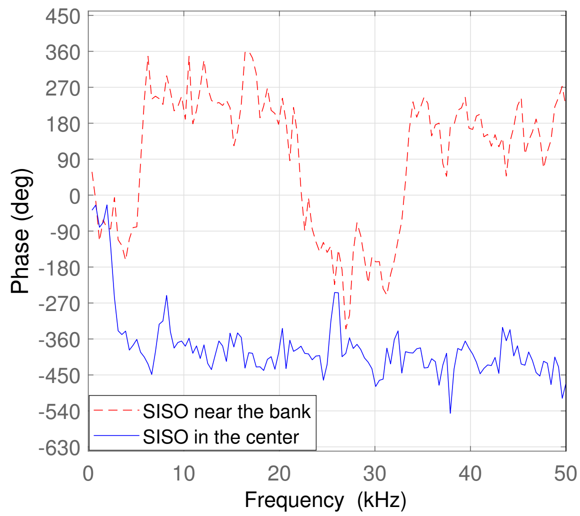



Figure 5.34: Channel response in the swimming pool: Phase of SISO.

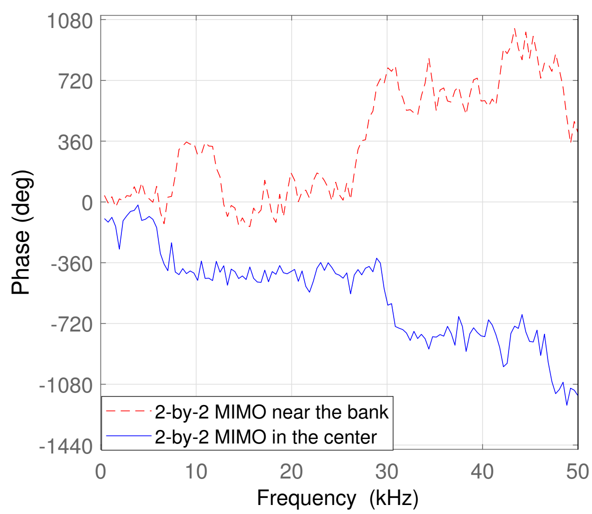



Figure 5.35: Channel response in the swimming pool: Phase of 2-by-2 MIMO.

The adaptivity of the proposed system was discussed under different scenarios and both objective and subjective metrics were considered to optimize the user Quality of Service (QoS) and Experience (QoE). Experimental results at Sonny Werblin Recreation Center, Rutgers University were presented that corroborated the analysis and intuitions.

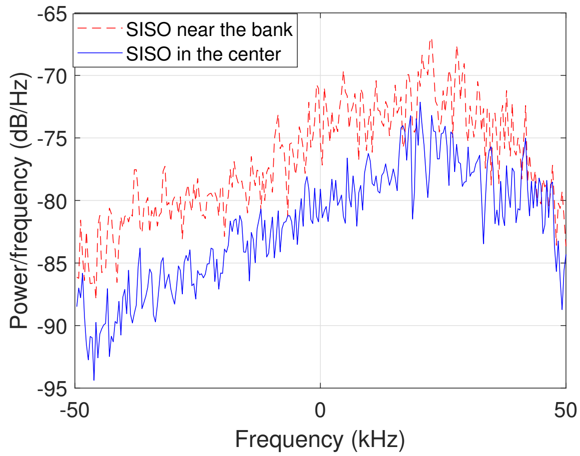



Figure 5.36: Channel response in the swimming pool: Power spectrum of SISO.

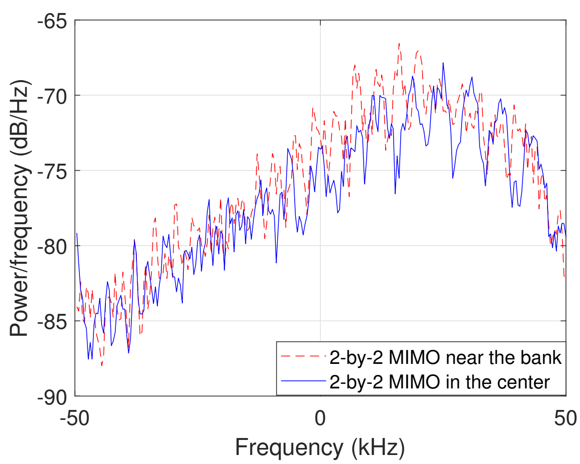



Figure 5.37: Channel response in the swimming pool: Power spectrum of 2-by-2 MIMO.

Table 5.3: Different Transmission Schemes for Node $R2$ with $1/3$ Channel Coding Rate.

| Scheme | SNR (dB) | PLT (kbps) | PSNR (dB) |
|---|---|---|---|
| SISO | $6-9$ | $28.57$ | $30.13-43.74$ |
| 1-by-2 SIMO | $2-5$ | $28.57$ | $30.94-43.74$ |
| 2-by-1 Alamouti | $4-7$ | $28.57$ | $28.47-43.74$ |
| 2-by-2 STBC | $1-5$ | $28.57$ | $26.22-43.74$ |
| 2-by-2 V-BLAST | $9-15$ | $57.14$ | $28.28-43.74$ |

Table 5.4: Different Transmission Schemes for Node $R2$ with Spatial Scalability Layer 1.

| Scheme | Channel Coding | SNR (dB) | PSNR (dB) |
|---|---|---|---|
| SISO | EEP $1/2$ | $7.5-8$ | $31.30-36.62$ |
| | EEP $1/3$ | $6.5-8$ | $30.45-42.15$ |
| | EEP $1/4$ | $5.5-8$ | $34.06-43.74$ |
| | UEP $1/3-1/2$ | $7-8$ | $30.40-36.97$ |
| | UEP $1/4-1/2$ | $7-8$ | $30.94-38.21$ |
| | UEP $1/4-1/3$ | $5.5-8$ | $30.15-43.74$ |
| 2-by-2 STBC | EEP $1/2$ | $2-4$ | $32.39-43.74$ |
| | EEP $1/3$ | $2-4$ | $39.34-43.74$ |
| | EEP $1/4$ | $0-4$ | $34.93-43.74$ |
| | UEP $1/3-1/2$ | $2-4$ | $34.36-43.74$ |
| | UEP $1/4-1/2$ | $2-4$ | $36.93-43.74$ |
| | UEP $1/4-1/3$ | $1-4$ | $30.26-43.74$ |
| 2-by-2 V-BLAST | EEP $1/2$ | $11-14$ | $33.51-38.14$ |
| | EEP $1/3$ | $10-14$ | $32.41-38.85$ |
| | EEP $1/4$ | $8-14$ | $30.15-43.74$ |
| | UEP $1/3-1/2$ | $10-14$ | $30.79-38.85$ |
| | UEP $1/4-1/2$ | $10-14$ | $33.18-40.23$ |
| | UEP $1/4-1/3$ | $9-14$ | $30.27-41.33$ |

# Chapter 6

# Scalable Video Coding Transmission for In-Network Underwater Imagery Analysis

Underwater imagery has enabled numerous civilian applications in various domains, ranging from academia to industry, and from industrial surveillance and maintenance to environmental protection and behavior of marine creatures studies. The accumulation of litter and plastic debris at the seafloor and the bottom of rivers are extremely harmful for the aquatic life. In this chapter, a solution is proposed for monitoring this problem using a team of Autonomous Underwater Vehicles (AUVs) to exchange the recorded video in order to reconstruct the map of regions of interest. However, underwater video transmission is a challenge in the harsh environment in which radio-frequency waves are absorbed for distances above a few tens of meters, optical waves require narrow laser beams and suffer from scattering and ocean wave motion, and acoustic waves—while long range—provide a very low bandwidth and unreliable channel for communication. In this solution, the scalable coded video of each vehicle is shared in-network with a selected group of receiving vehicles through the underwater acoustic channel. Presented evaluations, including both simulations and experiments, confirm the efficiency, reconfigurability , and flexibility of the proposed solution using acoustic software-defined modems.

## 6.1 Overview

Marine litter and debris, including both beached and floating objects, is one of the most serious and fast growing environmental threats in the oceans and seafloors. The negative impacts of litter accumulation on the aquatic life are unquestionable. Litter is spread widely throughout the seafloor, but its distribution is usually patchy with densities from $1$ item up to around 200 items per each $10\ \mathrm{m}$, as reported for Messina Strait's channels (one of geologically active areas of the Central Mediterranean Sea) [93]. Rivers are one of the main sources of entering litter to the seas, since they carry the litter with their currents to the sea or ocean. Deploying a team of Autonomous Underwater Vehicles (AUVs), equipped with down-looking cameras, can help in detecting these objects on the seafloor and riverbed, build a map of the pollution, and therefore, can issue early warnings so to reduce the damage to human and aquatic life. However, coordination among multiple AUVs is a challenge [113], specially when video is the subject of data exchange. AUVs should be able to encode the video, and to transmit it to other vehicles (generally to heterogeneous dynamic nodes) efficiently [95]. There are still many open problems in real-time and near-real-time underwater video processing and transmission.

To achieve these goals, novel efficient mechanisms and hardware should be utilized to make the video transmission feasible for underwater scenarios. Boosting the data rate and system reliability is possible if all the available domains are exploited in an efficient manner [110]. To stream and transmit underwater video, reliable and robust techniques are required in an environment, in which Radio Frequency (RF) waves are absorbed for distances above a few tens of meters, optical waves require narrow laser beams and suffer from scattering and ocean wave motions, and acoustic waves—while being able to propagate up to several tens of kilometers—lead to a communication channel that is very dynamic, prone to fading, spectrum limited with passband bandwidths of only a few tens of $\mathrm{kHz}$ due to high transmission loss at frequencies above $50\ \mathrm{kHz}$, and affected by the colored ambient noise.

**Motivation:** Traditional commercial acoustic modems with their fixed-hardware designs hardly meet the required data-rate and flexibility to support the futuristic underwater multimedia applications. Over the past few years, novel solutions based on adaptive and reconfigurable architectures—i.e., Software Defined Acoustic Radios (SDAR)—have been proposed. Using SDAR helps the scientists and engineers to explore different protocols and techniques on a single hardware, perform in-network analysis, and transmit the high-volume data, such as video, to a remote node depending on environment and system specifications. This concept is changing the business model of commercial acoustic modems in a near future since they are focusing more on efficient hardware/architectures and proprietary high-performance algorithms [39].

Furthermore, using conventional video compression/encoding techniques will not meet the requirements for these futuristic underwater video transmissions due to the need for higher data rate and more reliability. Therefore, more reconfigurable and flexible techniques should be utilized to

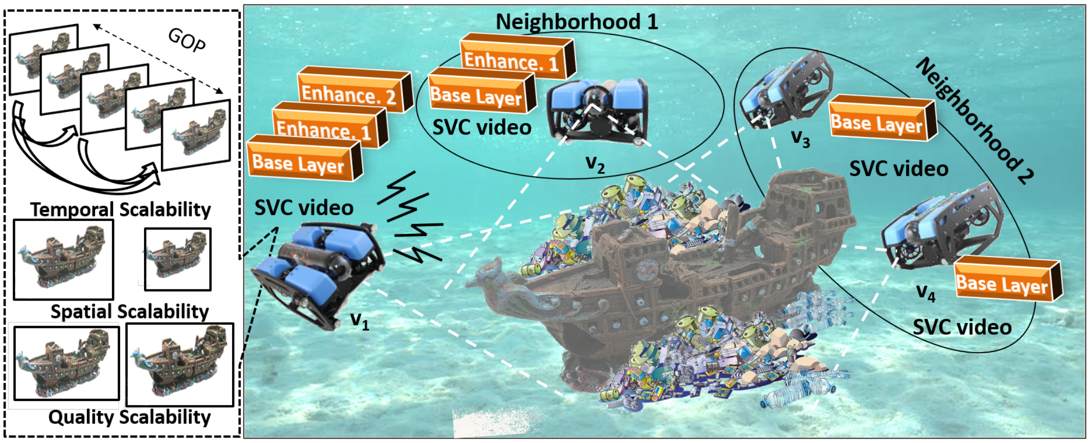



Figure 6.1: System model of the proposed SVC-based video transmission among a team of underwater vehicles, with the help of high-performance modified vehicles [1].

address this problem. In practice and in many underwater imagery/streaming applications, since the visual depth of the camera is limited in the water, the vehicle should get close enough to the target to be able to detect it, therefore, usually a single vehicle/camera can not cover the whole scene (because of the limitation in the field of view and visual depth) and can not create the global map of the environment. The coordination among the underwater vehicles is also addressed in this chapter.

A solution is proposed to encode and share the video among AUVs until the global information/reconstruction of the region of interest is achieved. Scalable Video Coding (SVC) [123], as the extension of H.264/MPEG-4 AVC, offers the required flexibility by encoding the chunks of video into a base layer and multiple enhancement layers given the requirements of the underwater channel. Fig. 6.1 shows our vision including multiple vehicles around a pile of objects. SVC base layer provides the minimum required quality, while enhancement layers offer a more enhanced quality based on different modalities–temporal scalability (frame rate), spatial scalability (frame size), and quality scalability (fidelity or SNR)–which makes this encoding a good choice for lossy video compression and erroneous transmission environments such as underwater. Here, a group of independent frames in the video structure is represented by a Group of Pictures (GOP) in the figure. Efficient video coding and reliable communications solutions are demanded for the coordination and communications among the vehicles. The reconstructed map can be used for in-network decision among the vehicles or can be transmitted to the buoy for further considerations.

**Contributions:** In many applications, more than one vehicle, due to the limited field of view and the visual depth of camera in the water, are needed to reconstruct the map of region of interest. We focus on in-network *scalable* underwater video sharing between AUVs and offer these contributions:

- A framework for underwater imagery analysis using partial information collected by various vehicles around the scene;
- An optimized solution to provide the maximum possible Quality of Service (QoS) via a proposed multicasting scalable coded video, while achieving the maximum Quality of Experience (QoE) for the scene reconstruction;
- Performance evaluation of this system under different scenarios using real videos captured from the Raritan river-New Jersey and through an SDAR testbed.

**Chapter Outline:** In Sect. 6.2, we go over the state of the art in underwater video transmission. In Sect. 6.3, we present our solution and discuss scalable video coding and the required optimizations. In Sect. 6.4, we evaluate the solution via the experiments and simulations, and then scale the results via simulations. Finally, in Sect. 6.5, we summarize the chapter.

## 6.2 Related Work

**Underwater Video Transmission:** There are several unique characteristics of underwater wireless networks that make Quality of Service (QoS) delivery of video content—ranging from delay sensitive to delay tolerant, and from loss sensitive to loss tolerant—a challenging task due to underwater acoustic frequency-dependent transmission loss, colored noise, multipath, Doppler frequency spread, high propagation delay as discussed in [95, 134]. The multiview video transmission in underwater acoustic path is discussed in [45] in which the authors propose time-shifted transmission slots to the encoder and other nodes to exchange control and video packets. The feasibility of transmitting video over short-length underwater links is investigated in [118, 148], where MPEG-4 video compression and a wavelet-based transmission method are tested on the coded Orthogonal Frequency Division Multiplexing (OFDM). Despite all these works, the problem of robust video transmission is still unsolved, and achieving high video quality is still a challenge when we consider the limited available bandwidth along with the harsh characteristics of the underwater acoustic channel, *which calls for novel high-spectral-efficiency in-network collaborative methods.* In the area of underwater video, [23] shows the feasibility of video streaming using currently commercially available hardware defined modems. The reconstructed objects can be used in Simultaneous Localization And Mapping (SLAM). SLAM is a widely used technique in ground robots, but less feasible in underwater environment specially in high turbidity situations and in the absence of reliable static landmarks. Some underwater visual SLAM solutions, such as in [105], create a sparse map for the navigation and localization in clear water.

**Scalable Video Coding (SVC):** SVC [123] outperforms the regular H.264 encoding when more flexibility and adaptation to the channel's condition are required [138]. In the area of SVC, previous research have touched on video sharing/multicasting in terrestrial context. A method for adapting the number of layers based on a fixed time allotment is proposed in [43]. This link-level method does not explore a multicast scenario. The authors in [151] explore dynamic layer adjustment in a content-delivery context where a direct-download system is paired with peer-to-peer. This sharing is top-down content delivery, rather than a scheme for in-group video sharing where each consumer is also a producer. A method for SVC video transmission is proposed in [61] using transmitter-side distortion estimates based on the channel state information. However, *none of these methods tackle the unique challenges faced in an underwater acoustic channel.* An adaptive distortion-rate tradeoff for underwater video transmission using a Multi-input Multi-output (MIMO)-based SDAR system is proposed in [104]. The scalability of the system is fulfilled using SVC compression standard. In [110] a new signaling for SVC-encoded underwater videos is proposed based on using non-contiguous OFDM and beamforming techniques with the help of Acoustic Vector Sensors (AVSs).

