Author:
Bocus, Junaid

Title:
Underwater Wireless Video Transmission for Remotely Operated Underwater Vehicles

Underwater Wireless Video Transmission for Remotely Operated Underwater Vehicles

By

MOHAMMUD JUNAID BOCUS

Department of Electrical and Electronic Engineering
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Engineering.

APRIL 2019

Word count: 37860

Abstract

Underwater Acoustic (UWA) communication has become an active area of research due to a growing range of applications for scientific, military and commercial purposes. Since many of these applications involve deep water intervention, underwater Remotely Operated Vehicles (ROVs) are deployed to avoid human casualties. Some applications also require a video to be transmitted in real-time to a surface vessel. The transmission of such data over the Underwater Acoustic Channel (UAC) is quite challenging due to its inherent characteristics, namely, frequency dependent attenuation, ambient noise, multipath distortion, propagation delay and Doppler effect. The bit rates that are currently achieved over long transmission distances are usually in the order of a few tens of kilobits per second (kbps). These are however not sufficiently high to transmit large data such as video in real-time.

In this thesis, the feasibility of real-time video transmission in both horizontally and vertically configured UACs is evaluated. In order to make maximum usage of the limited acoustic bandwidth and simultaneously overcome the harmful propagation phenomena in an UAC, physical layer waveforms such as Orthogonal Frequency Division Multiplexing (OFDM), Filterbank Multicarrier (FBMC) modulation and Orthogonal Time frequency Space (OTFS) modulation are evaluated for both Single-User (SU) and Multi-User (MU) scenarios. To further boost the achievable bit rates, Multiple-Input Multiple-Output (MIMO) and massive MIMO systems are considered for the SU and MU scenarios, respectively. In terms of video compression, the H.264 Advanced Video Coding (AVC) standard is considered due to its good error resilience compared to other standards.

It is demonstrated that FBMC provides robust performance in doubly-dispersive UACs and can even outperform OFDM. Furthermore, FBMC systems based on OFDM - Offset Quadrature Amplitude Modulation (OFDM-OQAM) achieve 100% bandwidth efficiency as no Cyclic Prefix (CP) is used, thus improving the overall bit rate as compared to OFDM. With the use of Forward Error Correction (FEC) codes such as Turbo codes, the error performance of the different systems is significantly improved in both horizontal and vertical UACs. It is further demonstrated that the massive MIMO technique allows multiple ROVs to communicate over the same time and frequency resources with a surface base-station (BS) with very good error and throughput performances. Finally, it is shown that OFDM-based OTFS systems outperform the conventional OFDM systems in a dynamic UAC. Fewer pilots are required in the delay-Doppler (DD) domain to estimate the sparse UAC as compared to the frequency domain, thereby improving the bandwidth efficiency of the OTFS system. Therefore, the OTFS modulation scheme not only achieves robust performance in UACs affected by severe Doppler effect but also enables a higher bit rate than conventional OFDM systems. In this research, theoretical bit rates between 70 kbps and as high as 240 kbps are reported for the various MIMO and massive MIMO systems operating over a 1 km time-varying UAC. It is shown that such bit rates are sufficiently high to transmit and receive video in real-time with adequate quality over the long distance UAC.

Dedication and acknowledgements

Firstly, I want to express my heartfelt gratitude and appreciation to my supervisors, Dr. Dimitris Agrafiotis and Professor Angela Doufexi, who have provided me with their many suggestions, support and continuous guidance during the course of this research project. I also want to express my deep appreciation to my internal reviewer, Professor Robert Piechocki for his many suggestions, valuable feedback and advice during the Ph.D. reviews.

I take this opportunity to acknowledge and honour the contribution of my dearest friends; namely Dr. Rui Fan, Dr. Ghaith Al-Juboori, Dr. Imad Al-Samman and Dr. William Andrew who have supported me in various ways throughout my studies and made my stay in Bristol enjoyable and memorable.

Studying in the United Kingdom has always been a dream for me and being part of a prestigious institution such as the University of Bristol would not have been possible without a sponsoring organisation. Therefore, I would like to express my sincere gratitude and thankfulness to the University of Bristol and also the Alumni and Friends Foundation for awarding me a scholarship to pursue my postgraduate research studies.

Last but not least, my deepest gratitude goes to my late mother, father, brother, sister and all close relatives and friends who stood by me through thick and thin and provided me with their continuous support and encouragement during my academic studies and also in my personal life. Being surrounded by loved ones is indeed a blessing and I pray to God that He grants my late mother the highest rank in paradise and strengthens our family bonds.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: .....

Table of Contents

LIST OF TABLES

LIST OF FIGURES

FIGURE **Page**

NOTATIONS

$[\cdot]_{\downarrow m}, [\cdot]_{\uparrow m}$	Down-sample, Up-sample by a factor of m
$\circledast$	Circular convolution
$\mathbf{0}_m$	All-zero matrix, dimension $m \times m$
$\mathbf{I}_m$	Identity matrix, dimension $m \times m$
$\mathbf{M} \in \mathbb{C}^{m_1 \times m_2}$	Matrix, boldface uppercase, dimension $m_1 \times m_2$
$\mathbf{v} \in \mathbb{C}^{v \times 1}$	Vector, boldface lowercase, dimension $v \times 1$
$\otimes$	Kronecker product operation between two matrices
$(\cdot)^*$	Complex-conjugate
$(\cdot)^{\mathsf{H}}$	Hermitian Transpose
$(\cdot)^{\mathsf{T}}$	Transpose
Z, Z^{-1}	Z-Transform, Inverse Z-Transform (unit delay)
$[\cdot]_M$	Modulo M operation
δ	Delta function
$\langle s_1, s_2 \rangle$	Projection of a signal s_1 onto a signal s_2 (inner product)
$\mathbb{E}\{\cdot\}$	Expectation
$\mathscr{A}(\cdot)$	Ambiguity function
$\mathscr{H}_i(\cdot)$	Hermite functions
$\mathscr{I}\{\cdot\}$	Imaginary part
$\mathscr{R}\{\cdot\}$	Real part
$\prod$	Product operator
$\propto$	is proportional to
$\star$	Linear convolution
$\mathrm{diag}[\cdot]$	Transforming a column or row vector into a diagonal matrix
$\mathrm{vec}(\mathbf{M})$	Transforming a matrix $\mathbf{M}$ into a column vector
$\mathbf{Pr}(A\|B)$	Conditional probability of A given that B has occured
j	$\sqrt{-1}$
s, S	Non-bold scalar representations
$\|\cdot\|$	Absolute value

ACRONYMS

2D	2-Dimensional
3D	3-Dimensional
3G	Third Generation
5G	Fifth Generation
AFB	Analysis Filter Bank
ASC	Autonomous Surface Craft
ASIMOV	Advanced System Integration for Managing the Co-ordinated Operation of Robotic Ocean Vehicles
AUV	Autonomous Underwater Vehicle
AVC	Advanced Video Coding
AWGN	Additive White Gaussian Noise
BCH	Bose-Chaudhuri-Hocquenghem
BER	Bit Error Rate
BS	base-station
CABAC	Context Adaptive Binary Arithmetic Coding
CAVLC	Context Adaptive Variable Length Coding
CDMA	Code Division Multiple Access
CFR	Channel Frequency Response
CIR	Channel Impulse Response
CMA	Constant Modulus Algorithm
CMT	Cosine Modulated Multitone
CP	Cyclic Prefix
CSF	Channel Scattering Function
CSI	Channel State Information
DCT	Discrete Cosine Transform
DD	delay-Doppler
DFE	Decision Feedback Equaliser
DFT	Discrete Fourier Transform
DPSK	Differential Phase Shift Keying
DS-CDMA	Direct Sequence-Code Division Multiple Access
DSB	Double Sideband Modulation
DSL	Digital Subscriber Line
DSP	Digital Signal Processing
DWMT	Discrete Wavelet Multitone
DWT	Discrete Wavelet Transform
FBMC	Filter Bank Multicarrier
FDM	Frequency Division Multiplexing

FEC	Forward Error Correction
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FMO	Flexible Macroblock Ordering
FMT	Filtered Multitone
fps	frames per second
FS	Frequency-Spreading
FT	Frequency-Time
GOP	Group of Pictures
HD	High-Definition
HDTV	High-Definition Television
HEVC	High Efficiency Video Coding
HFM	Hyperbolic Frequency Modulated
i.i.d	independent and identically distributed
IAM	Interference Approximation Method
ICI	Inter-Carrier Interference
IDFT	Inverse Discrete Fourier Transform
IDR	Instantaneous Decoding Refresh
IFFT	Inverse Fast Fourier Transform
IOTA	Isotropic Orthogonal Transform Algorithm
ISFFT	Inverse Symplectic Finite Fourier Transform
ISI	Inter-Symbol Interference
ISO/IEC	International Organisation for Standardisation/International Electrotechnical Commission
ITC	International Transducer Corporation
ITU-T	Telecommunication Standardisation Sector of the International Telecommunications Union
JPEG	Joint Photographic Experts Group
LC	Layer Coding
LDPC	Low-Density Parity-Check
LFM	Linear Frequency Modulated
LMS	Least Mean Squares
LPF	Low-pass Filter
LS	Least Square
LTE	Long Term Evolution
MAP	maximum a posteriori
MB	macroblock
MDC	Multiple Description Coding
MIMO	Multiple-Input Multiple-Output
MLD	Maximum Likelihood Detection
MLSE	Maximum Likelihood Sequence Estimation
MMSE	Minimum Mean Square Error
MPEG	Moving Picture Experts Group
MRC	Maximum-Ratio Combining
MSE	Mean Squared Error
MU	Multi-User
MV	Motion Vector

NAL	Network Abstraction Layer
NOMA	Non-Orthogonal Multiple Access
NPR	Near-Perfect Reconstruction
OCRR	Open Circuit Receiving Response
OFDM	Orthogonal Frequency Division Multiplexing
OFDM-OQAM	Orthogonal Frequency Division Multiplexing-Offset Quadrature Amplitude Modulation
OFDMA	Orthogonal Frequency Division Multiple Access
OMA	Orthogonal Multiple Access
OMP	Orthogonal Matching Pursuit
OQPSK	Offset Quadrature Phase Shift Keying
OTFS	Orthogonal Time Frequency Space Modulation
P/S	Parallel-to-Serial
PAM	Pulse Amplitude Modulation
PDF	Probability Density Function
PDP	Power Delay Profile
PER	Packet Error Rate
PHYDYAS	Physical Layer for Dynamic Spectrum Access and Cognitive Radio
PLR	Packet Loss Rate
pmf	probability mass function
POP	Pairs of Pilots
PPN	Polyphase Network
PPS	Picture Parameter Sets
PSD	Power Spectral Density
PSK	Phase Shift Keying
PSNR	Peak-Signal-to-Noise Ratio
QAM	Quadrature Amplitude Modulation
QP	Quantisation Parameter
QPSK	Quadrature Phase Shift Keying
RF	Radio-Frequency
RLS	Recursive Least Squares
RMS	Root Mean Square
ROV	Remotely Operated Vehicle
RRC	Root Raised Cosine
RS	Reed Solomon
RX	receiver
S/P	Serial-to-Parallel
SAD	Sum of Absolute Difference
SD	Standard-Definition
SDR	Software Defined Radio
SEI	Supplemental Enhancement Information
SFB	Synthesis Filter Bank
SFBC	Space-Frequency Block Coding
SFFT	Symplectic Finite Fourier Transform
SI	Switching Intra
SIC	Successive Interference Cancellation

SINR	Signal-to-Interference-plus-Noise Ratio
SIR	Signal-to-Interference Ratio
SISO	Single-Input Single-Output
SMT	Staggered Multitone
SNR	Signal-to-Noise Ratio
SP	Switching Predictive
SPS	Sequence Parameter Sets
SSE	Sum of Squared Error
STBC	Space-Time Block Codes
STC	Space-Time Codes
STTC	Space-Time Trellis Codes
SU	Single-User
SVC	Scalable Video Coding
TDD	Time-Division Duplex
TF	Time-Frequency
TFL	Time-Frequency Localisation
TVR	Transmitting Voltage Response
TX	transmitter
UAC	Underwater Acoustic Channel
UHD	Ultra High-Definition
UWA	Underwater Acoustic
V-BLAST	Vertical Bell Laboratories Layered Space-Time
VCL	Video Coding Layer
VLC	Variable Length Codes
VSB	Vestigial sideband
WHOI	Woods Hole Oceanographic Institution
WSSUS	Wide-Sense Stationary Uncorrelated Scattering
ZF	Zero-Forcing

CHAPTER 1

INTRODUCTION

The Earth's surface is covered by approximately 70% of water and many parts of the ocean are yet to be explored. The average depth of the ocean is around 3.6 km [1] and the deepest point below the water surface that a human scuba diver has managed to reach is 332 m [2]. Most experienced recreational scuba divers however do not dive beyond 40 m due to the severity in health risks involved. Using a deep water submersible named *Bathyscaphe Trieste* on 23 January 1960, two persons have managed to reach a record maximum depth of approximately 10,911 m in the Challenger Deep, which is the deepest part known in the ocean and is located in the Mariana Trench. In order to avoid human casualties at extreme depths, unmanned underwater vehicles are now being deployed. Recent technological advancements have enabled these vehicles to communicate wirelessly to a surface station, thus overcoming the limitations of tethered vehicles and paving the way for new opportunities in next generation underwater communication systems. However, underwater wireless communication presents some distinct and more complex challenges compared to terrestrial Radio-Frequency (RF) communication systems. This chapter outlines the need for underwater wireless communication and its corresponding challenges. The state-of-the-art in Underwater Acoustic (UWA) communication is provided for both horizontal and vertical transmission scenarios. Finally, this chapter gives an overview of the thesis structure and the key contributions in this field.

1.1 Motivation

Recently, a number of applications including military (e.g., mine reconnaissance, surveillance, intelligence collection), scientific (e.g., marine life exploration, pollution monitoring, oceanography and archaeology) and commercial (e.g., fishing, underwater construction, offshore oil and gas surveys) have initiated research in high data rate underwater communications [3, 4]. A number of these applications require deep sea intervention where human divers cannot reach due to the dangers associated with the

high water pressure. Therefore, autonomous underwater drones such as Remotely Operated Vehicles (ROVs) are used. These underwater vehicles are usually fitted with equipment such as robotic arms, cutting tools, measuring instruments, light sources and video cameras to help with the underwater expeditions. Most of the ROVs which are deployed at sea are connected to a surface station through an umbilical cable which is not only used to send command and control signals to the ROV but also for the latter to transmit data such as images or videos to the surface vessel. The tether cable however presents some disadvantages, namely, it is not cost-effective and limits the range of the ROV to a fixed depth due to the finite length of the cable. By resorting to wireless transmission, the drawbacks of the tethered underwater vehicles can be overcome, thereby offering new possibilities for underwater communication.

RF waves can support high data rate transmission over long distances in a terrestrial communication system but this is however not feasible in seawater due to the latter being conductive in nature. Hence, the RF waves will suffer strong attenuation and will typically operate at very low frequencies. This translates to large antennas being used with high power consumption. Conversely, the use of optical waves for underwater communication has shown promising results in terms of high data rates, for instance, [5] and references therein. However, optical waves require an accurate alignment between the optical transmitter and receiver and this may be impractical for a number applications. Another disadvantage of optical waves for underwater communication is that they are limited to short distances, typically <100 m, due to the scattering effects of optical waves with suspending underwater particles [6]. Hence, acoustic waves remain the most suitable medium for underwater communication, especially for long-range links where the use of cables is impractical.

1.2 Problem Definition

Underwater video transmission using acoustic waves is highly challenging and progress in this field, especially in terms of achievable bit rates, has been incremental. This is due to a number of limitations in the UWA domain which are stated next.

The Underwater Acoustic Channel (UAC) is regarded as one of the most challenging medium for wireless communication since it combines the worst properties of a terrestrial RF channel in terms of poor physical link quality and the high latency of a satellite link [7]. The acoustic bandwidth is extremely restricted for long-range acoustic links which vary from a few Hz to hundreds of kHz [4]. This is because the lower and upper bounds of the bandwidth are limited mainly by the ambient noise and frequency-dependent attenuation, respectively. Other limitations of the UAC include multipath distortion, Doppler effect and high propagation delays because of the low speed of acoustic waves in water ($\approx$ 1500 m/s). When the transmitter and receiver are separated horizontally, multipath distortion can be extreme and the delay spread can span over tens or even hundreds of milliseconds [7]. Vertical links, on the other hand, exhibit lower multipath distortion and the bit rate that can be achieved is mostly limited by the bandwidth of the system [4, 8]. Therefore, high data rates are possible through the use of bandwidth-efficient modulation schemes [8].

Nowadays, many underwater applications such as ocean exploration and monitoring require real-time high-quality video to be transmitted from an underwater vehicle to a surface vessel [8]. This is a challenging task given the severe limitation on the acoustic bandwidth. Due to the limited bandwidth of long-range acoustic links (≥ 1 km), bandwidth-efficient signalling becomes a necessity to achieve high bit rates and simultaneously minimise the physical propagation phenomena of the UAC. In this regard, multicarrier modulation techniques such as Orthogonal Frequency Division Multiplexing (OFDM) and more recently Filter Bank Multicarrier (FBMC) systems have drawn much attention for UWA communication.

Furthermore videos, especially high-quality ones, require a high bit rate for real-time transmission to be possible. Thus, suitable video compression techniques with lower time complexities are necessary to reduce the high bit rate requirement of real-time video streaming while maintaining the perceived visual quality. Moreover, transmission errors in the UAC need to be corrected for achieving a successful real-time video communication between a ROV and a surface station. A compressed video is composed of packets which consist of a series of bits. Compressed bit-streams are vulnerable to bursts of bit errors which cause video packet loss. Usually for underwater video transmission, high compression rates are used and this increases the sensitivity to transmission errors [9]. A single bit error in a video packet can cause it to be dropped [10], thus degrading the video quality. Hence efficient Forward Error Correction (FEC) codes are indispensable in UWA communication systems to achieve a target Bit Error Rate (BER) and reduce the number of retransmissions which may cause a delay and also increase the energy consumption [3]. However, FEC codes add undesired redundancy to the data and thus further decrease the useful data rate in an already bandwidth-limited channel. It is to be noted that some video compression standards usually include error resilience techniques to cope with packet or frame losses, thus minimising the degradation of the video quality.

1.3 State of the Art

A lot of research in UWA communication has taken place over the years employing non-coherent modulation, coherent modulation, Direct Sequence-Code Division Multiple Access (DS-CDMA), multicarrier techniques such as OFDM and spatial modulation which uses Multiple-Input Multiple-Output (MIMO) technologies [4]. Non-coherent modulation methods are used for applications where features such as robustness and low implementation complexity are desired rather than a high data rate transmission [4]. Coherent modulation techniques, on the other hand, were introduced to improve the overall bit rate in an UAC. With multicarrier techniques such as OFDM which can be used in conjunction with MIMO, enhanced data rate and better bandwidth efficiency can be achieved. OFDM exploits the diversity in the frequency domain using several subcarriers. It enables simple frequency-domain equalisation at the receiver using a Cyclic Prefix (CP) and it also provides robustness against Inter-Symbol Interference (ISI).

1.3.1 Vertical Transmission

The first underwater image transmission system was implemented in Japan in 1989 [11]. The Discrete Cosine Transform (DCT) was used as the image coding technique to reduce the colour information in 256×256 still images to 2 bits per pixel. Successful image transmission was achieved over a vertical path of 3.5 km using 4-Phase Shift Keying (PSK) modulation and a 25 kHz carrier for a system's bandwidth of 10 kHz. The bit rate achieved was 20 kbps and the BER was less than 10^{-4} at a Signal-to-Noise Ratio (SNR) of approximately 20 dB.

In 1991, some researchers implemented a vertical-path image acoustic transmission system which was successfully tested over a water depth of 2 km [12]. The system used a carrier frequency of 53 kHz and achieved a bit rate of 19.2 kbps at a BER of 10^{-4} using Differential Phase Shift Keying (DPSK) modulation. The Joint Photographic Experts Group (JPEG) standard was used as the image compression technique.

The Advanced System Integration for Managing the Coordinated Operation of Robotic Ocean Vehicles (ASIMOV) project is another underwater image transmission system that was implemented by Portuguese scientists in 2000 [13]. The system operated at a carrier frequency of 61 kHz and it was designed for strictly vertical transmission links whereby an Autonomous Surface Craft (ASC) moved at the surface alongside an Autonomous Underwater Vehicle (AUV). The ASC could in turn transmit the data it received from the AUV to a relay base-station (BS) via RF waves. Tests carried out at an approximate depth of 50 m showed that a sequence of still pictures could be transmitted at 2 frames per second (fps) at a bit rate of 30 kbps using 8-PSK modulation. The target BER was set below 10^{-3} for successful image transmission.

Another experimental underwater video transmission system based on the Moving Picture Experts Group (MPEG)-4 standard was developed in Japan in 2002 [14]. Video of resolutions 640×480 and 174×144 were transmitted at a frame rate of 10 fps . The system was tested using 4-PSK, 8-PSK and 16-Quadrature Amplitude Modulation (QAM) modulation schemes over shallow water of vertical link distance of 30 m and the maximum bit rate achieved was 128 kbps with 16-QAM for a system's bandwidth of 40 kHz (with a carrier frequency of 100 kHz). In terms of error rate, a BER of 10^{-5} was obtained with 16-QAM modulation at a SNR of approximately 21 dB.

One year later an experiment was performed to investigate the transmission of video sequences over a short vertical link of 10 m using 8-PSK and 16 to 64-QAM modulation schemes [8]. Two transducers operating in a 60 kHz to 90 kHz bandwidth were attached to a pole with a 10 m separation between them and the pole was vertically submerged in water. The receiver was submerged at a depth of 2 m under the sea surface while the transmitter was at the opposite lower end of the pole. The video compression method was based on DCT and Huffman for entropy coding. By transmitting at 25000 symbols per second and using a carrier frequency of 75 kHz, the results showed that the highest bit rate was achieved using 64-QAM (150 kbps) as compared to 8-PSK (75 kbps), 16-QAM (100 kbps), and 32-QAM (125 kbps). These rates were sufficient to transmit a video clip at a resolution of 144×176 pixels at 15 fps in real-time. For all the modulation schemes, a Decision Feedback Equaliser (DFE) based on the Recursive

Least Squares (RLS) algorithm allowed the video clip to be received without any bit error.

1.3.2 Horizontal Transmission

Over the past few years, experimental studies have mostly considered horizontally-configured UACs. Using Single-Input Single-Output (SISO)-OFDM, reliable communication with rates of up to tens of kbps have been achieved. For instance, in [15], an overall bit rate of 30 kbps (with zero symbol error) was achieved over a horizontal transmission range of 2500 m using a 1024-subcarrier OFDM system. Quadrature Phase Shift Keying (QPSK) modulation was used over a 24 kHz acoustic bandwidth (22 kHz to 46 kHz). The water depth was 12 m and the transmitter and 12-element receiver array were both submerged at a depth of 6 m.

Ribas et al. [16] designed a system based on the MPEG-4 compression standard and OFDM modulation. Using a system's bandwidth of 115 kHz, 8-PSK modulation, 16384 subcarriers and BCH (63,18) codes, a maximum bit rate of 90 kbps was achieved over a horizontal link distance of 200 m with no bit errors. A video of resolution 176×144 (compressed to 64 kbps) could be transmitted in real-time at 30 fps over the horizontal link. Vall et al. [17] later improved on the error robustness of the video transmission system in [16] by incorporating a Doppler compensation algorithm in the OFDM system and using the error resiliency tools offered by the MPEG-4 standard.

MIMO-OFDM was found to be an attractive solution for the dynamic and bandwidth-limited UACs to improve the reliability [18–21]) and data rate (e.g., [22–24]). For instance, the authors in [19] applied Alamouti space-frequency block coding to a MIMO-OFDM system (to exploit diversity gain) using a range of subcarriers from 64 up to 1024. Channel estimation was performed using two methods, namely, Orthogonal Matching Pursuit (OMP) and least squares with adaptive threshold. The horizontal transmission link distance varied between 3500 m and 7000 m. The water depth was 100 m and the 2-element array transmitter and 12-element array receiver were submerged at a depth of 40 m and 50 m, respectively. The system's performance was evaluated over a 5 kHz acoustic bandwidth (10-15 kHz) using QPSK modulation. A maximum bit rate of 8.7 kbps was achieved at a BER of approximately 10^{-4}.

In [22], a spatially multiplexed MIMO-OFDM system was used to improve the bit rate over an UAC. A real sea experiment was carried out over a horizontal transmission distance of 450 m. A 2-element array transmitter and 6-element array receiver were each submerged approximately 6 m below a surface buoy and boat, respectively. A high acoustic bandwidth of 62.5 kHz was used. The number of subcarriers in the MIMO-OFDM system was set at 4096. Pilot-based channel estimation was also considered. Using 16-QAM and $\frac{1}{2}$-rate Low-Density Parity-Check (LDPC) code, a bit rate as high as 125.7 kbps was achieved with a zero BER.

A few other studies which have considered horizontal transmission and have achieved reasonable bit rates over long range acoustic links are summarised in Table 1.1.

Table 1.1: Bit Rates Achieved in UWA Communication Studies over a Horizontal UAC.

Reference	Bandwidth (kHz)	Water Depth (m)	Range (m)	BER	Bit Rate (kbps)
[25]	40	1600	680	≈ 0	120
[24]	10	15	1000	$\approx 10^{-4}$	31.25
[26]	10	90	2000	$\approx 10^{-3}$	32
[20]	23	70	2000	≈ 0	48

1.3.3 Commercial and Scientific UWA Communication Systems

In November 2007, Woods Hole Oceanographic Institution (WHOI) tested an underwater communication system, entitled Nereus, which provided real-time control and video transmission for ROVs [27]. Nereus was able to operate at depths of up to 11 km. In fact, it was designed to perform an expedition in the Challenger Deep which is the deepest part surveyed in the ocean. The system used both acoustic and optical waves to achieve a data rate as high as 10 Mbps. Nereus was in fact an adaptive system which could switch between different bandwidths depending on the application it was used for. In an experiment, Nereus was submerged to a depth of 700 m using acoustic control and a tracking system was used [28]. An optical transmitter was attached on a cable beneath the surface vessel. Whenever Nereus came within the range of the optical transmitter, the optical receiver on the ROV automatically captured the signal and High-Definition (HD) video was transmitted to the surface ship. Through this video, the ship operator was able to manipulate the ROV's robotic arm using acoustic control. In May 2014 Nereus was lost at sea during an expedition in the Kermadec Trench at a depth of 9.9 km. It was later reported that the vessel imploded due to the extreme pressure based on the debris that were retrieved.

In an online article published in 2008, Sea-Eye Underwater Ltd claimed that they could achieve high data rate real-time underwater wireless video transmission over an initial target distance between 100 m and 200 m using an acoustic modem [29]. The latter was designed to cope with underwater noise, multipath distortion and Doppler effect. A maximum data rate of 200 kbps was reported using high frequencies which lie between 500 kHz and 1 MHz. The transmission distance was planned to be extended further in the range between 300 m to 500 m.

BaltRobotics company [30] developed an UWA video communication prototype for underwater vehicles that makes use of an efficient data compression technique as well as a highly efficient modulation method in order to achieve real-time video transmission. The system operates within a bandwidth of 80 kHz at a maximum depth of 200 m. The achievable bit rate is 128 kbps (with a bandwidth efficiency of 1.8 b/s/Hz). The system transmits Standard-Definition (SD) colour video at a resolution of 640×480 pixels at 15 fps. The raw video compression is performed at a rate of 0.02 bits per pixel. The system can also perform 3-Dimensional (3D) modelling and object recognition in an underwater environment.

The main players in the UWA communication market in the United Kingdom include companies such as Sonardyne International Ltd [31], Aquatec Group Ltd [32] and Tritech International Ltd [33]. The Aquatec Group designs and builds acoustic modems tailored to the customer's specifications in order to achieve optimised underwater communication performance and as such the acoustic modems are not

Table 1.2: State-of-the Art UWA Modems.

Manufacturer and Model	Range (m)	Operating Frequency (kHz)	Max. Bit Rate (kbps)
SDAM [37]	200	76.2-173.8	260
FAU-HERMES [38]	120	262-375	87.768
EvoLogics S2CM HS [39]	300	120-180	62.5
LinkQuest UWM2200 [40]	1000	53.55-89.25	35.7
EvoLogics S2CR 48/78 [41]	1000	48-78	31.2
Sercel MATS 3G [42]	5000	30-39	24.6

available as off-the-shelf products. The company also provides short range optical modems for command and data transfer as well as custom-made system and instrument designs for subsea communication and carrying out specific underwater measurements. Sonardyne International and Tritech International also provide custom engineering solutions for unmanned underwater vehicles in terms of navigation including obstacle avoidance, target recognition, tracking, positioning and acoustic communication. In terms of academic research, the University of York [34] and Newcastle University [35] have been very active in many aspects of UWA communication over the past few years. These include the UAC modelling, physical layer signalling, full-duplex communication, routing protocols and localisation of nodes in underwater networks, among others. The Newcastle University has even commercialised some of its pioneering research in UWA communication. For example, the university's latest research in spread spectrum acoustic technology has been incorporated in the MicronNav system, manufactured under license by Tritech International Ltd, for tracking underwater ROVs [36].

Most commercial acoustic modems provide bit rates that are adequate for navigation and control but are not sufficiently high for the transmission of real-time video with satisfactory quality. Table 1.2 summarises some of the state-of-the art acoustic modems which are currently available for commercial and non-commercial (research and experimental) purposes.

1.4 Thesis Organisation

Chapter 2 provides an overview of the UAC characteristics along with their mathematical expressions and describes how the channel is generated. This chapter also covers the basics of video compression using H.264/Advanced Video Coding (AVC).

In Chapter 3, new physical layer waveforms such as FBMC modulation are introduced and the benefits they can bring to UWA communication as compared to OFDM systems are investigated. In this chapter, the BER performance of a SISO system based on Filtered Multitone (FMT) and Orthogonal Frequency Division Multiplexing-Offset Quadrature Amplitude Modulation (OFDM-OQAM) is evaluated and compared to the conventional OFDM system in 200 m horizontally and vertically-configured time-invariant UACs. The maximum achievable bit rates with the different systems are also provided at the end of this chapter.