## 6.3 Proposed Solution for Video Sharing and Map Reconstruction

In this section, the solution for in-network video sharing and coordination among multiple AUVs is presented. First, the construction of SVC-encoded video streams and the proposed strategy to estimate the optimal parameters are discussed given underwater acoustic channel constraints as it will be explained in the optimization problems. Afterwards, the SVC-based multicasting solution to increase the overall quality of video is introduced. Finally, the proposed protocol will be presented for an efficient map reconstruction while multiple vehicles are involved in the merging process.

**Construction of SVC-encoded Video Streams:** Encoding the original video into several layers using SVC discards the need for transcoding or re-encoding the video. However, an efficient strategy is required to leverage the scalibiliteis of SVC and adapt the encoder to the receiver's status as well as the quality of acoustic channel.

**Video Sharing Setup:** Assume $V$ vehicles are deployed around a scene, as shown in Fig. 6.1, at time slot $t$ and form a wireless network of $(V, H)$, where $H$ stands for the point to point link between two vehicles, when vehicles are in the communications range of each other. Vehicles encode the initial video using SVC, and make it ready for broadcasting. To facilitate the communications, vehicles set up a basic Time Domain Multiple Access (TDMA) system and assign a time slot to each vehicle since the network size is small in underwater scenarios and the nodes are usually close together. The underwater acoustic channel presents problems for a coordinated and synchronized system such as TDMA, but due to the severe bandwidth constraint, it is important to use a Medium Access Control (MAC) that does not constrain vehicles to an even smaller slice of bandwidth, such as FDMA. Authors in [72] show that even in the underwater acoustic environment, and specially for multicast transmissions, TDMA can allow for efficient and collision-free communications. Other random- and controlled-access MAC solutions such as Carrier-sense Multiple Access (CSMA) transmit multiple packets through the same underwater channel, which might lead to packet collisions at the receiver [113]. To address the synchronization problem in TDMA (as the main weakness of using TDMA underwater), We use an unsynchronized MAC protocol, e.g., Tone Lohi (T-Lohi) [140], especially in sparse networks with limited number of nodes. The vehicles start contending any time they realize the channel is not occupied.

**Base-layer Video Sharing:** Assume each vehicle records the scene from its own angle and possibly it has an overlapping coverage with other vehicles. SVC-based video is segmented into $C$ chunks in each vehicle $j \in \{1, ..., V\}$ with a base layer $b_j$ (layer 0) with the rate $R(b_j)$ and $l_j \in 1, 2, ..., L_j$ enhancement layers with rate $R(l_j)$. Each node broadcasts the chunks of its base layer video through an acoustic channel. When a vehicle $i$ receives the base layer data of chunk $c \in C$ in time slot $t$ from transmitting vehicle $j \in \{1, ..., V\}$ and $j \neq i$ in the communication range, the received signal can be expressed as $y_i^c(t) = h_{ij}^c(\tau_{ij}; t) * b_j^c(t) + z_i(t)$, where $h_{ij}^c(\tau_{ij}^c; t)$ stands for the channel coefficient with delay $\tau_{ij}^c$ between vehicles $i$ and $j$, $y_i^c(t)$ represents the received signal, $*$ stands for the convolution operation, and $z_i(t)$ shows the background underwater colored noise. For a band-limited non-ideal underwater channel with the frequency response of $H_{ij}^c(f)$ and a Gaussian noise with the power spectral density of $S_i(f)$, the capacity $\mathcal{C}$ of each channel can be expressed as follows [100].

$$\mathcal{C}_{ij}^{c} = \frac{1}{2}\int_{-\infty}^{\infty} \log\Big(1 + \frac{P_j^c(f) \mid H_{ij}^c(f) \mid^2}{S_i(f)}\Big) df. \tag{6.1}$$

Here, $P_j^c(f)$ stands for the power spectral density of $b_j^c$ from transmitting vehicle $j$ in chunk $c$. We drop time index $t$ for the sake of simplicity and present our analysis for the time length of chunk $c$. Assume Channel State Information (CSI) is available at the transmitter and the channel is constant during broadcasting of a video stream in chunk $c$ and $B_W$ represents the channel bandwidth, which is assumed to be the same for all the users. The base layer data rate $R_{ij}(b_j^c)$ can be expressed as $R_{ij}(b_j^c) = B_W \mathcal{C}_{ij}^c$. We consider the tradeoff between the transmit power and data rate for a fixed bandwidth $B_W$ in each vehicle $j$ such that the outage does not occur. Since we assume each vehicle $j$ broadcasts its data to all other vehicles in its neighborhood through *independent channels*, the broadcast data rate $\boldsymbol{R}_j(b_j)_{BC}$ for all chunks can be bounded as follows.

$$\boldsymbol{R}_j(b_j)_{BC} = \{R_{ij}(b_j^c) : R_{m,j}^*(b_j) < R_{ij}(b_j^c) < \mathbb{E}[\mathcal{C}_{ij}]\}. \tag{6.2}$$

In this equation, $\mathbb{E}[.]$ represents the expectation operator, $\boldsymbol{R}_j(b_j)_{BC}$ stands for the practical transmission rate for broadcasting, and $R_{m,j}^*(b_j) \in \boldsymbol{R}_j^*(b_j) = [R_{1j}^*(b_j), ..., R_{V-1j}^*(b_j)]$ shows the minimum rate required in all fading situations [60] for $V-1$ receiving vehicles to avoid an outage.

In practical scenarios, in which the CSI is not fully known at the transmitting vehicle and channel gains are not known in advance, we assume that the transmitting vehicle $j$ statistically knows the ordering of the other vehicles for each chunk $c$ in time slot $t$ in terms of their instantaneous channel gains, i.e., $\mid h_{1j}^c \mid < \mid h_{2j}^c \mid < ... < \mid h_{3j}^c \mid$, for receiving vehicles $1, ..., V-1$, from weak to strong. The broadcast channel can be considered as a multiple-component channel such that a weaker component is a degraded version of the other component in a symmetric broadcast channel. It can be proved that the vehicles have the same channel quality and hence could decode the broadcast data. Here, the fading statistics are assumed to be symmetric. Considering the principle of ergodicity, if an arbitrary user $k$ can decode its data reliably, then we can conclude all the other users should be able to decode the broadcast data in the same way. This assumption breaks in the asymmetric fading case in which the users have different fading distributions. Therefore, sorting is not possible which leads to a non-degraded channel [145, Ch. 6].

We optimize the total rate for broadcasting from vehicle $j$ to other vehicles as follows.

$$\underset{p_j}{\text{maximize}} \quad \mathbb{E}\Big[\sum_{\substack{i=1\\i\neq j}}^{V} \alpha_i \log\Big(1+\frac{p_j^c |\, h_{ij}^c \,|^2}{s_i}\Big)\Big], \tag{6.3a}$$

$$\text{s.t.} \quad p_{th} \leq p_j^c \leq p_{max}, \quad \forall j \in \{1,...,V\}, \tag{6.3b}$$

$$\boldsymbol{R}_j(b_j)_{BC} \succeq R_j^*(b_j)\mathbf{1}, \tag{6.3c}$$

$$\boldsymbol{R}_j(b_j)_{BC} \preceq \mathbb{E}[\mathfrak{C}_j]. \tag{6.3d}$$

Here $\alpha_i \in \{0,1\}$ is the weighting factor, which is defined in the multicasting strategy, $p_{th}$ and $p_{max}$ show the minimum and maximum transmit power, respectively. $\mathbf{1}$ stands for an all-one vector, i.e., a vector whose entries are all equal to one, $\succeq$ and $\preceq$ represent the component-wise inequality. The capacity $\mathfrak{C}_j$ stands for the vector of all capacities to the receiving vehicles.

The optimization problem presented in (6.3) is a convex problem, since the objective function and the constraints are convex/concave; $\log(1+p_j^c|\, h_{ij}^c \,|^2/s_i)$ is concave because it is the composition of a concave function $(\log)$ with an affine mapping of $p_j^c$. Moreover, the non-negative weighted sum preserves the convexity (concavity) and the expectation of a convex (concave) function is convex (concave) [19]. Furthermore, the constraints are all affine.

In a broadcast scenario, each transmitting vehicle propagates its base layer video to all the receiving vehicles, since decoding the base layer is independent of other enhancement layers. However, the optimized data rate, calculated in (6.3), might not be sufficient for a higher quality video through the enhancement layers. Each enhancement layer $l_j$ with a defined encoding rate of $R(l_j)$ can be decoded when firstly it is received reliably and secondly its lower layer $l_j - 1$ is successfully decoded, i.e., in other words, unsuccessful decoding of the lower layers leads to a failure in decoding the current layer.

**Multicasting for Enhancement-layer Video Sharing:** In a multicast scenario and due to heterogeneity of underwater nodes, we assume the nodes with poor channel quality are able to decode the video with the base layer, while the nodes with a better communications channel quality can be served by a scalable video with a higher quality, i.e., with more enhancement layers. To be able to send the enhancement layers, a broadcasting strategy is proposed in which the vehicles with the

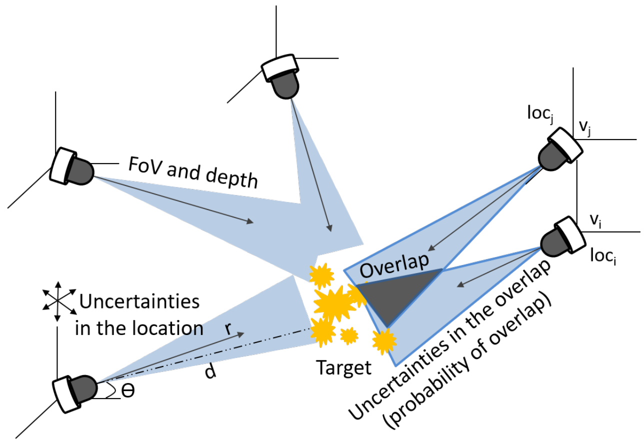



Figure 6.2: Schematics of the potential overlap between the vehicles considering the uncertainties in the location of vehicles.

worst channel are shut down in the broadcasting, i.e., $\alpha_i = 0$ in (6.6a), in order to increase the total transmission data rate. Therefore, a pseudo-multicasting network is created. Apparently, the more vehicles with impaired channels are shut down, the more enhancement layers can be transmitted to the remaining vehicles and therefore video QoS increases.

On the other hand, since the vehicles are at different locations around the scene with different viewpoints (as it is depicted in Fig. 6.2), shutting them down, leads to lack of observation and so it results in losing some information while the map is reconstructed. Map reconstruction requires a good amount of Fields of View (FoV) overlap among the vehicles. Assume the vehicles' cameras have some degrees of spatial correlation, as shown in Fig. 6.2, which is identified via the vehicles' configuration, i.e., area of overlap between FoVs of two cameras [84]. The FoV of cameras is limited to the area they observe, therefore, the information they get is directly related to the directional sensing and configuration of the vehicle. This overlap is used by the algorithm as a measure to shutdown the redundant vehicles if there exists a sufficient overlap for map reconstruction.

Let the FoV model of vehicle $i$, after 3-D to 2-D projection and calibration, be described by $(loc_i, r_i, \vec{D}_i, \beta_i)$ as in [32], in which $loc_i$ stands for the location of the vehicle, $r_i$ represents the sensing radius of the camera, $\vec{D}_i$ indicates the sensing direction (i.e., the center line of sight of the camera's FoV), and $\beta_i$ is the offset angle. A model for the spatial correlation can be derived based on the above parameters as follows. Suppose vehicles $i$ and $j$ are two arbitrary vehicles that observe an overlapped area of interest; their disparity function $\delta$ (complementary to the correlation coefficient $\eta$ as $\delta = 1-\eta$) is defined as follows [32]: $\delta = \frac{1}{4}\Big( \Big|\frac{d\sin\theta}{d+\cos\theta}\Big| + \Big|\frac{d\sin\theta}{d-\cos\theta}\Big| + \Big|\frac{d\cos\theta}{d+\sin\theta} - 1\Big| + \Big|\frac{-d\cos\theta}{d-\sin\theta} + 1\Big| \Big)$, where $d$ denotes the camera depth (here, the difference between the $loc_i$ and the target's location assuming the camera sensing direction $\vec{D}_i$ is headed to the target) and $\theta$ is the angle between the sensing direction and the $x$-axis, so that the location $loc_i$ can be expressed by $(-d\cos\theta, -d\sin\theta)$ after the 2-D projection. Specifically, for two vehicles $i$ and $j$ with parameters $(d_i, r_i, \theta_i)$ and $(d_j, r_j, \theta_j)$, respectively, the disparity between their images can be calculated as follows [32, 84],

$$\delta_{i,j} = \frac{1}{4}\Bigg(\left|\frac{-d_i\sin\theta_i - r_i\cos\theta_i}{d_i+\cos\theta_i} - \frac{-d_j\sin\theta_j - r_j\cos\theta_j}{d_j+\cos\theta_j}\right| + \left|\frac{d_i\sin\theta_i + r_i\cos\theta_i}{d_i-\cos\theta_i} - \frac{d_j\sin\theta_j + r_j\cos\theta_j}{d_j-\cos\theta_j}\right| +$$
$$\left|\frac{d_i\cos\theta_i - r_i\sin\theta_i}{d_i+\sin\theta_i} - \frac{d_j\cos\theta_j - r_j\sin\theta_j}{d_j+\sin\theta_j}\right| + \left|\frac{-d_i\cos\theta_i + r_i\sin\theta_i}{d_i-\sin\theta_i} - \frac{-d_j\cos\theta_j + r_j\sin\theta_j}{d_j-\sin\theta_j}\right|\Bigg). \quad (6.4)$$

However, finding the exact amount of correlation might not be feasible due to the position uncertainty of the vehicles and the effect of currents on the vehicles due to vehicle's drifting. Therefore, inaccuracies in position estimation increases and it becomes worse over time when the vehicle stays longer underwater, which leads to non-negligible drifts in the vehicle's position and thus making the camera overlap accurate calculations inapplicable.

In [113], an approach has been proposed to estimate vehicles' position through a statistical method based on the vehicles' confidence region. Assume each vehicle $i$ measures $N$ random samples of its location as $\{loc_i^{(n)}\}_{n=1}^{N}$. The measured locations are samples of a normal distribution $\mathcal{N}(\mu_i, \sigma_i^2)$ with the mean and variance $\mu_i$ and $\sigma_i^2$, respectively. The samples also follow a normal distribution with mean $\mu_i'$ and variance ${\sigma'}_i^2$. It can be inferred that $\dfrac{\mu_i' - \mu_i}{{\sigma'}_i^2/\sqrt{N}}$ is a pivot and it has a student's t-distribution with $N-1$ degrees of freedom. The mean $\mu_i' = \sum_{n=1}^{N} loc_i^{(n)}/N$ and the variance can be estimated as ${\sigma'}_i^2 = 1/(N-1)\sum_{n=1}^{N}\left(loc_i^{(n)} - \mu_i'\right)^2$ [113]. The uncertainty region, i.e., confidence interval, of this vehicle can be derived as $\Pr(L_i \leq \mu_i' \leq U_i) \geq 1-\gamma$. Here $\gamma$ is the confidence degree, $\Pr(.)$ represents the probability function, and $L_i$ and $U_i$ are the interval boundaries of vehicle $i$ and are estimated as $L_i = \mu_i' - \mathcal{T}_{(N-1,\alpha/2)}\sigma^2/\sqrt{N}$ and $U_i = \mu_i' + \mathcal{T}_{(N-1,\alpha/2)}\sigma_i^2/\sqrt{N}$. Here, $\mathcal{T}_{N-1,\alpha/2}$ is the t-distribution critical value with $N-1$ degrees of freedom. To estimate the amount of overlap between two vehicles $i$ and $j$, the probability of overlap is defined as $\Pr_{i,j}^{(o)} = \Pr(\eta_{i,j} > 0) = \Pr(\delta_{i,j} < 1)$, we have, $\mathrm{Pr}_{\mathrm{i,j}}^{(\mathrm{o})} = \int_0^\infty \mathrm{f}(\eta_{\mathrm{i,j}})\mathrm{d}\eta_{\mathrm{i,j}} = \frac{1}{\sigma_{\mathrm{i,j}}\sqrt{2\pi}}\int_0^\infty \exp\left\{-\frac{1}{2}\left(\frac{\eta_{\mathrm{i,j}}-\mu_{\mathrm{i,j}}}{\sigma_{\mathrm{i,j}}}\right)^2\right\}\mathrm{d}\eta_{\mathrm{i,j}}$. By defining the auxiliary variable $x = (\eta_{i,j} - \mu_{i,j})/\sigma_{i,j}$, we can conclude the following result.

$$\mathrm{Pr}^{(\mathrm{o})}_{\mathrm{i,j}} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(\mu_{\mathrm{i,j}}/\sigma_{\mathrm{i,j}})} \mathrm{e}^{-(\mathrm{x}^2/2)} \mathrm{dx} = \Phi(\frac{\mu_{\mathrm{i,j}}}{\sigma_{\mathrm{i,j}}}), \tag{6.5}$$

where $\Phi(.)$ is the Cumulative Distribution Function (CDF) of the standard normal distribution.

The following optimization problem in (6.6) justifies the discussion on the number of enhancement layers that the transmitter can handle on the top of the encoded base layer video. This is a knapsack program, which defines the enhancement layers of rate $R_j(l_j)$ that could be transmitted over the underwater channel with maximum achievable communication data rate $R_{max}$,

$$\underset{\lambda_k}{\text{maximize}} \quad \sum_{l=1}^{L} \lambda_l \lambda_{l-1} R_j(l_j), \tag{6.6a}$$

$$\text{s.t.} \quad \sum_{l=1}^{L} \lambda_l \lambda_{l-1} R_j(l_j) \leq R_{max}, \tag{6.6b}$$

$$\lambda_0 = 1,\ \lambda_l \in \{0,1\}, \forall l \in \{1, ..., L\}. \tag{6.6c}$$

We determine the minimum number of vehicles to shut down such that we achieve the required QoS in the received video with an acceptable Quality of Experience (QoE) in the reconstructed map of environment based on a defined amount of spatial correlation. Vehicles are eligible to transmit a video with higher enhancement layers while the layers bellow are successfully received/decoded. In this case, the following optimization problem can be presented for every chunk $c$ of the video, given the optimal power $P_j$ and the data rate $R_j$ calculated from (6.3),

$$\underset{\alpha_i}{\text{maximize}} \sum_{i=1}^{V-1} \alpha_i \tag{6.7a}$$

$$s.t.\ \alpha_i \in \{0,1\}, \tag{6.7b}$$

$$\mathbb{E}\Big[\sum_{\substack{i=1\\ i\neq j}}^{V} \alpha_i \log\Big(1 + \frac{p_j^c |\ h_{ij}^c\ |^2}{s_i}\Big)\Big] \geq \mathrm{QoS}_{th}(l_j), \tag{6.7c}$$

$$D_i < D_{th}, \tag{6.7d}$$

$$\mathrm{Pr}^{(\mathrm{o})}_{\mathrm{i,k}} \geq \mathrm{Pr}_{\mathrm{th}}, \quad \forall \mathrm{i,k} \in \{1, ..., \mathrm{V}-1\}, \tag{6.7e}$$

where the objective function (6.7a) is the total number of vehicles. Maximizing the total number of vehicles (i.e. minimizing the number of vehicles to shut down) ensures the QoE since more vehicles from different angles are present in the map reconstruction. On the other hand, to satisfy a threshold QoS, the proposed method will shut down the vehicles with the worst channel to keep the average broadcasting rate over a minimum value, as shown in (6.7c). The other metric for QOS is represented in constraint (6.7d) which is defined by the SVC encoder and depends on the scalability and the number of enhancement layers that the encoder uses. For an encoded video, we can write [137] $D_i = \hat{\theta}/(R_j - R_0) + D_0$, where $D_i$ represents the distortion of the video at the vehicle $i$ at the time of reconstruction and $R_j$ is the rate of the encoder at vehicle $j$; the other remaining variables $\hat{\theta}$, $R_0$, and $D_0$ depend on the encoded video and on the model, and are estimated empirically. The last constraint (6.7e) shuts down the vehicles which have a higher probability of overlap with the neighboring vehicles to have the minimum reduction in the QoE.

**In-network Marine Litter Map Reconstruction:** As it was discussed in the previous sections and due to the limited FoV of each single vehicle, a cooperation among the vehicles is required so that the required map can be reconstructed. Different cooperation strategies can be proposed based on the exchanged data, acoustic channel requirements, level of complexity (that the vehicles can handle to process the data locally) and the QoS/QoE requirements as follows: $(i)$ Vehicles exchange their local maps after each partial map is created. This strategy requires the minimum amount of data exchange since the merger creates the global map based on only a consensus on the exchanged local maps. $(ii)$ Vehicles exchange the SVC-based channel independent videos, i.e., base layers. $(iii)$ Vehicles exchange SVC adaptive channel dependent video, i.e., base and enhancement layers. This is the most desirable strategy that is also adaptive with the channel quality. $(iv)$ Vehicles exchange the high quality video considering the acoustic channel bandwidth and the channel fading. This strategy is usually not feasible underwater due to the bandwidth limitation and time-varying nature of the underwater acoustic channel. Fig. 6.3 depicts the strategy shown in this chapter. After sharing the base layer, as discussed in the previous sections, the vehicles with unreliable channels are shut down to be able to reach the required rate for sending the enhancement layers. We lose some part of the scene from those nodes which experience the shut down. Therefore, the vehicles should reach a consensus to decide on the node who finally reconstructs the global map.

**Local Map Reconstruction:** With the base layer video received at each node, along with that

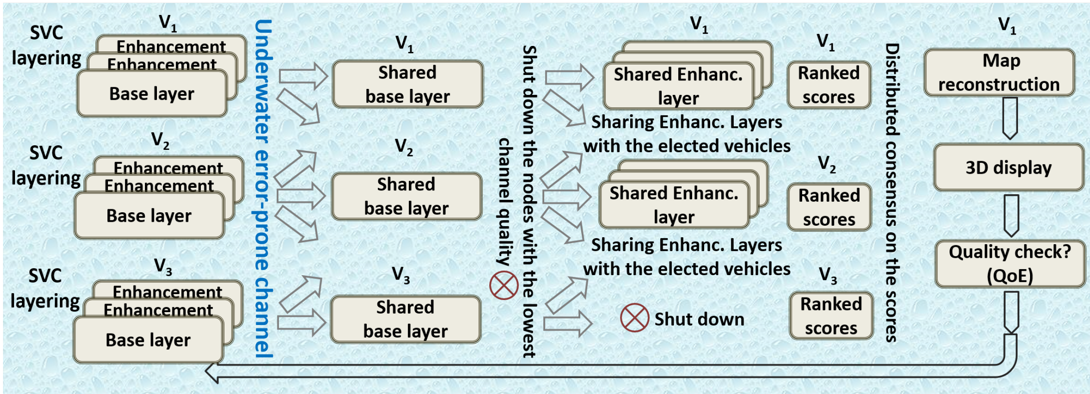



Figure 6.3: Map construction flowchart. Vehicles broadcast the base layer, while enhancement layers are shared with vehicles with better acoustic channel quality.

node's own high quality 4K original video, each node can perform a quick attempt at the map reconstruction. First, images are compared pairwise using SIFT/ORB to determine feature matches. Some of these pairwise matches will be false, and will appear in some pairwise comparisons but not in others that show similar perspectives on the scene. Because all nodes have some versions of the video, from different angles, the quality of reconstruction (measured by number of feature matches) should relate to two factors. Firstly, it depends on the amount of error-induced distortion in the base layer videos received from the other nodes. Secondly, it depends on the utility the locally stored 4K quality video on the reconstructing vehicle provides to the map reconstruction. Therefore, a vehicle that makes many feature matches in the intermediate local reconstruction attempt is a good candidate to share its recorded video at a higher quality in the next phase, because its video is a valuable part of the reconstruction and easy to match with the other videos. The underwater environment poses additional challenges in recording good video for the purposes of map reconstruction. While it can be shown that water itself is not a barrier to getting a good reconstruction, there are serious problems with lighting, scattering, turbidity, and clarity when taking underwater video.

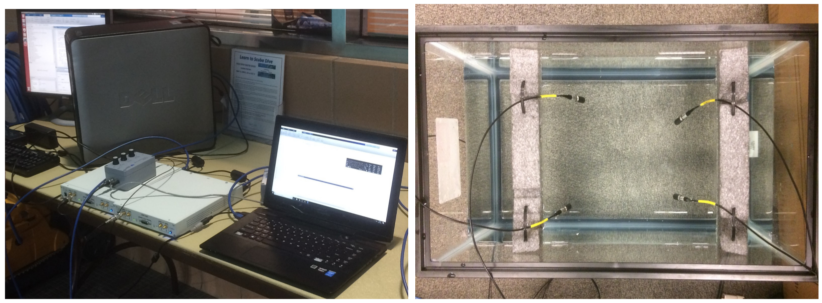
**Scoring and Sharing:** Using the optimizations described in the previous sections, each transmitting node decides on the set of nodes to shut down before broadcasts its higher quality layers, i.e., enhancement layers. Therefore, some nodes miss some portions of video from some other angles since they did not receive them. A Reconstruction Score (RS) is formed which is taken as a metric for how successful this vehicle would be at performing the later final reconstruction, as well as how valuable its local video is. This RS is shared in the following step to elect the Final

**Algorithm 4** SVC-based Map Reconstruction.

```
while reconstruction is NOT satisfactory do
    Layers = ScalableVideoCoder(localVideo)
    EstablishMACchedule()
    $\tau_b$ ← allotted time for base layer sharing
    while $t < \tau_b$ do
        Share(Layers.LayerIndex(0))        % broadcasting
        Receive(ExternalVideo)
    end while
    receivedframes ← extractframes(receivedVideos)
    SIFT/ORBmatch(receivedframes)
    Reconstruct(matchedframes)
    RS ← score(reconstruct)
    random_broadcast_max(RS)
    $\tau_l$ ← allotted time for enhancement layer sharing
    while $t < \tau_l$ do
        if $v$ is not FRV then
            Shut down the vehicles with the weakest channel
            Share(Layers.LayerIndex($L$))        % multicasting
            Receive(ExternalVideo)
        else
            Reconstruct(Dataset)
        end if
    end while
end while
```

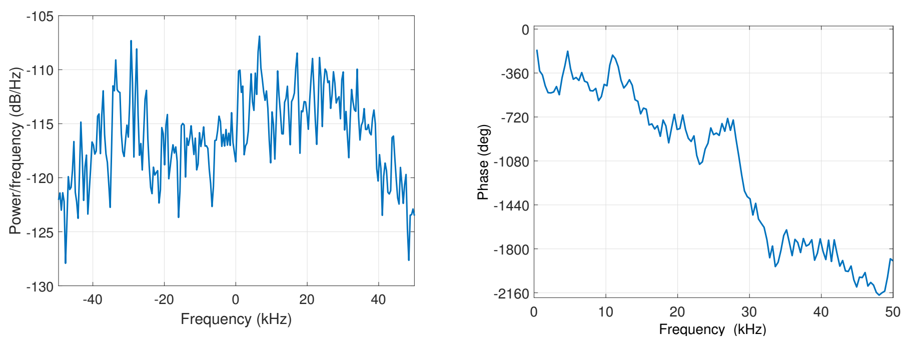
Reconstructing Vehicle (FRV). Each node will share its RS to the group, such that at the end of this step all vehicles should have a list of each other vehicle's RS. As the process continues, nodes will become more aware of their position relative to other nodes. Since the RS is a very small amount of data, each vehicle can also share in the packet a map of camera positions (past vehicle positions) it has matched with. An average of these maps can be used to inform the vehicle's navigation in the time before the final reconstruction can be performed.

**Consensus Algorithm on the Scores:** To select the vehicle with the highest score for the final reconstruction, vehicles form the communication primitive to their neighboring vehicles. In particular, consensus is an iterative process where the nodes communicate with their neighbors to exchange their scores for a fixed number of iterations or until convergence [84]. As the output of this process, the best vehicle is selected for final reconstruction. Asynchronous broadcasting-based consensus method proposed in [84] is to achieve the average value of the initial measurements. However, we wish to sort the scores to find the maximum in each iteration of the process. Each



(a) (b)

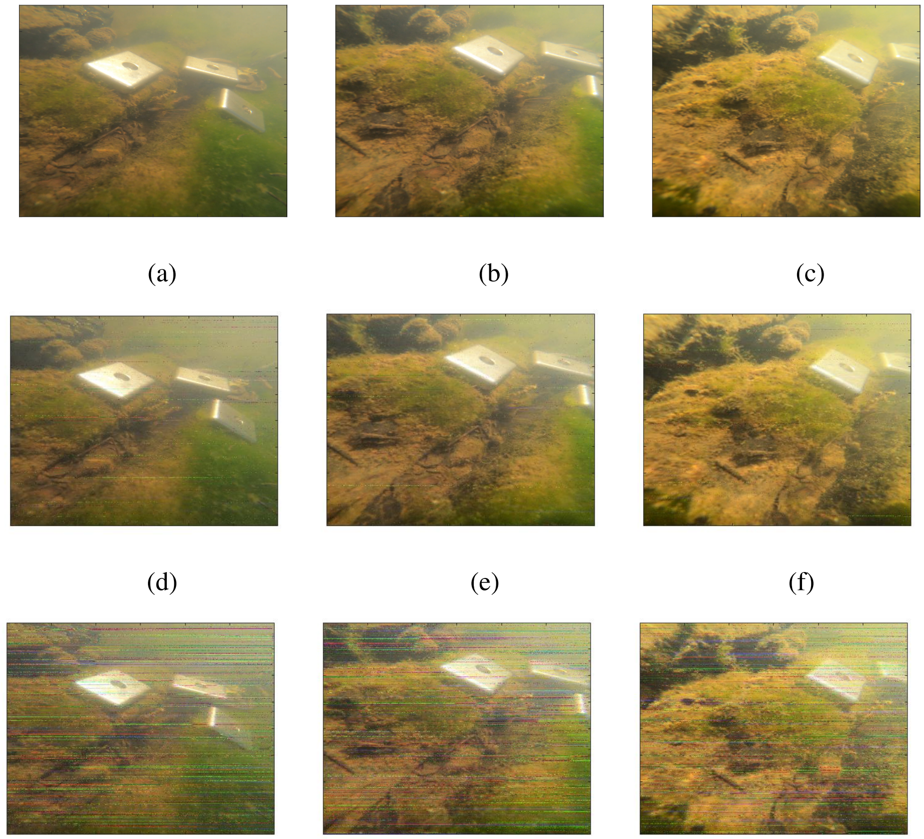
Figure 6.4: (a) Software-defined testbed; (b) Water tank with TC4013 Teledyne transducers.




(a) (b)

Figure 6.5: Channel response in the water tank shows (a) Power spectrum; (b) Phase.

node $v$ broadcasts its own score to its $\mathcal{N}_v$ neighboring nodes within its communication range [59]. The neighbors, such as $w$, which received the data, update their data according to $y_w(t_c+1) = \max\big(\mathrm{y_v(t_c), y_w(t_c)}\big), \forall \mathrm{w} \in \mathcal{N}_\mathrm{v}$, where $N_v$ stands for the neighborhood of transmitting node $v$. The remaining nodes in the network update their values as $y_w(t_c+1) = y_w(t_c), \forall w \notin \mathcal{N}_v$. This algorithm keeps the maximum value and so does not show an undesirable behavior in terms of convergence. After consensus, each vehicle should know the maximum RS among them and the vehicle that has it. The vehicle who has the highest score will transmit a final packet indicating its RS and intent to become the FRV. If there is no reply within the time limit, it is the FRV and the SVC enhancement layer sharing will commence. Algorithm 4 represents the solution in a sequential procedure for a specific coded video while the encoding and reconstruction is performed through the mentioned steps. Vehicles share their encoded base-layer and enhancement layers videos (after shutting down the vehicles with a low quality channel). After local reconstruction, matching and ranking the scores, the node with the highest score will be elected to perform the reconstruction.

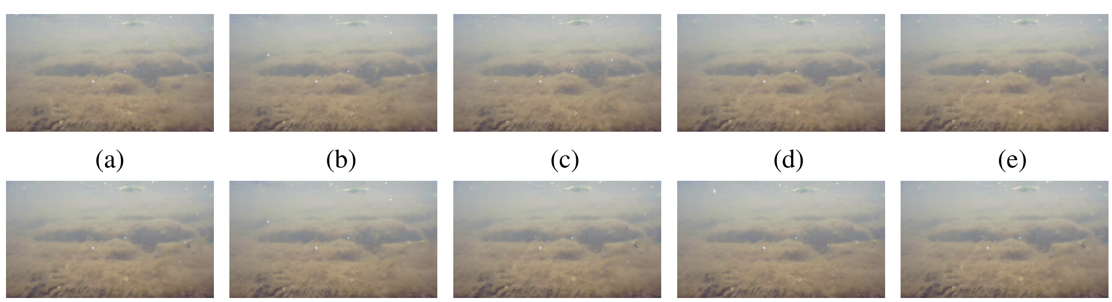


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 6.6: (a)-(c) Frames from original video; (d)-(f) Frames of video received/reconstructed in a vehicle with a good channel; (g)-(i) Frames of video received/reconstructed at a vehicle with an average to low channel quality.



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Figure 6.7: SVC layers for a selected frame; (a) Base layer of original video; (b)-(e) Base layer and $1-4$ enhancement layers of original video; (f) Base layer of received video; (g)-(j) Base layer and enhancement layers of received video.