Chapter 4 presents the performance evaluation of coded and un-coded spatially-multiplexed MIMO-FBMC based on OFDM-OQAM modulation and MIMO-OFDM over 1 km time-varying horizontal and vertical UACs. Preamble-based channel estimation based on the Interference Approximation Method (IAM) is also considered in the simulations. Furthermore, video transmission is evaluated over the 1 km vertical UAC using the Turbo-coded MIMO OFDM-OQAM systems.

In Chapter 5, massive MIMO reception is proposed to improve the communication reliability over a 1 km vertically-configured time-varying UAC and at the same time boost the achievable bit rate in both Single-User (SU) and Multi-User (MU) systems. For the SU scenario, the performance of Turbo-coded massive MIMO systems based on FBMC modulation and OFDM are evaluated over the channel. To overcome the intrinsic imaginary interference in the conventional OFDM-OQAM system and making its application to MIMO straightforward, a modified FBMC system is considered whereby complex data symbols are transmitted instead of real-valued ones. The transmission of a 480p video which is compressed using the H.264/AVC standard is also evaluated over the UAC using the systems that achieve the highest bit rates. In some applications, several underwater vehicles may need to be deployed. To this end, massive MIMO is also investigated for a 4-user scenario in this chapter, where all the users (ROVs) share the same time and frequency resources to transmit their information to the surface receiver over the 1 km vertical UAC.

In the context of MU UWA communication, Appendix-A investigates spatially multiplexed MIMO Non-Orthogonal Multiple Access (NOMA) systems over a 1 km vertically-configured time-varying UAC. The performance evaluation of Turbo-coded NOMA-OFDM and NOMA-FBMC systems are evaluated for a 2-user scenario where both underwater ROVs use the same bandwidth to transmit their information to a surface vessel. The FBMC waveform is implemented as in Chapter 5 due to its low complexity and ease of application to MIMO.

In Chapter 6, video transmission is investigated using OFDM-based Orthogonal Time Frequency Space Modulation (OTFS) systems in an UAC. This chapter presents the mathematical formulations for both SU MIMO and MU massive MIMO OFDM-based OTFS systems where the ROVs are equipped with multiple transmitting elements. The performance evaluation of coded and un-coded spatially multiplexed MIMO and massive MIMO systems for SU and MU scenarios, respectively, are evaluated over a 1 km vertically-configured time-varying UAC. Furthermore, the performance of the systems is assessed in terms of their maximum achievable bit rates. The transmitted videos are compressed using the H.264/AVC standard and the received video quality is analysed in terms of Peak-Signal-to-Noise Ratio (PSNR). Channel estimation is first performed by multiplexing pilots in the frequency domain and the performance of the OTFS systems are evaluated using frequency domain equalisation. At the end of the chapter, it is also proposed to use pilots in the delay-Doppler (DD) domain and perform symbol detection in the same domain to further boost the bit rates over the UAC.

Finally, Chapter 7 summarises the thesis and provides recommendations for future work.

1.5 List of Publications and Submissions

1.5.1 Publications

- M. J. Bocus, A. Doufexi, and D. Agrafiotis, "MU-massive MIMO for UWA communication," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, August 2018, pp. 1–5
- M. J. Bocus, D. Agrafiotis, and A. Doufexi, "Non-orthogonal multiple access (NOMA) for underwater acoustic communication," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, August 2018, pp. 1–5
- M. J. Bocus, D. Agrafiotis, and A. Doufexi, "Underwater acoustic video transmission using MIMO-FBMC," in *2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*, May 2018, pp. 1–6
- M. J. Bocus, D. Agrafiotis, and A. Doufexi, "Real-time video transmission using massive MIMO in an underwater acoustic channel," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2018, pp. 1–6
- M. J. Bocus, A. Doufexi, and D. Agrafiotis, "Performance evaluation of MIMO-OFDM/OQAM in time-varying underwater acoustic channels," in *OCEANS 2017 - Anchorage*, Sept 2017, pp. 1–6
- M. J. Bocus, A. Doufexi, and D. Agrafiotis, "Performance evaluation of filterbank multicarrier systems in an underwater acoustic channel," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2016, pp. 1–6

1.5.2 Submissions

- M. J. Bocus, D. Agrafiotis, and A. Doufexi, "Investigating video transmission using OFDM-based MIMO OTFS systems for UWA communication," in *IEEE Journal of Oceanic Engineering*, 2019
- M. J. Bocus, A. Doufexi, and D. Agrafiotis, "On the performance of OFDM-based massive MIMO OTFS systems for UWA communication," in *IET Communications*, 2019

CHAPTER 2

OVERVIEW OF UNDERWATER COMMUNICATIONS AND VIDEO COMPRESSION

The UAC is characterised by roughly the same factors as in a terrestrial RF channel. However, the degree of severity of the factors is different between these two domains. This chapter provides an overview of the UAC characteristics along with their mathematical expressions and describes how the channel is generated. This chapter also covers the basics of video compression using H.264/AVC. The chapter is organised as follows: Section 2.1 presents a brief overview of the terrestrial RF channels. Section 2.2 describes the main characteristics of the UAC and gives a description of the UAC generation, which is used in most parts of this thesis. Section 2.3 introduces the basics of video coding, focusing on the H.264/AVC standard. A comparison between H.264/AVC and High Efficiency Video Coding (HEVC) standards for video transmission in an UAC is also carried out. Finally, Section 2.4 concludes this chapter.

2.1 Brief Overview of RF Channels

The terrestrial RF channel is affected by several propagation phenomena. The main one is the path loss which is broadly defined as the loss in signal power as it propagates through the wireless medium. This loss depends on factors such as distance and wavelength. The greater the transmission distance, the lower is the received power. Signals with small wavelengths (i.e., high-frequency signals) suffer from more attenuation and can thus travel over a shorter distance than higher wavelengths signals.

The types of fading that occur in a RF channel can be grouped into fast fading and slow fading. The latter is mainly caused by distant objects, terrain configuration such as mountains or man-made structures such as tunnels and buildings. This fading occurs over relatively long periods of time and may cause the signal to be absorbed or reflected thereby reducing its strength in addition to the path loss. Short-term

Table 2.1: Typical Bandwidth Values for Different Transmission Ranges in an UAC [53].

	Transmission distance (km)	**Bandwidth (kHz)**
Very Short	<0.1	>100
Short	between 0.1 and 1	20-50
Medium	between 1 and 10	10
Long	between 10 and 100	2-5
Very Long	>100	<1

fading refers to the rapid fluctuations in the signal (fast fading) due to multipath effect [51]. The main causes of fading include multipath reflections, diffraction and scattering. This means that the RF signal arrives at the receiver through many paths (with different phases and amplitudes) at different instants in time (delays) [52]. Multipath propagation of a signal causes delay spread and frequency-selectivity in the time and frequency domains, respectively. When either or both the transmitter and receiver are moving or the objects lying in the path of the signal are in motion, then the multipath arrivals change over time. This is referred to as time-selectivity and Doppler shift/spread which occur in the time and frequency domains, respectively. Rayleigh and Rician statistics are commonly used to model the wireless channel in terrestrial RF communication systems without and with a line of sight, respectively. Noise in a terrestrial RF channel is often modelled as Additive White Gaussian Noise (AWGN).

2.2 Underwater Acoustic Channel Characteristics

Depending on the range of transmission, an UAC can be classified as very short, short, medium, long and very long [4] and the bandwidth will typically vary from a few Hz to hundreds of kHz. The typical available bandwidth for different transmission distances is shown in Table 2.1. The UAC can also be classified as horizontal or vertical depending on the direction of transmission of the acoustic waves with respect to the sea floor [4]. Horizontal channels may exhibit long multipath delay spreads, in the order of tens to even hundreds of milliseconds (ms) [7], depending on the water depth and distance between the transmitter and receiver. On the other hand, carefully positioned vertical channels tend to suffer less from multipath propagation [4, 8]. A water depth under 100 m is usually referred to as shallow water whereas deep water refers to higher depths [4]. Depending on the type of channel configuration, factors that influence an UWA communication system include transmission loss, propagation delay, ambient noise, multipath propagation and Doppler effect.

2.2.1 Path Loss

The path loss is dependent on both the transmission distance and frequency of communication. It is mainly caused by geometrical spreading, scattering and absorption. Geometric spreading can be broadly defined as the loss in power as an acoustic wave propagates away from the source [54]. As the distance of propagation increases, the sound energy covers a larger surface area causing the sound intensity

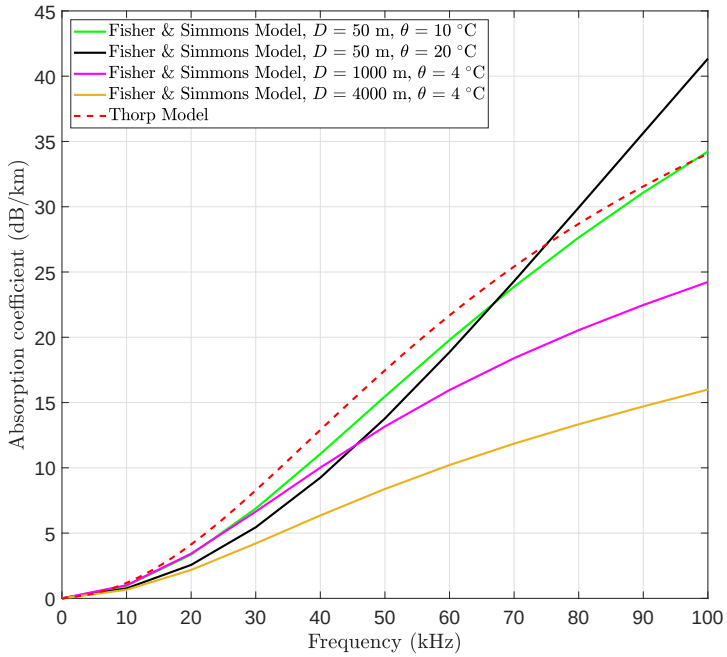


Figure 2.1: Plot of absorption coefficient in dB/km.

to decrease. For spherical waves emitted from a point source, the loss in power due to geometrical spreading is proportional to the square of the distance whereas for cylindrical waves which are radiated from a long line source, the loss in power is proportional to the distance travelled [54]. Scattering refers to the phenomenon whereby an incident ray is reflected in many directions due to uneven surfaces. In an underwater environment, scattering mainly occurs when the acoustic waves interact with obstacles such as sea creatures, bubbles and uneven sea surfaces and sea floors [54]. The scattering loss is dependent on the acoustic wavelength, whereby the loss increases with decreasing wavelength [54]. High frequency acoustic signals are attenuated to a larger extent compared to low frequency signals for a given distance. The transmission loss (in dB) can be expressed as [3]

$$10\log A(d,f) = k_{\text{gs}} \times 10\log d \;+\; d \times 10\log \alpha(f) \tag{2.1}$$

where $\alpha(f)$ is the absorption coefficient in dB/km for a frequency f, d is the transmission link distance, k_{gs} denotes the frequency-independent geometrical spreading factor which has a value of 1 and 2 for cylindrical and spherical spreading, respectively [7]. A practical value of $k_{\text{gs}} = 1.5$ is often considered [4]. $\alpha(f)$ can be computed in dB/km using a simple model known as the Thorp model [55] which only depends on the frequency and is given by (for f in kHz)

$$\alpha(f) = \frac{0.11 f^2}{1+f^2} + \frac{44 f^2}{4100+f^2} + 2.75\times10^{-4} f^2 + 0.003. \tag{2.2}$$

Equation (2.2) is valid for frequencies above a few hundreds Hz [4]. For lower frequencies, the absorption coefficient $\alpha(f)$ is expressed as [4]

$$\alpha(f)=\frac{0.11f^2}{1+f^2}+0.011f^2+0.002. \tag{2.3}$$

The Fisher and Simmons [56] model can alternatively be used to calculate the absorption coefficient. This model takes into account the effects of pressure, pH, salinity, temperature and relaxation frequencies due to the chemical components, namely boric acid and magnesium sulphate, and is valid for the frequency range of 100 Hz$<f<$1 MHz [57]. The complete expression for computing $\alpha(f)$ (in dB/km) using the Fisher and Simmons model is

$$\alpha(f)=\frac{\mathbb{A}_1\mathbb{P}_1\mathbb{F}_1f^2}{\mathbb{F}_1^2+f^2}+\frac{\mathbb{A}_2\mathbb{P}_2\mathbb{F}_2f^2}{\mathbb{F}_2^2+f^2}+\mathbb{A}_3\mathbb{P}_3f^2, \tag{2.4}$$

where the terms with subscripts 1, 2 and 3 account for the dependence of the chemical relaxation of boric acid, magnesium sulphate and the water viscosity on the absorption coefficient, respectively. $\mathbb{F}_1$ and $\mathbb{F}_2$ correspond to the relaxation frequencies of boric acid and magnesium sulphate, respectively, while $\mathbb{P}_1$, $\mathbb{P}_2$ and $\mathbb{P}_3$ denote the pressure dependencies of the three corresponding components. $\mathbb{A}_1$, $\mathbb{A}_2$ and $\mathbb{A}_3$ were originally assumed to be constants but were later demonstrated through experimentations in [58, 59] that they can vary with the water properties. The coefficients in equation (2.4) are computed as follows [58, 59]

Boric Acid Contribution

$$\mathbb{A}_1=\frac{8.86}{c}\times10^{(0.78\text{pH}-5)} \text{ [in dB/(km. kHz)]} \tag{2.5}$$

$$\mathbb{P}_1=1 \text{ [in atm]} \tag{2.6}$$

$$\mathbb{F}_1=2.8(S/35)^{0.5}\times10^{4-\frac{1245}{273+\theta}} \text{ [in kHz]} \tag{2.7}$$

Magnesium Sulphate Contribution

$$\mathbb{A}_2=21.44\frac{S}{c}(1+0.025\theta) \text{ [in dB/(km. kHz)]} \tag{2.8}$$

$$\mathbb{P}_2=1-1.37\times10^{-4}D+6.2\times10^{-9}D^2 \text{ [in atm]} \tag{2.9}$$

$$\mathbb{F}_2=\frac{8.17^{(8-\frac{1990}{273+\theta})}}{1+0.0018(S-35)} \text{ [in kHz]} \tag{2.10}$$

Pure Water Contribution

$$\mathbb{P}_3=1-3.83\times10^{-5}D+4.9\times10^{-10}D^2 \text{ [in atm]} \tag{2.11}$$

$$\mathbb{A}_3=4.937\times10^{-4}-2.59\times10^{-5}\theta+ 9.11\times10^{-7}\theta^2-1.50\times10^{-8}\theta^3 \text{ (for } \theta\leq20^\circ\text{C) [in dB km}^{-1}\text{ kHz}^{-2}] \tag{2.12}$$

$$\mathbb{A}_3=3.964\times10^{-4}-1.146\times10^{-5}\theta+1.45\times10^{-7}\theta^2- 6.50\times10^{-10}\theta^3 \text{ (for } \theta>20^\circ\text{C) [in dB km}^{-1}\text{ kHz}^{-2}] \tag{2.13}$$

where c is the speed of sound in water in m/s, D is the water depth in m, S is the salinity of sea water in parts per thousand, θ is the temperature in °C, and f is the frequency in kHz. Figure 2.1 shows the graph of the absorption coefficient for a given frequency range, which is computed using both the Thorp and Fisher and Simmons models. It can be seen that the absorption coefficient increases with increasing frequency. Furthermore, it can be observed from Figure 2.1 that factors such as water depth and temperature have an effect on the absorption coefficient.

2.2.2 Propagation Speed

The speed of sound in water is approximately five orders of magnitude less than the speed of electromagnetic waves in air. Therefore the UAC experiences a longer propagation delay/latency compared to terrestrial RF links. Moreover, in higher layer protocols where there needs to be a handshaking or acknowledgement process between the transmitter and receiver, the low speed of sound may cause a loss of transmission efficiency when retransmissions occur and this severely impacts the system's throughput [4]. A reasonable approximation of the speed of sound (c) in seawater as a function of the depth (D) in meters, temperature (θ) in °C, and salinity (S) in parts per thousand is as follows [60]

$$\begin{aligned} c &= 1448.96 + 4.591\theta - 0.05304\theta^2 + 0.0002374\theta^3 \\ &+ 1.340(S–35) + 0.0163D + 1.675\times10^{-7}D^2 \\ &- 0.01025\theta(S–35) - 7.139\times10^{-13}\theta D^3. \end{aligned} \tag{2.14}$$

Equation (2.14) is valid for -2 $\leq \theta \leq$ 30°C, 25 $\leq S \leq$ 40 in parts per thousand and 0 $\leq D \leq$ 8000 m. Both the pressure and temperature are normally constant in shallow water, and so is the speed of sound [7]. However, as the water depth increases in a region known as the main thermocline (c.f., Figure 2.2), the temperature falls rapidly and the speed of sound decreases. This is because in the thermocline region, the increase in pressure is not high enough to actually affect the speed of sound. Beyond the thermocline region, which is known as the deep isothermal layer, the pressure increases further and the water temperature reaches a constant value of approximately 4°C [7]. In this region, the increasing pressure causes the speed of sound to increase.

2.2.3 Underwater Ambient Noise

Noise which occurs underwater can be either man-made or natural [4]. The latter is mainly caused by factors such as waves, wind, currents and tides or biological and seismic activities. Man-made noise is due to shipping activities or machinery noise [4]. The ambient noise sources include turbulence, breaking waves, distant shipping and thermal noise. These are defined by the Empirical formulae (in dB re μPa per Hz where f is in kHz) [3]:

$$10\log N_{\text{tb}}(f) = 17 - 30\log(f)$$

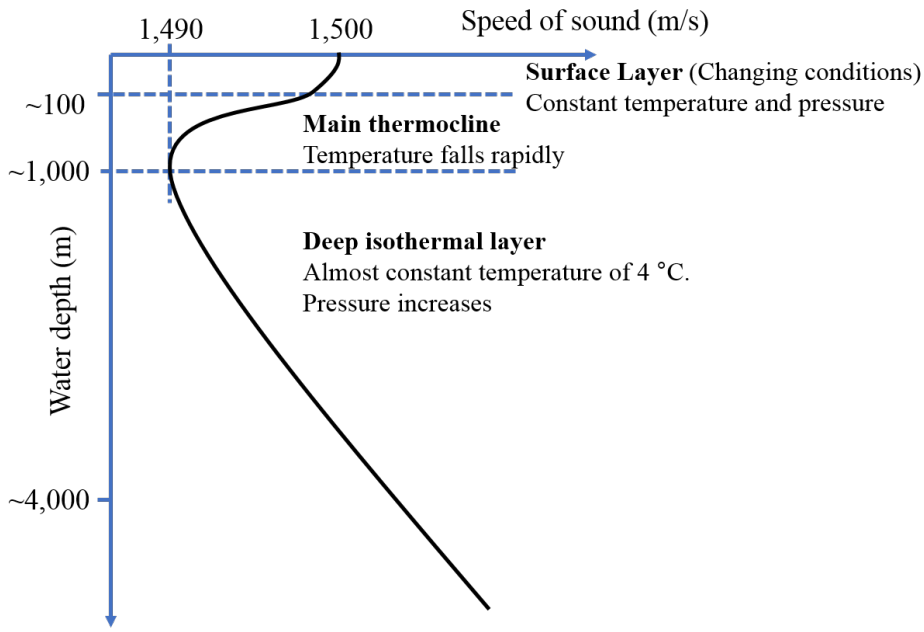


Figure 2.2: Underwater sound speed profile.

$$
\begin{aligned}
10\log N_{\mathrm{s}}(f) &= 40 + 20(s - 0.5) + 26\log(f) - 60\log(f+0.03) \\
10\log N_{\mathrm{w}}(f) &= 50 + 7.5w^{0.5} + 20\log(f) - 40\log(f+0.4) \\
10\log N_{\mathrm{th}}(f) &= -15 + 20\log(f),
\end{aligned} \tag{2.15}
$$

where N_{tb} denotes the turbulence noise, N_{s} represents the shipping noise (s is the shipping factor which has a value between 0 and 1), N_{w} is the noise due to breaking waves which are caused by wind (w is the wind speed in m/s) and N_{th} is the thermal noise. The overall Power Spectral Density (PSD) of the ambient noise is given by (in μPa) [3]

$$N_{\mathrm{all}}(f)=N_{\mathrm{tb}}(f)+N_{\mathrm{s}}(f)+N_{\mathrm{w}}(f)+N_{\mathrm{th}}(f). \tag{2.16}$$

Figure 2.3 shows the overall PSD of ambient noise under different conditions of shipping factor and wind speed. While AWGN has a constant PSD over a given frequency range, the PSD of ambient noise varies over the frequency range as seen in Figure 2.3.

2.2.4 Signal-to-Noise Ratio (SNR)

The available bandwidth and transmission range for UWA communication depend on the SNR which is determined by the attenuation and noise level. The fact that the attenuation increases with frequency while the ambient noise decays with increasing frequency makes the SNR vary over the signal bandwidth [7]. The SNR can be computed as [3]

$$\mathrm{SNR}(d,f)=\frac{S_{\mathrm{tx}}(f)}{N_{\mathrm{all}}(f)A(d,f)}, \tag{2.17}$$

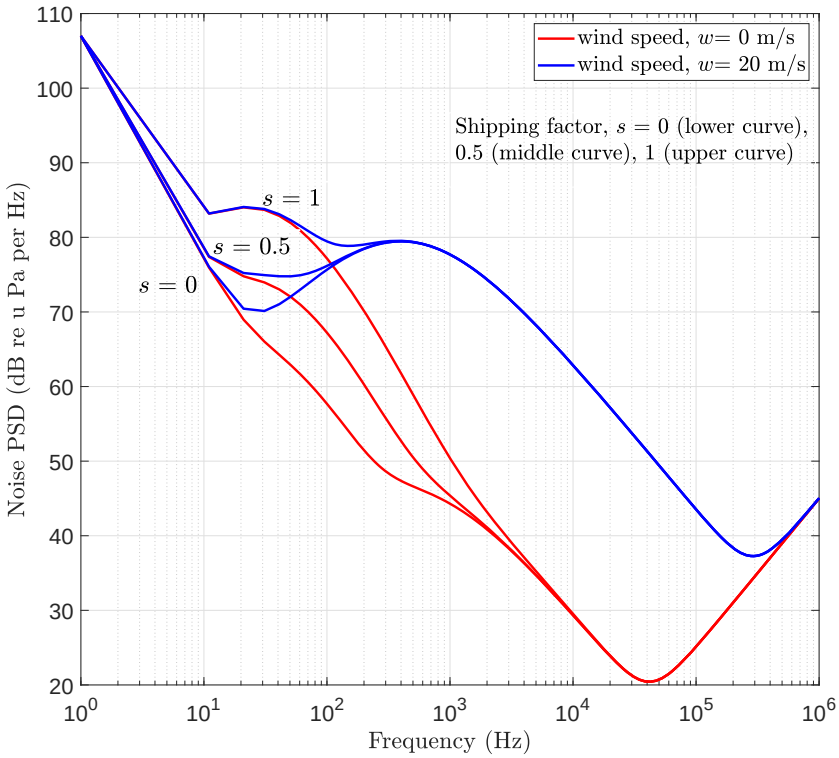


Figure 2.3: PSD of underwater ambient noise.

where $S_{\text{tx}}(f)$ is the PSD of the transmitted signal whose power can be adjusted to achieve the desired SNR level [7]. There is an optimum frequency for each transmitter-receiver separation at which the receive SNR is maximum. This can be observed in the graph of the inverse of the AN product as shown in Figure 2.4 for various transmission distances.

2.2.5 Multipath Propagation

Multipath propagation depends on the location of the transmitter and receiver in the transmission scenario. In deep water, multipath is primarily due to ray bending (or refraction because of the spatial variations in acoustic velocity [7]) while in shallow water multiple arrivals of the same signal are mainly due to bottom and surface reflections as illustrated in Figure 2.5. Multipath arrivals can severely affect the performance of an UWA communication system since they cause ISI. As can be observed in the horizontal channel configuration in Figure 2.5, the signal which is emitted from the transmitter (TX) arrives at the receiver (RX) in the form of a direct ray (denoted as DR) as well as multiple reflected rays which can be divided into four categories based on the type of reflection with the sea surface and/or sea bottom and the order of reflections, $\mathfrak{n}$ [61, 62]. In Figure 2.5 where $\mathfrak{n}$=1, SS_1 and BB_1 correspond to an acoustic ray which is reflected at the surface and sea floor, respectively, before it is captured by

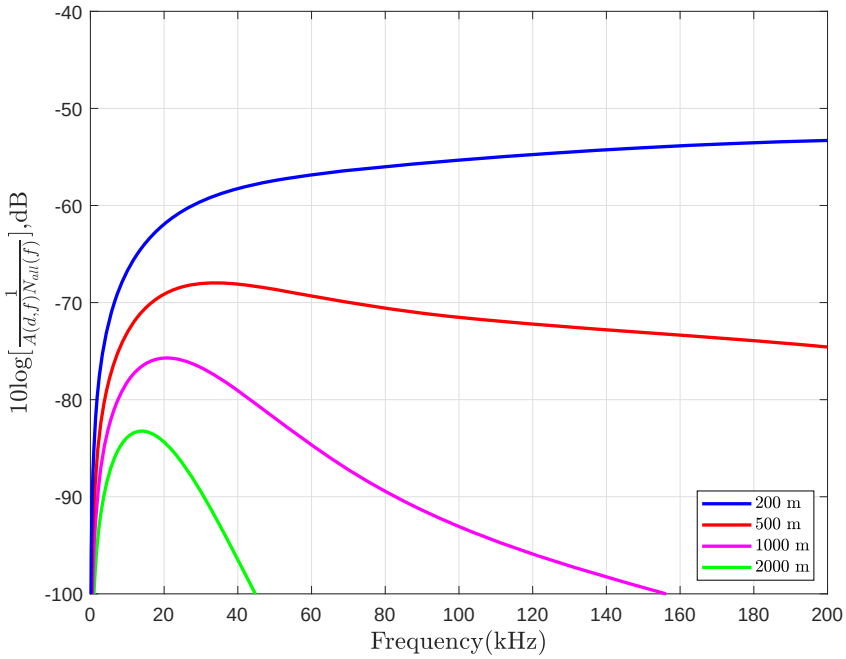


Figure 2.4: Inverse *AN* product graph to show how the SNR depends on both distance and frequency in an UAC.