## 6.4 Performance Evaluation

In this section, the experiments and simulation results are presented.

**Testbed Setup:** The proposed approach is evaluated by conducting preliminary field experiments. A video feed, captured by our underwater vehicles in the Raritan river, New Jersey, is passed to the SDAR and an acoustic transducer in a water tank. A high-performance and scalable platform with a programmable Kintex-7 FPGA, called X-300 designed by Ettus Research Group [5], is exploited as SDAR in this research, as the testbed shown in Fig. 6.4(a). It contains a mainboard to provide basic functionalities of the modem, while the daughter-boards take care of up/down conversions and of the other required bandpass signal processing procedures. Teledyne Marine TC4013 transducers [3] with a frequency range of $170$ kHz are used in the proposed testbed, shown in Fig. 6.4(b). Fig. 6.5 represents the channel response experienced in this testbed, containing the power spectrum of the channel in 6.5(a) and its phase in 6.5(b). The video was collected from the bottom of the Raritan river, New Jersey, using multiple cameras. The Joint Scalable Video Mode (JSVM) software is used as the reference package for implementing SVC. Using the FixedQPEncoder program, test videos were down-sampled and then encoded into multiple layers of different qualities. Each layer has a target fixed bit rate, and the Quantization Parameter (QP) is varied in order to optimize the Peak Signal-to-Noise Ratio (PSNR) metric while staying under the target bitrate.

**Results:** Fig. 6.6 shows the effect of the acoustic communication channel on the quality of the received video. The passband channel bandwidth is $100$ kHz with carrier frequency of $100$ kHz and the sampling rate is $200$ kHz. In Figs. 6.6(a)-(c), the original successive frames are shown, while in Figs. 6.6(d)-(f) the quality of the received signal through a good channel is compared to the quality of the received signal through a low to average channel in Figs. 6.6(g)-(i).

Fig. 6.7 depicts different SVC layers of a selected frame from the captured video. Fig. 6.7(a) shows the base layer and Fig. 6.7(b)-(e) represent the base and $1$ to $4$ enhancement layers of the original captured video. The corresponding frame rates for these layers are $1.8750, 3.75, 7.5, 15, 30$, respectively with the minimum bit rates of $100.9, 179.4, 293.3, 415.3, 517.5$ kbps. The corresponding PSNR values are $45.1, 44.14, 43.31, 42.68$ and $42.19$ dB, respectively. Fig. 6.7(f)-(j) show the associated base and enhancement layers of the same frame, when passed through our testbed. Note that the difference between number of enhancement layers can be distinguished better in the video.

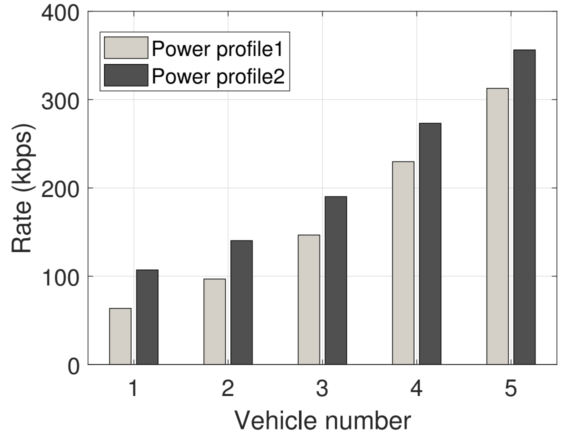



Figure 6.8: Optimal received rate at different vehicles, which are sorted based on their channel quality for two power profiles when all vehicles are active.

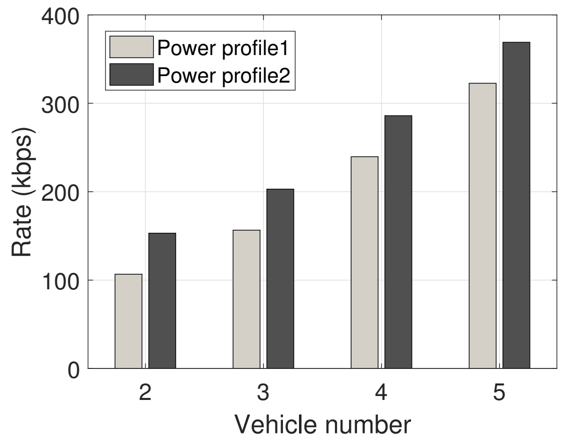



Figure 6.9: Optimal received rate at different vehicles which are sorted based on their channel quality when the vehicle with the worst channel quality is shut down.

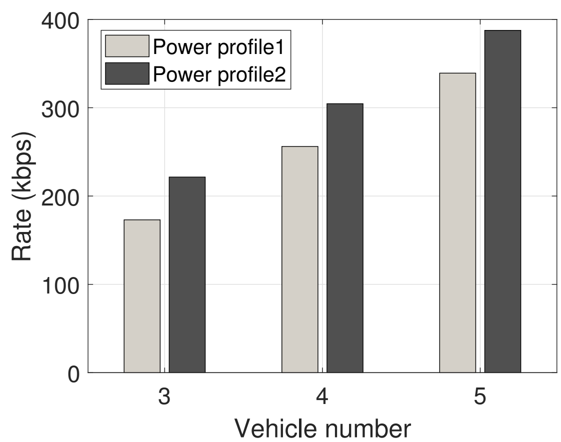



Figure 6.10: Optimal received rate at different vehicles which are sorted based on their channel quality for two defined power profiles when two vehicles are shut down.

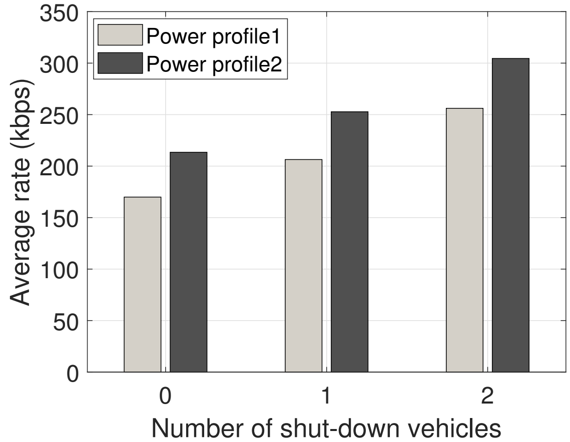
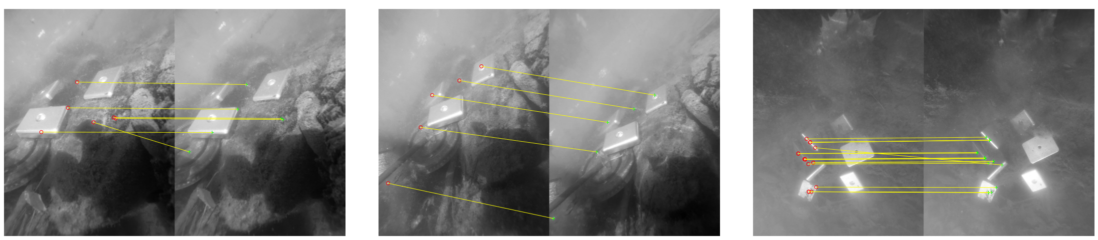
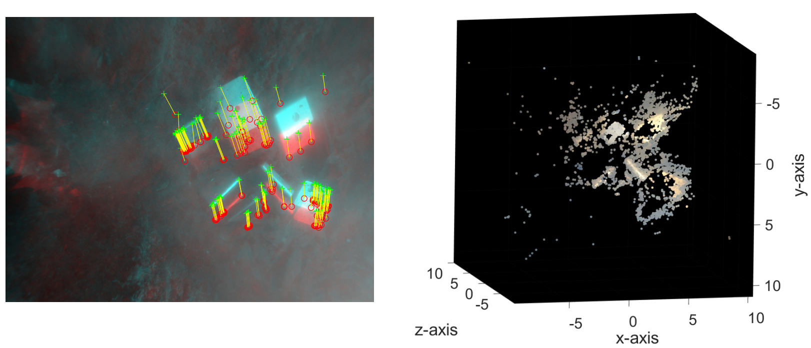



Figure 6.11: Optimal received rate which are sorted based on their channel quality for two power profiles when the number of shut down vehicles changes.



(a) (b) (c)

Figure 6.12: Feature matching for different vehicles.




(a) (b)

Figure 6.13: (a) Tracked points; (b) Reconstructed map.

Figs. 6.8-6.10 demonstrate the optimal received rates at different vehicles as a result of solving the proposed optimization problems. The vehicles are sorted based on their channel quality for two different power profiles. In Fig. 6.8, all the vehicles which are able to receive the base layer video are assumed in active mode. The vehicle which experiences a better channel receives the video with a higher rate. Figs. 6.9-6.10 show the vehicles with the worst channel quality are shut down (one vehicle and two vehicles in these two figures, respectively). Fig. 6.11 represents the proposed solution for the broadcast rate when variable number of vehicles are shut down. Two different power profiles are considered. By shutting down the vehicles with a low channel quality, the average broadcast rate is improved as shown in this figure. However, QoE in the result decreases since less vehicles are involved in the procedure, as explained in the solution.

Figs. 6.12(a)-(c) show the output of the feature matching and reconstruction based on the proposed algorithm. Each vehicle observes the scene partially since there are serious problems with lighting, scattering, turbidity, and clarity when taking underwater videos. In Figs. 6.12(a)-(b), the vehicles detect three objects, while from other perspective, as shown in Fig. 6.12(c), six objects are detected. Fig. 6.13 shows the final steps towards map reconstruction. Fig. 6.13(a) represents the tracked features in the shared images and Fig. 6.13(b) is the reconstructed map of the region. The map can be used as a QoE metric to evaluate how accurate the desired map should be.

## 6.5 Summary

In this chapter, a novel in-network coordination that employed Scalable Video Coding (SVC) was introduced. Large amounts of data such as videos underwater is not easy to transmit due to the error-prone underwater channel. This research investigated sharing SVC streams among AUVs in a multicast manner in which the vehicles with different capabilities/channel can be served by a single scalable stream to perform in-network map reconstruction. After sharing base layers, enhancement layers are shared with the vehicles with a better acoustic channel quality (lower quality channels are shut down). Final reconstruction is performed after a consensus on the highest-rank vehicle after it receives a high-quality video with higher QoS and QoE from other eligible nodes. Performance evaluation was presented based on experiments using video captured from the Raritan River, New Jersey and transmitted through our software-defined acoustic testbed, in addition to simulation.

# Chapter 7

# Reliable Data Transmission in Underwater Internet of Things

Achieving reliable and persistent environmental field estimation in Underwater Internet of Things (UW IoT) is a challenging problem, due to its unexplored and unpredictable nature in addition to the the limited-bandwidth and error-prone acoustic channel as well as the harsh and unpredictable underwater environment. Given the need for high-resolution spatio-temporal sensing in such environment, traditional digital sensors are not suitable due to their high cost, high power consumption, and non-biodegradable nature. Further, reliable communication techniques that avoid retransmissions are crucial for reconstructing the phenomenon in a timely manner at the fusion center such as a drone. To address the above challenges, a novel architecture is proposed consisting of a substrate of densely deployed underwater all-analog sensors that continuously transmit data to the surface digital buoys. Furthermore, a correlation-aware Hybrid Automatic Repeat Request (HARQ) technique is presented to transmit data from the surface buoys to the fusion center. Such HARQ technique leverages redundancy in the buoy data (arising from the correlation of the phenomenon at the analog nodes) to avoid retransmissions, thus saving energy and time. The performance of the proposed correlation-aware HARQ technique has been evaluated via simulations and shown to achieve the desired behavior.

## 7.1 Overview

The Underwater Internet of Things (UW IoT) [40] is a novel class of IoTs enabling various practical applications in aqueous environments such as oceanographic data collection, pollution and environmental monitoring, tsunami detection/disaster prevention, assisted navigation, and tactical surveillance [110, 114]. A new design has to be envisioned for sensors/things in UW IoT as traditional digital sensors are expensive (cannot be deployed in high density), high-power consuming (need to be put to sleep, thus losing temporal granularity), and finally pollute the environment.

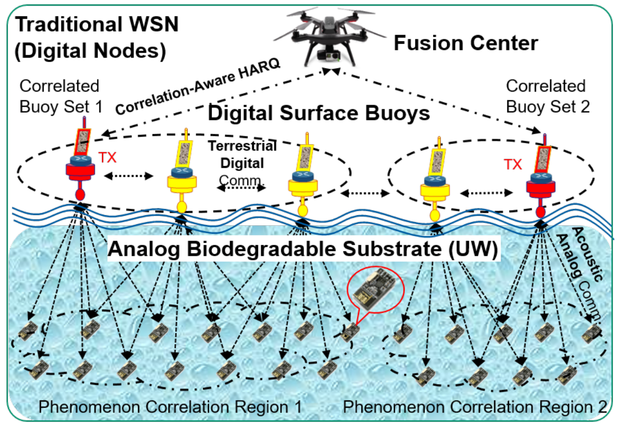



Figure 7.1: A novel sensing architecture for real-time, persistent water monitoring using analog sensors as substrate above which lies a Wireless Sensor Network (WSN) consisting of digital surface buoys communicating among themselves and occasionally to a fusion center (e.g., drone) using correlation-aware HARQ technique.

Moreover, similarity/correlation can generally be observed both in the underwater phenomenon as well as in the channel used for communication, which can be leveraged to improve efficiency.

**Motivation:** Firstly, for the UW IoTs to be a successful technology, the "things" or sensors should be able to capture high temporal and spatial variations of multiple manifestations—such as temperature, salinity, potential Hydrogen (pH)—of the phenomenon in the underwater environment. This requires high-resolution (in both time and space) sensing. Traditional digital sensors may not be the right candidates for such scenarios, as they: (i) Have high power consumption, because of which they are put to sleep based on specific duty cycles; moreover, existing sensor-encoding solutions use all-digital hardware, which demands high power and circuit complexity; as such, when the phenomenon exhibits high temporal variation, their batteries drain fast. (ii) Are expensive, making them a costly choice for high-density deployment, which is needed to track a phenomenon with high spatial variation. (iii) May pollute when deployed in water bodies as the material used in the manufacturing of such sensors is not biodegradable; currently, most electronics are typically made with nondecomposable, nonbiocompatible, and sometimes even toxic materials, leading to various ecological challenges.

Secondly, when the sensors are densely deployed, the recorded values may be spatially and/or temporally correlated. Since communication of a large amount of measured data between the nodes results in large overhead (in terms of energy, time, and bandwidth), conventional point-to-point communication techniques at the physical and Multiple Access Control (MAC) layers generally fail to provide the required functionalities for such scenarios [113]. A reliable communication technique that takes into account the spatial and temporal correlations of the phenomenon to avoid costly retransmissions and thereby save energy and time resources is greatly needed.

To address the above challenges, we envision an architecture for the UW IoT system, as shown in Fig. 7.1, where the analog nodes in the underwater biodegradable substrate transmit data continuously to the digital surface buoys in the traditional Wireless Sensor Network (WSN), which aggregate and transmit the data to the fusion center (e.g., a drone). The underwater analog biodegradable substrate consists of wirelessly-transmitting all-analog sensors with Shannon-mapping [126] capabilities, a low-complexity technique for Analog Joint Source-Channel Coding (AJSCC) [52]. The sensors employ Frequency Position Modulation and Multiplexing (FPMM) [156], which allocates a specific frequency to a specific value of a specific node to communicate with the surface buoys. Biodegradable Micro Electro-Mechanical Systems (MEMS)-based acoustic transceivers with ranges of few meters are a perfect fit to our scenario. The digital surface buoys decode the values received from the analog sensors, i.e., they perform the reverse operation of Shannon mapping. Since the data received at the surface buoys could be redundant (due to correlation in the underlying phenomenon), as shown in Fig. 7.1, the digital buoys elect some of them to be "representatives" of the other buoys. Only these representative buoys transmit data to the drone (indicated in red as TX in Fig. 7.1). Further, the representative buoys employ a novel correlation-based closed-loop Hybrid Automatic Repeat Request (HARQ) solution to transmit data to the drone. Such a technique leverages the similarity and correlation of the data to avoid costly retransmissions and thereby save energy, time, and bandwidth resources, thus enabling a timely reconstruction of the phenomenon at the fusion center.

**Contributions:** The contributions of this work are as follows. A reliable correlation-based HARQ is introduced to transmit data between the buoys and the drone that leverages the correlation of the data to avoid costly retransmissions; chaotic Direct Sequence Spread Spectrum (DS-SS) is adopted to guarantee secure buoy-drone transmissions. This saves the energy, bandwidth, and communications overhead by defining the role of a representative (REP) for a group of correlated nodes to report the data to the upper level in a self-configurable and scalable cluster-based architecture. The proposed sensor-encoding and correlation-aware HARQ techniques are validated in terms of functionality using simulations.