RX. BS_1 represents a ray which is reflected off the sea floor and then the surface before it reaches RX. Lastly, SB_1 defines a ray which makes a surface reflection followed by a reflection off the sea floor and then terminates its course at the receiver. Assuming a constant speed of sound, the path length, angle of arrival and path delay of each ray can be calculated using the Lloyd's mirror geometry. The path length, d, of each ray can be computed as [61]

$$d=\sqrt{L^2+\mathscr{L}^2}, \tag{2.18}$$

and the angle of arrival of each ray at the receiver is given by

$$\psi=\aleph\tan^{-1}\left(\frac{\mathscr{L}}{L}\right), \tag{2.19}$$

where L is the horizontal separation distance between the transmitter (TX) and receiver (RX) as shown in Figure 2.5 and

$$\mathscr{L}=H_{\mathrm{RX}}-H_{\mathrm{TX}}, \qquad \aleph=-1, \qquad \text{for DR}, \tag{2.20a}$$

$$\mathscr{L}=2\mathfrak{n}D-H_{\mathrm{TX}}-H_{\mathrm{RX}}, \qquad \aleph=1, \qquad \text{for SS}_{\mathfrak{n}}, \tag{2.20b}$$

$$\mathscr{L}=2\mathfrak{n}D-H_{\mathrm{TX}}+H_{\mathrm{RX}}, \qquad \aleph=-1, \qquad \text{for SB}_{\mathfrak{n}}, \tag{2.20c}$$

$$\mathscr{L}=2\mathfrak{n}D+H_{\mathrm{TX}}-H_{\mathrm{RX}}, \qquad \aleph=1, \qquad \text{for BS}_{\mathfrak{n}}, \tag{2.20d}$$

$$\mathscr{L}=2(\mathfrak{n}-1)D+H_{\mathrm{TX}}+H_{\mathrm{RX}}, \qquad \aleph=-1, \qquad \text{for BB}_{\mathfrak{n}}, \tag{2.20e}$$

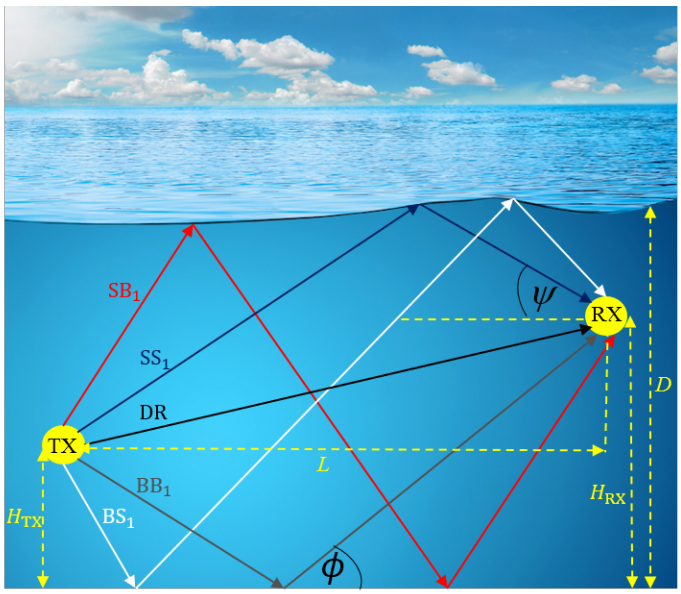


Figure 2.5: Illustration of multipath propagation in a horizontally-configured UAC.

where D is the water depth, H_{TX} and H_{RX} are the transmitter and receiver height from the sea floor, respectively. Substituting equations (2.20) into equation (2.18), the path length of each ray is given by

$$d_{\mathrm{DR}}=\sqrt{L^2+(H_{\mathrm{RX}}-H_{\mathrm{TX}})^2}, \tag{2.21a}$$

$$d_{\mathrm{SS_n}}=\sqrt{L^2+[2\mathfrak{n}D-H_{\mathrm{TX}}-H_{\mathrm{RX}}]^2}, \tag{2.21b}$$

$$d_{\mathrm{SB_n}}=\sqrt{L^2+[2\mathfrak{n}D-H_{\mathrm{TX}}+H_{\mathrm{RX}}]^2}, \tag{2.21c}$$

$$d_{\mathrm{BS_n}}=\sqrt{L^2+[2\mathfrak{n}D+H_{\mathrm{TX}}-H_{\mathrm{RX}}]^2}, \tag{2.21d}$$

$$d_{\mathrm{BB_n}}=\sqrt{L^2+[2(\mathfrak{n}-1)D+H_{\mathrm{TX}}+H_{\mathrm{RX}}]^2}. \tag{2.21e}$$

The frequency response of the ith path is given by

$$H_i(f)=\Gamma_i\Big/\sqrt{A(d_i,f)}, \tag{2.22}$$

where d_i is the length of the ith propagation path with an associated delay of $\tau_i=(d_i/c)$ (where c is the speed of sound in water) and Γ_i is the cumulative reflection coefficient for surface and sea floor reflections which can be computed as follows,

$$\Gamma_i=\gamma_{\mathrm{s}}^{n_{\mathrm{sr}}}\gamma_{\mathrm{b}}^{n_{\mathrm{br}}}(\phi_i), \tag{2.23}$$

where ϕ_i is the grazing angle of the ith propagation, n_{sr} and n_{br} denote the number of surface and bottom reflections, respectively, while γ_{s} and γ_{b} are the surface and bottom reflection coefficients, respectively. The surface reflection γ_{s} can be assumed to be -1 under ideal sea conditions (calm flat sea surface) [7]. For other sea conditions where scattering of the incident ray occurs, the surface reflection can be computed using the Beckmann-Spizzichino model as [63]

$$\gamma_{\mathrm{s}}=\frac{\left[1+\left(\frac{f}{\sqrt{10}\times 378w^{-2}}\right)^2\right]\Big/\left[1+\left(\frac{f}{378w^{-2}}\right)^2\right]}{10^{\frac{\left(1+\frac{(90-w)}{60}\right)\left(\frac{\phi}{30}\right)^2}{10}}}, \tag{2.24}$$

where w denotes the wind speed in Knots (1 Knot = 0.5144 m/s) , f is the carrier frequency in kHz and ϕ is the grazing angle at the surface in degrees. The bottom reflection coefficient γ_{b} is dependent on the grazing angle and the type of the sea floor (hard or soft) [7]. Assuming ideal plane-wave propagation conditions, the bottom reflection coefficient can be computed as [64]

$$\gamma_{\mathrm{b}}=\begin{cases}\frac{\rho_{\mathrm{b}}\sin\phi-\rho\sqrt{(c/c_{\mathrm{b}})^2-\cos^2\phi}}{\rho_{\mathrm{b}}\sin\phi+\rho\sqrt{(c/c_{\mathrm{b}})^2-\cos^2\phi}}, & \cos\phi\leq c/c_{\mathrm{b}}\\ 1, & \text{otherwise}\end{cases} \tag{2.25}$$

where ρ is the density of water while ρ_{b} and c_{b} are the density and speed of sound at the sea bottom respectively. Nominally, ρ=1000 g/m^3, ρ_{b}=1800 g/m^3 and c_{b}=1300 m/s [64]. The overall channel response in the frequency domain is given by [65]

$$H(f)=\sum_i H_i(f)e^{-j2\pi f\tau_i}, \tag{2.26}$$

and the corresponding impulse response is given by

$$h(t)=\sum_i h_i(t-\tau_i), \tag{2.27}$$

where $h_i(t)$ is the Inverse Discrete Fourier Transform (IDFT) of $H_i(f)$. Each path typically behaves as a low-pass filter [7]. A baseband model of the UAC with discrete multipath components can be represented as [3]

$$h(\tau,t)=\sum_i \mathfrak{A}_i(t)\delta(\tau-\tau_i(t)), \tag{2.28}$$

where δ is a Dirac delta function and $\mathfrak{A}_i(t)$ is the time-varying amplitude of the ith path. Equation (2.28) is characterised by frequency spreading and time spreading. The latter is due to multipath propagation which occurs in the channel and the total number of paths is determined by the reflections which arrive at the receiver with significant energy [66]. Frequency spreading is due to the time-variability of the quantities, $\mathfrak{A}_i(t)$ and $\tau_i(t)$. If the paths' amplitudes remain constant within a short data block, then $\mathfrak{A}_i(t)\approx\mathfrak{A}_i$. Motion at the transmitter and/or receiver results in a Doppler shift which is proportional to the ratio, $\xi=v_{\mathrm{rel}}/c$ [7]. ξ denotes the Doppler scale factor and v_{rel} is the relative velocity between the mobile nodes. The relative motion between the transmitter and receiver causes the path lengths in a multipath channel to vary and the time-variability of the path delays results in either a time dilation or compression of the signal, which is characterised by the time scaling factor $1-\xi$ ("+" sign if transmitter and receiver are moving towards each other) [66]. Thus, if we consider an input baseband signal of bandwidth B which is centred around 0 Hz, in the presence of motion the received signal (after carrier demodulation) will be observed as having a bandwidth $(1-\xi)\times B$ (Doppler spreading) centred around $-v_{\mathrm{rel}}\times f_c/c$ (Doppler shifting), where f_c is the carrier frequency [66]. Therefore, motion-induced Doppler shifting

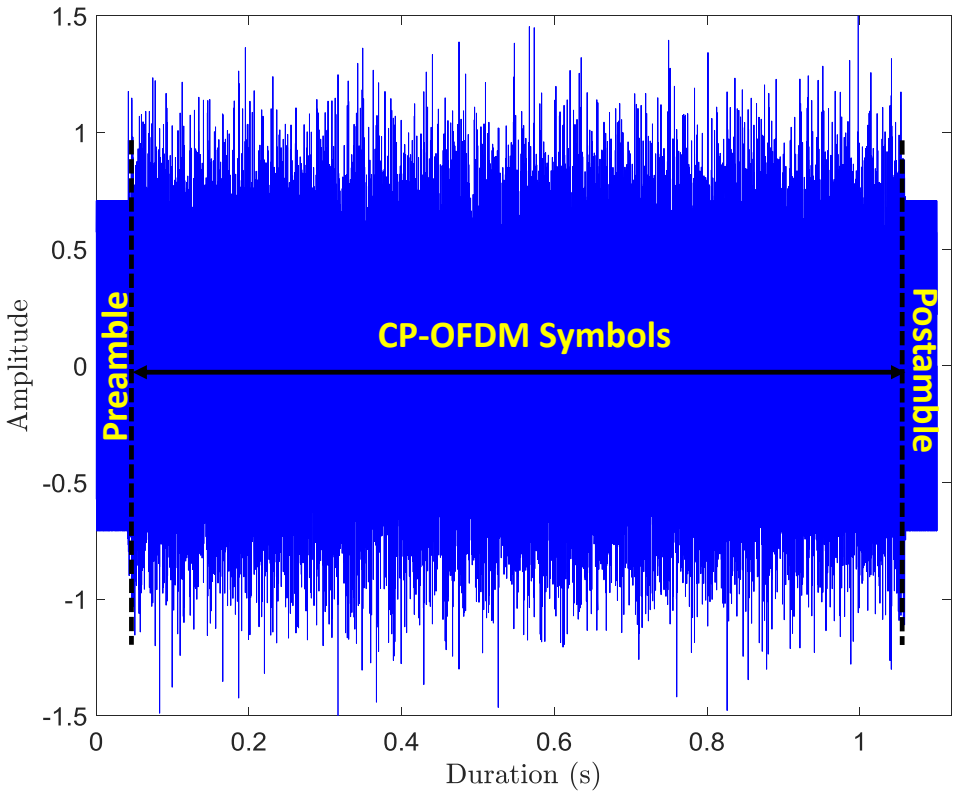


Figure 2.6: Transmitted CP-OFDM signal.

and spreading are both characterised by the factor, $\xi = v_{\text{rel}}/c$. The speed of moving ROVs is usually in the order of a few m/s. However, even when they are stationary, currents or tides can force them to move resulting in slight Doppler shifts. The low speed of sound in water makes motion-induced Doppler distortion non-negligible. Hence, for N_{mp} discrete paths, equation (2.28) can be re-written as [3]

$$h(\tau,t) = \sum_{i=1}^{N_{\text{mp}}} \mathfrak{A}_i \delta(\tau - [\tau_i - \xi_i t]). \tag{2.29}$$

If an underwater vehicle moves at several m/s, ξ_i will be in the order of 10^{-3}, which is a non-negligible value as compared to a highly mobile terrestrial system where the factor can be as low as 10^{-7} [7]. If all the propagation paths are assumed to have the same Doppler scale factor, i.e., $\xi_i \approx \xi$, then the expression in (2.29) can be simplified to

$$h(\tau,t) = \sum_{i=1}^{N_{\text{mp}}} \mathfrak{A}_i \delta(\tau - [\tau_i - \xi t]). \tag{2.30}$$

In this case, the common Doppler scale factor ξ can be simply removed through a resampling process at the receiver as described in the following subsection.

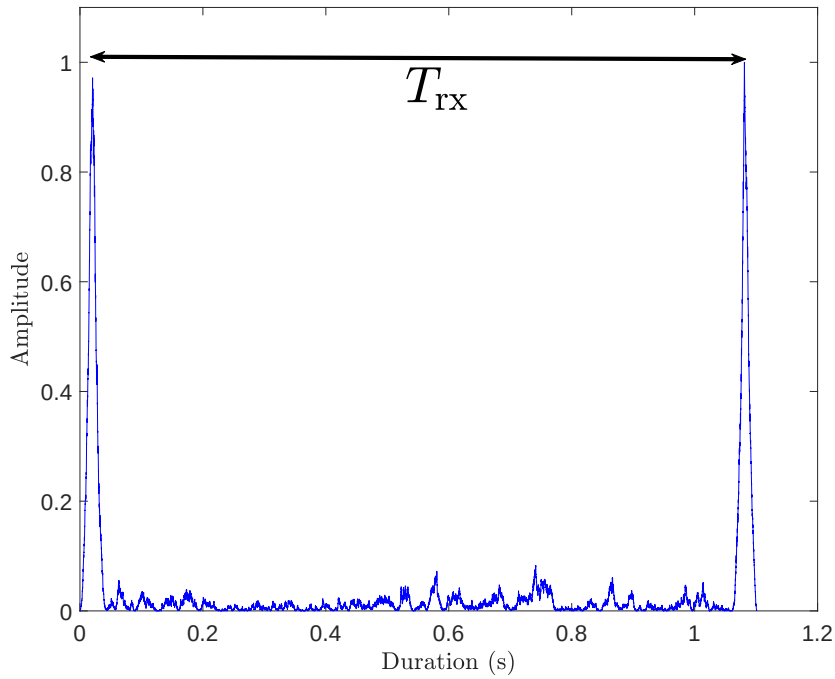


Figure 2.7: Cross-correlation result between received OFDM signal and known transmitted preambles.

2.2.5.1 Doppler Scale Estimation and Compensation

The multiple reflected copies of the transmitted signal, especially in a horizontal channel, are subjected to noticeable Doppler shifts when either or (both) the transmitter and receiver are moving [67]. Doppler shifts reduce the channel coherence time or in other words cause the channel to fluctuate more rapidly [68]. Doppler scale estimation can be achieved by inserting waveforms in the transmitted signal which are known at the receiver. One of the approaches is to use a training sequence (or preamble) which can be created using either Doppler-insensitive waveforms such as Hyperbolic Frequency Modulated (HFM) or Linear Frequency Modulated (LFM) waveforms or Doppler-sensitive waveforms such as Costa, m-sequence or polyphase sequence waveforms [69].

An example of a transmitted OFDM frame of duration 1.1 seconds is shown in Figure 2.6 where the sampling frequency is 100 kHz. In this example, 1024 subcarriers occupy a bandwidth of 25 kHz (OFDM symbol duration of 40.96 ms) and the CP duration is 10 ms. The number of OFDM symbols between the preamble and postamble is 20. The preamble/postamble duration is 40.96 ms. The preamble/postamble is made up of two identical halves which are generated using a pseudo-random m-sequence generator with a primitive polynomial of x^9+x^4+1. The maximal length sequence (also known as m-sequence) is a class of non-linear shift register sequences that has excellent autocorrelation characteristics. It is assumed that the receiver knows the initial time (or number of samples), T_{tx}, between the preamble and postamble. The OFDM signal is then transmitted through a multipath channel where all paths are

assumed to have a common Doppler scale factor ξ. This means that the signal is scaled in time from T to $T(1+\xi)$, where T is the OFDM symbol duration [69]. Time synchronisation is achieved through the cross-correlation of the received OFDM signal with the known transmitted preambles [69]. The preambles are usually chosen such that they show a high peak when the autocorrelation is performed, thereby allowing the receiver to identify the exact timing instant at which the preambles are received. The cross-correlation between the transmitted preambles and received OFDM signal is shown in Figure 2.7. The position of the peaks with the highest amplitudes are the timing instants at which the preambles are located. The time between the received preambles can be denoted as T_{rx}. The receiver then compares T_{rx} with T_{tx} to find an estimate of the Doppler scale factor as follows

$$\hat{\xi}=\frac{T_{\text{rx}}}{T_{\text{tx}}}-1. \tag{2.31}$$

The main Doppler effect is eliminated by resampling the received signal with the factor $(1+\hat{\xi})$, i.e., $y\left(\frac{t}{1+\hat{\xi}}\right)$, where $y(t)$ is the received passband OFDM signal. If ξ and $\hat{\xi}$ are not equal, each subcarrier will have a residual Doppler shift which can be denoted as $\hat{e}=\frac{\xi-\hat{\xi}}{1+\hat{\xi}}f_c$ where f_c is the carrier frequency [70]. So after compensating for the main Doppler scale factor and residual Doppler shifts, the signal can be expressed as [54]

$$\tilde{y}(t)=y\left(\frac{t}{1+\hat{\xi}}\right)e^{-j2\pi\hat{e}t}. \tag{2.32}$$

The above technique is suitable when the speed of motion between adjacent preambles is almost constant [17]. For sudden change in speed, the Doppler scale factor may vary across an OFDM frame [17], In this case an average of the Doppler scale factor is computed at the expense of a degraded performance. The method of transmitting a preamble and postamble has low complexity but the disadvantage is that the whole data frame should be received/buffered before the Doppler scale can be estimated and hence may cause a delay in receiver processing [54].

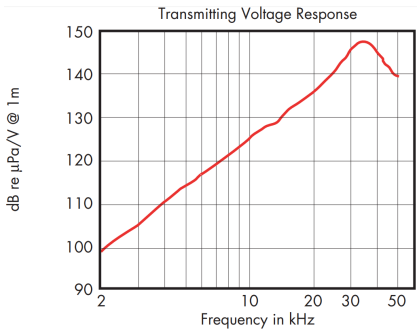
2.2.6 Transducer Characteristics

In most active UWA communication systems, the transducer can be used as a sound projector (transmitter) only, where the electrical signals are converted into pressure (acoustic) waves, or as a transmitter/receiver combination to both transmit and detect acoustic waves. When used at the receiver, the transducers serve as hydrophones to sense the transmitted sound waves and convert them into electrical energy. The underwater acoustic transducers are usually built using piezoelectric materials such as ceramics, polymers, single crystals and composites [71]. Different UWA transducers have different operational frequency range, but they are usually designed to operate at their resonance frequency to achieve the most efficient electrical to acoustical energy coupling [72], thereby maximising the output power. Therefore, for a target transmission scenario, after finding the optimum operating frequency band, a suitable transducer needs to be selected to ensure its frequency response is compatible with the chosen channel bandwidth in order to achieve the highest transmit/receive gain for the acoustic modem [37]. Note that an acoustic modem consists of three main parts; a transducer, an analogue transceiver system

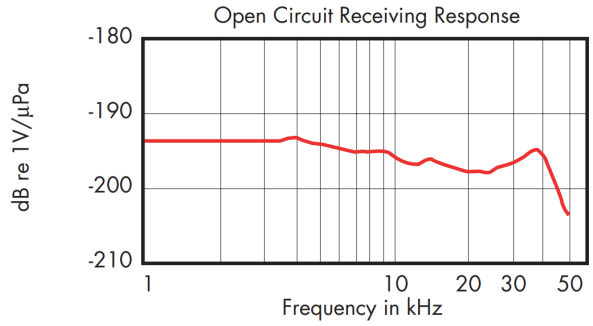
with an amplifier and pre-amplifier at the transmitter and receiver, respectively, and a Digital Signal Processing (DSP) platform, which is the core of the communication system. The transducer's electro-mechanical specifications usually include the Transmitting Voltage Response (TVR), Open Circuit Receiving Response (OCRR) and directivity pattern. An example of these specifications for the spherical ITC-1032 transducer, manufactured by Gavial-International Transducer Corporation (ITC), is shown in Figure 2.8. The other nominal specifications for the ITC-1032 transducer are provided in Table 2.2. The TVR (transmitting sensitivity) is basically defined as the measured sound pressure level (as a function of frequency) at a distance of 1 m away from the transducer (projector), per 1 V of input voltage generated by the transducer [72]. The TVR is measured in dB with a reference pressure value of 1 μPa and therefore its unit is specified as dB re μPa/V. The OCRR (receiving sensitivity) is defined as the output voltage (V) generated by the transducer when a pressure of 1 μPa is applied to it. The OCRR is expressed in dB re 1V/μPa and is a function of frequency. As can be observed in Figure 2.8(a), the performance of the transducer (projector) peaks at its resonance frequency (33 kHz) and declines rapidly as the transmission frequency moves away from the resonance point. Therefore, only a narrow frequency range can be used to achieve the best transmission performance. Similar to the TVR plot in Figure 2.8(a), a peak can be observed at the resonance frequency in the OCRR plot shown in Figure 2.8(b). However, unlike the TVR curve which has a steep slope below the resonance frequency, the OCRR frequency response, on the other hand, is relatively flat all the way down to ~0 Hz. Therefore, when the transducer is used as a hydrophone, the entire frequency range below the resonance point can be used. However, if a flatter frequency response is required over the entire operating frequency range for the communication scenario of interest, then another model (or brand) of transducer that satisfies this requirement may be used as hydrophone. Note that when choosing a hydrophone, it is desired that it has a flat frequency response over a wide frequency range, that is, the aim is to obtain the same amount of voltage per amount of acoustical excitation, regardless of the frequency. The ITC-1032 transducer provides uniform omnidirectional characteristics as shown in the directivity pattern in Figure 2.8(c). The transducers that are used in UWA communication systems usually have an omni-directional beam pattern in the horizontal plane to minimise reflections from the surface and sea-floor, especially in shallow water channels [72]. Once the most suitable hardware (in terms of transducer, power amplifier, voltage pre-amplifier, etc.) has been selected for a given UWA communication scenario, bandwidth-efficient physical layer schemes can be implemented at the DSP unit to maximise the use of the available bandwidth, especially for long-range links, and achieve a high bit rate while coping with the channel impairments.

2.2.7 Underwater Acoustic Channel Generation

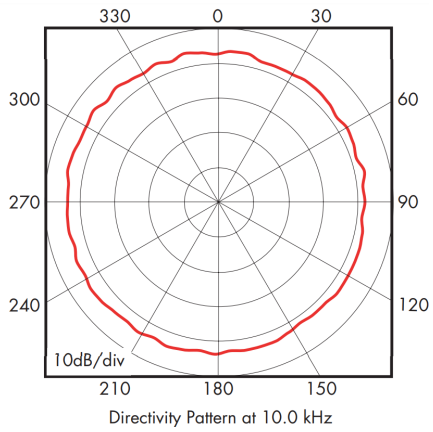
A multipath generator [74] is used to compute the number of reflections and their nominal path lengths, angles of arrival, cumulative reflection coefficient, path gains and delays for a given channel geometry (see Section 2.2.5). The lower band-edge frequency is used to determine the maximum number of paths as long as the path gains remain above a pre-defined threshold. Or alternatively, the total number of



(a) TVR



(b) OCRR



(c) Beam pattern

Figure 2.8: ITC-1032 transducer specifications [73].

Table 2.2: Nominal Specifications for ITC-1032 Transducer [73].

Type	Projector/Hydrophone
Resonance frequency, f_r	33 kHz
Depth rating	1250 m
TVR at f_r	149 dB re μPa/V @1m
Midband OCV	-194 dB re 1V/μPa
Suggested Band	0.01 - 50 kHz
Beam Type	Spherical
Input Power	800 watts

paths can be generated as long as the power difference between the strongest path and the weakest one is equal to or below a pre-defined value. The Bellhop acoustic software [75] can alternatively be used to generate the multipath profile for a given channel configuration as shown in Figure 2.11. Unlike RF channels, there is no standardised channel model for an UAC. Also research in this field depends heavily on expensive and time-consuming sea experiments [76]. Hence, a suitable channel model which is consistent with the real-world UAC should be used in order to design an UWA communication system.

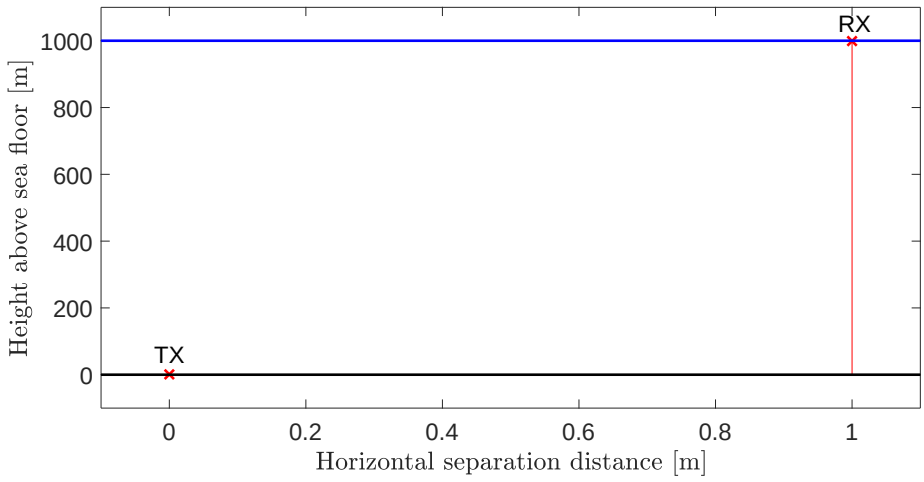


Figure 2.9: Vertical channel geometry.

In this regard, statistical modelling is often used since it considers the fast time fluctuations that occur in an UAC as a result of phenomena such as random movement between the transmitter and receiver or scattering [74, 76]. In this research work, a statistical channel model based on the principle of maximum entropy [76] is used to generate the time-varying channel coefficients, unless otherwise stated. This model is known to be accurate and consistent with a real UAC even if only partial information about the channel is known such as the channel Power Delay Profile (PDP) and Doppler spread. Thus, using the maximum entropy principle an accurate UAC model can be obtained without requiring comprehensive knowledge about the channel propagation phenomena. The statistical UAC model in [76] is described next.

A linear time-varying channel can be modelled by its impulse response in terms of a delay τ and time t, i.e., $h(\tau,t)$ or by its spreading function in terms of a delay τ and Doppler ν, i.e., $\mathcal{S}(\tau,\nu)$. The latter characterises the scatterer reflectivity and complex attenuation for all the paths with delay τ and Doppler ν [76]. For an input signal $x(t)$, the output $y(t)$ of the system can be expressed as

$$y(t)=\int_{-\infty}^{\infty} h(\tau,t)x(t-\tau)d\tau \tag{2.33}$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{S}(\tau,\nu)x(t-\tau)e^{j2\pi\nu t}d\tau d\nu, \tag{2.34}$$

where $\mathcal{S}(\tau,\nu)=\int_{-\infty}^{\infty} h(\tau,t)e^{-j2\pi\nu t}dt$. For a Wide-Sense Stationary Uncorrelated Scattering (WSSUS) channel, the scatterers with different delay or different Doppler are uncorrelated, i.e.,

$$\mathbb{E}\big[\mathcal{S}(\tau,\nu)\mathcal{S}^*(\tau',\nu')\big]=\mathcal{C}(\tau,\nu)\delta(\tau-\tau')\delta(\nu-\nu'), \tag{2.35}$$

where $\mathcal{C}(\tau,\nu)$ is known as the scattering function which can be expressed as

$$\mathcal{C}(\tau,\nu)=\mathbb{E}\big[|\mathcal{S}(\tau,\nu)|^2\big] \tag{2.36}$$

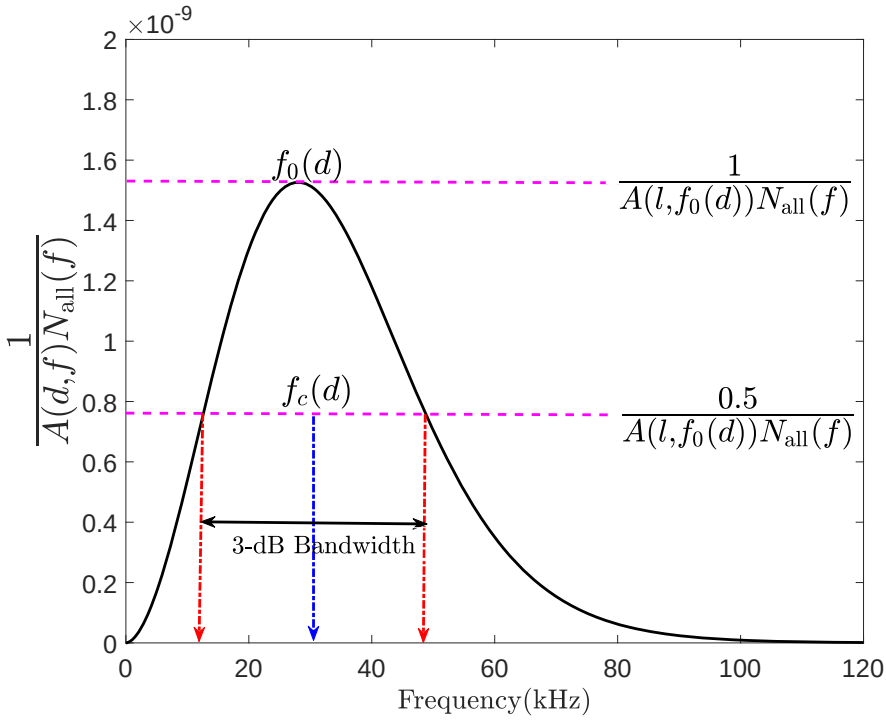


Figure 2.10: Operating frequency range and centre frequency for an UWA transmission distance of 1 km.

$$=\int_{-\infty}^{\infty}\mathbb{E}\big[h(\tau,t)h^*(\tau,t+\Delta t)\big]e^{-j2\pi\nu\Delta t}d\Delta t\cdot \tag{2.37}$$

The discrete-time baseband equivalent of the channel input-output relation in equation (2.33) can be written as

$$y[n]=\sum_{i=0}^{N_{\text{mp}}-1}h_i[n]x[n-i], \tag{2.38}$$

where $h_i[n]$, $x[n]$ and $y[n]$ are the sampled versions of $h(\tau,t)$, $x(t)$ and $y(t)$, respectively, and N_{mp} denotes the total number of discrete paths (taps) in the channel. The scattering function in equation (2.37) is expressed in discrete-time as follows

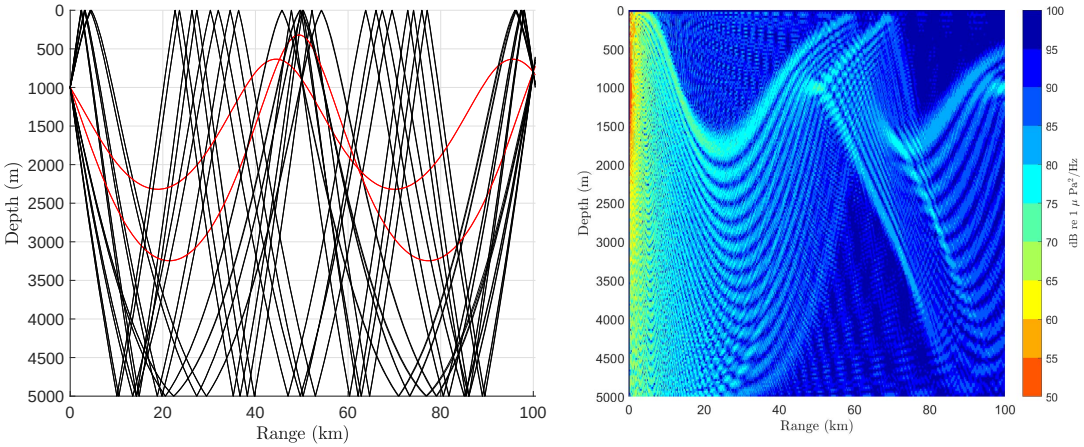
$$C_i(\nu)=\sum_{u=-\infty}^{+\infty}\mathbb{E}\big[h_i[n]h_i^*[n+u]\big]e^{-j2\pi\nu u}. \tag{2.39}$$

Note that $C_i(\nu)$ is discrete along the delay axis but continuous in Doppler. Assuming that the channel is modelled as a multi-variate random process where $H(n)=\big[h_0[n],\cdots,h_{N_{\text{mp}}-1}[n]\big]$, and considering that the only known information about the channel is that it has a total finite energy equal to σ^2, then the channel Probability Density Function (PDF) can be derived by maximising the entropy [76]

$$-\int \log P(H)dP(H), \tag{2.40}$$

given the constraints

$$\int dP(H)=1, \tag{2.41}$$



(a) (b)

Figure 2.11: Illustration of a Bellhop multipath profile for a given transmission scenario.

$$\int \sum_{i=0}^{N_{\text{mp}}-1} |h_i|^2 dP(H) = \sigma^2. \tag{2.42}$$

Using the Lagrange function, the maximisation problem can be solved as follows

$$\mathrm{L}(P) = -\int \log P(H) dP(H) + \alpha\left(1 - \int dP(H)\right) + \lambda\left(\sigma^2 - \int \sum_{i=0}^{N_{\text{mp}}-1} |h_i|^2 dP(H)\right), \tag{2.43}$$

and finding the functional derivative of $\mathrm{L}(P)$ and setting it zero yields the following expression [76]

$$-\log P(H) - 1 - \alpha - \lambda \sum_{i=0}^{N_{\text{mp}}-1} |h_i|^2 = 0, \tag{2.44}$$

where α and λ are known as the Lagrange multipliers. Using equations (2.41), (2.42) and (2.44), the PDF of the channel is given by

$$P(H) = \frac{1}{(\pi\sigma^2/N_{\text{mp}})^{N_{\text{mp}}}} \prod_{i=0}^{N_{\text{mp}}-1} e^{-\frac{N_{\text{mp}}|h_i|^2}{\sigma^2}}. \tag{2.45}$$

The expression in (2.45) represents independent and identically distributed (i.i.d) Gaussian distributed channel taps [76]. If the PDP of the channel is known such that each tap has an average power $\mathbb{E}\left[|h_i|^2\right] = \sigma_i^2 + |s_i|^2$ where $s_i = \mathbb{E}[h_i]$ (specular component), then the expression in (2.45) becomes [76]

$$P(H) = \prod_{i=0}^{N_{\text{mp}}-1} \frac{1}{(\pi\sigma_i^2)} e^{-\frac{|h_i - s_i|^2}{\sigma_i^2}}. \tag{2.46}$$

The Rice factor in this channel model is given by $|s_i|^2/\sigma_i^2$ and is exponentially decaying with the tap delay [76]. From (2.45) and (2.46), it can be inferred that the joint entropy of H can be formulated as the sum of the entropy of each h_i, and therefore the entropy maximisation of H is equivalent to the entropy

maximisation of each h_i separately [76]. Thus, the scattering function that maximises the entropy can be obtained by deriving the Doppler spectrum of each channel tap. For a discrete-time Gaussian process $h[n]$, the entropy rate $\bar{\mathbf{h}}$ is given by

$$\bar{\mathbf{h}}=\frac{1}{2}\log(2\pi e)+\frac{1}{2}\int_{-1/2}^{1/2}\log \mathcal{C}(\nu)d\nu, \tag{2.47}$$

where $\mathcal{C}(\nu)$ is the Doppler power spectrum with bounded normalised support $[\nu_{\min},\nu_{\max}]$ and $-1/2\leq \nu_{\min}\leq\nu\leq\nu_{\max}\leq 1/2$. The aim is to obtain the $\mathcal{C}(\nu)$ that maximises $\bar{\mathbf{h}}$ for each tap, subject to some constraints which express the state of knowledge about the time variations in the channel [76]. For each tap, this knowledge may be formulated in terms of a set of spectral moments, μ_p, of known functions $m_p(\nu)$, i.e.,

$$\mu_p=\int_{\nu_{\min}}^{\nu_{\max}} m_p(\nu)\mathcal{C}(\nu)d\nu, \tag{2.48}$$

where $p\in\mathcal{P}$ and $\mathcal{P}$ denotes a set of constraint orders. It is desirable to limit the parameters of the model to statistics of order 2 or less, i.e., $p\leq 2$, because such statistics become physically simple to interpret given that their order of magnitude is normally known in a particular environment [76]. It has been found in [76] that limiting the channel knowledge to second order statistics is enough to generate an accurate UAC model. In this case, μ_0 represents the total average power of $h[n]$, μ_1 denotes the barycentre of $\mathcal{C}(\nu)$ and μ_2 expresses the Root Mean Square (RMS) Doppler spread, D_s, of $h[n]$ such that $D_s=\sqrt{\mu_2/\mu_0-(\mu_1/\mu_0)^2}$ [76]. Geometrical moments can alternatively be used to express the state of knowledge about the time variations in the channel such that $m_p(\nu)=\nu^p$. Thus, using (2.47) and (2.48), the model for the Channel Scattering Function (CSF) can be derived by solving, for each tap, the following optimisation problem [76]

$$\begin{aligned} \max_{\mathcal{C}(\nu)} \quad & \int_{\nu_{\min}}^{\nu_{\max}} \log \mathcal{C}(\nu)d\nu, \\ \text{subject to} \quad & \int_{\nu_{\min}}^{\nu_{\max}} \nu^p \mathcal{C}(\nu)d\nu=\mu_p, \forall p\in\mathcal{P}. \end{aligned} \tag{2.49}$$

Similar to (2.43), the optimisation problem in (2.49) can be solved using the Lagrange function as follows

$$\mathrm{L}(\mathcal{C})=\int \log \mathcal{C}(\nu)d\nu-\sum_{p\in\mathcal{P}}\lambda_p\left(\mu_p-\int \nu^p\mathcal{C}(\nu)d\nu\right). \tag{2.50}$$

Differentiating $\mathrm{L}(\mathcal{C})$ with respect to $\mathcal{C}$ yields the following expression [76]

$$\mathcal{C}(\nu)=\frac{1}{\sum_{p\in\mathcal{P}}\lambda_p\nu^p}, \tag{2.51}$$

where the multipliers λ_p are obtained by solving the set of non-linear equations

$$\int_{\nu_{\min}}^{\nu_{\max}} \frac{\nu^q}{\sum_{p\in\mathcal{P}}\lambda_p\nu^p}d\nu=\mu_q, \forall q\in\mathcal{P}. \tag{2.52}$$

Consider a vertical channel geometry as shown in Figure 2.9 where the transmitter and receiver heights from the sea floor are 1 m and 999 m, respectively. The water depth is 1000 m. Based on the

transmission distance of approximately 1000 m, a suitable bandwidth can be obtained using the plot in Figure 2.10. $f_0(d)$ denotes the optimum at which maximum SNR is achieved at the receiver. The transmission bandwidth is chosen to be around $f_0(d)$ based on the 3-dB bandwidth [17]. In Figure 2.10, the carrier frequency is denoted as $f_c(d)$. The multipath calculator in [74] is used to generate the PDP and the statistical UAC model in [76] is used to generate the baseband Channel Impulse Response (CIR) for the channel configuration in Figure 2.9. The corresponding channel responses in terms of PDP, band-limited CIR, time-variant response and CSF are shown in Figure 2.12. The CSF characterises the average output power in the channel as a function of the delay τ and Doppler ν. In other words, it gives an indication of the rate of channel variations at different multipath delays. In this example, the maximum RMS Doppler spread in the UAC is 2 Hz. It should be noted that throughout this thesis, the simulations of the different systems are performed in a complex discrete-time baseband setting. Furthermore, the frequency response of the transducers (projectors/hydrophones) and their corresponding beam patterns are not considered in the simulations.

2.3 Video Coding and Transmission

Next-generation UWA communication systems are expected to support real-time video transmission from an unmanned underwater vehicle to a surface vessel [3]. It thus becomes increasingly important to reduce the gap between the supported bit rate of the UAC and that required for achieving video transmission with acceptable quality [3, 8]. To this end, efficient video compression techniques are necessary to fit the transmitted videos in the limited acoustic bandwidth without causing a severe degradation on the overall video quality. Some studies considering underwater image transmission have adopted the approach of transmitting a sequence of still images over a vertically-configured UAC (see [11–13] and references therein), where each image in the sequence is usually independently encoded using the transform domain based on DCT. Sub-band coding such as the Discrete Wavelet Transform (DWT) can alternatively be used for image coding [8]. In these methods, the DCT or DWT coefficients are first quantised and then the quantised levels are efficiently encoded such that each image is represented with as few bits as possible to achieve an acceptable frame rate. Real-time low resolution video transmission with satisfactory quality has also been achieved over a horizontal acoustic link of a few hundred metres using MPEG-4 standard (e.g., [16, 17]). Nowadays, better standards in the form of H.264 (MPEG-4 Part 10 AVC) and HEVC are available. The latter two standards improve on the previous MPEG-4 standard with higher bit rate savings and improved visual quality and are typically used in HD and Ultra High-Definition (UHD) applications. For similar subjective quality, HEVC aims to achieve a reduction in bit rate by approximately 50% compared to its predecessor H.264/AVC [77]. This aim is attainable since HEVC brings further enhancements and flexibility compared to previous standards in terms of partitioning, prediction of Motion Vectors (MVs), interpolation, de-blocking filters and parallel processing structures [78]. The interested reader is referred to [78] for an overview of HEVC.

In the following subsections, the basics of video coding and compression are covered and an overview

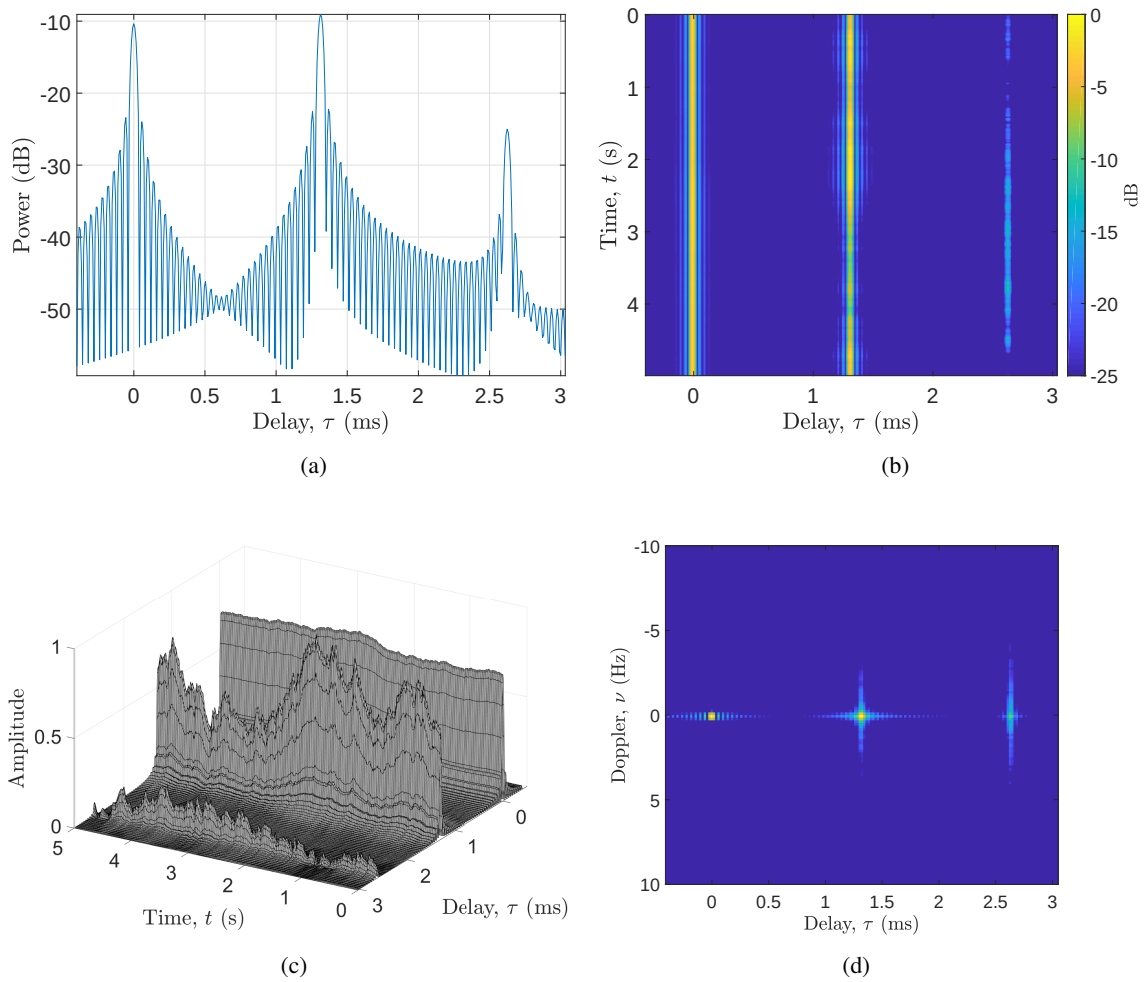


Figure 2.12: Example of channel responses for a vertical UAC of 1 km (a) PDP (b) band-limited CIR (c) time-variant channel response (d) CSF.

of the main features of the H.264/AVC standard is provided.

2.3.1 Basics of Video Coding

Video compression is the process whereby an original uncompressed video is processed in such a way that it occupies less storage space and minimises both the transmission cost and time. Transmission also involves sending the video and successfully delivering it to the receiver side by using loss protection and recovery techniques [79]. Lossy video coding standards exploit the redundancies that exist in the original video scene and also the imperfections in the human visual system to provide a representation such that the changes made to the video are imperceptible to the human eye. The main terminologies associated with a video are defined next.

The *resolution* is a common term associated with images and videos. It defines the number of

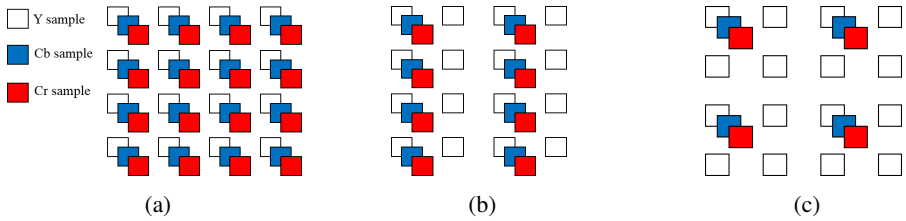


Figure 2.13: Illustration of chroma sub-sampling formats in YUV space (a) 4:4:4 (b) 4:2:2 (c) 4:2:0.

pixels (basic unit of composition or building blocks of a digital graphic) that can be displayed on the horizontal (width) and vertical (height) axes of a digital display device. For example, a computer display of resolution 800×600 is represented by 480,000 pixels, each of which is characterised by its (x,y) coordinates. Furthermore, in a colour image, a number of bits is used to represent each pixel and this is typically known as the *colour depth*. If, for example, a colour image is represented by 24 bits per pixel (bpp), the total number of different colours available is equal to 2^{24}=16,777,216. Therefore, the higher the resolution and the number of bpp in a colour image, the better is the quality at the expense of a larger image size. A video consists of a sequence of images which are displayed in a timely fashion to provide the effect of motion. The number of images or frames that are displayed in one second (fps) is known as the *frame rate* and this value dictates the smoothness of the motion. A high value implies a smoother motion at the expense of a lower compression rate. The term *bit rate*, measured in bits per second (bps), is also associated with a video and it gives an indication of the compression level of the video [80]. In other words, it defines the quality and size of the video, with a high value implying a higher video quality but large file size. The main principles involved in video coding and compression are introduced next.

2.3.1.1 Colour Representation

The human perception of colours is stimulated by electromagnetic waves where the visible spectral range include values from approximately 380 nm to 780 nm [79]. Colours which correspond to a single wavelength are known as *primary colours* and the cone cells in the human visual system are most sensitive to the three primary colours, namely, red, blue and green (RGB). The perception of different colours is achieved through the combination of these three primary colours. Since the human visual system is more sensitive to luminance (brightness) than chrominance (colours) [81], a more efficient way of representing colours is required. In this regard, the YCbCr colour space, also known as the YUV space, is used by most lossy image and video compression techniques, where more chroma components are discarded than luma ones without causing a reduction in the perceived quality. The Y-component (luma) in the YCbCr colour space corresponds to brightness whereas the components Cb (blue chrominance) and Cr (red chrominance), represent the amount of colour deviation from grey towards blue and red, respectively [82]. Video coding standards such as the MPEG-4 and H.264/AVC support three types of chroma sub-sampling formats, namely, 4:4:4, 4:2:2 and 4:2:0. These formats are illustrated in Figure

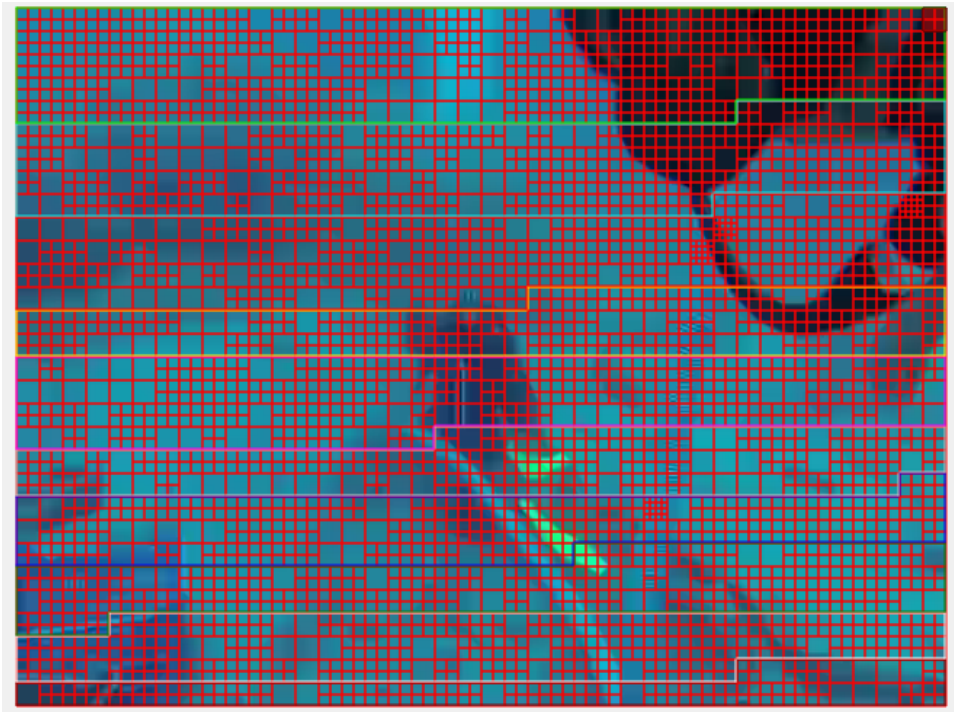


Figure 2.14: Illustration of macroblocks in a video frame.

2.13. The 4:4:4 format in Figure 2.13(a) shows that the luma and two chroma components, Cb and Cr, all have the same resolution. The 4:2:2 format in Figure 2.13(b) shows that the two chroma components are sub-sampled by a factor of two only in the horizontal dimension, while the 4:2:0 format in Figure 2.13(c) shows that the chroma components are sub-sampled by a factor of two in both the horizontal and vertical dimensions. Hence, by using the YCbCr colour space instead of the RGB system, video coding standards are able to represent a scene with less amount of data and this may be regarded as the simplest means to achieve compression.

2.3.2 Fundamentals of Video Compression

Over the last three decades, several video coding standards ranging from H.261 to H.265/HEVC, including the MPEG standards, have been proposed for video compression and each standard is expected to bring a number of improvements over its predecessor, especially in terms of compression efficiency. While these standards also have different computational complexity and implementation procedure, they are nevertheless made up of the same basic building blocks. Video coding aims to decrease the amount of redundancy that exists in an uncompressed video stream. To this end, various tools are used during the video encoding process, namely, prediction control, transform and quantisation, motion estimation and compensation and entropy coding. These operations are performed on a block-by-block basis, more specifically, each frame is partitioned into block-shaped units called MBs, as shown in Figure 2.14. The MB has a typical size of 16×16 pixels and it represents the basic processing unit in the MPEG-x and

(a) (b)

Figure 2.15: Two consecutive frames in an underwater video scene (a) Frame 1 (b) Frame 2.

H.26x video compression standards. These standards are also known as *block-based* video encoders.

2.3.2.1 Prediction

Usually a video scene contains a lot of temporal and spatial redundancies, where consecutive frames or neighbouring pixels within the same frame show a lot of similarity. The correlation between successive frames is referred to as *temporal redundancy* while the correlation between neighbouring pixels within the same frame is referred to as *spatial redundancy* [81]. For example, two consecutive frames in an underwater scenery are shown in Figure 2.15. As can be seen from the latter figure, Frame 1 and Frame 2 seem to be identical and therefore, in this case, it is more efficient to encode only the difference between them and then use Frame 1 as a reference frame. The basic concept of prediction coding is to obtain a differential signal (also known as the residual) from the original video signal and its prediction [81]. At the receiver, the video decoder tries to reconstruct the original signal by summing the residual and the predicted signals.

Motion estimation and compensation are part of the prediction coding stage in block-based video encoders. Motion estimation consists of predicting a current block by finding the best matching block, within a given search range, in one or more reference frames [82]. When the video encoder has more than one reference frames at its disposal, a further improvement in the prediction process can be achieved [82]. The best matching block is usually determined based on a cost criterion such as the Mean Squared Error (MSE), Sum of Squared Error (SSE) or Sum of Absolute Difference (SAD) [83]. The reference block which results in the smallest error or difference is then chosen. The relative position of the chosen reference block with respect to the current block gives rise to the *motion vector*. Figure 2.16 shows an example of a video frame displaying motion vectors indicated by the arrows. The difference signal (residual) between the reference block and the current block, along with the motion vector, serve as input to the next stage in the video encoding process. Encoding only the residual between two frames without motion compensation does not lead to efficient compression. When there is relative motion of objects

Figure 2.16: Motion vectors in a video frame.

between successive frames, simple differencing will result in large residual values [84], especially in scenes where there is a lot of motion, e.g., sport scenes. In this case, it is more efficient to estimate the relative motion of objects between consecutive frames and then compensate for it. The residual frame will therefore contain less amount of data that needs to be encoded, as compared to the case where no motion compensation is used. The decoder is informed about the motion compensation through the motion vectors. These specify how an object in a scene moves in the horizontal and vertical directions and thus the decoder can align the blocks and reconstruct the frame [84].

In prediction coding, a sequence of frames in a video sequence can be encoded in such a way that they make up a Group of Pictures (GOP), making them independent of other frames in the video sequence. Three types of frames are used during video coding, namely, intra (I), predictive (P) and bidirectional predictive (B) frames. The first frame in a video sequence is always an I-frame which is independent of all other frames, that is, it does not need other frames to decode [82]. It uses intra-prediction whereby the current MB is predicted from previously coded neighbouring blocks within the same frame [85]. While the frequent use of I-frames limits the propagation of errors when frames are corrupted or lost, they however tend to have a high bandwidth consumption [86]. A P-frame is an inter-coded frame whereby only the information that has changed from the previous I-frame is stored [86]. Therefore, P-frames use fewer bits than I-frames. It is encoded using either a previous I-frame or P-frame as reference (unidirectional prediction). Since the P-frame depends on a previous reference frame, it is more likely to be affected by errors. The B-frame is also an inter-coded frame that provides the highest compression efficiency as compared to I- and P- frames. The B-frames reference the previous I-frames and future P-frames (bidirectional) and they store only the image differences between them [86]. A sequence of I-, P- and B-frames is illustrated in Figure 2.17 where the arrows indicate the inter-frame dependencies.

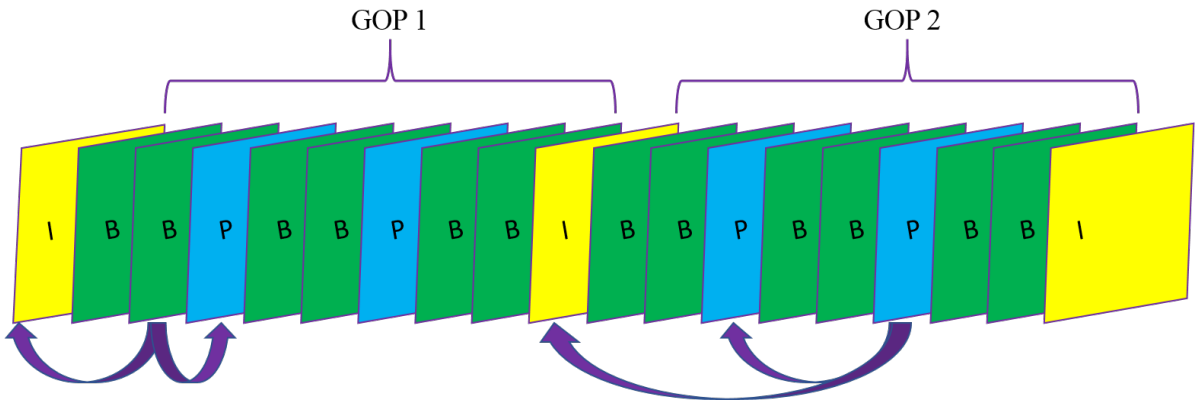


Figure 2.17: A sequence of I-, B- and P-frames.

2.3.2.2 Transform Coding

In addition to prediction coding, transform coding is used to decrease the amount of correlation that exists in the residual signal by converting the spatial domain signal to the frequency domain [81]. The DCT is a widely used mathematical transform in block-based video encoders and decoders since it allows for an efficient implementation. In a broad sense, DCT converts an input data into the frequency domain such that most of the information is concentrated on a few coefficients only [82]. The transform coding operation can be mathematically represented as

$$\mathbf{Y} = \mathbf{A}\,\mathbf{D}\,\mathbf{A}^{\mathsf{T}}, \tag{2.53}$$

where $\mathbf{D}$ is the block of residual data of size $\mathcal{K}\times\mathcal{K}$ and $\mathbf{A}$ is the DCT matrix, with the element on the ith row and jth column given by

$$\mathbf{A}_{i,j} = f_i \cos\frac{i(2j+1)\pi}{2\mathcal{K}} \tag{2.54}$$

for

$$f_i = \begin{cases} \frac{1}{\sqrt{\mathcal{K}}} & , i = 0 \\ \sqrt{\frac{2}{\mathcal{K}}} & , i \neq 0\,. \end{cases} \tag{2.55}$$

As an example, consider an 8×8 block of pixels in Frame 1 of Figure 2.15, which is given by

$$\mathbf{D} = \begin{bmatrix} 160 & 160 & 160 & 160 & 160 & 160 & 160 & 160 \\ 105 & 105 & 105 & 105 & 105 & 105 & 105 & 105 \\ 115 & 115 & 114 & 114 & 114 & 114 & 115 & 115 \\ 115 & 115 & 115 & 115 & 114 & 114 & 115 & 116 \\ 115 & 115 & 115 & 115 & 114 & 115 & 115 & 115 \\ 115 & 115 & 115 & 115 & 114 & 115 & 115 & 114 \\ 114 & 115 & 114 & 114 & 115 & 114 & 115 & 115 \\ 114 & 114 & 115 & 115 & 115 & 114 & 115 & 115 \end{bmatrix}. \tag{2.56}$$

Although transform coding is performed on the block of residual data following the prediction coding, the original frame is considered in this example for illustration purposes only. The 8×8 DCT matrix is given by

$$\mathbf{A} = \begin{bmatrix} 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 \\ 0.4904 & 0.4157 & 0.2778 & 0.0975 & -0.0975 & -0.2778 & -0.4157 & -0.4904 \\ 0.4619 & 0.1913 & -0.1913 & -0.4619 & -0.4619 & -0.1913 & 0.1913 & 0.4619 \\ 0.4157 & -0.0975 & -0.4904 & -0.2778 & 0.2778 & 0.4904 & 0.0975 & -0.4157 \\ 0.3536 & -0.3536 & -0.3536 & 0.3536 & 0.3536 & -0.3536 & -0.3536 & 0.3536 \\ 0.2778 & -0.4904 & 0.0975 & 0.4157 & -0.4157 & -0.0975 & 0.4904 & -0.2778 \\ 0.1913 & -0.4619 & 0.4619 & -0.1913 & -0.1913 & 0.4619 & -0.4619 & 0.1913 \\ 0.0975 & -0.2778 & 0.4157 & -0.4904 & 0.4904 & -0.4157 & 0.2778 & -0.0975 \end{bmatrix}, \tag{2.57}$$

while output of the transform coding operation is given by

$$\mathbf{Y} = \begin{bmatrix} 3.7377 & -0.0009 & 0.0039 & -0.0031 & -0.0005 & 0.0013 & -0.0021 & -0.0001 \\ 0.2022 & 0.0014 & 0.0019 & 0.0013 & 0.0004 & -0.0003 & 0.0012 & -0.0012 \\ 0.2092 & -0.0020 & -0.0044 & 0.0010 & 0.0012 & -0.0017 & -0.0008 & 0.0029 \\ 0.2209 & 0.0022 & -0.0023 & 0.0025 & -0.0027 & 0.0008 & 0.0013 & -0.0019 \\ 0.2181 & -0.0009 & -0.0008 & -0.0031 & 0.0015 & 0.0013 & 0.0019 & -0.0001 \\ 0.1912 & -0.0010 & 0.0028 & -0.0009 & 0.0010 & -0.0022 & -0.0022 & 0.0008 \\ 0.1452 & 0.0017 & 0.0001 & 0.0010 & -0.0010 & 0.0023 & 0.0005 & -0.0005 \\ 0.0772 & -0.0009 & 0.0015 & 0.0017 & -0.0003 & -0.0017 & -0.0013 & 0.0002 \end{bmatrix}. \tag{2.58}$$

As can be seen in (2.58), the top-leftmost element contains the most energy and this element is commonly known as the 'DC' coefficient [79]. The main objective of transform coding is to locate those elements that can be discarded without severely impacting on the reconstructed image/video quality.

2.3.2.3 Quantisation

Quantisation is the process whereby the transformed coefficients are divided by a step size, commonly referred to as the *Quantisation Parameter (QP)*, and then the results are rounded to the closest integer so that they can be encoded using fewer bits [82]. The inverse operation is performed by the decoder whereby the quantised coefficients are multiplied by the same values that were used at the encoder. Video coding is lossy in nature because of the irreversible quantisation process where the rounding errors result in some loss of information, which is however, generally imperceptible to the human eyes [82]. The quantisation process can be mathematically represented as

$$\mathbf{Y}_{\text{quantised}} = QP \times \text{round}\left(\frac{\mathbf{Y}}{QP}\right). \tag{2.59}$$

The value of QP can be adjusted to control the number of transformed coefficients that are discarded, with a high value implying a high compression but lower reconstructed image/video quality.

2.3.2.4 Entropy Coding

Following the quantisation process, the quantised coefficients are fed to an entropy encoder for further compression. Entropy coding is a lossless compression technique which is used to compress digital data wherein the frequently occurring symbols are encoded with fewer bits than the rarely occurring ones. Some common strategies that are used by an entropy encoder to achieve lossless compression include the arithmetic codes and Variable Length Codes (VLC) such as the Huffman codes.