**Chapter Outline:** In Sect. 7.2, we go over the state of the art and similar research in the literature. In Sect. 7.3, we discuss the proposed solution including the novel correlation-aware HARQ technique for IoT scenarios. In Sect. 7.4, we present the simulation results and discuss the benefits of this solution. Finally, in Sect. 7.5, we summarize the chapter.

## 7.2 Related Work

This section positions our work with respect to state-of-the-art research in reliable UW and terrestrial communication/channel coding techniques.

To improve the accuracy and efficiency of a system that exhibits spatial and temporal correlations [67], an Error Control (EC) strategy with acknowledgment such as Hybrid Automatic Repeat Request (HARQ) [31, 109] can be exploited. Hybrid ARQ (HARQ)—as a combination of ARQ and Forward Error Correction (FEC)—reduces the retransmissions and increases the system reliability in poor channel conditions [109]. A type-I HARQ discards the erroneous received packet and repeats the same packet retransmission until the error is corrected. However, if the channel is not in good condition, e.g., when in deep fade, the predefined FEC might not be adequate and the throughput may drop again because of multiple retransmissions [133]. While more efficient than type-I, a type-II HARQ requires a larger buffer size and has a higher complexity. It adapts itself with the channel in such a way that it first transmits the packet along with the error detection bits when the channel is good. When the channel becomes worse and after detecting the erroneous packet, a NACK is sent back and—rather than retransmitting the same packet as type-I does—FEC information is transmitted to help decode the stored packet in the receiver's buffer. If the error persists, a second NACK is issued and the same FEC might be retransmitted or extra FEC might be added depending on the coding strategy. Incremental Redundancy (IR) HARQ, which shows a higher throughput efficiency in terrestrial time-varying channels, adds extra redundant information in each round of retransmission after receiving the NACK message [130]. Terrestrial standards such as High Speed Packet Access (HSPA) and Long Term Evolution (LTE) have exploited HARQ synchronously for the uplink, and asynchronously in the downlink direction. The requirements for designing a network-optimized HARQ for the fifth generation (5G) of mobile communications is discussed in [86]. Given the necessity of supporting futuristic applications such as UW IoT, we believe that a new design for HARQ that leverages the correlation in the data is essential.

## 7.3 Proposed Correlation Aware HARQ

In this section, The important constituents of the proposed architecture is explained, and the correlation-aware HARQ technique that leverages the correlation in the buoy data to avoid retransmissions between the buoys and the drone The communication technique, adopted by analog sensors, is described and followed by decoding of the data at the buoy. The proposed architecture which enables high-resolution (both in space and time) sensing of the phenomenon is shown in Fig. 7.1. Assume a dense network of underwater nodes is randomly deployed in a large scale underwater region for a time-dependent data acquisition from a phenomenon of interest which is characterized by its manifestations. Nodes are equipped with sensors and transceivers for communications with limited predefined sensing and communication range due to their low-power nature. A suggested architecture for this system is shown in Fig. 7.1, in which heterogeneous underwater nodes can measure noisy data of the physical phenomenon manifestations. A low-power and low complexity design allows sensors to be deployed in high density (so as to enable high spatial resolution) and be on continuously (i.e., enable high temporal resolution as they need to be put to sleep to save power). The buoys transmit data (which was received from the analog sensors at the lower layer) to the fusion center such as a naviator/drone. Since the data received at the buoys could be correlated, we elect some of them to be a representative (REP) of the other buoys. We propose a novel correlation-based closed-loop HARQ solution that leverages the similarity and correlation of the data to avoid costly (in terms of energy and time) retransmissions. Here only these REP (instead of all) buoys communicate with the fusion center, such as a drone or a base station, to help reconstruct the phenomenon. The communications between the buoys and the fusion center is delay tolerant since the latter may not be always available.

**Spatial and Temporal Correlations:** Let $n_i, \{i = 1, ..., N\}$ denote $N$ subregions with the location index $L_i \in \mathbb{L} \subset \mathbb{R}^3$, where $\mathbb{L}$ denotes the 3D environment's space. The data is shown by matrix $\mathcal{P} = [P_1, ..., P_N]$. The $i\text{th}$ column of $\mathcal{P}$, corresponding to subregion $L_i$, consists of data from the $K$ manifestations, i.e., $[\mathcal{P}]_i = P_i = [\psi_i^{(1)}(t), ..., \psi_i^{(k)}(t), ..., \psi_i^{(K)}(t)]$.

*Definition* 3. For the subset of *interconnected* subregions $\mathcal{L}_i,\ \{i = 1, ..., \mathcal{N}\} \subset \mathbb{L}$, let the spatial

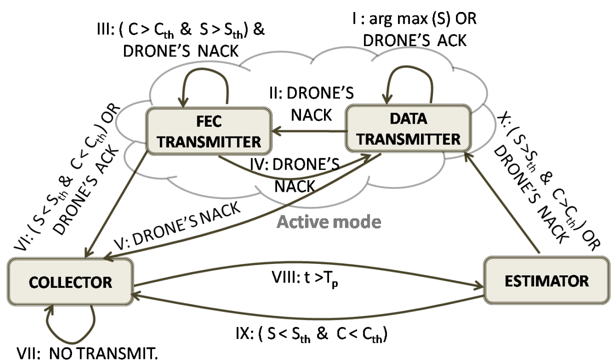



Figure 7.2: State transition diagram for a buoy in a correlated set. In the active mode $(I)-(IV)$, HARQ is initiated with the highest similarity. If a NACK is issued, the FEC transmitter is chosen via $(II)$ or it goes to $(V)$. If FEC fails, the data is dropped and the next packet is transmitted via $(IV)$. If the correlation drops below a threshold, it demotes to a collector via $(VI)$, until the estimator is activated via $(VIII)$.

correlation between two sampled values $\psi_i^{(k)}(t)$ and $\psi_j^{(k)}(t)$ with means $\underline{\psi}_i^{(k)}$ and $\underline{\psi}_j^{(k)}$ and standard deviations $\sigma_i^{(k)}$ and $\sigma_j^{(k)}$, at time $t$, be as

$$C_{i,j}^{(k)} = \frac{\mathbb{E}\big[(\psi_i^{(k)} - \underline{\psi}_i^{(k)})(\psi_j^{(k)} - \underline{\psi}_j^{(k)})\big]}{\sigma_i^{(k)}\sigma_j^{(k)}}, \tag{7.1}$$

where $\mathbb{E}[.]$ represents the expectation value and the time notation is dropped for simplicity.

*Definition* 4. Let the spatial correlation be a function $\mathcal{F}(.)$ of the distance between two locations $\mathcal{L}_i$ and $\mathcal{L}_j$ as in $\mathbb{E}\big[(\psi_i^{(k)} - \underline{\psi}_i^{(k)})(\psi_j^{(k)} - \underline{\psi}_j^{(k)})\big] = \mathcal{F}(\mathcal{L}_i - \mathcal{L}_j)$. The subregions are correlated when $C_{i,j}^{(k)} > C_{th}^{(k)}$, where $C_{th}^{(k)}$ is the spatial correlation threshold. The correlation matrix for the manifestation $k$ is symmetric and is defined by $[\mathbb{C}^{(k)}]_{i,j} = C_{i,j}^{(k)}$.

*Definition* 5. The measured data of a manifestation $k$ in subregions $i$ and $j$ are said to be similar if the normalized difference of their means is less than a threshold $(1 - S_{th}^{(k)})$, where $S_{th}^{(k)}$ is the similarity threshold, i.e., $1 - \mid \underline{\psi}_i^{(k)} - \underline{\psi}_j^{(k)} \mid / \underline{\psi}_j^{(k)} > S_{th}^{(k)}$.

*Definition* 6. Temporal correlation is defined as the degree of correlation for which two consecutive sampled data at the buoy are correlated. In other words, the amount of correlation does not change within the time $\tau < T_{\mathcal{P}}^k$, where $T_{\mathcal{P}}^k$ is the temporal correlation of the $k$-th manifestation. It can be concluded that $C_{i,j}^{(k)}(t+\tau) = C_{i,j}^{(k)}(t)$ and so the expectation of the random variable $\psi_i$ is constant.

Each buoy at a time has one of the roles of a *data transmitter/FEC transmitter* (active mode), a *collector*, or an *estimator*. The transition among the roles is decided by the phenomenon's correlation/similarity and also the feedback command given by the drone as shown in the state diagram,

Fig. 7.2. Data transmitter role is decided based on two factors: the highest similarity, $\arg\max(\mathrm{S_{i,j}})$, and spatial correlation greater than a threshold $C_{i,j} > C_{th}$; the control command from the drone. If the data transmission is not successful and the drone issues a NACK, the buoy's state will change from data transmitter to FEC transmitter as shown in $(II)$. Drone could turn the buoy's role to a collector if it decides to not receive any data from the same buoy, as shown in $(V)$ and $(VI)$. An estimator is a collector that has started to evaluate whether its newly received data from analog nodes is still correlated and similar to that of its data transmitter buoy or not. It changes its state to become a data transmitter if there is a significant variation in the spatial distribution of its data over time as shown in $(X)$. The two main communications aspects, called *intra-cluster* communications (among the buoys) and *inter-cluster* (between the data/FEC transmitters and the drone), are discussed in details in the following sections.

**Intra-cluster Chaotic-based Spread Spectrum:** Code Division Multiple Access (CDMA), as both physical-layer and multiple-access techniques, can be beneficial to handle the destructive effect of frequency-selective fading, as well as the simultaneous reception from multiple transmitting devices by using an appropriate spreading, especially in an IoT network. Furthermore, given the security dedicated in chaotic CDMA's nature, the jamming attacks as a critical malicious threat can be satisfied [114]. Although Pseudo-Noise (PN) sequences have been extensively employed in DS-SS, considering their limitation in the number of sequences and their cross-correlation properties, [50] proposed using chaotic sequences through an uncomplicated deterministic dynamic map such as in Logsitic map [13]. Chaotic systems can produce an infinite set of uncorrelated sequences and can provide secure communication. Similar to the PN sequences, they look like noise but unlike PN, chaotic codes are not binary and are different for every bit of transmitted data which makes it much harder for an eavesdropper to regenerate the sequence. The use of a distributed CDMA scheme, supporting an adaptive EC strategy, can therefore increase the channel reuse and reduce packet retransmissions in scenarios with a large number of buoys, thus increasing network reliability while decreasing the energy consumption. While conventional EC strategies consider only point-to-point data protection, more efficient techniques are required in such applications in which $(i)$ the buoys have some sort of similarities and correlations in time or space in the measured data; $(ii)$ the importance of phenomenon monitoring is higher than protecting of each buoy alone; $(iii)$ the communications overhead is huge—since multiple buoys communicate with each other and with the

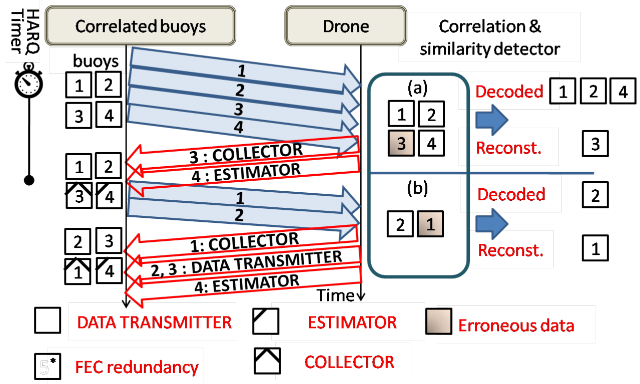



Figure 7.3: Proposed protocol for data transmission to the drone via a correlated set of buoys—shown with numbers $1, 2, ...$. Spatial and temporal correlations are considered in decoding. Conservative and borderline approaches are compared in (a) and (b).

drone—and so a scheduling is required. To support reconfigurable and flexible IoT applications in which the number of simultaneous transmissions is not known in advance, chaotic sequences can be a good candidate to support any number of transmitters. Chaotic sequences not only provide the security in the channel, but also possess a considerable robustness against the multipath effect due to their good auto-and cross-correlation properties.

**Inter-cluster Correlation-based HARQ:** The transmission process is conducted in an environment in which the phenomenon's manifestations change from time to time as represented by temporal and spatial correlations of the phenomenon. Therefore, an appropriate multi-point EC strategy should $(i)$ take advantage of the defined correlations of data/FEC transmitter buoys versus collector and estimator buoys for efficient decoding by increasing the probability of successful decoding in each round; $(ii)$ reduce the probability of retransmission and also the communications overhead with a correlation-based coding, while the temporal correlation of the phenomenon is long enough; $(iii)$ pause the decoding process and go for the retransmission, while the temporal correlation of the phenomenon can not be ignored.

As portrayed in Fig. 7.3, buoys are shown on the left with numbers $1, 2, ...$ while the drone stands on the right side of the diagram. Each data transmitter buoy broadcasts its packets through independent channels to the other buoys and to the drone (within a single-hop distance). As an example, a conservative approach is taken in case (a), since no a-priori information is available beforehand. Assume the decoder successfully decodes all the packets, except 3. Here—instead of issuing a NACK as in the conventional HARQ—buoy 3's data can be reconstructed using the correlation among the other decoded data. The cost is the extra error in estimating the corrupted data; however, it can be ignored if buoys are highly correlated. In case (b), the borderline approach is considered to avoid the excessive redundancy (communications overhead) of case (a) and to reduce the interference by using only 1 and 2 as the data transmitter while 3 and 4 are changed to collector and estimator, respectively. Assume the data from 1 is erroneous while 2 is still decodable. Again, 1 is reconstructed without any extra FEC with the cost of more reconstruction error.

Fig. 7.4 presents two cases (c) and (d) in which both the received data—such as those of 2 and 3—are erroneous. The corrupted data is recoverable, if the temporal correlation of the phenomenon is valid at that time instant and available at the drone, as shown in case (c). If this condition does not hold—which is the scenario in case (d)—since the overall information is not enough for making the decision, 3 is notified to send the FEC information by switching to the FEC transmitter role. The extra information in $2^*$ and $3^*$, combined with the data from 1 and 4, help the decoder to have both the conventional HARQ and the spatial correlation properties. Therefore, the decision is made on 2 and 3 based on the tolerable amount of reconstruction error in 3. We define three different notions of time in our coding scheme, as shown in Fig. 7.4. *HARQ timer*, $T_{HARQ}$, is used to show the transmissions/FEC transmission time (for erroneous packets) for every round of communications. *Correl. timer*, $T_{Corr}$, represents the time within which data transmission/decoding can be done consecutively based on the temporal correlation of the phenomenon. This time is less than or equal to $T_{\mathcal{P}}$. If the data is not acknowledged in $T_{HARQ}$, but $T_{Corr} >> T_{HARQ}$, then the data can be recovered without retransmission as explained in case (c). *Struc. timer*, $T_{Stru}$, is the time in which the structure is almost constant. Therefore, outside of this time period, the correlated sets should be reconsidered, because of the analog node movement. In this case, the extra FEC or using the correlation might not be helpful; therefore, the retransmission of the original/new data using new set of buoys will be the solution for time $t > T_{Stru}$.