2.3.3 Video Quality Assessment

Video compression is described as being lossy since partial or redundant information in an original video sequence is discarded during the encoding process and this cannot be recovered at the receiver following the decoding process. A quality assessment tool is therefore required to evaluate the effects of changing or tweaking the encoding parameters on the resulting encoded video quality. It should be noted that compression is not the only means by which the original image or video quality is affected. Further distortions may occur, for example, during storage or transmission over an erroneous channel. The two methods that are used to assess the image/video quality include the subjective or the objective approach [87]. In the former method, a group of observers rate or score the quality of videos based on their visual perception. Ideally, the subjective method would be the most suitable for assessing the video quality but is however, often expensive and time-consuming [87].

The objective method, on the other hand, uses some algorithms to quantify the distortions in the encoded video. The quantitative values are generally calculated using the original video as a reference and they indicate whether distortions are present in the encoded video sequence [88]. The PSNR is a widely used objective video quality metric due to its low complexity [89]. Although the objective method may not always reflect the human viewer's judgement of the video quality, the PSNR is found to be a valid metric for evaluating the quality of a compressed video, as investigated in [89]. Therefore, throughout this research work, the PSNR metric is used for measuring the quality of the processed videos. In practice, only the luminance component is considered in the metric computation since, as mentioned previously, the human eyes are more sensitive to luminance than chrominance. The PSNR metric is defined as (in dB)

$$PSNR = 10\log_{10}\frac{(2^u - 1)^2}{MSE}, \tag{2.60}$$

where u denotes the number of bits per sample and the MSE between the original (reference) frame and the processed frame is computed as

$$MSE = \frac{1}{N_{\text{pixels}}}\sum_i\sum_j\left(Y_{\text{original}}(i,j) - Y_{\text{processed}}(i,j)\right)^2 \tag{2.61}$$

where N_{pixels} is the total number of pixels in the frame, $Y_{\text{original}}(i,j)$ and $Y_{\text{processed}}(i,j)$ represent the luminance value at the ith row and jth column in the original and processed frames, respectively. The higher the PSNR value, the better is the quality of the processed video.

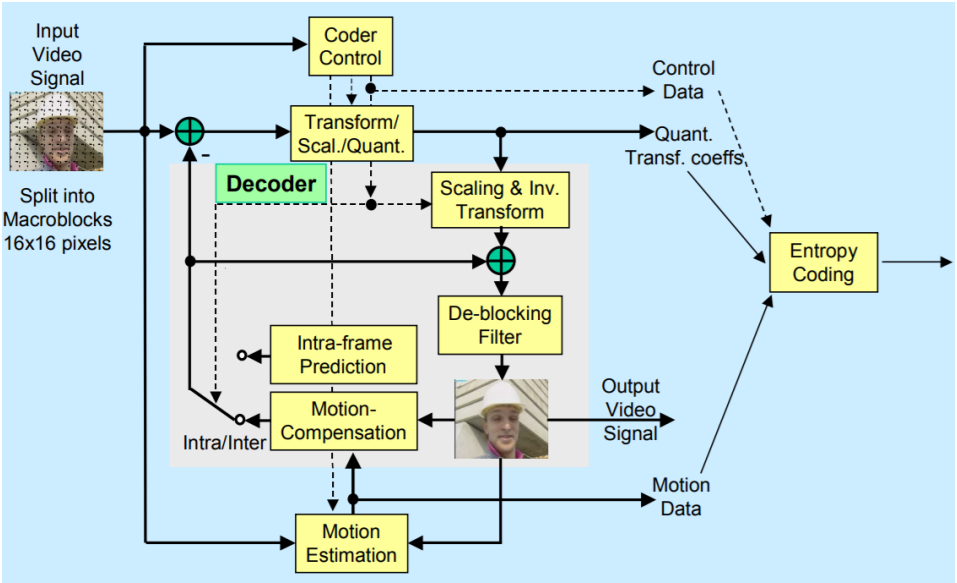


Figure 2.18: Block-based hybrid coding structure for H.264/AVC [90].

2.3.4 H.264/AVC Coding Standard

The H.264/AVC standard was implemented and published in 2003 through the cooperation of two international standards bodies, namely, the Telecommunication Standardisation Sector of the International Telecommunications Union (ITU-T) and the International Organisation for Standardisation/International Electrotechnical Commission (ISO/IEC). Compared to its predecessors such as MPEG-2 and MPEG-4 Visual, H.264/AVC achieves a better compression efficiency, with bit rate savings as high as 50% for the same perceived video quality [90]. Even with the emergence of HEVC, the H.264/AVC standard is still widely used in many applications including High-Definition Television (HDTV), Blu-Ray discs, internet video streaming, video conferencing and surveillance, multimedia mobile phone applications, among others.

Figure 2.18 shows the block-based hybrid coding structure of the H.264/AVC standard which creates an H.264-compliant bitstream from an input video sequence. The following subsections outline some of the main tools and features introduced in the H.264/AVC standard that distinguish it from previous video coding standards. The interested reader is referred to [91] for more in-depth information on the H.264/AVC standard.

2.3.4.1 Layers

In order to handle a wide range of applications and ensure interoperability across both wired and wireless networks, the H.264/AVC standard specifies two layers, namely, the Network Abstraction Layer (NAL) and the Video Coding Layer (VCL). The latter was implemented to represent the encoded video content in an efficient way while the NAL encapsulates the VCL data and specifies header information so that

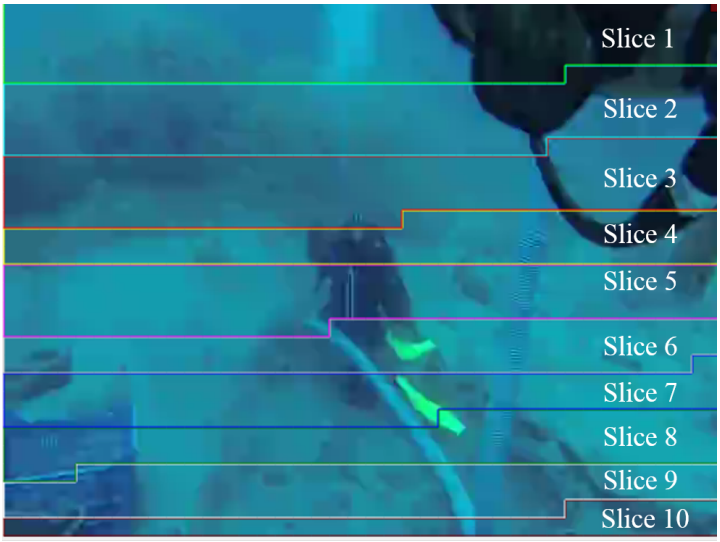


Figure 2.19: Illustration of slices in a video frame.

the video data can be easily stored or transmitted over a communication channel [91].

The encoded video data is organised into NAL units, which can be regarded as video packets that are transmitted through the channel. The type of data in a NAL unit is specified by the header, which is the first byte of the NAL unit and the remaining bytes consist of the actual payload data [90]. The NAL units can be further categorised into two groups, namely, the VCL NAL units and non-VCL NAL units. The VCL NAL units contain the coded slice data or slice partitions for a given frame, where a *slice* can be broadly defined as a grouping of MBs. As for the non-VCL NAL units, they hold information on how the encoded video should be decoded and this include the parameter sets and Supplemental Enhancement Information (SEI) [90]. The latter refers to additional information, such as picture timing, that can help improve the usability of the decoded video sequence [90]. The parameter set consists of information that is likely to change infrequently and it allows the decoding of several VCL NAL units. [90] The two types of parameter sets include the Sequence Parameter Sets (SPS) and the Picture Parameter Sets (PPS). The SPS refers to a coded video sequence while the PPS relates to the decoding of one or more frames in the sequence [90]. Furthermore, the Instantaneous Decoding Refresh (IDR) access unit marks the beginning of each coded video sequence and this consists of an intra-coded picture that can be decoded independently without using any previous pictures as reference [90].

2.3.4.2 Slices and Slice Groups

Video packets are very sensitive to errors in the sense that if one or more bit errors are detected in a packet at the receiver, the whole corrupted packet is dropped [10]. Packet loss affects the overall video quality and hence in addition to FEC coding, suitable techniques should be used during the video coding process to provide enhanced error robustness, for examples using slices, frequent coding of I-frames, and so on. A video frame can be divided into *slices* as shown in Figure 2.19 and each slice can be

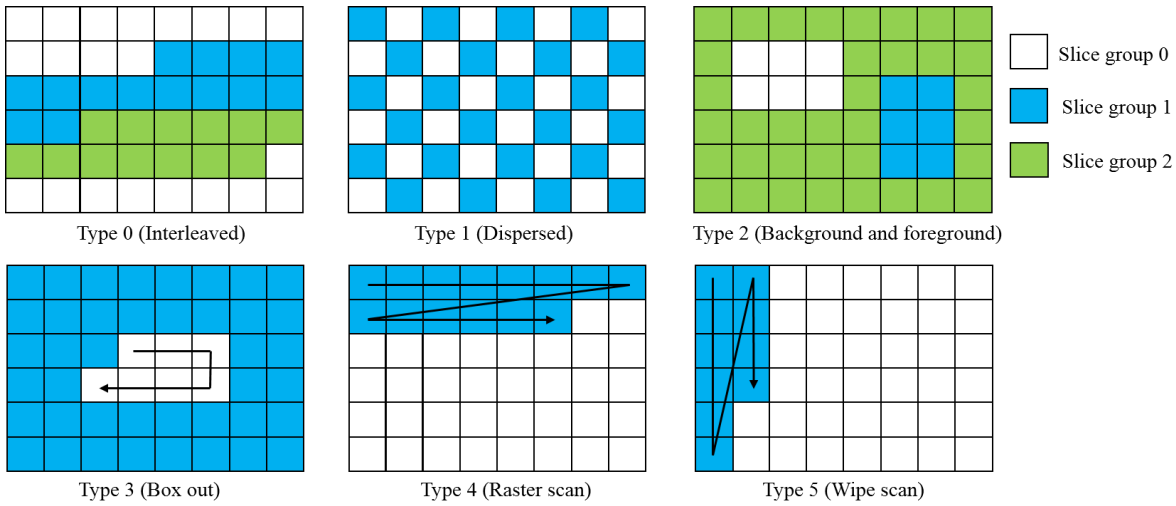


Figure 2.20: Types of FMO.

independently decoded. This increases the error resilience since errors within a given slice do not spread to other slices.

Slices can be further grouped into *slice groups*, where each slice within a group is processed in a raster scan order [90]. The H.264/AVC standard introduced the Flexible Macroblock Ordering (FMO) feature in its *baseline* profile to further enhance the error resiliency. Note that the *baseline* profile targets applications which require low delay, low complexity and maximum error robustness [92]. Using FMO, the slice groups can be grouped in such a way that, if for example, one slice group is missing at the decoder side, the MBs which are lost can be reconstructed using interpolation based on surrounding samples of MBs from other slice groups [93]. The H.264/AVC standard defines seven types of FMO mapping, six of which are illustrated in Figure 2.20. The seventh FMO mapping is completely arbitrary and the user can specify his own slice grouping in the encoder.

The H.264/AVC standard supports different coding types for each slice, regardless of the slice grouping. The different slice types are as follows [90]

- **I-slice**: Intra-prediction is used to code the MBs where the current MB is predicted from previously coded MBs in the same slice.
- **P-slice**: Some of the MBs can be coded using inter-prediction as well as the coding types present in an I-slice, with at most one motion-compensated prediction signal per block.
- **B-slice**: Some MBs can be coded using inter-prediction as well as the coding types present in a P-slice, with two motion-compensated prediction signals per block.

For the P- and B-slices, previously coded pictures can be used to create the prediction signal for the MBs [92]. While previous video coding standards use the I-, P- and B- slice types only, the H.264/AVC standard introduced two additional slice types, namely, the Switching Predictive (SP) and Switching

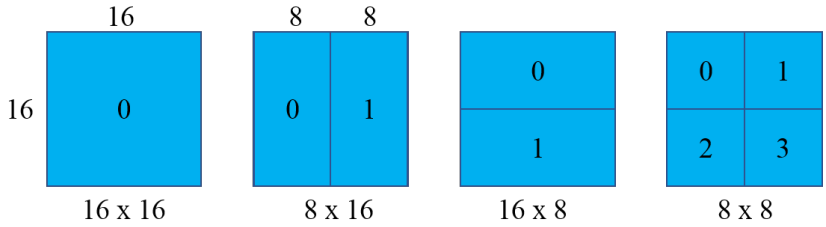


Figure 2.21: Partitioning of a MB for motion-compensated prediction.

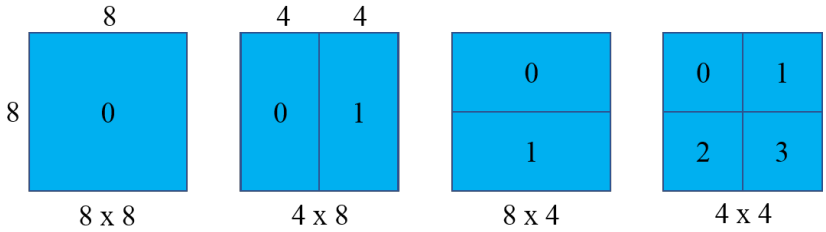


Figure 2.22: Partitioning of 8×8 sub-macroblock.

Intra (SI) slices. These slices facilitate switching between streams coded at different bit rates and provide error resilience [90]. SP-slices can be coded using two different reference pictures and exactly the same data can be decoded from the SP-slices. An SI-slice is coded in such a way that it decodes exactly the same data as a corresponding SP-slice. SP-slices are coded at regular intervals and for each SP-slice, there exists at least one additional SP-slice using another reference picture, or an SI-slice.

2.3.4.3 Prediction in H.264/AVC Standard

Unlike previous standards, H.264/AVC supports inter-prediction with different block sizes ranging from 16×16 down to 4×4. *Inter-prediction* refers to the process of predicting a MB in a current frame from one or more MBs in other frame(s) [82]. The motion prediction for P- and B-slices is performed by partitioning the MB into fixed size blocks as shown in Figure 2.21. Moreover, when the MB is split into the four 8×8 sub-macroblocks, each of these can be further partitioned as shown in Figure 2.22. Each MB consists of its own motion vectors (maximum of 16) which are accurate to a quarter-sample resolution in the luma component, implying that they can point to locations between pixels [82]. The different partition sizes allow for a better motion compensation as large and small partition sizes can be used for smooth and detailed areas in a picture, respectively. A better motion compensation also translates to a higher compression efficiency. The main difference between B- and P- slices in the H.264/AVC standard is that the MBs in B-slices employ a weighted average of two different motion-compensated prediction values for finding the prediction signal [92]. In fact, B-slices use two reference picture lists while P-slices use a single list for prediction [90].

As mentioned in section 2.3.2.1, the I-frame uses *intra-prediction*, i.e., previously coded MBs

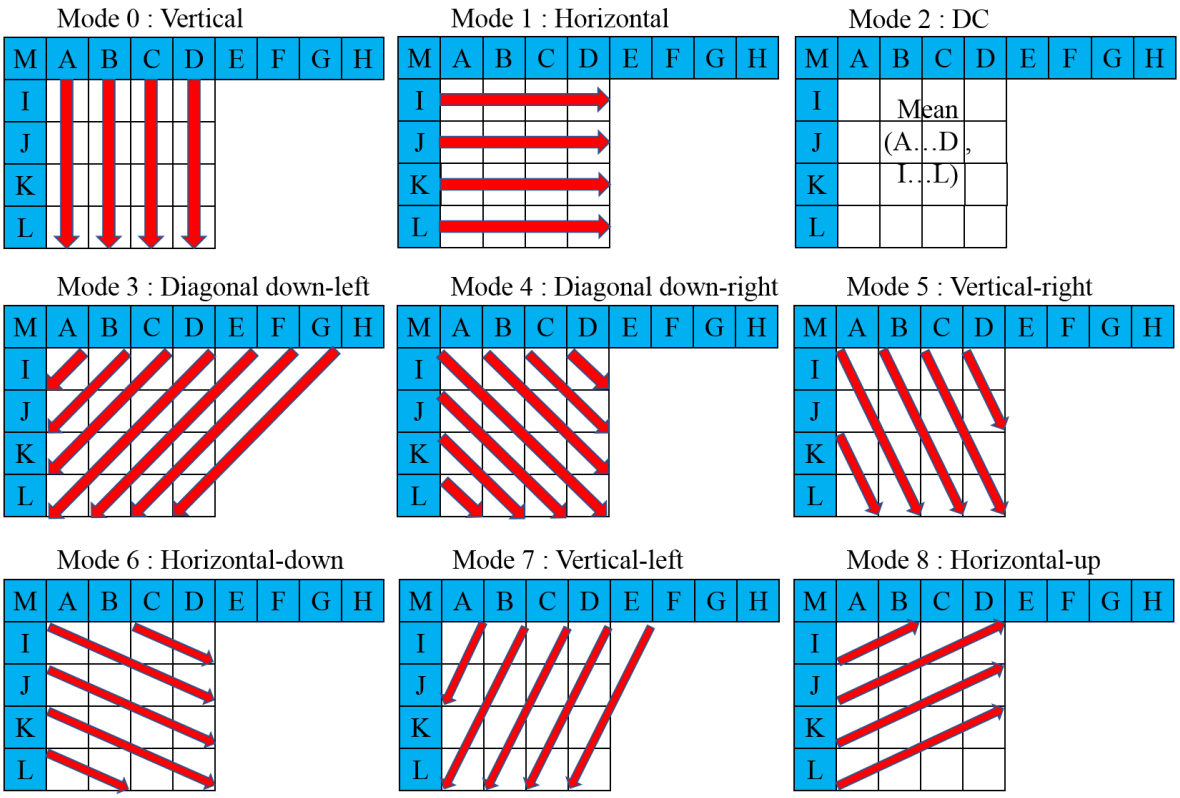


Figure 2.23: Intra_4×4 prediction modes.

within the same frame are used to predict the current MB [85]. As compared to previous standards, the intra-prediction in H.264/AVC is performed in the spatial domain instead of the transform domain [92]. The intra-prediction is performed on luma blocks of sizes 16×16, 8×8 or 4×4, commonly denoted as Intra_16×16, Intra_8×8, and Intra_4×4, respectively. Nine prediction modes are available for the 4×4 luma blocks, as shown in Figure 2.23. The Intra_4×4 mode is mostly used to code parts of an image which have a high detail level whereas the Intra_16×16 mode is mostly used to code smooth regions in an image [90]. Note that the Intra_16×16 block supports only four prediction modes, as shown in Figure 2.24. The Intra_8×8 mode is supported only in the *High* profile of the H.264/AVC encoder, which targets mostly HD applications. All its prediction modes are similar to the Intra_4×4 mode, except that a low-pass filter is used to achieve a better prediction performance [82].

2.3.4.4 Transform and Quantisation in H.264/AVC Standard

As compared to previous video coding standards which use a 2D DCT on residual blocks of size 8×8, the H.264/AVC standard introduces a group of integer transforms of different block sizes [92]. H.264/AVC can choose between 4×4 and 8×8 transforms for the MBs. The integer DCT transformation is often applied to 4×4 blocks to help reduce ringing artefacts (bands which appear near the edges in an image) [82]. Furthermore, the 4×4 integer transform allows for less computations on the residual data and

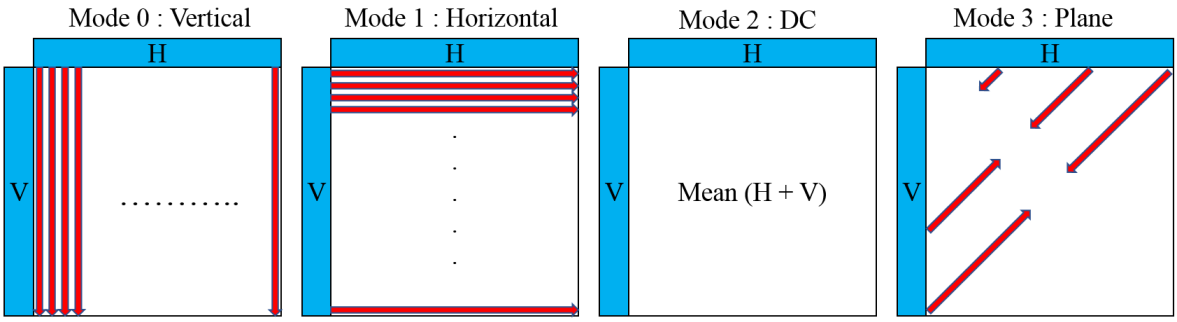


Figure 2.24: Intra_16×16 prediction modes.

also prevents decoding mismatch since an exact inverse transform is possible at the decoder. The 4×4 transform matrix as used by the H.264/AVC standard is given by

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}. \tag{2.62}$$

The 8×8 transform, on the other hand, is mostly used to preserve the fine details and textures in HD videos [82].

For the quantisation of the transform coefficients, a total of 52 different *QP* values are specified for the encoder in the H.264/AVC standard. This large range of QP allows a trade-off to be achieved between the bit rate and reconstructed image quality. The quantisation step size is doubled for every increment of six in the value of *QP* [92]. In other words, a unitary increment in the value of *QP* results in an increase in the quantisation step size by approximately 12%, which also causes a bit rate reduction by approximately 12% [90].

2.3.4.5 De-blocking Filter

One of the problem that arises in block-based video coding techniques, especially at low bit rates, is the appearance of block artefacts when a frame is processed using MB structures. This is due to the block edges that are not reconstructed as accurately as the pixels within the blocks. To address this issue, H.264/AVC specifies an adaptive in-loop de-blocking filter to make the block transitions smoother and therefore improve the appearance of the decoded frames. In the encoder, the filter is used after the inverse transform operation but before the MB is reconstructed and stored for future predictions, thereby improving the efficiency of the motion compensation [82]. In the decoder, the de-blocking filter is applied before reconstructing and displaying the MB.

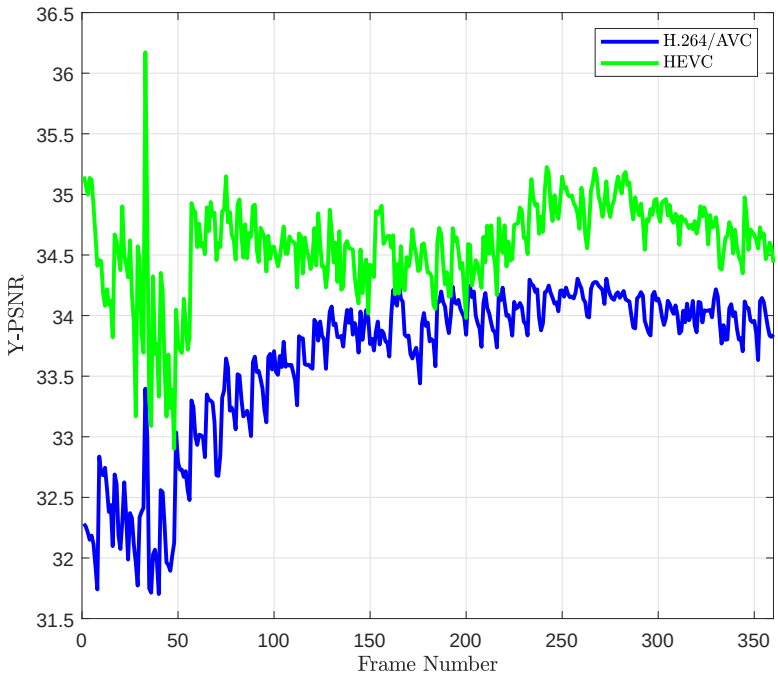


Figure 2.25: PSNR comparison between compressed H.264/AVC and HEVC video streams prior to transmission.

2.3.4.6 Entropy Coding in H.264/AVC Standard

H.264/AVC supports two different methods for entropy encoding, which is the final step in the video encoding process. The first method is the Context Adaptive Variable Length Coding (CAVLC) which is based on an adaptive variant of the Huffman coding [82] in the sense that several VLC code tables are used instead of a single one, allowing for a higher compression to be achieved. CAVLC is used in applications which require fast decoding and where compression efficiency is not a priority. The second method is the Context Adaptive Binary Arithmetic Coding (CABAC) which is based on the arithmetic coding technique. CABAC typically achieves a bit rate reduction between 5% and 15% as compared to CAVLC [90], and is therefore used for HD video compression. The H.264/AVC standard includes CABAC only in its *Main* and higher profiles since CABAC requires a larger amount of time to decode. On the other hand, CAVLC is supported in all the profiles of H.264/AVC .

2.3.5 Comparison between H.264/AVC and HEVC for UWA Video Transmission

In this section, the video quality comparison in terms of PSNR is evaluated for HEVC and H.264/AVC encoded streams. A SD video of resolution 640×480 with a frame rate of 30 fps is considered. The total number of frames in both streams is 360. The HEVC and H.264/AVC are both compressed at the same bit rate of 125 kbps, i.e., they both have the same file sizes. For the purpose of simulation a GOP size of

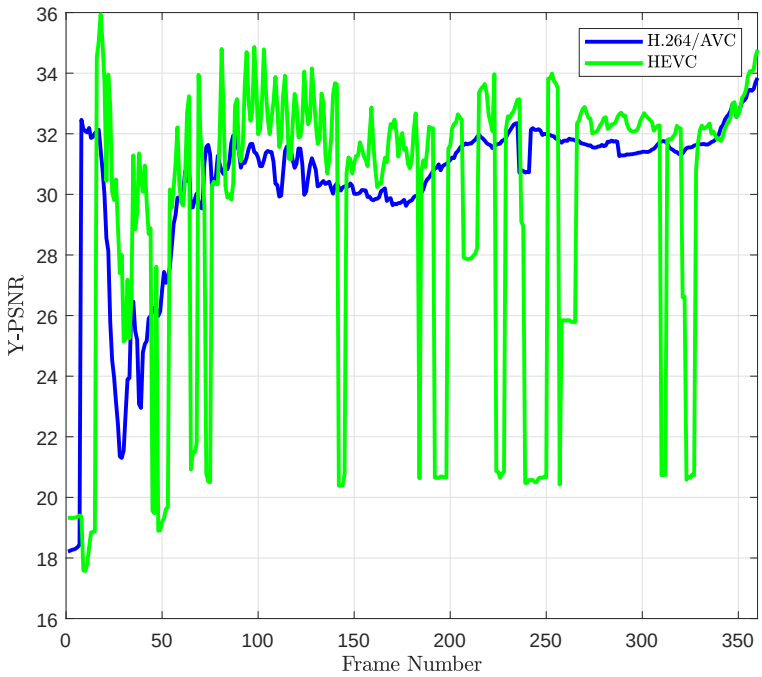


Figure 2.26: PSNR plot for received H.264/AVC and HEVC video streams (without error concealment).

8 is chosen. IDR frames are also used during the HEVC encoding to avoid propagation of errors between frames. The IDR frame can be regarded as a special type of I-frame (as used in H.264/AVC) which allows the decoding of subsequent frames without using any previously decoded frame before the IDR frame as reference. The *Low Delay Main* configuration file is used in the HEVC encoder since it is more suited for applications requiring low latency [94]. A maximum slice size of 680 bytes is selected for both encoded streams. The PSNR comparison between the HEVC and H.264/AVC video streams (prior to transmission) is shown in Figure 2.25. As expected, the HEVC stream has a better video quality than the H.264/AVC stream for similar compression rate. Note that a video having a PSNR value in the range between 31 dB and 37 dB is considered to be of good quality while a PSNR value in the range between 25 dB and 31 dB is considered to be of satisfactory quality according to the ITU quality scale [95].

Figure 2.26 shows the PSNR plot between the raw uncompressed 480p video and the received HEVC and H.264/AVC video streams (after passing through a channel). The Packet Loss Rate (PLR) for both video streams is assumed to be 5.6%. The received bitstreams were decoded using the FFMPEG software without using any error concealment technique. As can be observed in Figure 2.26, the received H.264/AVC stream seems to be more resilient to packet loss than the HEVC encoded stream. In fact, HEVC is known to have a lower error resilience due to the increase in the temporal dependency [96], especially when the data is compressed at low bit rates. Hence, in addition to FEC and video coding using IDR frames and slices, other error resilience techniques such as Multiple Description Coding (MDC)

or Layer Coding (LC) may be required in HEVC [96]. Furthermore, HEVC usually involves more computational overheads than H.264/AVC due to the complex prediction process to remove redundant information in the video [96]. For the above mentioned reasons, H.264/AVC is considered for video compression in this research work.

2.4 Chapter Summary

This chapter presented a brief introduction to UAC characteristics, including path loss, propagation speed, ambient noise, multipath and Doppler effect. For simulating such channels, these characteristics are taken into account when computing the path lengths, angles of arrival, path gains and delays for a given channel geometry. Finally, the basics of video coding using H.264/AVC were covered and a comparison between H.264/AVC and HEVC standards for video transmission in an UAC was also carried out.

CHAPTER 3

PERFORMANCE EVALUATION OF SISO-FBMC SYSTEMS FOR UWA COMMUNICATION

3.1 Introduction

OFDM is currently the most commonly used multicarrier technique due to its robustness against ISI -caused by delay spread in frequency selective channels-, higher bandwidth efficiency compared to traditional Frequency Division Multiplexing (FDM) systems and low complexity frequency domain equalisation. OFDM, however, does have some limitations. The CP is a key element in an OFDM system, however its use results in a loss of useful bandwidth that could have been otherwise used for data transmission. Additionally OFDM exhibits out-of-band power leakage [97] and is sensitive to frequency dispersion. In the case of UACs, the CIR can span over tens to even hundreds of milliseconds [7]. In order to cope with the extended CIR in UACs, the duration of the CP and consequently the OFDM symbol should be increased to preserve the bandwidth efficiency [98]. The problem that may arise in fast varying UACs is that Doppler distortion caused by motion may cause severe Inter-Carrier Interference (ICI) in the OFDM system and hence degrade its performance.

One technique that can overcome the drawbacks of OFDM in an UAC and simultaneously improve the bandwidth efficiency is FBMC. In FBMC transmission where a CP is not used, robust performance in channels characterised by both time and frequency dispersions (doubly-dispersive) can still be achieved through the use of suitable pulse shaping filters such as Isotropic Orthogonal Transform Algorithm (IOTA) [99] and Hermite function based prototype filters [100].

FBMC systems were developed prior to OFDM. In the mid 1960's Chang [101] and Saltzberg [102] laid the foundation for multicarrier techniques. In his work Chang presented the procedures for transmitting Pulse Amplitude Modulation (PAM) symbols in a bandwidth-efficient way by using a bank of overlapping filters, namely Vestigial sideband (VSB) modulated filters [103]. Saltzberg modified

Chang's method to transmit QAM symbols using Double Sideband Modulation (DSB) and both methods achieve the same bandwidth efficiency. Saltzberg's method is often referred to as Staggered Multitone (SMT) or OFDM-OQAM because of the time offset by half a symbol duration between the in-phase and quadrature components in each QAM symbol. Authors that have managed to efficiently implement Saltzberg's system digitally using polyphase structures include Bellanger [104] and Hiroshaki [105]. The Digital Subscriber Line (DSL) evolution in the 1990's prompted research on other FBMC techniques such as the Discrete Wavelet Multitone (DWMT) and FMT. Research has demonstrated that DWMT is very similar to Chang's method and it was re-branded as Cosine Modulated Multitone (CMT) since it uses cosine modulated filterbanks [103]. The only difference between SMT and CMT is that the data symbols are PAM modulated in CMT and QAM modulated in SMT. The application of CMT or SMT to a particular scenario is just a matter of choice. While in Chang's and Saltzberg's methods the adjacent subcarriers overlap to achieve high bandwidth efficiency, in FMT they are separated by guard bands in the same way as conventional FDM (hence lower bandwidth efficiency than other FBMC systems).

Recent studies have addressed the feasibility of FBMC systems for UWA communication (see e.g., [106–108] and references therein). For instance in [106] the authors proposed to use an FMT system with relatively wideband subcarriers. In this work, a novel channel-estimation based DFE is proposed for the FMT system. Their design required a multi-tap equaliser per subcarrier. Simulation results showed that FMT may outperform OFDM in the presence of channel variation but the experimental data were not conclusive in this respect since the channel was almost stable. FMT was also investigated in [107] where a prototype filter was designed for symbols arranged in a hexagonal lattice structure to address doubly-dispersive channels. The filter is essentially a more robust version of the Haas and Belfiore [100] method which is based on Hermite functions. FMT was chosen due to its simpler structure and the fact that it is easily applicable to MIMO systems. In this modified FMT system, the subcarriers were allowed to overlap to a smaller extent. In terms of Signal-to-Interference-plus-Noise Ratio (SINR), the modified FMT system exhibited around 2 dB better performance than OFDM and conventional non-overlapping FMT in a high mobility environment. Each subcarrier was assumed to experience flat fading and therefore the channel at each subcarrier was estimated as a complex valued gain using pilot symbols. A one-tap equaliser was used in this work. CMT was investigated for UWA communication in [108]. The work focused on Doppler compensation using a low-complexity algorithm based on the frequency spreading technique. The choice of CMT was based on the fact that it can achieve a bandwidth efficiency of 100%. Frequency spreading multi-tap equalisation was used in order to have a wider bandwidth for each subcarrier and achieve a more precise equalisation. A channel that introduced only a time scaling to the transmit signal was considered. The Signal-to-Interference Ratio (SIR) of a 64-subcarrier system was analysed by varying the Doppler scaling factor. It was shown that for Doppler scaling factor values ranging from 0 to 0.05, the SIR remained constant at about 65 dB.

In this chapter, the performance of systems based on FBMC modulation is evaluated and the benefits they can bring to UWA communication as compared to the traditional OFDM system are investigated. Firstly, the BER performance of a SISO system based on FMT and OFDM-OQAM is evaluated in

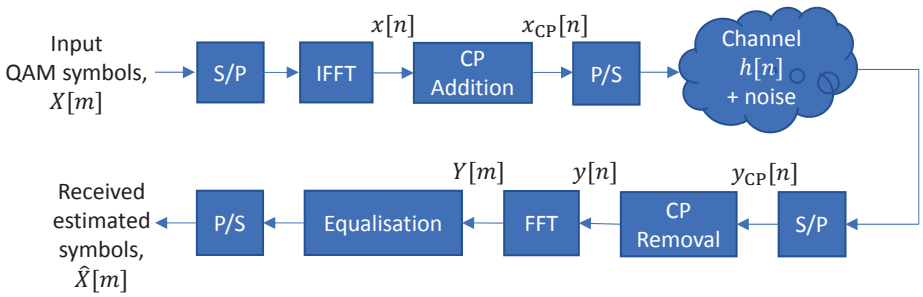


Figure 3.1: OFDM basic block diagram.

AWGN, Rayleigh and UACs without using a CP. Assuming a time-invariant channel and perfect channel estimation, it is shown that the FBMC systems provide the same performance as OFDM. Since OFDM-OQAM can achieve 100% bandwidth efficiency as compared to FMT (where the subcarriers are non-overlapping), it is evaluated in two channel scenarios whereby data is transmitted over a 200 m horizontally and vertically-configured UACs. It is found that the horizontal channel suffers from greater time dispersion than a vertical channel and hence the probability of bit errors is higher. The theoretical bit rate that can be achieved over the 200 m vertical and horizontal UWA links with Turbo-coded SISO OFDM-OQAM system is 138 kbps as compared to 121 kbps for the SISO OFDM system.

The main contributions of this chapter are published in [48]. The rest of this chapter is organised as follows. Section 3.2 provides the mathematical formulations for both OFDM and FBMC systems, namely FMT and OFDM-OQAM. Section 3.3 presents the simulation results of the multicarrier systems in AWGN, Rayleigh fading and horizontally and vertically configured UACs. Finally, some concluding remarks are provided in Section 3.4.

3.2 OFDM and FBMC Systems

3.2.1 Introduction to OFDM

The general block diagram of an OFDM transceiver system is shown in Figure 3.1. At the transmitter, an input bitstream is mapped to QAM symbols followed by a Serial-to-Parallel (S/P) conversion. The modulated symbols can be represented as $X[0], X[1], \cdots, X[M-1]$ where M is the number of subcarriers. Inverse Fast Fourier Transform (IFFT) is then applied to the frequency-domain symbol block of length M to yield the time-domain OFDM waveform. The time-domain OFDM signal prior to CP addition can be represented in baseband discrete-time as

$$x[n] = \text{IFFT}\{X[m]\} = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} X[m] e^{j\frac{2\pi}{M}nm}, \quad 0 \leq n \leq M-1, \tag{3.1}$$

where n is the sample index and $X[m]$ denotes the modulated QAM symbol on the mth subcarrier. In order to mitigate ISI between consecutive OFDM symbols, a CP with a length at least equal to the channel delay spread is appended to each symbol. When a CP is appended to each OFDM symbol, the

resulting signal can be expressed as $x[n]_M = x[n \bmod M]$ (denoting a modulo M operation). $x[n]_M$ is a periodic version of $x[n]$ with a period M [109]. The time-domain signal at the channel output is given by

$$y_{\mathrm{CP}}[n] = h[n] \star x_{\mathrm{CP}}[n] + \omega[n], \tag{3.2}$$

where $x_{\mathrm{CP}}[n]$ is the OFDM signal after CP addition, $h[n]$ is the linear time-invariant channel and $\omega[n]$ is the additive noise. At the receiver, following the CP removal, the time-domain signal can be represented as

$$y[n] = h[n] \circledast x[n] + \omega[n]. \tag{3.3}$$

The frequency-domain signal following the Fast Fourier Transform (FFT) demodulation can be expressed as

$$Y[m] = \frac{1}{\sqrt{M}} \sum_{n=0}^{M-1} y[n] e^{-j\frac{2\pi}{M}nm}, \quad 0 \leq m \leq M-1 \tag{3.4}$$

$$= H[m] X[m] + \Omega[m], \tag{3.5}$$

where $H[m]$, $X[m]$, $\Omega[m]$ and $Y[m]$ are the M-point FFTs of $h[n]$, $x[n]$, $\omega[n]$ and $y[n]$, respectively. Assuming that the Channel Frequency Response (CFR), $H[m]$, is known, the estimated transmitted symbols can be obtained using Zero-Forcing (ZF) equalisation as

$$\hat{X}[m] = \frac{Y[m]}{H[m]}. \tag{3.6}$$

In practice, $\hat{X}[m]$ is imperfect due to factors such as the additive noise and imperfect channel estimation.

3.2.2 Introduction to OFDM-OQAM

OFDM-OQAM can be implemented using the Frequency-Spreading (FS) technique [110] or the Polyphase Network (PPN) technique. In the FS approach, the OQAM symbols are filtered in the frequency domain followed by an IFFT, the size of which is given by the product of the prototype filter overlapping factor (K) and the number of subcarriers (M) [110]. Then an overlap and sum operation is performed on the IFFT output. At the receiver, a sliding window is used in the time domain followed by KM-point FFT and filtering using a matched prototype filter. The PPN technique, on the other hand, reduces the complexity introduced by additional filtering operations at the transmitter and receiver [97]. The PPN simply refers to a set of subchannel digital filters, the coefficients of which are obtained from a single prototype filter [111]. A polyphase implementation of OFDM-OQAM is shown in Figure 3.2. The input bitstream is first mapped to complex QAM modulated symbols $a_m[\bar{l}]$ ($\bar{l}$ is the sample index at the OQAM pre-processing input and the OQAM post-processing output) followed by an OQAM pre-processing operation. The latter consists of a complex-to-real conversion where the real and imaginary parts of each QAM symbol is separated by half of a symbol duration ($T/2$). This conversion increases the sampling rate by a factor of two. To fulfil the real orthogonality condition between adjacent subcarriers, the real-valued symbols $b_m[\tilde{l}]$ are multiplied by $\varphi_m[\tilde{l}] = j^{m+\tilde{l}}$, where $\tilde{l}$ is

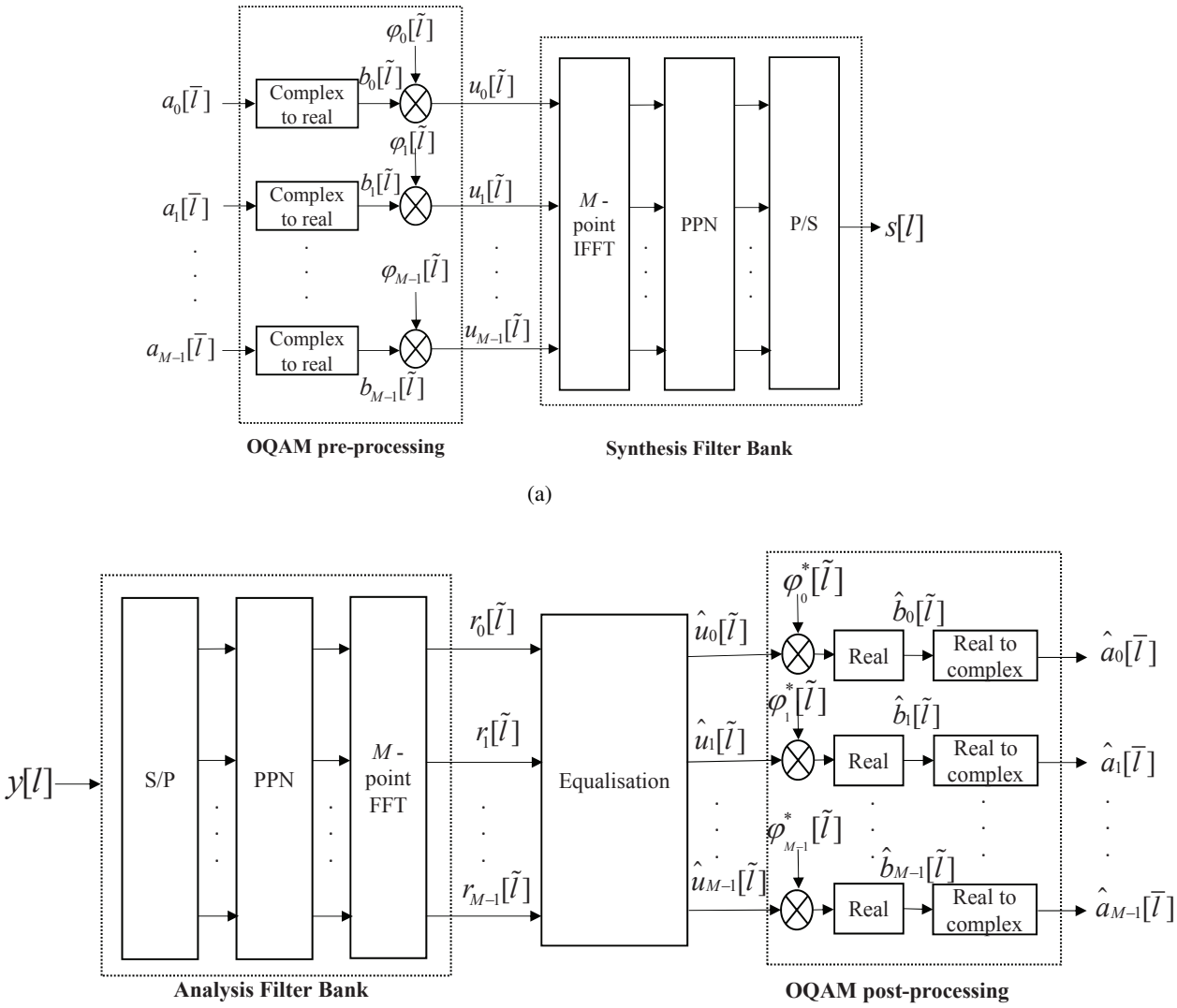


Figure 3.2: OFDM-OQAM block diagram (a) Transmitter (b) Receiver.

the sample index at the pre-processing output and post-processing input and m is the subcarrier index [112]. In OFDM-OQAM, odd (or even) subcarriers are not overlapped. If only QAM symbols were to be used then it would only be possible to use the alternate subcarriers. This would however reduce the capacity of the system by a factor of two [97]. In order to achieve the maximum capacity, all subcarriers have to be used and therefore real orthogonality is required between adjacent subcarriers [113]. This is the reason why OQAM modulation is used.

The discrete-time baseband transmitted OFDM-OQAM signal can be expressed as

$$s[l]=\sum_{m=0}^{M-1}\sum_{\tilde{l}=-\infty}^{\infty} u_m[\tilde{l}]\,g\left[l-\frac{\tilde{l}M}{2}\right]e^{j\frac{2\pi}{M}m\left(l-\frac{Lg-1}{2}\right)}, \tag{3.7}$$

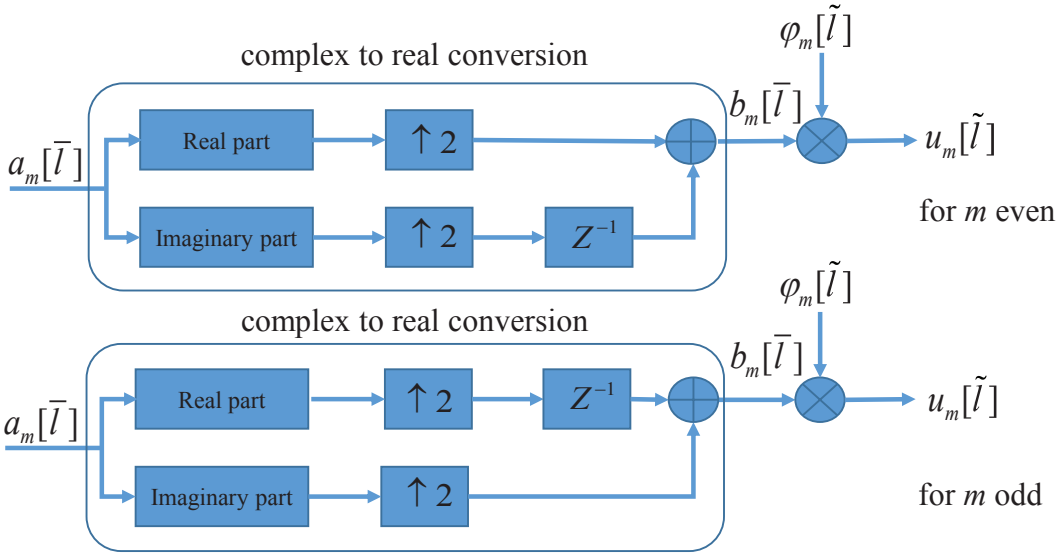


Figure 3.3: OQAM pre-processing block diagram.

where l is the sample index at the Synthesis Filter Bank (SFB) output and the Analysis Filter Bank (AFB) input, $g[l]$ is the prototype filter of length L_g, and $u_m[\tilde{l}]$ is given by

$$u_m[\tilde{l}]=b_m[\tilde{l}]\varphi_m[\tilde{l}], \tag{3.8}$$

where $b_m[\tilde{l}]$ represents the real-valued symbols for the mth subcarrier which are transmitted at a rate of $2/T$ (T is the symbol duration which is defined as $T=1/\Delta f$ where Δf is the subcarrier spacing). The complex-to-real conversion can be represented as [112]

$$b_m[\tilde{l}]=\begin{cases}\mathscr{R}\{a_m[\bar{l}]\} & \text{for} \quad m \quad \text{even,}\\ \mathscr{I}\{a_m[\bar{l}]\} & \text{for} \quad m \quad \text{odd,}\end{cases} \tag{3.9}$$

$$b_m[\tilde{l}+1]=\begin{cases}\mathscr{I}\{a_m[\bar{l}]\} & \text{for} \quad m \quad \text{even,}\\ \mathscr{R}\{a_m[\bar{l}]\} & \text{for} \quad m \quad \text{odd,}\end{cases} \tag{3.10}$$

where the symbols $b_m[\tilde{l}]$ and $b_m[\tilde{l}+1]$ are respectively the in-phase and quadrature components of the complex symbol $a_m[\bar{l}]$. At the receiver side, the inverse operation is performed, that is, real-to-complex conversion which decreases the sampling rate by a factor of 2. The block diagrams of the OQAM pre-processing and post-processing are shown in Figure 3.3 and Figure 3.4, respectively. The prototype filter $g[l]$ is shifted in frequency to produce the subchannels which cover the whole bandwidth [112]. Therefore, the synthesis filter for the mth subcarrier can be expressed as

$$g_m[l]=g[l]\exp\left(j\frac{2\pi m}{M}\left(l-\frac{L_g-1}{2}\right)\right), \tag{3.11}$$

where $l=0,1,...,L_g-1$. L_g depends on the size of the filter bank (M subcarriers) and the number of OFDM-OQAM symbol waveforms K that overlap in the time domain as $L_g=KM$ [112]. Assuming

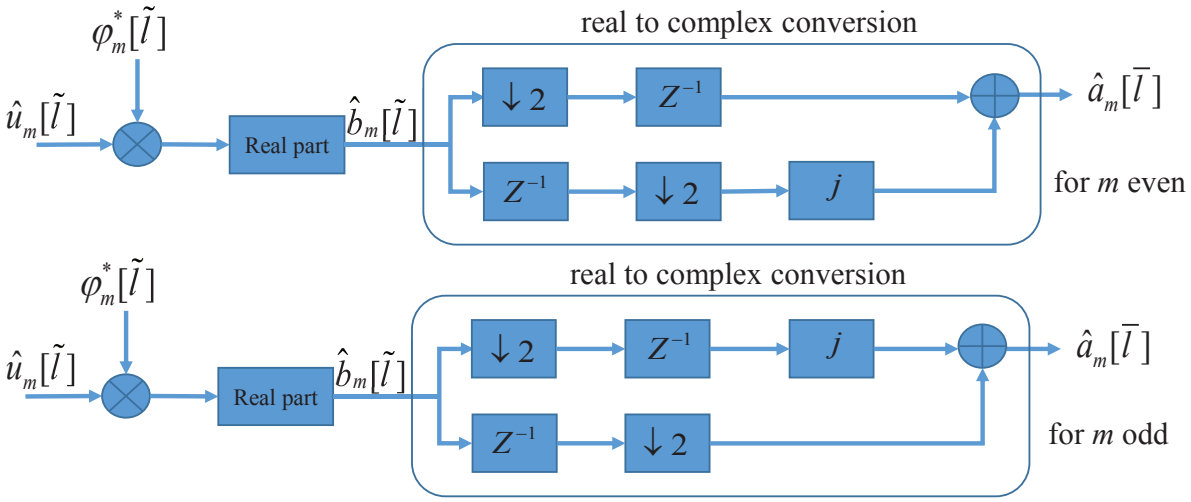


Figure 3.4: OQAM post-processing block diagram.

perfect reconstruction (as in an ideal channel), the mth analysis filter is simply a complex-conjugated and time-reversed version of the corresponding synthesis filter which can be defined as [112]

$$g_m^{\text{an}}[l]=g_m^*[L_g-1-l]. \tag{3.12}$$

However, since a real wireless channel is not ideal, the prototype filter is usually designed such that it guarantees Near-Perfect Reconstruction (NPR). This means that the output signals are approximately delayed versions of the input ones and a small amount of filter bank induced distortion can be tolerated provided it is smaller than the distortion introduced by the channel [114]. The prototype filter can be designed using the frequency sampling method or window-based technique [112]. In the former method, the impulse response of the filter coefficients are obtained by taking the IFFT of the KM uniformly spaced samples from the desired frequency response of the filter. The window-based technique typically consists of multiplying the impulse response of a low-pass filter by a window function to obtain the prototype filter. In this work, a frequency sampling-based prototype filter is considered as proposed in the Physical Layer for Dynamic Spectrum Access and Cognitive Radio (PHYDYAS) project [113]. The frequency domain coefficients of the Finite Impulse Response (FIR) prototype filter for an overlapping factor of K=4 are as follows

$$G_0=1; \quad G_{\pm1}=0.97196; \quad G_{\pm2}=\sqrt{2}/2; \quad G_{\pm3}=0.235147. \tag{3.13}$$

For this value of K, a highly frequency selective filter is achieved with almost no out of band leakage. The frequency response of the filter is given by [113]

$$G(f)=\sum_{i=-(K-1)}^{K-1} G_i \frac{\sin\left(\pi\left(f-\frac{i}{MK}\right)MK\right)}{MK\sin\left(\pi\left(f-\frac{i}{MK}\right)\right)}, \tag{3.14}$$

and the corresponding impulse response (after IFFT) is [97]

$$g[l]=1+2\sum_{i=1}^{K-1}(-1)^i G_i \cos\left(\frac{2\pi i}{MK}l\right). \tag{3.15}$$

For an OFDM-OQAM system implemented using polyphase filtering structures with the IFFT and FFT pairs, the mth synthesis polyphase filter is given by

$$f_m[l]=g[m+lM], \tag{3.16}$$

and the mth analysis polyphase filter is expressed as [112]

$$f_m^{\text{an}}[l] = f_{M-1-m}[l] = g[M-1-m+lM]. \tag{3.17}$$

After the synthesis filtering process, the OFDM-OQAM signal undergoes a Parallel-to-Serial (P/S) conversion and is then transmitted over the wireless channel. The received signal over the mth subcarrier and lth OFDM-OQAM symbol can be expressed as

$$y_m[l]=s_m[l]h_m[l]+\omega_m[l], \quad m=0,1,\cdots,M-1, \tag{3.18}$$

where $h_m[l]$ and $\omega_m[l]$ are, respectively, the complex-valued channel gain and AWGN experienced by the mth subcarrier of the lth symbol. The received signal $y_m[l]$ then goes through the AFB and the resulting signal before the equalisation process in Figure 3.2(b) is given by [112]

$$\begin{aligned} r_m[\tilde{l}] &= \left[y_m[l]\star g_m^{\text{an}}[l]\right]_{\downarrow M/2} \\ &= u_m[\tilde{l}].\chi_m[\tilde{l}]+\eta_m[\tilde{l}], \end{aligned} \tag{3.19}$$

where $\chi_m[\tilde{l}]=\left[h_m[l]\ g_m[l]\star g_m^{\text{an}}[l]\right]_{\downarrow M/2}$ and $\eta_m[\tilde{l}]$ represents Gaussian noise with variance σ^2. Assuming the prototype filter can guarantee NPR, then $\chi_m[\tilde{l}]=\left[h_m[l]\ g_m[l]\star g_m^{\text{an}}[l]\right]_{\downarrow M/2}\cong h_m[\tilde{l}]$ [112]. In this case, equation (3.19) can be simplified to

$$r_m[\tilde{l}]=u_m[\tilde{l}]h_m[\tilde{l}]+\eta_m[\tilde{l}], \quad m=0,1,\cdots,M-1. \tag{3.20}$$

Since in OFDM-OQAM the real and imaginary parts of the symbols are transmitted alternatively, equalisation has to be performed so that ISI due to the frequency-selective nature of the channel does not occur between the adjacent symbols. If the number of subcarriers in the system is large enough each subcarrier will experience flat-fading and hence a single-tap equaliser per subcarrier will be adequate to remove any distortion. A number of per-subcarrier equalisation methods have been proposed in [115] ranging from a classical one-tap approach, frequency sampling, fractionally-spaced Minimum Mean Square Error (MMSE), simplified MMSE and fractionally-spaced adaptive Least Mean Squares (LMS). Multiband band MMSE equalisation is also proposed. Assuming perfect channel estimation, it was shown that the multiband band MMSE equaliser provides the best performance at high SNR values in highly frequency selective channels but it has a higher complexity. This equalisation method also

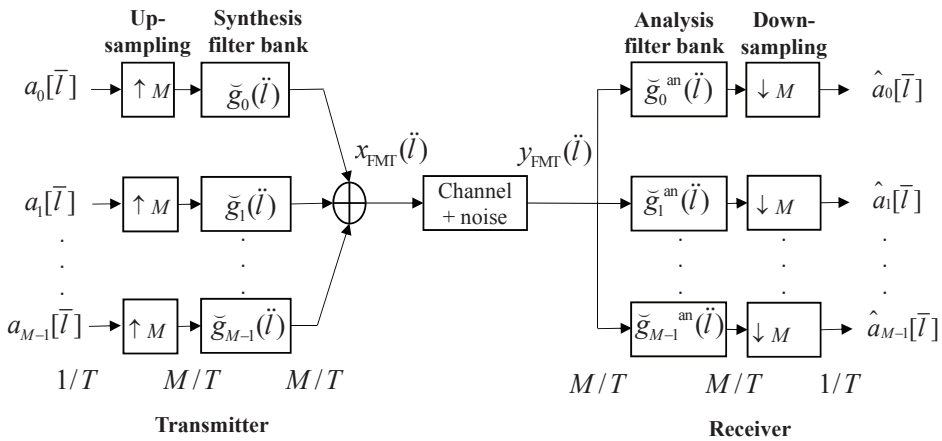


Figure 3.5: Direct implementation of FMT.

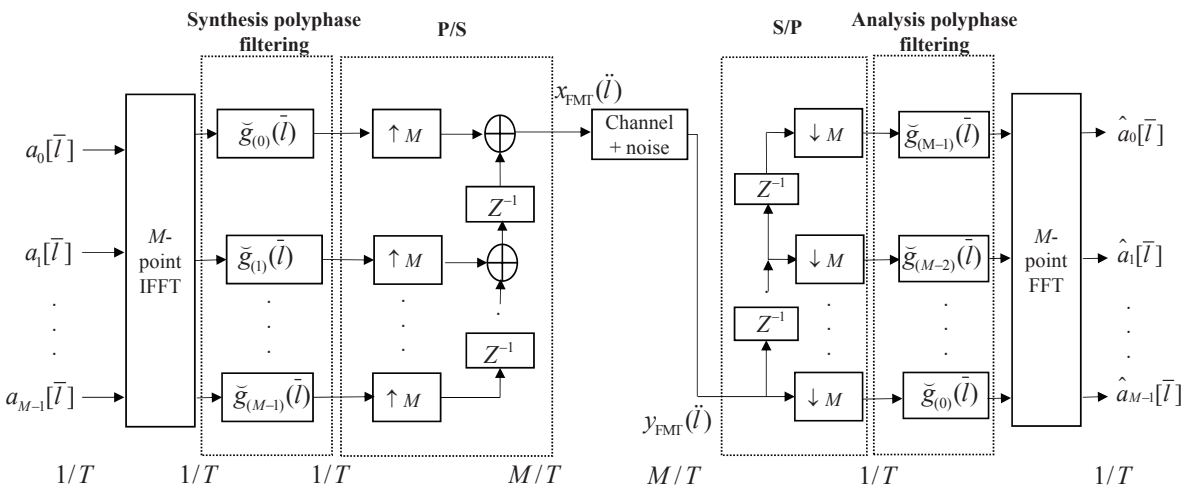


Figure 3.6: Efficient implementation of FMT.

enables the OFDM-OQAM system to outperform the conventional CP-OFDM system. The equalisation process in Figure 3.2(b) can be performed as follows [112]

$$\hat{u}_m[\tilde{l}] = r_m[\tilde{l}]\left(\frac{\hat{h}_m^*[\tilde{l}]}{\left|\hat{h}_m[\tilde{l}]\right|^2}\right), \tag{3.21}$$

where $\hat{u}_m[\tilde{l}]$ and $\hat{h}_m[\tilde{l}]$ are the estimates of the data symbols $u_m[\tilde{l}]$ and channel $h_m[\tilde{l}]$, respectively. Following the equalisation process, OQAM post-processing is performed to obtain estimates of the transmitted complex-valued symbols, $\hat{a}_m[\bar{l}]$.