**Data Reconstruction:** Based on the renewal-reward theorem [159], let the HARQ long-term throughput for buoy $i$ be $\eta_i = \mathbb{E}[\tilde{X}_i]/\mathbb{E}[\tilde{T}_i]$, where $\mathbb{E}[\tilde{X}_i]$ and $\mathbb{E}[\tilde{T}_i]$ represent the number of decoded information nats and the number of attempts for channel use during a packet transmission period for buoy $i$, respectively. $\mathbb{E}[\tilde{X}_i]$ is defined as $\mathbb{E}[\tilde{X}_i] = X_i(1 - \overline{\mathrm{Pr}_{o(i)}})$, in which $X_i$ is the number of information nats for a packet and $\overline{\mathrm{Pr}_{o(i)}}$ denotes the probability that the data is not decodeable during a packet transmission period. We consider an alternative packet which comes from a correlated buoy $j$ with the correlation coefficient $C_{i,j}$, and $\overline{\mathrm{Pr}_{o(j)}} < \overline{\mathrm{Pr}_{o(i)}}$, i.e., the communication channel

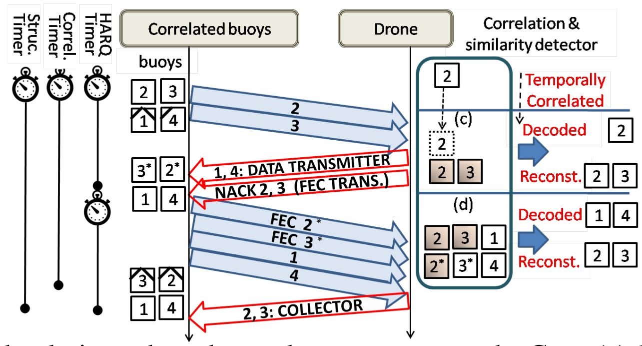
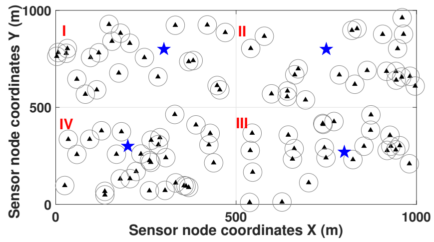


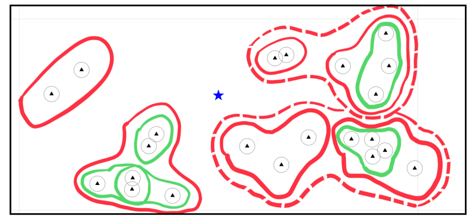

Figure 7.4: Proposed solution when the packets are corrupted. Case (c) leverages the temporal correlation while case (d) uses spatial correlation via combined data and FEC.

that $j$ experiences is better than $i$'s channel. We conclude that $R_j < R_i$, where $R$ stands for the maximum number of transmission rounds in the HARQ. The probability of decoding in round $R_i$ for buoy $i$ given that the data has not been decoded in the previous $R_i - 1$ rounds is equivalent to $\Pr\left(\mathrm{NACK_1, ..., NACK_{R_i-1}, ACK_{R_i}}\right)$. Assume at every round $r$, for all the acknowledged buoys in the same correlated set, there exists at least a buoy $j$, where $R_j < R_i$. $j$ is chosen as $\arg\max \mathrm{C_{i,j}}$. Therefore, $R_i' = \min[\mathrm{R_i, R_j}]$ and so it is a function of the correlation between $i$ and $j$. Then, $\mathbb{E}[\tilde{T}_i]$ can be defined as $\mathbb{E}[\tilde{T}_i] = \sum_{r=1}^{R_i'} X_{i\,r} \Pr\left(\mathrm{NACK_1, ..., NACK_{r-1}, ACK_r}\right)$, where $X_{i\,r}$ stands for the part of HARQ data which is transmitted at round $r$ from buoy $i$. $\mathbb{E}[\tilde{T}_i]$ can be simplified in two terms as $\mathbb{E}[\tilde{T}_i] = \sum_{r=1}^{R_i'-1} X_{i\,r} \Pr\left(\mathrm{NACK_1, ..., NACK_{r-1}}\right) + \mathbf{X}_{\mathrm{j}\,\mathrm{R_i'}} \Pr\left(\mathrm{NACK_1, ..., ACK_{R_i'}}\right)$ [75]. Here, the total data transmission from buoy $j$ up to the end of round $R_i'$ is shown with $\mathbf{X}_{j\,R_i'} = \sum_{r=1}^{R_i'} X_{j\,r}$. The content of $\mathbf{X}_j$ is correlated with $X_i$ within the temporal correlation $T_{\mathcal{P}}$. Since their data are not exactly equal, this leads to an uncertainty related to the amount of $C_{i,j}$ with the benefit of avoiding $R_i - R_i'$ rounds of HARQ retransmissions and having a better long-term throughput.

When the data is transmitted from buoy $i$ to the drone, it is polluted with the noise and affected by the fading and the interference. We assume the received signal $\hat{X}_i$ at the drone follows an Autoregressive time series model of order one, i.e., AR(1). That is at time instant $t_m$, $\hat{X}_i[t_m] = \mu + \rho(\hat{X}_i[t_m - 1] - \mu) + \epsilon_i$, where $\mu$ is the mean of $\hat{X}_i$, $\epsilon_i$ is a zero mean unit variance independent Gaussian process, and $\rho$ is the autoregression parameter which is related to the temporal correlation of the phenomenon. Assume we take $\mu = 0$ for the sake of simplicity. For $-1 < \rho < 1$ and for any discrete time lag $t_0$, $\mathrm{Cov}(\hat{\mathrm{X}}_\mathrm{i}[\mathrm{t_m + t_0}], \hat{\mathrm{X}}_\mathrm{i}[\mathrm{t_m}]) = \rho^{\mathrm{t_0}}/(1-\rho^2)$ [127]. Here $\hat{X}_i[t_m], \hat{X}_i[t_m - 1], ...$ form a Markov process given that channel coherence time ($T_c$) is greater than the time window between two successive received samples $t_m$ and $t_{m-1}$. Note that in addition to the noisy measured data by analog nodes, other layers of error are added to the data when it is received




(a)



(b)

Figure 7.5: (a) Random distribution of nodes (▲). Drone (★) passes by each region (shown by Roman numbers) to fuse the data; (b) Magnified view of region $II$ with different degrees of correlation. Different colors show how the buoys are correlated.

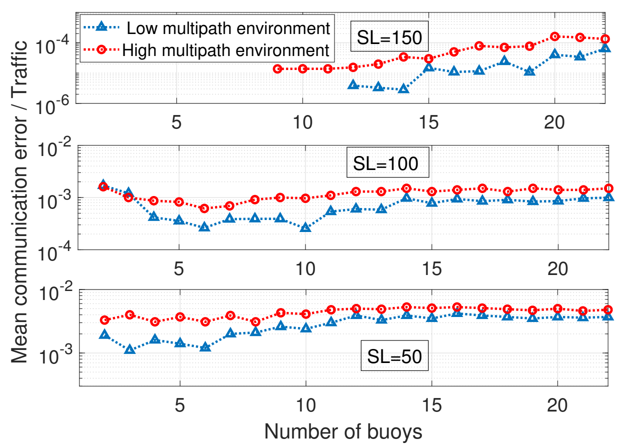



Figure 7.6: Mean communication error per traffic for buoys inside one region in low and high multipath while three different spreadings for the chaotic code are considered.

by the drone. Therefore, the difference between the measured value of the transmitter buoy and other buoys in the same correlated set is not zero. The transmitted data from $i$th transmitter buoy can be written as $\mathcal{P}+\mathcal{E}_{s,i}+\mathcal{E}_{c,i}$, where $\mathcal{E}_{s,i}$ and $\mathcal{E}_{c,i}$ represent the sensing error and communications error related to the transmitter buoy $i$, respectively. This signal is transmitted through channel $h_i$ and is received at the drone as $\hat{X}_i$. The phenomenon can be estimated as $\widehat{\mathcal{P}} = F\big(\sum_{i=1}^{R} \hat{X}_i + \mathcal{N}\big)$, where $F(.)$ represents a data extraction function, and $\mathcal{N}$ is the background AWGN noise present in the environment. Therefore, the reconstruction error, $\mathcal{E}_{rec}$, is defined as $\mathcal{E}_{rec} = \left|\mathcal{P} - \widehat{\mathcal{P}}\right|$. Considering the number of buoys transmitting in each round of transmission, the goal is to increase the performance by reducing the communications overhead and destructive effect of other buoys, i.e., to minimize $\left\|\widehat{\mathcal{P}} - \mathcal{P}\right\|^2$, using appropriate channel coding.

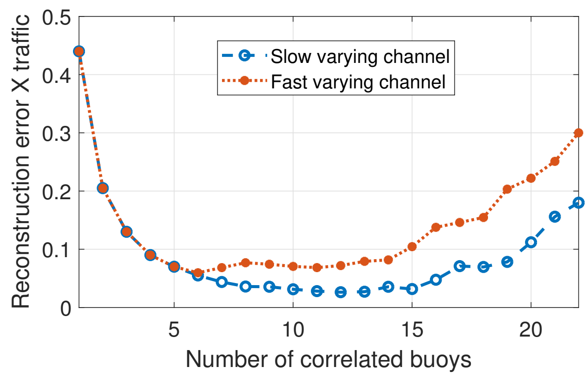



Figure 7.7: Reconstruction error × traffic for buoys inside a region while considering two different fading channels.

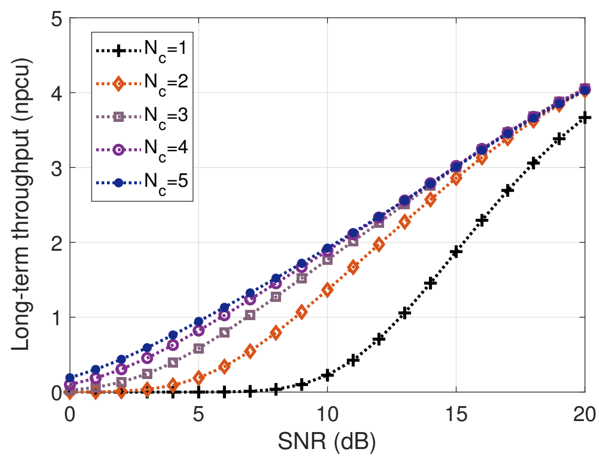



Figure 7.8: Long-term throughput of correlated HARQ, in nats-per-channel-use (npcu), for different number of data transmitter buoys.

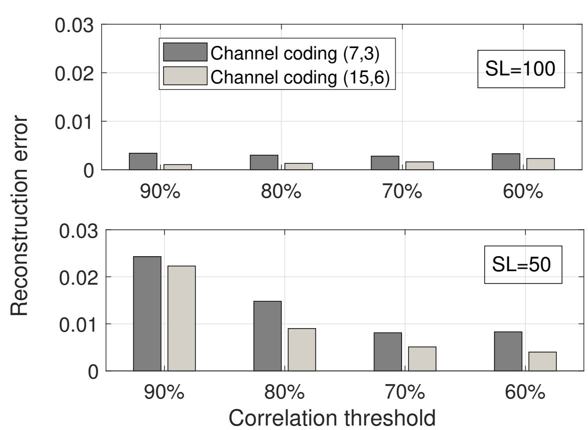



Figure 7.9: Reconstruction error for different correlation thresholds and channel coding.

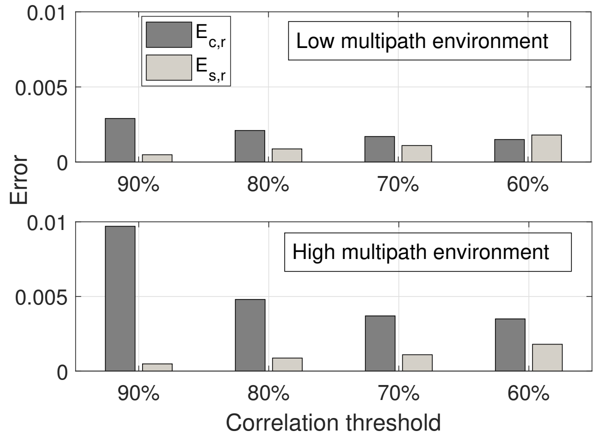



Figure 7.10: Sensing error and communication error for different correlation thresholds, in low and high multipath environments for chaotic code ($SL = 100$).

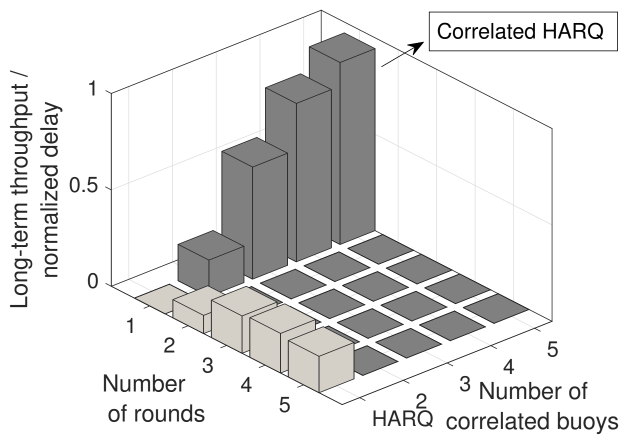



Figure 7.11: Long-term throughput per normalized delay for correlated HARQ versus number of rounds compared to the conventional HARQ.

## 7.4 Performance Evaluation

**Buoy and Fusion Center:** We consider an area of study with $100$ randomly deployed buoys for the simulation, as represented in Fig. 7.5(a). We divide this area into four regions of interest; each of them shows a fusion center, the drone, passes by the region to communicate with the buoys and to fuse the collected data. The buoys are grouped based on their similarity and correlation. A magnified version of region II is displayed in Fig. 7.5(b). This figure also shows the various possible sets of correlated buoys based on different correlation thresholds. Data transmission is performed using Binary Phase Shift Keying (BPSK) modulation and Reed-Solomon coding $(7, 3)$ or $(15, 6)$

and CDMA spreading sequence with logistic map. As a general rule, spreading sequences should have minimal cross-correlation to minimize the interference between the buoys and a delta-function shape autocorrelation to maximize the detection accuracy of the desired buoy.

Fig. 7.6 investigates the performance of our method in terms of the mean communication error normalized by total traffic for different number of transmitter buoys. It compares the effect of changing the spreading length and presents the results when communication takes place in channels with different multipath effects. For $SL = 150$ when the number of buoys are small, the error is less than $10^{-6}$ (they are not depicted in the figure). The curves show communication error increases when the number of buoys grows (as the result of multi-user interference); however codes with larger $SL$ are more successful in combating the multi-user interference. In Fig. 7.7, reconstruction error is presented for total traffic of the transmitter buoys. It also compares the effect of transmitting under slow and fast fading channels. Fig. 7.8 calculates the long-term throughput of the proposed HARQ for different number of correlated buoys. The correlation threshold is a determinant factor in evaluating the performance of the system. In Figs. 7.9- 7.10, we investigate the effect of different correlation thresholds. In Fig. 7.9, reconstruction error is represented with HARQ channel coding $(7, 3)$ and $(15, 6)$. Fig. 7.10 presents two components of reconstruction error (sensing and communication errors). Fig. 7.11 shows the long-term throughput per normalized delay for correlated HARQ compared to the conventional HARQ.

## 7.5 Summary

In this chapter, a novel architecture for UW IoT was proposed consisting of a substrate deployed underwater in high-density and transmitting data continuously to digital surface buoys. A novel multi-point correlation-aware hybrid ARQ technique was designed to transfer data from digital surface buoys to the fusion center that leverages the correlation in the data to avoid costly retransmissions and thereby enable timely reconstruction of the phenomenon. The proposed techniques have been evaluated via MATLAB simulations and shown to provide the desired performance. The dynamic change in the structure and its effect on the proposed HARQ solution were also presented and discussed in the protocol.

# Chapter 8

# Conclusion and Future Directions

This chapter summarizes the main contributions of this dissertation and briefly explains the next steps and future research directions that can be pursued as the outcome of this dissertation.

There are novel underwater applications that should be able to capture real-time, or near-real-time, multimedia data and store/process/compress it while it is being transmitted. The terrestrial and conventional communication techniques fail to provide the required robustness and reliability for these futuristic underwater applications due to the characteristics of the underwater environment. Therefore, this dissertation discussed and introduced novel solutions in different layers (i.e., physical-, MAC-, and link-layer, and also cross-layer designs) to support robust, reliable, and high-data rate underwater multimedia transmission. In particular, this dissertation focused more on the following aspects:

(1) An Acoustic Vector Sensor (AVS)-based solution, called Signal-Space-Frequency Beamforming (SSFB), was designed to transmit underwater videos at high data rates using acoustic waves for short/medium distances. Data was modulated and transmitted via NC-OFDM, and detected via beam's angle of arrival at the receiver side. The receiver (buoy) was equipped with AVS—hydrophones that measures acoustic particle velocity in addition to scalar pressure—in a multiple-antenna-array structure, while the transmitter (vehicle) was equipped with a circular array of transducers. This method enables video transmission in applications such as coastal and tactical surveillance, which require multimedia acquisition and classification.

(2) A novel probabilistic-based Medium Access Control (MAC) solution was proposed based on Space Division Multiple Access (SDMA) to share reliably the space among the steered vehicles so as to reduce the acoustic interference in sparse networks. This method leverages the inherent position uncertainty of the moving vehicles. A two stage estimation technique was presented based on interval estimation and Unscented Kalman Filtering (UKF) to estimate the position of the vehicle

and focus the beam, respectively. An optimization problem was solved to minimize the statistical interference. The method was extended to the scenario of non-separable vehicles via a hybrid T-SDMA solution. Since the proposed approach could handle the interference while the vehicles were moving, it could achieve a high data rate and reliability.

(3) The reliability and the quality of multimedia delivery were improved by proposing a robust closed-loop hybrid Automatic Repeat Request (ARQ) technique that was specifically designed for the harsh underwater environment. A collaborative strategy was introduced for a CDMA-based underwater hybrid ARQ to increase the overall throughput of the network by adjusting physical- and link-layer parameters and compensating for the poor underwater acoustic communication links. The node with a low-quality communication link piggybacks on its neighboring nodes' transmissions when protecting its data against errors. The solution achieves higher network reliability and throughput by allocating an appropriate share of system resources to different nodes and lower latency caused by the conventional HARQ retransmission strategy.