3.2.3 Overview of Filtered Multitone (FMT)

The basic concept behind FMT is that the total bandwidth, B, is divided into M non-overlapping subchannels, where each one has a bandwidth of B/M. Hence, FMT is not as bandwidth-efficient as OFDM-OQAM. In order to achieve high spectral containment pulse shaping filters are used. Since the number of subcarriers in FMT is usually much less than in OFDM, the subcarrier bandwidth is wider, i.e., a shorter symbol duration. The wide subcarrier bandwidth coupled with the ability to contain the spectrum make FMT very robust against ICI [106]. The filter bank is implemented from the frequency shifted versions of a low-pass prototype filter $\breve{g}[\ddot{l}]$ as follows [116]

$$\breve{g}_m[\ddot{l}]=\frac{1}{\sqrt{M}}\breve{g}[\ddot{l}]e^{j2\pi\frac{m}{M}\ddot{l}}, \tag{3.22}$$

where M is the number of subcarriers, $\ddot{l}$=0,1,...,$M\zeta-1$, m=0,1,...,$M-1$ and ζ is the overlapping factor. The length of the prototype filter is given by $M\zeta$. The block diagram of a direct implementation of an FMT transceiver system is shown in Figure 3.5 [116]. $a_m[\bar{l}]$ are PSK or QAM symbols which are up-sampled by a factor of M with each symbol being filtered at a rate M/T (where T is the FMT symbol period). The M filtered signals that have been appropriately shifted in frequency are then summed to give the overall transmit signal $x_{\text{FMT}}[\ddot{l}]$. Each subcarrier m is centred at a frequency $f_m=m/T$. Note that the sample indices $\bar{l}$ and $\ddot{l}$ denote a sampling period of T and T/M, respectively. The transmitted FMT signal is expressed as [116]

$$x_{\text{FMT}}[\ddot{l}]=\sum_{\bar{l}=-\infty}^{\infty}\breve{g}(\ddot{l}-\bar{l}M)\frac{1}{\sqrt{M}}\sum_{m=0}^{M-1}a_m(\bar{l})e^{j2\pi m\ddot{l}/M}. \tag{3.23}$$

At the receiver, the filters are matched to those of the transmitter. The analysis filter is thus expressed as

$$\breve{g}_m^{\text{an}}[\ddot{l}]=\frac{1}{\sqrt{M}}\breve{g}(\ddot{l}-1)e^{j2\pi\frac{m}{M}\ddot{l}}, \quad \ddot{l}=1,2,...,M\zeta. \tag{3.24}$$

Hence, the output symbol on the mth subcarrier is given by

$$\begin{aligned}\hat{a}_m[\bar{l}]&=\sum_{n=1}^{M\zeta}y_{\text{FMT}}(\bar{l}M-\ddot{l})\breve{g}_m^{\text{an}}[\ddot{l}]\\&=\frac{1}{\sqrt{M}}\sum_{\ddot{l}=1}^{M\zeta}y_{\text{FMT}}(\bar{l}M-\ddot{l})\breve{g}(\ddot{l}-1)e^{j2\pi\frac{m}{M}\ddot{l}},\end{aligned} \tag{3.25}$$

where y_{FMT} is the received FMT signal. An efficient implementation of FMT using IFFT and FFT operations can be derived by introducing the polyphase components of the filter where the filtering is performed at a rate of $1/T$ instead of M/T as shown in Figure 3.6. In FMT the filters are not designed to satisfy the perfect reconstruction condition [116]. Hence, while the system is robust against ICI, it will introduce ISI for each subcarrier in a multipath channel. Equalisation can be performed in the time or frequency domain using a per-subcarrier MMSE based DFE [116].

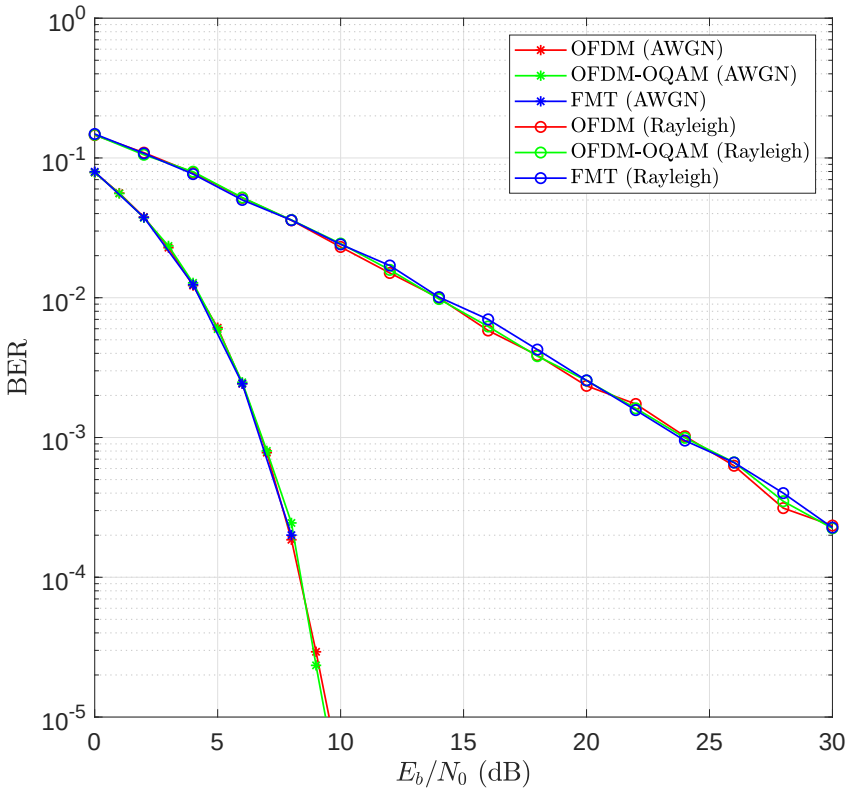


Figure 3.7: Uncoded BER performance of OFDM and FBMC systems in AWGN and 5-tap Rayleigh fading channels using QPSK modulation.

3.3 Simulation Results

Figure 3.7 shows the BER performance for un-coded OFDM and FBMC (FMT and OFDM-OQAM) systems in AWGN channel and a 5-tap Rayleigh fading channel. QPSK modulation and 8192 subcarriers were used for both OFDM and FBMC systems. The prototype filter for the OFDM-OQAM system is based on equation (3.14) with an overlapping factor of K=4 while the prototype filter for the FMT system is an FIR Hamming filter with an overlapping factor of ζ=10. It was assumed that the receiver has perfect channel knowledge and each subcarrier experiences only flat-fading in the Rayleigh channel scenario. Hence a single-tap equaliser was used for all systems where the received signal on each subcarrier was divided by the frequency domain channel response to obtain the estimated symbols. As can be observed from Figure 3.7, the FBMC systems do not require a CP to provide comparable performance to OFDM both under AWGN and multipath fading channels.

The prototype filter in FMT is usually designed to have a sharp cut-off in the frequency domain. As a result the filters will be longer and have larger sidelobes in the time domain. So even in an ideal channel scenario, ISI is likely to occur and per-subcarrier adaptive equalisation will be required to cancel the

interference. On the other hand, since in OFDM-OQAM each subcarrier is modulated with a real-valued symbol instead of complex one, pulse shape filters with good time-frequency localisation can be used. These reduce both ISI and ICI without the need of introducing the CP [117]. It should be noted, however, that according to Balian Low Theorem [118], for a multicarrier system the following desirable features cannot be achieved simultaneously:

1. Maximum spectrum efficiency,
$$TF=1, \tag{3.26}$$
T and F are the time and subcarrier spacings, respectively.

2. Time localisation
$$\varsigma_t=\sqrt{\int_{-\infty}^{\infty}\left(t-\bar{t}\right)^2|g(t)|^2\,\mathrm{d}t}<\infty, \tag{3.27}$$

3. Frequency localisation
$$\varsigma_f=\sqrt{\int_{-\infty}^{\infty}\left(f-\bar{f}\right)^2|G(f)|^2\,\mathrm{d}f}<\infty, \tag{3.28}$$

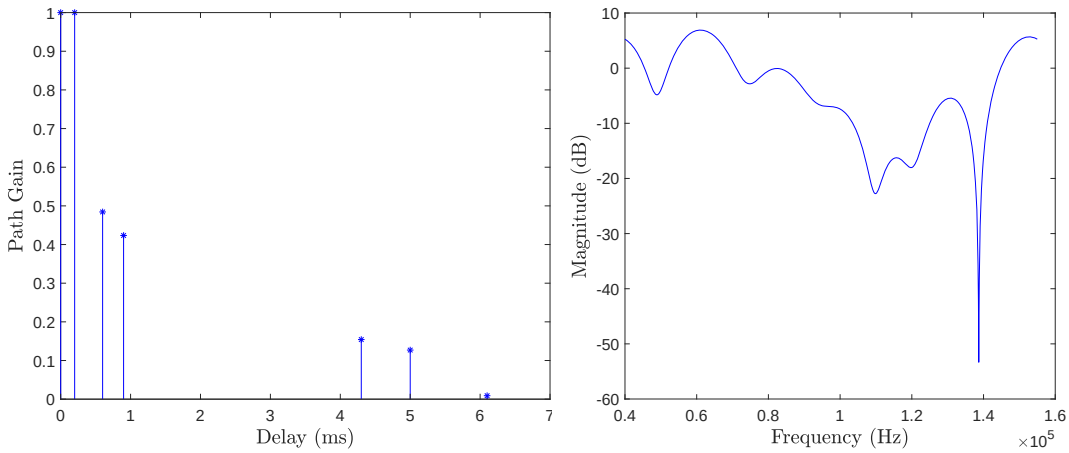
4. Orthogonality
$$\langle g_{m_1,\tilde{l}_1}(t)\ ,\ g^{\mathrm{an}}_{m_2,\tilde{l}_2}(t)\rangle = \delta_{(m_1-m_2),(\tilde{l}_1-\tilde{l}_2)} \tag{3.29}$$
i.e., the projection of the transmitted basis pulse, $g_{m,l}(t)$, onto the received basis pulse, $g^{\mathrm{an}}_{m,l}(t)$, and
$$\underbrace{\mathscr{A}\left(T(\tilde{l}_1-\tilde{l}_2)\ ,\ F(m_1-m_2)\right)}_{\text{Ambiguity function}} = \delta_{(m_1-m_2),(\tilde{l}_1-\tilde{l}_2)}, \tag{3.30}$$

where δ denotes the Kronecker delta function, the pulse $G(f)$ is the Fourier transform of the prototype filter $g(t)$, $\bar{t}=\int_{-\infty}^{\infty}t|g(t)|^2\mathrm{d}t$ and $\bar{f}=\int_{-\infty}^{\infty}f|G(f)|^2\mathrm{d}f$ are the mean time and mean frequency of the pulse, respectively. Note that for a multicarrier system using the same prototype filter, $g(t)$ both at the transmitter and receiver (matched filtering), the ambiguity function which is used to assess the time and frequency properties of the filter is defined as [119]
$$\mathscr{A}(\tau,\nu)=\int_{-\infty}^{\infty}g\left(t-\frac{\tau}{2}\right)g^*\left(t+\frac{\tau}{2}\right)e^{j2\pi\nu t}\mathrm{d}t. \tag{3.31}$$

Therefore, when designing a multicarrier waveform, at least one of the desirable features needs to be sacrificed [120]. For instance, while OFDM reduces the computational complexity using rectangular prototype filters at the transmitter and receiver and the use of a CP helps to preserve orthogonality in frequency-selective channels, it has poor spectral localisation (due to the rectangular prototype filters) and lower spectral efficiency (due to the addition of the CP). As for OFDM-OQAM, the complex orthogonality condition in (3.29) no longer holds as real-valued symbols are transmitted [120]. Equation (3.7) can be re-written as
$$s[l]=\sum_{m=0}^{M-1}\sum_{\tilde{l}=-\infty}^{\infty}b_m[\tilde{l}]g\left[l-\frac{\tilde{l}M}{2}\right]e^{j\frac{2\pi}{M}m\left(l-\frac{L_g-1}{2}\right)}e^{j\beta_m[\tilde{l}]}, \tag{3.32}$$

Table 3.1: Spectral Efficiency of OFDM-OQAM, FMT and OFDM Systems.

Multicarrier System	Complex Symbol Period	Subcarrier Spacing	Spectral Efficiency
CP-OFDM	$T+T_{\mathrm{CP}}$	$1/T$	$T/(T+T_{\mathrm{CP}})$
OFDM-OQAM	T	$1/T$	1
FMT	T	$(1+\epsilon)/T$	$1/(1+\epsilon)$



(a) (b)

Figure 3.8: Channel responses of 200 m horizontally-configured UAC (a) Power Delay Profile (b) Channel Frequency Response.

where $\beta_m[\tilde{l}]$ is a phase term and $e^{j\beta_m[\tilde{l}]} = e^{j\frac{\pi}{2}(m+\tilde{l})} = \varphi_m[\tilde{l}] = j^{m+\tilde{l}}$. Due to the phase shift of $\frac{\pi}{2}(m+\tilde{l})$, the interference which occurs in OFDM-OQAM is purely imaginary, making channel estimation and the application of MIMO to OFDM-OQAM particularly challenging [121].

Table 3.1 shows the spectral efficiency of OFDM, FMT and OFDM-OQAM, where T is the symbol period, T_{CP} is the CP duration in OFDM and ϵ is the roll-off factor of the prototype filter in FMT [106]. Since OFDM-OQAM achieves the maximum bandwidth efficiency of 100%, it is a very good candidate for UWA video transmission. For this reason, this technology is considered for the following UAC simulation instead of FMT. Two UAC transmission scenarios are investigated, namely, a horizontally-configured channel and a vertically-configured one. For a shallow water horizontally-configured multipath channel, the transmitter and receiver are assumed to be submerged at a depth of 20 m and 2 m, respectively, and the water depth is assumed to be 25 m. The transmitter and receiver are separated by a horizontal distance of 200 m. The frequency ranges from 40 kHz to 155 kHz with a centre frequency of 97.5 kHz. The maximum delay spread in this scenario is 6.1 ms. The PDP and CFR for the horizontal UAC is shown in Figure 3.8. As for the vertical transmission scenario, the transmitter-receiver pair are submerged at a depth of 195 m and 2 m, respectively, in a water depth of 200 m. In this channel configuration, the multipath propagation is less harsh and therefore the achievable

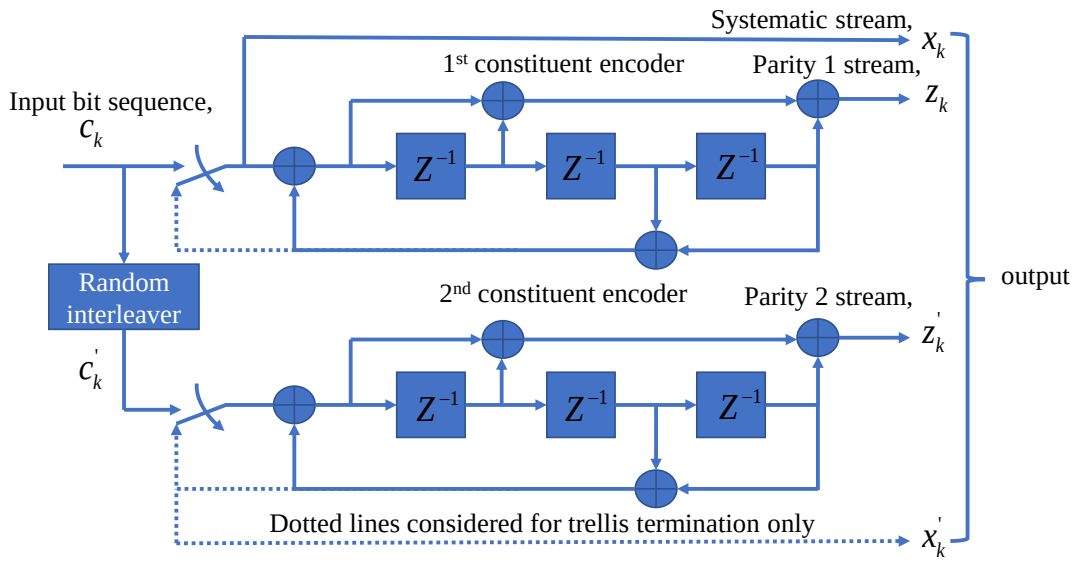


Figure 3.9: Block diagram of rate-1/3 Turbo encoder.

bit rate is mostly limited by the available bandwidth [8]. Hence, a single path which suffers from a delay and attenuation is considered for the vertical UWA link. Figure 3.10 shows the BER performance of OFDM and OFDM-OQAM (referred to as FBMC) in both channel scenarios. For the purpose of simulation, 16-QAM modulation was used with 8192 subcarriers based on the maximum delay spread of the channel. The complex symbol duration is 71.2 ms and a CP duration of 10 ms is considered for the OFDM system in both the horizontal and vertical UACs. In the legend of Figure 3.10, "H" refers to the shallow water horizontally-configured multipath channel while "V" refers to the vertically-configured channel. It is to be noted that Doppler effect has not been considered and perfect channel estimation is assumed. Furthermore, underwater noise is modelled as AWGN in both transmission scenarios. Reed Solomon (RS) and Turbo codes are used as error correction codes.

The block diagram of the Turbo encoder used in the simulations is shown in Figure 3.9 [122]. The structure of the encoder consists of a parallel concatenation of two identical 8-state constituent convolutional encoders separated by a random interleaver. The latter can be implemented for input bits of arbitrary lengths and it basically creates a mapping between the input and output bit positions. If the input bit sequence has a length of N, then $N!$ combinations are possible for the mapping but only one combination is chosen to permute the input bits [123]. The main parameters of the Turbo encoder used in the simulations throughout this thesis, unless otherwise stated, are as follows:

- Generator polynomials of 8-state convolutional encoders given by $G=[1,\frac{g_0}{g_1}]$, where g_0=[1 0 1 1] and g_1=[1 1 0 1].
- Random interleaver
- Coding rate = $\frac{1}{3}$.

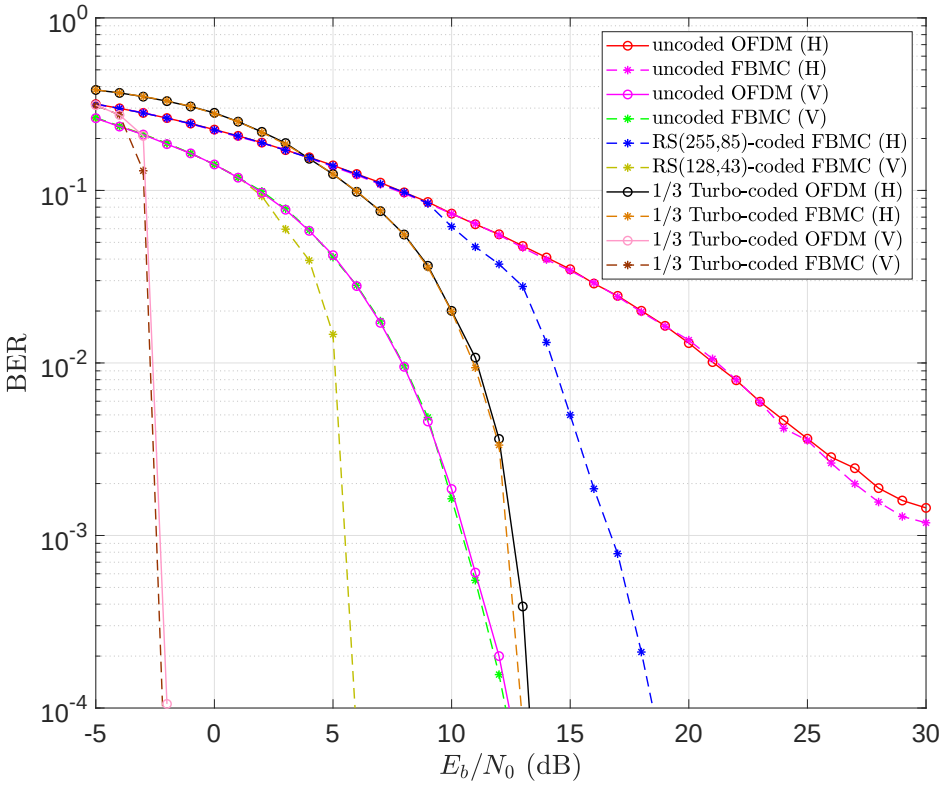


Figure 3.10: BER performance of uncoded and coded OFDM and OFDM-OQAM systems in horizontal and vertical UACs using 16-QAM.

- Number of decoding iterations = 8.

As can be observed in Figure 3.10, the uncoded CP-OFDM and OFDM-OQAM systems achieve similar performance in both channels. For the same code rate, Turbo codes with 8 decoding iterations greatly outperform the RS codes. For instance at a BER of 10^{-4}, the Turbo-coded FBMC system outperforms the RS-coded FBMC system by 8.1 dB in the vertically-configured channel and 5.6 dB in the horizontally-configured UAC. The useful bit rate of an OFDM system can be computed as

$$\text{bitrate} = r_{\text{FEC}} \times \frac{M}{T+T_{\text{CP}}} \times \text{no. of bits/symbol}, \tag{3.33}$$

where r_{FEC} is the FEC code rate, M is the number of subcarriers, T is the OFDM symbol duration and T_{CP} is the CP duration. Assuming M=8192, T_{CP}=10 ms for OFDM, $r_{\text{FEC}}=\frac{1}{3}$ and 16-QAM modulation, the theoretical bit rates for the OFDM and OFDM-OQAM systems are given in Table 3.2. For the same transmission bandwidth of 115 kHz, OFDM-OQAM achieves around 14% higher bit rate than OFDM. For the vertical UAC scenario, higher FEC code rates can be used together with higher modulation orders to achieve improved bit rates since the number of bit errors in this channel configuration is less

Table 3.2: Theoretical Achievable Bit Rate for SISO OFDM-OQAM and OFDM Systems.

	Complex Symbol Period (ms)	Bandwidth (kHz)	No. of Subcarriers	Bandwidth Efficiency (sps/Hz)	Bitrate (kbps)
CP-OFDM	81.2	115	8192	0.88	121
OFDM-OQAM	71.2	115	8192	1	138

than that in a horizontal UAC for a given E_b/N_0 value. However, the number of errors should be kept as low as possible for video transmission since only a few bit errors can severely degrade the video quality.

3.4 Conclusion

OFDM is a very popular technique that has been used for terrestrial wireless communication as well as UWA communication. However, it has some drawbacks which potentially limits its usage in future generation wireless networks where higher data rates are desirable. In this chapter, it was shown that FBMC systems can provide the same performance as the traditional OFDM system in AWGN, Rayleigh and time-invariant UACs without the use of a CP. It was also shown that with the use of robust FEC codes such as Turbo codes, very good BER performance can be obtained in the UACs. Finally, it was shown that the Turbo-coded SISO OFDM-OQAM system achieves around 14% higher bit rate than its OFDM counterpart over the 200 m UACs. This makes FBMC, more specifically OFDM-OQAM, attractive for UWA video transmission where high data rates are desirable for delivering good quality video, especially given the very limited acoustic bandwidth.

CHAPTER 4

PERFORMANCE EVALUATION OF MIMO OFDM-OQAM SYSTEMS IN TIME-VARYING UACS

4.1 Introduction

MIMO systems can be employed to either increase the link reliability (diversity gain) or enhance the data rate (spatial multiplexing gain). Spatial diversity involves encoding the information across space and time such that several copies of a signal are transmitted and the receiver then combines them to achieve a diversity gain. One of the methods to achieve diversity gain is to use Space-Time Codes (STC). These can be categorised as Space-Time Block Codes (STBC) and Space-Time Trellis Codes (STTC). In STTC, multiple replicas of a trellis code are transmitted in space and time dimensions. These codes are optimally decoded using a Viterbi decoder based on Maximum Likelihood Sequence Estimation (MLSE). STTC typically offer both coding and diversity gain. Increased performance can be achieved using more trellis states at the expense of a larger decoding complexity [124]. A very simple STBC technique for two transmit antennas and one receive antenna was proposed by Alamouti in 1998 [125]. The aim was to improve signal quality at the receiver by transmitting independent replicas of the signal over narrowband frequency-flat channels. Hence, the probability that the entire signal will fade at the same time is reduced significantly. The Alamouti STBC is orthogonal and therefore allows a simple decoding principle to be be used. Alamouti STBC can be easily applied to OFDM to yield a diversity gain in frequency-selective channels. However, in Alamouti-coded OFDM systems, two main impairments exist, namely, ISI which is caused by time dispersion and ICI which results from frequency dispersion. ISI can be mitigated by using a CP. The Alamouti scheme assumes that the channel remains unchanged over the two adjacent OFDM symbols. However, for time-varying or fast-fading channels, this assumption may no longer hold as the channel variation over each symbol becomes significant. Hence, in this case encoding can be performed in the frequency domain on adjacent

subcarriers for a given OFDM symbol instead of adjacent OFDM symbols for a given subcarrier [126]. This is known as Space-Frequency Block Coding (SFBC). The Alamouti coherence assumption is expected to hold in this case since the subcarriers in an OFDM system are closely packed together such that each subcarrier experiences only flat-fading. However when Doppler shift is high, ICI will occur between adjacent subcarriers. This destroys the orthogonality of the Alamouti scheme resulting in significant performance degradation. Some UWA studies which have considered diversity gain over shallow-water horizontally-configured UACs include [18–21] where very low BERs have been achieved over long distance acoustic links (> 1 km).

In order to increase the bit rate, MIMO systems based on spatial multiplexing can be used, where parallel streams of information are transmitted simultaneously over multiple transmitting elements without incurring a power penalty. Several detection techniques can be used at the receiver to recover the transmitted symbols for example, ZF, MMSE and Maximum Likelihood Detection (MLD). The latter is known to be an optimal detection technique but it suffers from an exponential increase in complexity with increasing modulation order and number of transmit antennas. Vertical Bell Laboratories Layered Space-Time (V-BLAST) is a popular scheme that transmits independent encoded streams over multiple antennas to exploit the spatial multiplexing gain.

The sole use of the MIMO scheme may not provide much benefits in the presence of channel impairments. The combination of MIMO with OFDM helps to achieve enhanced spectral efficiency in a frequency-selective channel while maintaining simple equalisation. SU MIMO-OFDM has been widely investigated for UWA communication to increase the data rate, see [22–24], and references therein. For instance in [22], a maximum bit rate of 125.7 kbps was reported over a horizontal transmission distance of 450 m.

In contrast, the application of MIMO to FBMC systems such as OFDM-OQAM or CMT is challenging because of the intrinsic imaginary interference which makes channel estimation difficult. Channel estimation can be carried out using preamble-based (e.g., [127–129]) or scattered pilot-based techniques [130]. The preamble-based channel estimation methods include Pairs of Pilots (POP), IAM and its variants. FMT-based MIMO systems can be easily implemented assuming that each subcarrier experiences flat-fading but is however less bandwidth-efficient than the other categories of FBMC systems.

In this chapter the BER performances of coded and un-coded spatially-multiplexed MIMO-FBMC based on OFDM-OQAM modulation and MIMO-OFDM are evaluated over 1 km time-varying horizontal and vertical UACs. Preamble-based channel estimation using the IAM is considered. The theoretical achievable bit rates of the MIMO systems are also presented for both channel configurations. It is shown that the Turbo-coded MIMO OFDM-OQAM systems not only achieve a better error performance but also a higher bit rate than the Turbo-coded MIMO-OFDM systems in both the horizontal and vertical channel scenarios. Furthermore, the transmission of a SD video is evaluated over the 1 km vertically-configured time-varying UAC using the systems that achieve the highest bit rates, that is, the Turbo-coded MIMO OFDM-OQAM. The video stream is compressed using the H.264/AVC standard. Theoretical bit rates of 73.7 kbps and 105.3 kbps obtained using the 2×12 and 4×12 OFDM-OQAM systems, respectively,

imply that an acceptable video quality can be achieved over the long-range vertical acoustic link in real-time (based on the computed video frame rates after compression). The main contributions of this chapter are published in [45, 47].

This remaining part of this chapter is structured as follows. Section 4.2 provides the discrete-time baseband models for SISO and MIMO OFDM-OQAM systems. Section 4.3 presents the simulation results of the systems in horizontally and vertically-configured UACs. The transmission of a SD video over the 1 km vertically-configured UAC is carried out in Section 4.4. Finally, Section 4.5 concludes this chapter.