(4) A framework was introduced based on Scalable Video Coding (SVC) H.264/MPEG-4 AVC compression standard in which the coded underwater videos leveraged the multiplexing-diversity tradeoff in an MIMO-based Software-Defined Acoustic Modem (SDAM) structure to balance the transmission data rate and reliability. Multiple optimization problems were proposed that provided the scalability in the video bitstream processing to adapt to the preference of end-users as well as to the varying characteristics of the network. The video quality level was determined by the best communication link while the transmission scheme was decided based on the worst communication link, which guarantees an appropriate data rate and quality for each user.

(5) A framework and a protocol for underwater imagery analysis were proposed based on partial information, collected by various vehicles around the scene, using Scalable Video Coded (SVC) multicasting and in-network coordination. This contribution has many applications in underwater imagery when more than one vehicle, due to the limited field of view and the visual depth of camera in the water, are needed to merge the video from different angles so as to reconstruct the map of region of interest. An optimized solution was presented to provide the maximum possible Quality of Service (QoS) via a proposed scalable multicasting strategy, while achieving the maximum Quality of Experience (QoE) for the scene reconstruction. The reconstructed map can be used for in-network decision among the vehicles or can be transmitted to the buoy for further considerations.

(6) A novel multi-point correlation-aware hybrid ARQ technique was designed to transfer data between digital surface buoys and the fusion center in the futuristic applications such as Underwater Internet of Things (UW IoT) with high-density deployed nodes in the shallow body of waters such as in the lakes and rivers. Given the need for high-resolution spatio-temporal sensing in such environment, novel reliable communication techniques are required that avoid retransmissions while reconstructing the phenomenon in a timely manner at the fusion center. The proposed solution leveraged the redundancy in the data arising from the spatial and temporal correlations of the measured phenomenon, thus it saved the energy and time.

The key open directions for future research can be highlighted as follows.

**Low-latency and High Quality Underwater Video Encoding and Transmission for Software Defined Acoustic Transceivers:** In the future, the proposed solutions in this dissertation can be extended to other efficient encoding schemes such as High Efficiency Video Coding (HEVC) and H.265. The goal could be to design a new encoder that maximizes the quality of video and optimizes the compression rate, particularly for underwater videos. This goal should be pursued given the limited bandwidth constraints in underwater from one hand, and latency in video encoding/transmission, due to the inherent propagation delay in the underwater channel from the other hand. Different tuning solutions can be generated to factor in the lighting, back scattering, and the turbidity of the water and their effect on this video codec.

**Communications and Coordination Among Multiple Heterogeneous Underwater Vehicles:** Due to the low bandwidth of the underwater acoustic channel—which leads to low data rate transmissions—and the time overhead imposed by both the channel propagation delay and the processing delay in the acoustic modems, novel strategies, algorithms and protocols are required for coordination and communication between multiple heterogeneous vehicles from video processing to decision making and to communications. Many robotics challenges need to be confronted and new path planning, obstacle avoidance, and navigation algorithms should be generated towards developing futuristic autonomous underwater vehicles.

# References


[1] BlueRobotics. http://www.bluerobotics.com/. Last time accessed: June 2020.

[2] Laboratory for image and video engineering - the university of Texas at Austin. http://live.ece.utexas.edu/research/Quality/. Last time accessed: April 2020.

[3] RESON TC4013 Hydrophone Product Information . http://www.teledynemarine.com/reson-tc4013. Last time accessed: May 2020.

[4] Stream Monitoring Program, Raritan Headwaters. https://www.raritanheadwaters.org/protect/stream-monitoring-program. Last time accessed: October 2019.

[5] USRP X Series. https://www.ettus.com/. Last time accessed: April 2020.

[6] M. Agiwal, A. Roy, and N. Saxena. Next generation 5G wireless networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 18(3):1617–1655, 2016.

[7] I. Ahamed and M. Vijay. Comparison of different diversity techniques in mimo antennas. In *Proceedings of Conference on Communication and Electronics Systems (ICCES)*, pages 47–50, October 2017.

[8] R. Ahmed and M. Stojanovic. Grouped packet coding: A method for reliable communication over fading channels with long delays. *IEEE Journal of Oceanic Engineering*, (99):1–11, 2018.

[9] I. F Akyildiz, D. Pompili, and T. Melodia. Challenges for efficient communication in underwater acoustic sensor networks. *ACM Sigbed Review*, 1(2):3–8, 2004.

[10] I. F. Akyildiz, D. Pompili, and T. Melodia. Underwater acoustic sensor networks: research challenges. *Ad Hoc Networks*, 3(3):257–279, March 2005.

[11] J. Alves, J. Potter, P. Guerrini, G. Zappa, and K. Le Page. The LOON in 2014: Test bed description. In *Proceedings of Underwater Communications and Networking (UComms)*, pages 1–4. IEEE, 2014.

[12] N. Amdouni, O. Ben Rhouma, and A. Bouallegue. Alamouti scheme's generalization. In *Proceedings of Mediterranean Microwave Symposium (MMS)*, pages 169–172, 2011.

[13] S. Azou, G. Burel, and C. Pistre. A chaotic direct-sequence spread-spectrum system for underwater communication. In *Proceedings of Oceans Conference*, volume 4, pages 2409–2415. MTS/IEEE, 2002.

[14] F. Babich, M. Comisso, and R. Corrado. Fast distortion estimation based on structural similarity for h.264/svc encoded videos. In *Proceedings of Vehicular Technology Conference (VTC Spring)*, pages 1–5. IEEE, 2015.

[15] S. V. Bana and P. Varaiya. Space division multiple access (SDMA) for robust ad-hoc vehicle communication networks. In *Proceedings of IEEE Intelligent Transportation Systems*, pages 962–967. IEEE, 2001.

[16] E. Basar. Index modulation techniques for 5g wireless networks. *IEEE Communications Magazine*, 54(7):168–175, 2016.

[17] T. Basten, H. De Bree, W. Druyvesteyn, and J. Wind. *Multiple Incoherent Sound Source Localization Using a Single Vector Sensor*. Auburn, AL: International Institute of Acoustics and Vibration, 2009.

[18] B. Binnerts, I. Mulders, K. Blom, M. Colin, and H. Dol. Development and demonstration of a live data streaming capability using an underwater acoustic communication link. In *Proceedings of MTS/IEEE OCEANS*, pages 1–7, May 2018.

[19] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[20] L. M. Brekhovskikh and I. P. Lysanov. *Fundamentals of Ocean Acoustics*. Springer Science & Business Media, 2003.

[21] D.S. Broomhead, J.P. Huke, and M.R. Muldoon. Codes for spread spectrum applications generated using chaotic dynamical systems. *Dynamics and stability of systems*, 14(1):95–105, 1999.

[22] J. L. Butler and C. H. Sherman. *Transducers and arrays for underwater sound*. Springer, 2016.

[23] F. Campagnaro, R. Francescon, D. Tronchin, and M. Zorzi. On the feasibility of video streaming through underwater acoustic links. In *2018 Fourth Underwater Communications and Networking Conference (UComms)*, pages 1–5, Aug 2018.

[24] P. Casari, M. Rossi, and M. Zorzi. Towards optimal broadcasting policies for HARQ based on fountain codes in underwater networks. In *Proceedings of the Conference on Wireless on Demand Network Systems and Services (WONS)*, pages 11–19. IEEE, 2008.

[25] G. Casella and R. L. Berger. *Statistical inference*, volume 70. Duxbury Press Belmont, CA, 2002.

[26] D. Caute, O. Axelsson, and K. Kenny. Multi-function broadband phased-array software defined sonar system and method, 2018. US Patent App. 15/713,411.

[27] B. Chen, P. C. Hickey, and D. Pompili. Trajectory-aware communication solution for underwater gliders using WHOI micro-modems. In *Proceedings of IEEE Conference on Sensor Mesh and Ad Hoc Communications and Networks (SECON)*, pages 511–519, 2010.

[28] Y. Chen, Z. Wang, L. Wan, H. Zhou, S. Zhou, and X. Xu. OFDM-modulated dynamic coded cooperation in underwater acoustic channels. *IEEE Journal of Oceanic Engineering*, 40(1):159–168, 2015.

[29] X. Cheng, L. Yang, and X. Cheng. *Cooperative OFDM underwater acoustic communications*. Springer, 2016.

[30] J. H. Chong, C. K. Ng, N. K. Noordin, and B. M. Ali. A low computational complexity V-BLAST/STBC detection mechanism in MIMO system. *Human-centric Computing and Information Sciences*, 4(1):2, 2014.

[31] D. Costello and S. Lin. *Error control coding: Fundamentals and Applications*. Prentice Hall, second edition, 2004.

[32] R. Dai and I. F. Akyildiz. A spatial correlation model for visual information in wireless multimedia sensor networks. *IEEE Transactions on Multimedia*, 11(6):1148–1159, 2009.

[33] H. E. De Bree, E. Tijs, and T. Akal. The Hydroflown: MEMS-based Underwater Acoustical Particle Velocoty Sensor, Results of lab tests and sea trials. In *Proceedings of the European Conference on Underwater Acoustics (ECUA)*, volume 2, pages 886–889, 2010.

[34] E. Demirors, J. Shi, R. Guida, and T. Melodia. SEANet G2: Toward a High-data-rate Software-defined Underwater Acoustic Networking Platform. In *Proceedings of the ACM International Conference on Underwater Networks & Systems*, page 12, 2016.

[35] R. Diamant, F. Campagnaro, M. D. F. De Grazia, P. Casari, A. Testolin, V. S. Calzado, and M. Zorzi. On the relationship between the underwater acoustic and optical channels. *IEEE Transactions on Wireless Communications*, 16:8037–8051, 2017.

[36] R. Diamant and L. Lampe. Adaptive error-correction coding scheme for underwater acoustic communication networks. *IEEE Journal of Oceanic Engineering*, 40(1):104–114, 2015.

[37] S. Dimitropoulos. Going fishing for ghost gear: Scientists are developing tracking tags to help catch lost fishing gear. `https://www.hakaimagazine.com/news/going-fishing-for-ghost-gear/`. Last time accessed: August 2019.

[38] W. Dinkelbach. On nonlinear fractional programming. *Management science*, 13(7):492–498, 1967.

[39] H. S. Dol, P. Casari, T. Van Der Zwan, and R. Otnes. Software-defined underwater acoustic modems: Historical review and the nilus approach. *IEEE Journal of Oceanic Engineering*, 42(3):722–737, 2017.

[40] M. C. Domingo. An overview of the internet of underwater things. *Journal of Network and Computer Applications*, 35(6):1879–1890, nov 2012.

[41] J. Edward. Sustainable raritan river initiative, Rutgers Edward J. Bloustein school of planning and public policy. `http://eac.rutgers.edu/srri`. Last time accessed: September 2018.

[42] M. Erol, L. F. M. Vieira, and M. Gerla. AUV-aided localization for underwater sensor networks. In *Proceedings of IEEE International Conference on Wireless Algorithms, Systems and Applications (WASA)*, pages 44–54, 2007.

[43] Y. Pourmohammadi Fallah, H. Mansour, S. Khan, P. Nasiopoulos, and H. M. Alnuweiri. A link adaptation scheme for efficient transmission of h.264 scalable video over multirate wlans. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(7):875–887, July 2008.

[44] N. Farr, A. Bowen, J. Ware, C. Pontbriand, and M. Tivey. An Integrated, Underwater Optical/Acoustic Communications System. In *Proceedings of IEEE/MTS OCEANS Conference*, pages 1–6, 2010.

[45] T. Fujihashi, S. Saruwatari, and T. Watanabe. Multi-view video transmission over underwater acoustic path. *IEEE Transactions on Multimedia*, 2018.

[46] B. Gerling. Raritan river impaired by numerous toxins, 2014, May 21.

[47] J. Gordon, D. Gillespie, J. Potter, A. Frantzis, M. P Simmonds, R. Swift, and D. Thompson. A review of the effects of seismic surveys on marine mammals. *Marine Technology Society Journal*, 37(4):16–34, 2003.

[48] S. Han, R. Chen, Y. Noh, and M. Gerla. Real-time video streaming from mobile underwater sensors. In *Proceedings of the ACM International Conference on Underwater Networks & Systems*, pages 21–28, 2014.

[49] J. Hayya, D. Armstrong, and N. Gressis. A note on the ratio of two normally distributed variables. *Management Science*, 21(11):1338–1341, 1975.

[50] G. Heidari-Bateni and C. D. McGillem. A chaotic direct-sequence spread-spectrum communication system. *IEEE Transactions on communications*, 42(234):1524–1527, 1994.

[51] J. Heidemann, M. Stojanovic, and M. Zorzi. Underwater sensor networks: Applications, advances, and challenges. *Royal Society*, 370(1958):158–175, May 2012.

[52] F. Hekland, G.E. Oien, and T.A. Ramstad. Using 2:1 Shannon mapping for joint source-channel coding. In *Proceedings of Data Compression Conference (DCC)*, pages 223–232, 2005.

[53] R. P. Hodges. *Underwater Acoustics: Analysis, Design, and Performance of Sonar*. John Wiley & Sons, 2011.

[54] W.S. Hodgkiss, HC Song, M Badiey, A Song, and M Siderius. Kauai acomms muri 2008 (kam08) experiment. Technical report, 2008.

[55] K. Huang, R. W. Heath Jr, and J. G. Andrews. Space division multiple access with a sum feedback rate constraint. *IEEE Transactions on Signal Processing*, 55(7):3879–3891, 2007.

[56] J. Imam. Deepest part of the world's ocean is incredibly noisy scientists say, cnn. `http://www.cnn.com/2016/03/05/world/deep-sea-audio-recording/index.html/`. Last time accessed: March 5, 2016.

[57] P. Ioannides and C. A. Balanis. Uniform Circular Arrays for Smart Antennas. *IEEE Antennas and propagation magazine*, 47(4):192–206, 2005.

[58] ITC. ITC3013 transducer specifications. `https://docs.wixstatic.com/ugd/8aa48d_4185da8a232645a89ce4e94801432299.pdf`. Last time accessed: December 2018.

[59] F. Iutzeler, P. Ciblat, and J. Jakubowicz. Analysis of max-consensus algorithms in wireless channels. *IEEE Transactions on Signal Processing*, 60(11):6103–6107, 2012.

[60] N. Jindal and A. Goldsmith. Capacity and optimal power allocation for fading broadcast channels with minimum rates. *IEEE Transactions on Information Theory*, 49(11):2895–2909, 2003.

[61] M. K. Jubran, M. Bansal, and L. P. Kondi. Low-delay low-complexity bandwidth-constrained wireless video transmission using svc over mimo systems. *IEEE Transactions on Multimedia*, 10(8):1698–1707, Dec 2008.

[62] A. A. Kalachikov and N. S. Shelkunov. Performance evaluation of the detection algorithms for MIMO spatial multiplexing based on analytical wireless MIMO channel models. In *Proceedings of International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE)*, pages 180–183, October 2018.

[63] S. M. Kay. Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory. 1993.

[64] J. C. Kinsey, R. M. Eustice, and L. L. Whitcomb. A survey of underwater vehicle navigation: Recent advances and new challenges. In *Proceedings of IFAC Conference of Manoeuvering and Control of Marine Craft*, volume 88, pages 1–12, 2006.

[65] M. Kountouris and J. G. Andrews. Downlink SDMA with limited feedback in interference-limited wireless networks. *IEEE Transactions on Wireless Communications*, 11(8):2730–2741, 2012.

[66] T. H. Le Dao, P. V. Giap, and H. V. Xiem. Adaptive long-term reference selection for efficient scalable surveillance video coding. In *Proceedings of International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC)*, pages 69–73. IEEE, 2018.

[67] E. K. Lee, H. Viswanathan, and D. Pompili. Distributed data-centric adaptive sampling for cyber-physical systems. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 9(4):21, 2015.

[68] J-S. Lee, F. De Simone, and T. Ebrahimi. Subjective quality evaluation via paired comparison: Application to scalable video coding. *IEEE Transactions on Multimedia*, 13(5):882–893, 2011.

[69] C. Li, Y. Xu, C. Xu, Z. An, B. Diao, and X. Li. DTMAC: A delay tolerant MAC protocol for underwater wireless sensor networks. *IEEE Sensors Journal*, 16(11):4137–4146, 2016.

[70] J. Li and P. Stoica. *Robust Adaptive Beamforming*. John Wiley & Sons, 2005.

[71] T. Li and A. Nehorai. Maximum likelihood direction finding in spatially colored noise fields using sparse sensor arrays. *IEEE Transactions on Signal Processing*, 59(3):1048–1062, 2011.

[72] S. Lmai, M. Chitre, C. Laot, and S. Houcke. Throughput-efficient super-tdma mac transmission schedules in ad hoc linear underwater acoustic networks. *IEEE Journal of Oceanic Engineering*, 42(1):156–174, 2017.

[73] S. Lu, Z. Wang, Z. Wang, and S. Zhou. Throughput of underwater wireless ad hoc networks with random access: A physical layer perspective. *IEEE Transactions on Wireless Communications*, 14(11):6257–6268, 2015.

[74] X. Lurton. *An Introduction To Underwater Acoustics: Principles and Applications*. Springer Science & Business Media, 2002.

[75] B. Makki and T. Eriksson. On the Performance of MIMO-ARQ Systems with Channel State Information at the Receiver. *IEEE Transactions on Communications*, 62(5):1588–1603, 2014.

[76] G. Marsaglia. Ratios of normal variables. *Journal of Statistical Software*, 16(4):1–10, 2006.

[77] T. Melodia, H. Kulhandjian, L.-Chung Kuo, and Emrecan Demirors. *Advances in underwater acoustic networking*, volume 852, pages 804–852. John Wiley & Sons, Inc., 2013.

[78] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun. Spatial Modulation. *IEEE Transactions on Vehicular Technology*, 57(4):2228–2241, 2008.

[79] R. R. Murphy, T. Nomura, A. Billard, and J. L. Burke. Human–robot interaction. *IEEE robotics & automation magazine*, 17(2):85–89, 2010.

[80] R. Narasimhan. Finite-SNR diversity–multiplexing tradeoff for correlated rayleigh and rician mimo channels. *IEEE Transactions on Information Theory*, 52(9):3965–3979, 2006.

[81] Oceansonics. Iclisten smart hydrophones. `http://oceansonics.com/iclisten-smart-hydrophones/`. Last time accessed: December 2018.

[82] S. J. Orfanidis. *Electromagnetic waves and antennas*. Rutgers University, NJ, 2002.

[83] S Pagadarai, A Kliks, H Bogucka, and AM Wyglinski. Non-contiguous Multicarrier Waveforms in Practical Opportunistic Wireless Systems. *IET radar, sonar & navigation*, 5(6):674–680, 2011.

[84] P. Pandey, M. Rahmati, D. Pompili, and W. U. Bajwa. Robust distributed dictionary learning for in-network image compression. In *Proceedings of International Conference on Autonomic Computing (ICAC)*, pages 61–70. IEEE, 2018.

[85] L. Paull, S. Saeedi, M. Seto, and H. Li. AUV navigation and localization: A review. *IEEE Journal of Oceanic Engineering*, 39(1):131–149, 2014.

[86] K. I. Pedersen, S. R. Khosravirad, G. Berardinelli, and F. Frederiksen. Rethink hybrid automatic repeat request design for 5G: Five configurable enhancements. *IEEE Wireless Communications*, 24(6):154–160, 2017.

[87] C. Pelekanakis, M. Stojanovic, and L. Freitag. High Rate Acoustic Link For Underwater Video Transmission. In *Proceedings of IEEE OCEANS Conference*, volume 2, pages 1091–1097, 2003.

[88] J. Pérez, J. Ibánez, L. Vielva, and I. Santamaria. Closed-form approximation for the outage capacity of orthogonal stbc. *IEEE Communications Letters*, 9(11):961–963, 2005.

[89] C. Petrioli, R. Petroccia, D. Spaccini, A. Vitaletti, T. Arzilli, D. Lamanna, A. Galizial, and E. Renzi. The SUNRISE GATE: Accessing the SUNRISE federation of facilities to test solutions for the internet of underwater things. In *Proceedings of Underwater Communications and Networking (UComms)*, pages 1–4. IEEE, 2014.

[90] C. Petrioli, R. Petroccia, and M. Stojanovic. A comparative performance evaluation of MAC protocols for underwater sensor networks. In *Proceedings of MTS/IEEE OCEANS Conference*, Quebec City, Quebec, Canada, September 15–18 2008.

[91] R. Petroccia. A distributed ID assignment and topology discovery protocol for underwater acoustic networks. In *Proceedings of the International Conference on Underwater Communications and Networking (UComms)*, pages 1–5. IEEE, 2016.

[92] R. Petroccia, C. Petrioli, and J. Potter. Performance evaluation of underwater medium access control protocols: At-sea experiments. *IEEE Journal of Oceanic Engineering*, 43(2):547–556, 2018.

[93] M. Pierdomenico, D. Casalbore, and F. L. Chiocci. Massive benthic litter funnelled to deep sea by flash-flood generated hyperpycnal flows. *Scientific Reports*, 9(1):5330, 2019.

[94] D. Pompili and I. F. Akyildiz. Overview of Networking Protocols for Underwater Wireless Communications. *IEEE Communications Mag.*, 47(1):97–102, 2009.

[95] D. Pompili and I. F. Akyildiz. A Multimedia Cross-Layer Protocol for Underwater Acoustic Sensor Networks. *IEEE Transactions on Wireless Communications*, 9(9):2924–2933, 2010.

[96] D. Pompili, T. Melodia, and I. F. Akyildiz. A CDMA-based medium access control for underwater acoustic sensor networks. *IEEE Transaction on Wireless Communications*, 8(4):1899–1909, 2009.

[97] M. Porter et al. Bellhop Code. `http://oalib.hlsresearch.com/Rays`. in the Acoustics Toolbox, Last time accessed: 2013.

[98] J. Potter, J. Alves, T. Furfaro, A. Vermeij, N. Jourden, D. Merani, G. Zappa, and A. Berni. Software defined open architecture modem development at cmre. In *Proceedings of Underwater Communications and Networking Conference (UComms)*, pages 1–4. IEEE, 2014.

[99] J. Potter, J. Alves, D. Green, G. Zappa, I. Nissen, and K. McCoy. The JANUS underwater communications standard. In *Proceedings of Underwater Communications and Networking Conference (UComms)*, pages 1–4. IEEE, 2014.

[100] J. G. Proakis and M. Salehi. *Digital communications*. McGraw-Hill, fifth edition, 2008.

[101] M. O. Pun, V. Koivunen, and V. Poor. Performance analysis of joint opportunistic scheduling and receiver design for MIMO-SDMA downlink systems. *IEEE Transactions on Communications*, 59(1):268–280, 2011.

[102] P. Qarabaqi and M. Stojanovic. Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels. *IEEE Journal of Oceanic Engineering*, 38(4):701–717, 2013.

[103] D. Rahamim, J. Tabrikian, and R. Shavit. Source Localization Using Vector Sensor Array in a Multipath Environment. *IEEE Transactions on Signal Processing*, 52(11):3096–3103, 2004.

[104] M. Rahmati, A. Gurney, and D. Pompili. Adaptive underwater video transmission via software-defined mimo acoustic modems. In *Proceedings of OCEANS Conference*, pages 1–6. MTS/IEEE, 2018.

[105] M. Rahmati, S. Karten, and D. Pompili. Slam-based underwater adaptive sampling using autonomous vehicles. In *Proceedings of MTS/IEEE OCEANS Conference*, pages 1–7, 2018.

[106] M. Rahmati, P. Pandey, and D. Pompili. Separation and classification of underwater acoustic sources. In *Proceedings of IEEE Underwater Communications and Networking (UComms)*, pages 1–5, 2014.

[107] M. Rahmati, R. Petroccia, and D. Pompili. In-network collaboration for cdma-based reliable underwater acoustic communications. *IEEE Journal of Oceanic Engineering*, 44(4):881–894, 2019.

[108] M. Rahmati and D. Pompili. Interference cancellation in multiuser acoustic underwater networks using probabilistic SDMA. In *Proceedings of IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pages 307–315, 2015.

[109] M. Rahmati and D. Pompili. uwmimo-harq: Hybrid arq for reliable underwater acoustic mimo communications. In *Proceedings of ACM International Conference on Underwater Networks and Systems (WUWNET)*, pages 1–8, 2015.

[110] M. Rahmati and D. Pompili. SSFB: Signal-space-frequency beamforming for underwater acoustic video transmission. In *Proceedings of International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pages 180–188. IEEE, 2017.

[111] M. Rahmati and D. Pompili. UNISeC: Inspection, separation, and classification of underwater acoustic noise point sources. *IEEE Journal of Oceanic Engineering*, 43:777–791, 2018.

[112] M. Rahmati and D. Pompili. Uw-svc: Scalable video coding transmission for in-network underwater imagery analysis. In *Proceedings of IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pages 1–8, 2019.

[113] M. Rahmati and D. Pompili. Probabilistic spatially-divided multiple access in underwater acoustic sparse networks. *IEEE Transactions on Mobile Computing*, 19(2):405–418, 2020.

[114] M. Rahmati, D. Pompili, and R. Petroccia. Collaborative hybrid arq for cdma-based reliable underwater acoustic communications. In *Proceedings of IEEE Underwater Communications and Networking Conference (UComms)*, pages 1–5, 2018.

[115] M. Rahmati, Z. Qi, and D. Pompili. Underwater adaptive video transmission using mimo-based software defined acoustic modems. *submitted to IEEE Transactions on Multimedia*, May 2020.

[116] M. Rahmati, V. Sadhu, and D. Pompili. Eco-uw iot: Eco-friendly reliable and persistent data transmission in underwater internet of things. In *Proceedings of International Conference on Sensing, Communication, and Networking (SECON)*, pages 1–9. IEEE, 2019.

[117] T. S. Rappaport. *Wireless communications: Principles and practice*. Prentice Hall PTR New Jersey, 1996.

[118] J. Ribas, D. Sura, and M. Stojanovic. Underwater wireless video transmission for supervisory control and inspection using acoustic ofdm. In *Proceedings of MTS/IEEE OCEANS Conference*, pages 1–9, 2010.

[119] W J. Richardson, C. R. Greene Jr, C. I. Malme, and D. H Thomson. *Marine Mammals and Noise*. Academic press, 2013.

[120] R. Roy and T. Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7):984–995, 1989.

[121] S. Schaible and T. Ibaraki. Fractional Programming. *European Journal of Operational Research*, 12(4):325–338, 1983.

[122] R. Schmidt. Multiple Emitter Location and Signal Parameter Estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.

[123] H. Schwarz, D. Marpe, and T. Wiegand. Overview of the scalable video coding extension of the h. 264/avc standard. *IEEE Transactions on circuits and systems for video technology*, 17(9):1103–1120, 2007.

[124] S. Shahabudeen, M. Motani, and M. Chitre. Analysis of a High-Performance MAC Protocol for Underwater Acoustic Networks. *IEEE Journal of Oceanic Engineering*, 39(1):74–89, 2014.

[125] S. Shahhosseini and S. A. Hosseini. A visual base image coding using unequal error protection in high error rate channels. In *Proceedings of International Conference on Computer Applications and Industrial Electronics*, pages 537–542, Dec 2010.

[126] C.E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, 1949.

[127] M. Sherman. *Spatial statistics and spatio-temporal data: Covariance functions and directional properties*. John Wiley & Sons, 2011.

[128] M. T. Silvia and R. T. Richards. A Theoretical and Experimental Investigation of Low-frequency Acoustic Vector Sensors. In *Proceedings of MTS/IEEE OCEANS Conference*, volume 3, pages 1886–1897, 2002.

[129] K. D. Singh, A. Ksentini, and B. Marienval. Quality of experience measurement tool for svc video coding. In *Proceedings of International Conference on Communications (ICC)*, pages 1–5. IEEE, 2011.

[130] E. Soljanin, R. Liu, and P. Spasojevic. Hybrid ARQ with random transmission assignments. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 66:321–334, 2004.

[131] Neptune sonar. Neptun T313 transducer specifications. `http://www.neptune-sonar.co.uk/wp-content/uploads/2016/03/Model-T313.pdf`. Last time accessed: December 2018.

[132] P. Stoica and A. Nehorai. MUSIC, Maximum Likelihood, and Cramer-Rao Bound. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(5):720–741, 1989.

[133] M. Stojanovic. Optimization of a data link protocol for an underwater acoustic channel. In *Proceedings of Oceans-Europe Conference*, volume 1, pages 68–73. IEEE, 2005.

[134] M. Stojanovic. On the relationship between capacity and distance in an underwater acoustic communication channel. *ACM SIGMOBILE Mobile Computing and Communications Review*, 11(4):34–43, 2007.

[135] M. Stojanovic and L. Freitag. Multichannel detection for wideband underwater acoustic CDMA communications. *IEEE Journal of Oceanic Engineering*, 31(3):685–695, 2006.

[136] M. Stojanovic and J. Preisig. Underwater acoustic communication channels: Propagation models and statistical characterization. *IEEE Communications Magazine*, 47(1):84–89, 2009.

[137] K. Stuhlmuller, N. Farber, M. Link, and B. Girod. Analysis of video transmission over lossy channels. *IEEE Journal on Selected Areas in Communications*, 18(6):1012–1032, 2000.

[138] T. Stutz and A. Uhl. A survey of H.264 avc/svc encryption. *IEEE Transactions on circuits and systems for video technology*, 22(3):325–339, 2012.

[139] M. Suzuki, T. Sasaki, and T. Tsuchiya. Digital acoustic image transmission system for deep-sea research submersible. In *Proceedings of IEEE OCEANS, Mastering the Oceans Through Technology Conference*, volume 2, pages 567–570, 1992.

[140] A. A. Syed, W. Ye, and J. Heidemann. Comparison and evaluation of the T-Lohi MAC for underwater acoustic sensor networks. *IEEE Journal on Selected Areas in Communications*, 26(9), 2008.

[141] H-P Tan, R. Diamant, W. KG Seah, and M. Waldmeyer. A Survey of Techniques and Challenges in Underwater Localization. *IEEE Journal of Oceanic Engineering*, 38(14):1663–1676, 2011.

[142] J. Tao, Y. R. Zheng, C. Xiao, and T.C. Yang. Robust MIMO Underwater Acoustic Communications Using Turbo Block Decision-Feedback Equalization. *IEEE Journal of Oceanic Engineering*, 35(4):948–960, 2010.

[143] S. Thrun. Toward a framework for human-robot interaction. *Human–Computer Interaction*, 19(1-2):9–24, 2004.

[144] D. Torres, J. Friedman, T. Schmid, M. B. Srivastava, Y. Noh, and M. Gerla. Software-defined underwater acoustic networking platform and its applications. *Ad Hoc Networks*, 34:252–264, 2015.

[145] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge university press, 2005.

[146] C. C. Tsimenidis, O. R. Hinton, A. E. Adams, and B. S. Sharif. Underwater acoustic receiver employing direct-sequence spread spectrum and spatial diversity combining for shallow-water multiaccess networking. *IEEE Journal of Oceanic Engineering*, 26(4):594–603, 2001.

[147] UNESCO. Blueprint for the future we want: Marine pollution. `http://www.unesco.org/new/en/natural-sciences/ioc-oceans/focus-areas/rio-20-ocean/blueprint-for-the-future-we-want/marine-pollution/facts-and-figures-on-marine-pollution`. Last time accessed: September 2019.

[148] L. D. Vall, D. Sura, and M. Stojanovic. Towards underwater video transmission. In *Proceedings of the International Workshop on Underwater Networks*, page 4. ACM, 2011.

[149] E. A. Wan and R. Van Der Merwe. The unscented kalman filter for nonlinear estimation. In *Proceedings of IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium (AS-SPCC)*, pages 153–158, 2000.

[150] Y. Wang, J. Zhou, and H. Kikuchi. Performance of a three-dimensional antenna array and its application in doa estimation. *Wireless Personal Communications*, pages 1–17, 2016.

[151] M. Wichtlhuber, J. Rckert, D. Winter, and D. Hausheer. How to adapt: Svc-based quality adaptation for hybrid peercasting systems. In *Proceedings of IFIP/IEEE International Symposium on Integrated Network Management (IM)*, pages 174–182, May 2015.

[152] S.-H. Wu, L.-K. Chiu, K.-Y. Lin, and T.-H. Chang. Robust hybrid beamforming with phased antenna arrays for downlink SDMA in indoor 60 ghz channels. *IEEE Transactions on Wireless Communications*, 12(9):4542–4557, 2013.

[153] Y. Xiao. *Underwater acoustic sensor networks*. CRC Press, 2010.

[154] K. Xiong, H.Y. Zhang, and C.W. Chan. Performance Evaluation of UKF-Based Nonlinear Filtering. *Automatica*, 42(2):261–270, 2006.

[155] N. Yilmazer, J. Koh, and T. K. Sarkar. Utilization of a Unitary Transform for Efficient Computation in the Matrix Pencil Method to Find the Direction of Arrival. *IEEE Transactions on Antennas and Propagation*, 54(1):175–181, 2006.

[156] X. Zhao, V. Sadhu, and D. Pompili. Analog signal compression and multiplexing techniques for healthcare internet of things. In *Proceedings of IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pages 398–406, 2017.

[157] S. Zhou and Z. Wang. *OFDM For Underwater Acoustic Communications*. John Wiley & Sons, 2014.

[158] Y. Zhu, Z. Peng, J-H Cui, and H. Chen. Toward Practical MAC Design for Underwater Acoustic Networks. *IEEE Transactions on Mobile Computing*, 14(4):872–886, 2015.

[159] M. Zorzi and R. R Rao. On the use of renewal theory in the analysis of ARQ protocols. *IEEE Transactions on Communications*, 44(9):1077–1081, 1996.