4.2 System Model

4.2.1 SISO OFDM-OQAM

In OFDM-OQAM real symbols are transmitted at twice the rate of OFDM-QAM [128] and a prototype filter with good Time-Frequency Localisation (TFL) property can be used to combat both ISI and ICI, while still ensuring maximum bandwidth efficiency. However, the orthogonality condition in OFDM-OQAM only holds in the real field and therefore there will always be some kind of interference between neighbouring subcarriers and symbols that makes channel estimation difficult [129]. OFDM-OQAM is analogous to CP-OFDM, the only difference being the introduction of filter banks at the transmitter and receiver and the removal of the CP block. Re-writing equation (3.32), the discrete-time signal at the SFB output can be expressed as [131]

$$s[l]=\sum_{m=0}^{M-1}\sum_{\tilde{l}} b_{m,\tilde{l}} g_{m,\tilde{l}}[l], \tag{4.1}$$

where $b_{m,\tilde{l}}$ are the real OQAM symbols and

$$g_{m,\tilde{l}}[l]=g\left[l-\tilde{l}\frac{M}{2}\right]e^{j\frac{2\pi}{M}m(l-\frac{L_g-1}{2})}e^{j\beta_{m,\tilde{l}}}, \tag{4.2}$$

where m and $\tilde{l}$ represent the subcarrier and OQAM symbol time indices, respectively, M is the number of subcarriers, g is the real symmetric prototype filter of length $L_g=KM$ (K is the filter overlapping factor) and $e^{j\beta_{m,\tilde{l}}} = j^{m+\tilde{l}}$, as defined previously in equation (3.32). The prototype filter g is implemented in such a way that the functions $g_{m,\tilde{l}}$ have real orthogonality [129]. Thus, assuming perfect time and frequency synchronisation in an ideal channel (i.e., no channel impairments such as noise, fading, etc), imaginary interference will still occur at the AFB output, which is represented as

$$\sum_{l} g_{m,\tilde{l}}(l) g^*_{p,q}(l)=j\langle g\rangle^{p,q}_{m,\tilde{l}}, \tag{4.3}$$

and is referred to as *intrinsic* interference [129]. If the number of subcarriers is large enough such that each subcarrier experiences flat-fading and assuming that the channel is constant over the duration of the prototype filter, then the received signal for the pth subcarrier and qth OFDM-OQAM symbol can

be expressed as [129]

$$y_{p,q}=H_{p,q}b_{p,q}+\underbrace{j\underbrace{\sum_{m=0}^{M-1}\sum_{\tilde{l}}}_{(m,\tilde{l})\neq(p,q)}H_{m,\tilde{l}}b_{m,\tilde{l}}\langle g\rangle_{m,\tilde{l}}^{p,q}}_{I_{p,q}}+\eta_{p,q}, \tag{4.4}$$

where $H_{p,q}$ is the M-point CFR, $I_{p,q}$ and $\eta_{p,q}$ are the interference and noise terms, respectively. If a prototype filter with a good TFL property is used, it can be assumed that the interference is due to only the first-order neighbours of the Frequency-Time (FT) points (p,q), i.e., $\mathcal{N}_{p,q}=\{(p\pm1,q\pm1),(p,q\pm1),(p\pm1,q)\}$. Moreover, if the CFR is constant over the first-order neighbours, equation (4.4) can be approximated as [129]

$$y_{p,q}\approx H_{p,q}\tilde{b}_{p,q}+\eta_{p,q}, \tag{4.5}$$

where

$$\begin{aligned}\tilde{b}_{p,q}&=b_{p,q}+\underbrace{j\sum_{(m,\tilde{l})\in\mathcal{N}_{p,q}}b_{m,\tilde{l}}\langle g\rangle_{m,\tilde{l}}^{p,q}}_{\upsilon_{p,q}}\\&=b_{p,q}+j\upsilon_{p,q},\end{aligned} \tag{4.6}$$

and $\upsilon_{p,q}$ is the imaginary part of the interference from the neighbouring FT points. By transmitting known pilots at a given FT point and its neighbourhood $\mathcal{N}_{p,q}$, the CFR estimate is obtained as follows [129]

$$\hat{H}_{p,q}=\frac{y_{p,q}}{\tilde{b}_{p,q}}\approx H_{p,q}+\frac{\eta_{p,q}}{\tilde{b}_{p,q}}. \tag{4.7}$$

The preambles are designed such that the pilots have maximum magnitude and thus the training symbols around $b_{p,q}$ should allow all the terms in $\upsilon_{p,q}$ to have the same sign in order for them to add together for all frequencies p [129]. In this respect the interference weights $\langle g\rangle_{m,\tilde{l}}^{p,q}$ need to be computed for neighbours $(m,\tilde{l})\in\mathcal{N}_{p,q}$ at each FT point (p,q). For all symbols q, the weights follow the pattern [129]

$$\begin{matrix}(-1)^p\Theta & -\Phi & (-1)^p\Theta\\ -(-1)^p\Xi & b_{p,q} & (-1)^p\Xi\\ (-1)^p\Theta & \Phi & (-1)^p\Theta\end{matrix}, \tag{4.8}$$

where the vertical and horizontal directions represent frequency and time, respectively. Θ, Φ, Ξ have positive values less than 1 and generally Φ, $\Xi>\Theta$. A few IAM preamble variants have been proposed in the literature, namely, IAM-R (real pilot symbols) [127], IAM-I (presence of imaginary pilot symbols) [128] and IAM-C [132]. The latter is a complex-based preamble and is considered in this work. In [132], it was shown that the IAM-C preamble structure provides good performance in channels characterised by both delay and Doppler spreads. Also, the IAM-C preamble slightly outperforms the IAM-R and IAM-I preambles. The three preamble structures are shown in Figure 4.1 for the SISO case where M=8 and OQPSK modulation is considered [129]. The preamble duration for the SISO system spans over

$$\begin{matrix} 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{matrix} \qquad \begin{matrix} 0 & b_0 & 0 \\ 0 & -jb_0 & 0 \\ 0 & -b_0 & 0 \\ 0 & -b_1 & 0 \\ 0 & jb_1 & 0 \\ 0 & b_1 & 0 \\ 0 & -b_0 & 0 \\ 0 & jb_0 & 0 \end{matrix} \qquad \begin{matrix} 0 & 1 & 0 \\ 0 & -j & 0 \\ 0 & -1 & 0 \\ 0 & j & 0 \\ 0 & 1 & 0 \\ 0 & -j & 0 \\ 0 & -1 & 0 \\ 0 & j & 0 \end{matrix}$$

(a) (b) (c)

Figure 4.1: SISO preambles (a) IAM-R (b) IAM-I (b_0,b_1 $\in\{1,-1\}$) (c) IAM-C.

three OFDM-OQAM symbols. As can be seen in Figure 4.1, the pilot vector lies between two zero vectors. This is required to prevent the pilot vector from being affected by interference from the data symbols (from both the past and current frames) [129].

4.2.2 MIMO OFDM-OQAM

Extending the formulations in the previous subsection to a MIMO system with n_t transmitting elements and n_r receiving hydrophones, the signal at each receiving hydrophone can be written as [129]

$$y^r_{p,q}=\sum_{t=1}^{n_t} H^{r,t}_{p,q}\tilde{b}^t_{p,q}+\eta^r_{p,q}, \quad r=1,2,\cdots,n_r, \tag{4.9}$$

where $H^{r,t}_{p,q}$ is the M-point CFR form the tth transmit antenna to the rth receive antenna, $\tilde{b}^t_{p,q}$ and $\eta^r_{p,q}$ are the corresponding virtual transmitted symbol and noise term, respectively. An overall input-output equation for the MIMO OFDM-OQAM system can be written as

$$\mathbf{y}_{p,q}=\mathbf{H}_{p,q}\tilde{\mathbf{b}}_{p,q}+\boldsymbol{\eta}_{p,q}, \tag{4.10}$$

where

$$\mathbf{y}_{p,q}=\begin{bmatrix} y^1_{p,q} & y^2_{p,q} & \cdots & y^{n_r}_{p,q} \end{bmatrix}^{\mathsf{T}},$$

$$\boldsymbol{\eta}_{p,q}=\begin{bmatrix} \eta^1_{p,q} & \eta^2_{p,q} & \cdots & \eta^{n_r}_{p,q} \end{bmatrix}^{\mathsf{T}},$$

$$\mathbf{H}_{p,q}=\begin{bmatrix} H^{1,1}_{p,q} & H^{1,2}_{p,q} & \cdots & H^{1,n_t}_{p,q} \\ H^{2,1}_{p,q} & H^{2,2}_{p,q} & \cdots & H^{2,n_t}_{p,q} \\ \vdots & \vdots & \ddots & \vdots \\ H^{n_r,1}_{p,q} & H^{n_r,2}_{p,q} & \cdots & H^{n_r,n_t}_{p,q} \end{bmatrix},$$

$$\tilde{\mathbf{b}}_{p,q}=\mathbf{b}_{p,q}+j\mathbf{v}_{p,q},$$

with

$$\mathbf{b}_{p,q}=\begin{bmatrix} b^1_{p,q} & b^2_{p,q} & \cdots & b^{n_t}_{p,q}, \end{bmatrix}^{\mathsf{T}},$$

0	1	0	1	0		0	1	0	-1	0
0	-j	0	-j	0		0	-j	0	j	0
0	-1	0	-1	0		0	-1	0	1	0
0	j	0	j	0		0	j	0	-j	0
0	1	0	1	0		0	1	0	-1	0
0	-j	0	-j	0		0	-j	0	j	0
0	-1	0	-1	0		0	-1	0	1	0
0	j	0	j	0		0	j	0	-j	0
	(a)						(b)			

Figure 4.2: 2 x n_r IAM-C preamble (a) Transmit transducer 1 (b) Transmit transducer 2.

and $\mathbf{v}_{p,q}$ is defined similarly. At least $2n_t$+1 OFDM-OQAM symbol durations are required to estimate the CFR [129]. For illustration, a $2\times n_r$ MIMO system is considered. The IAM-C preamble structure in this case is shown in Figure 4.2 (M=8 and Offset Quadrature Phase Shift Keying (OQPSK) modulation assumed). Considering only the pilot vectors at timing instants q=1,3, equation (4.10) can be re-written as [129]

$$\begin{bmatrix}\mathbf{y}_{p,1} & \mathbf{y}_{p,3}\end{bmatrix}=\mathbf{H}_{p,1}\begin{bmatrix}\tilde{b}^1_{p,1} & \tilde{b}^1_{p,3}\\ \tilde{b}^2_{p,1} & \tilde{b}^2_{p,3}\end{bmatrix}+\begin{bmatrix}\boldsymbol{\eta}_{p,1} & \boldsymbol{\eta}_{p,3}\end{bmatrix}. \tag{4.11}$$

Analysing the preamble in Figure 4.2, one can notice that $\tilde{b}^1_{p,1}=\tilde{b}^1_{p,3}=\tilde{b}^2_{p,1}=-\tilde{b}^2_{p,3}\equiv\tilde{b}_p$ and thus

$$\begin{bmatrix}\boldsymbol{y}_{p,1} & \boldsymbol{y}_{p,3}\end{bmatrix}=\mathbf{H}_{p,1}\begin{bmatrix}\tilde{b}_p & \tilde{b}_p\\ \tilde{b}_p & -\tilde{b}_p\end{bmatrix}+\begin{bmatrix}\boldsymbol{\eta}_{p,1} & \boldsymbol{\eta}_{p,3}\end{bmatrix} \tag{4.12}$$

$$=\mathbf{H}_{p,1}\tilde{b}_p\mathbf{Q}_2+\begin{bmatrix}\boldsymbol{\eta}_{p,1} & \boldsymbol{\eta}_{p,3}\end{bmatrix}, \tag{4.13}$$

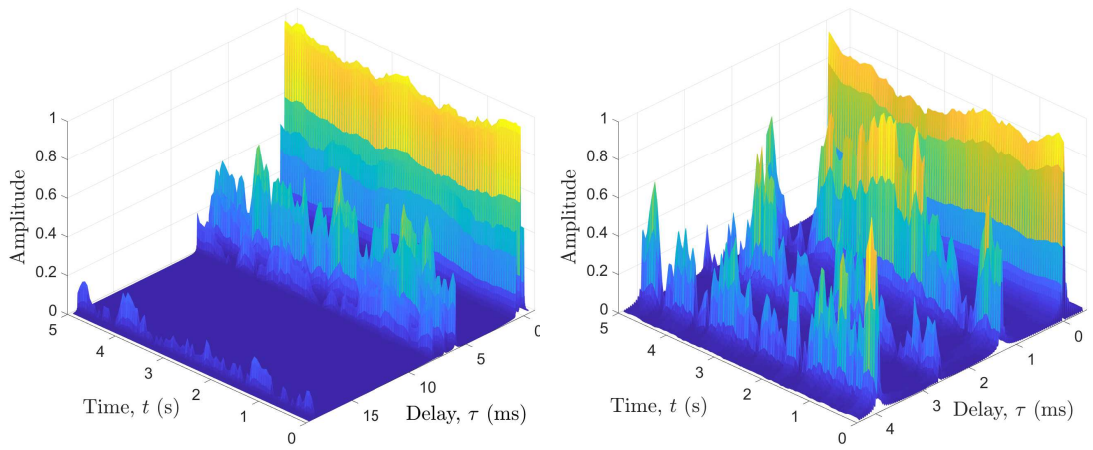
where $\mathbf{Q}_2$ is a Hadamard matrix of order n_t. The CFR estimate at the pth subcarrier is obtained as follows

$$\hat{\mathbf{H}}_{p,1}=\begin{bmatrix}\mathbf{y}_{p,1} & \mathbf{y}_{p,3}\end{bmatrix}\frac{1}{\tilde{b}_p}\mathbf{Q}_2^{-1} \tag{4.14}$$

$$=\mathbf{H}_{p,1}+\frac{1}{2\tilde{b}_p}\begin{bmatrix}\boldsymbol{\eta}_{p,1} & \boldsymbol{\eta}_{p,3}\end{bmatrix}\mathbf{Q}_2. \tag{4.15}$$

4.3 Simulation Results

Initially, a horizontally-configured time-varying channel of link distance 1000 m and a water depth of 50 m (shallow water) is considered. The transmitter and receiver are located at a depth of 48 m and 6 m, respectively. The bandwidth of the system is 25 kHz with a carrier frequency, f_c of 32.5 kHz. A statistical model of the UAC is used where the channel coefficients are obtained using the maximum entropy principle [76]. A maximum RMS Doppler spread of 2 Hz is considered. A typical horizontal



(a) (b)

Figure 4.3: Time-variant channel impulse responses (a) Horizontal UAC (b) Vertical UAC.

CIR as observed between a given transmitting element and receiving element is shown in Figure 4.3(a) where the maximum delay spread is 18.6 ms. The BER performance for the MIMO-OFDM and MIMO OFDM-OQAM systems in the horizontal UAC is shown in Figure 4.7. Coloured noise is considered in the simulations to better reflect the real-world underwater ambient noise. For this purpose, a Low-pass Filter (LPF) is used to filter white noise such that its response approximates the decaying PSD of underwater ambient noise [133], as shown in Figure 4.4. The transfer function of the LPF is given by

$$\mathcal{H}(Z)=\frac{1}{Z^6-0.9Z^5+0.4Z^4-0.2Z^3+0.1Z^2-0.1Z}. \tag{4.16}$$

The other parameters include 16-QAM modulation and 1024 subcarriers. The CP duration for the OFDM system is 20.48 ms. The preambles for the MIMO-OFDM and MIMO OFDM-OQAM systems span over 2 complex symbol durations and 5 real symbol durations, respectively, for n_t=2. For n_t=4, the preamble lengths are 4 complex symbol durations and 9 real symbol durations for the MIMO-OFDM and MIMO OFDM-OQAM systems, respectively.

The OFDM-OQAM system is based on a Hermite prototype filter with an overlapping factor of K=4. Haas and Belfiore [100] pointed out that the prototype filter $g(t)$ can be implemented using a set of Hermite functions $\mathcal{H}_i(\cdot)$ as follows [130]

$$g(t)=\frac{1}{\sqrt{T_0}}e^{-2\pi\left(\frac{t}{T_0}\right)^2}\sum_{i=\{0,4,8,12,16,20\}}a_i\mathcal{H}_i\left(2\sqrt{\pi}\frac{t}{T_0}\right), \tag{4.17}$$

where T_0 is a parameter that performs a scaling in time and which is dependent on the required subcarrier (or time) spacing, and the coefficients are given by [130]

$$\begin{aligned} a_0&=1.412692577, & a_{12}&=-2.2611.10^{-9},\\ a_4&=-3.0145.10^{-3}, & a_{16}&=-4.4570.10^{-15},\\ a_8&=-8.8041.10^{-6}, & a_{20}&=1.8633.10^{-16}. \end{aligned} \tag{4.18}$$

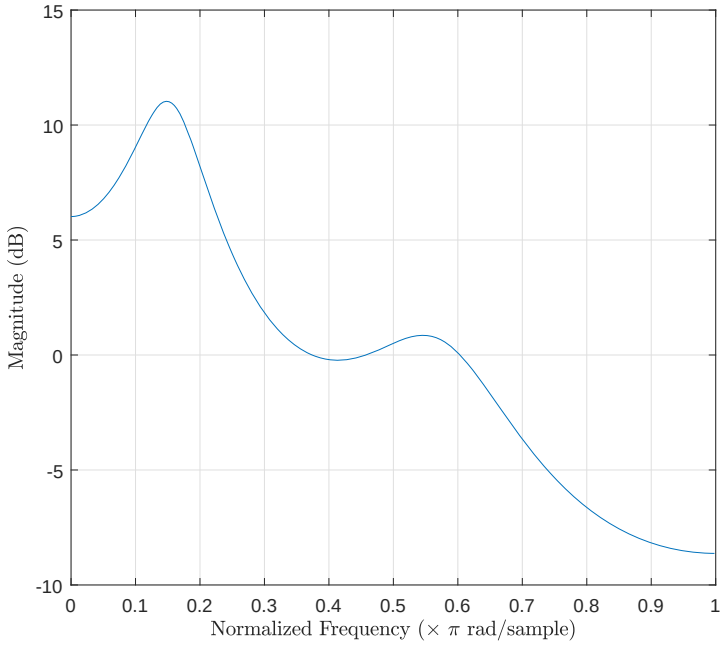
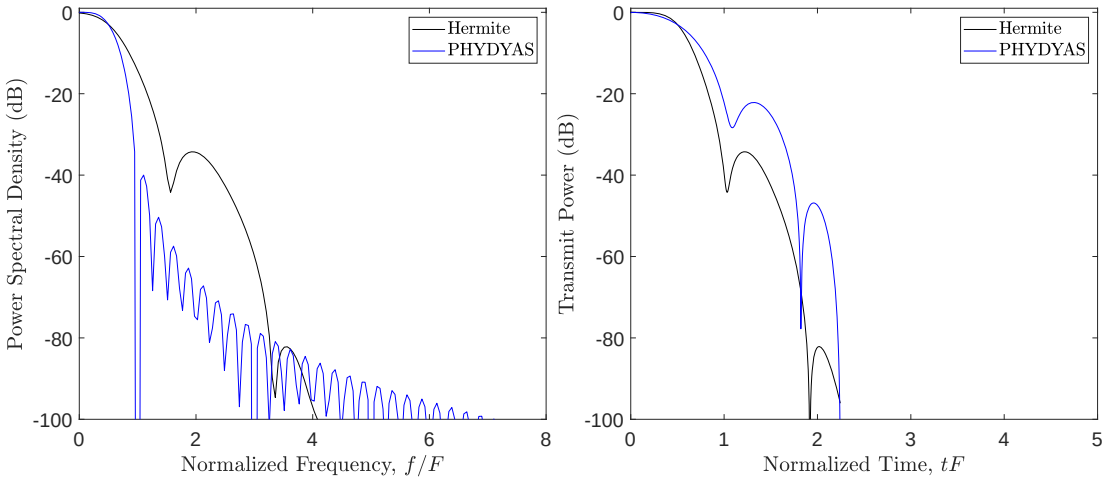


Figure 4.4: Low-pass filter response used to generate coloured noise.

The prototype filter defined in equation (4.17) guarantees orthogonality for frequency and time spacings of F=2/T_0 and T=T_0, respectively [130]. The time and frequency localisation parameters (see equations (3.27) and (3.28)) for the Hermite filter are given as ς_t=0.2015T_0 and ς_f=0.403/T_0, respectively, (compared to ς_t=0.2745T_0 and ς_f=0.328/T_0 for the PHYDYAS filter) [130]. The Hermite filter is based on the Gaussian function and is symmetrical both in the time and frequency domains. The Hermite pulse has a joint TFL of $\varsigma_t\varsigma_f$=0.0812045 and therefore satisfies the bound of $\varsigma_t\varsigma_f \geq 1/4\pi$ which is achieved by the Gaussian pulse [130]. This makes the Hermite filter suitable for channels characterised by both time and frequency dispersions (doubly-dispersive). It is to be noted that the PHYDYAS filter has a better frequency localisation than the Hermite filter but a worse time localisation and a worse joint TFL [120]. The comparison of the frequency and impulse responses between the Hermite and PHYDYAS prototype filters for an overlapping factor of K=4 are shown in Figure 4.5.

A single-tap equaliser is used for the MIMO-OFDM system while the receiver structure for the MIMO OFDM-OQAM system consists of a 3-tap equaliser which is implemented using the frequency-sampling approach [115]. The latter technique allows the subcarrier frequency response to be equalised at a number of frequencies which correspond to the number of equaliser coefficients as shown in Figure 4.6. The subcarrier spacing is denoted as Δf. For a single-tap equaliser, the coefficient is simply computed as $EQ(m)$=1/$\mathscr{C}(m)$ (where m is the subcarrier index and $\mathscr{C}(m)$ is the measured value of the CFR at the mth subcarrier). As for a multi-tap equaliser, the coefficients are computed from multiple values of the CFR. Recall that at the AFB in the receiver, there is an oversampling by a factor of 2 and hence the subcarrier sampling frequency is effectively $2\Delta f$. For the 3-tap equalisation process shown



(a) (b)

Figure 4.5: Comparison of the frequency and impulse responses of Hermite and PHYDYAS prototype filters (a) Frequency Response (b) Impulse Response.

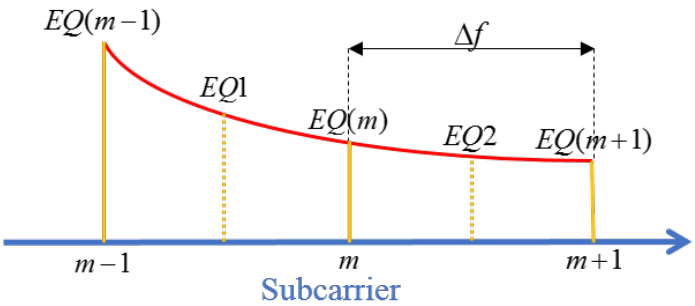


Figure 4.6: 3-tap equalisation.

in Figure 4.6, three frequencies are of interest: the centre frequency $m\Delta f$ and the two intermediate frequencies $(m\pm 1/2)\Delta f$. The equaliser coefficients are then computed such that its frequency response passes through $EQ1$, $EQ(m)$ and $EQ2$ [115]. Since the frequency sampling step is $\Delta f/2$, the total number of samples is equal to $2\Delta f/(\Delta f/2)$=4, implying that one value (denoted as U) is unknown. The equaliser coefficients can be computed using the IFFT of the set $\{U,EQ1,EQ(m),EQ2\}$ [115]. Let the coefficients be denoted by $\{eq_0,eq_1,eq_2,eq_3\}$ where eq_0 is the centre coefficient which is surrounded by eq_{-1}=eq_3 and eq_1. For a 3-tap equaliser, eq_2 is set to zero to compute the unknown value U using the relation [115]

$$U-EQ1+EQ(m)-EQ2=0. \tag{4.19}$$

The centre coefficient eq_0 is then given by

$$eq_0=U+EQ1+EQ(m)+EQ2. \tag{4.20}$$

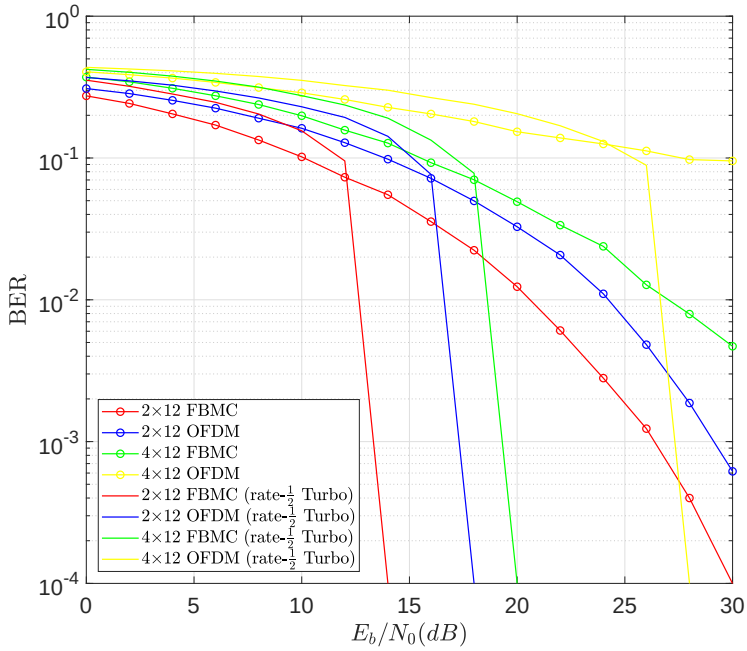


Figure 4.7: BER performance of uncoded and coded $n_t \times n_r$ OFDM & FBMC systems using 16-QAM in horizontal UAC.

Assuming linear interpolation, the interpolated values $EQ1$ and $EQ2$ are given by

$$EQ1=[EQ(m-1)+EQ(m)]/2, \tag{4.21}$$

$$EQ2=[EQ(m)+EQ(m+1)]/2, \tag{4.22}$$

and the coefficients are computed as follows

$$eq_0=[EQ(m-1)+EQ(m+1)+2EQ(m)]/4, \tag{4.23}$$

$$eq_1=[EQ(m-1)+EQ(m+1)-2EQ(m)+j(EQ(m-1)-EQ(m+1))]/8, \tag{4.24}$$

$$eq_{-1}=[EQ(m-1)+EQ(m+1)-2EQ(m)-j(EQ(m-1)-EQ(m+1))]/8. \tag{4.25}$$

Considering the characteristics of the OQAM scheme, the signs of the coefficients eq_1 and eq_{-1} need to be changed according to the subcarrier index m and this is done through a multiplication by $(-1)^m$ [115].

As can be observed in Figure 4.7, the MIMO OFDM-OQAM systems achieve better performance than the MIMO-OFDM systems in the horizontal UAC for the same transmission time. This shows the robustness of OFDM-OQAM against both time and frequency dispersions. It should be noted that the subcarriers in a conventional OFDM-OQAM system may not necessarily experience flat-fading since the lack of a cyclic prefix implies that the condition for circular convolution is not satisfied. Hence, it is desirable to have a large number of subcarriers so that they can experience flat-fading. Furthermore,

the approximation in (4.5) is not accurate, especially for channels with high time dispersions [129]. When the number of transmitting elements is increased from 2 to 4, the performance is worse because of cross-talk between parallel channels.

The data rate can increase linearly with the number of transmitting elements in a MIMO system. However, in fast-time varying UACs, the number of transmitting elements is kept low to achieve a good trade-off between data rate and error performance. For a given bandwidth B, M subcarriers and CP duration T_{CP}, the bandwidth efficiency, $\mathscr{B}_{\mathrm{raw}}$ of a MIMO-OFDM system in symbols/s/Hz/transmitter is given by [23]

$$\mathscr{B}_{\mathrm{raw}}=\frac{1}{1+BT_{\mathrm{CP}}/M}. \tag{4.26}$$

Thus, for n_t transmitting elements and modulation level $\mathscr{M}$, the effective bandwidth efficiency is given by

$$\mathscr{B}_{\mathrm{eff}}=n_t\times\mathscr{B}_{\mathrm{raw}}\times\log_2\mathscr{M}. \tag{4.27}$$

In a system where overheads such as OFDM CP duration, FEC code rate, pilot and null subcarriers used for channel and Doppler estimation, respectively, are considered, the effective bandwidth efficiency in bits/s/Hz is given by

$$\mathscr{B}_{\mathrm{eff}}=n_t\times r_{\mathrm{FEC}}\times\frac{T}{T+T_{\mathrm{CP}}}\times\frac{M_{\mathrm{data}}}{M}\times\log_2\mathscr{M}, \tag{4.28}$$

where r_{FEC} is the FEC code rate, M_{data} is the number data-carrying subcarriers and T is the OFDM symbol duration. For a given system's bandwidth B, the number of subcarriers M is limited by the channel coherence time, t_{coh} [23]. When designing an OFDM system, it is often assumed that the channel remains approximately constant over one block, i.e., $T=M/B\ll t_{\mathrm{coh}}$ [23]. However, the OFDM symbol duration increases with increasing number of subcarriers and hence it becomes increasingly difficult to track the channel variations which may result in ICI. A shorter block duration prevents ICI and ensures minimum system's complexity but represents an under-usage of the system's resources since each block is followed by a CP. Furthermore, the number of transmitting elements cannot be increased indefinitely due to cross-talk between parallel channels. In [23], the maximum number of transmitting elements is stated to satisfy the expression $n_t\leq M/N_{\mathrm{mp}}$ where N_{mp} is the number of channel impulse response coefficients (channel span). Therefore, the bandwidth efficiency is limited as [23, 134]

$$n_t/(1+\tau B/M)\leq M^2/N_{\mathrm{mp}}(M+N_{\mathrm{mp}}), \tag{4.29}$$

where τ is the channel delay spread. Ideally, $t_{\mathrm{coh}}\gg\tau$ and hence the system can be designed such that $M\gg N_{\mathrm{mp}}$. In this case, the bandwidth efficiency will behave as $M/N_{\mathrm{mp}}\ll t_{\mathrm{coh}}/\tau$ [23].

Considering the rate-$\frac{1}{2}$ Turbo-coded MIMO systems in Figure 4.7, at a BER of 10^{-4}, the 2×12 and 4×12 OFDM-OQAM systems yield 4 dB and 8 dB better performance than the 2×12 and 4×12 OFDM systems, respectively. Note that puncturing is applied to the original $\frac{1}{3}$-rate Turbo encoded bit sequence to obtain a sequence with a code rate of $\frac{1}{2}$. This process consists of omitting some of the coded bits during transmission to reduce the amount of redundancy (therefore better bandwidth efficiency is achieved) but this however reduces the effectiveness of the error correcting capability of the coding

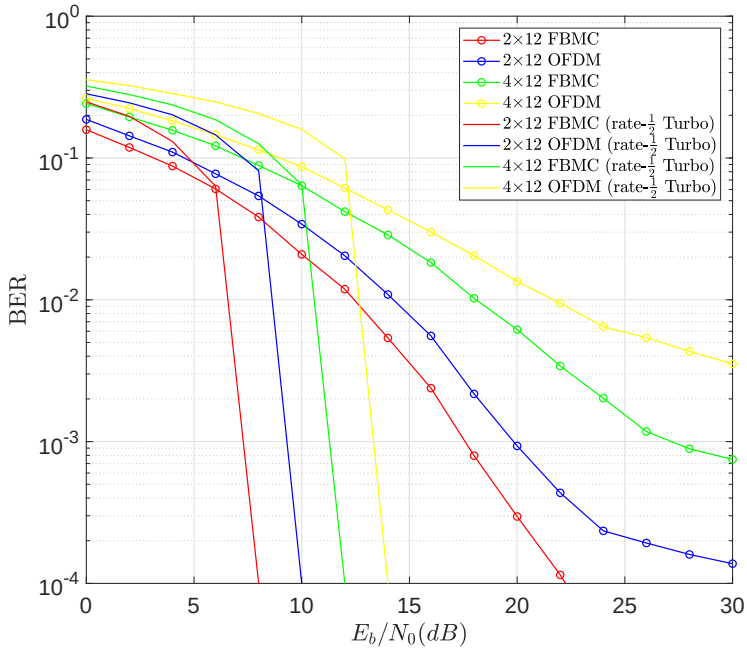


Figure 4.8: BER performance of uncoded and coded $n_t \times n_r$ OFDM & FBMC systems using 16-QAM in vertical UAC.

technique. For the vertically-configured channel, the transmitter and receiver are submerged at a depth of 998 m and 1 m, respectively. The horizontal separation between them is 1 m. All other parameters are the same as the horizontal channel scenario except for the number of subcarriers and CP duration which are set to 512 and 5.12 ms, respectively. A typical CIR as observed between a given transmitting element and receiving hydrophone in the vertical channel scenario is shown in Figure 4.3(b) where the delay spread is 3.9 ms. As can be observed in Figure 4.8, the MIMO OFDM-OQAM systems once again outperform the MIMO-OFDM systems. Considering the same transmission time, at a BER of 10^{-4}, the Turbo-coded 2×12 and 4×12 OFDM-OQAM systems outperform the coded OFDM systems with the same MIMO configuration by 2 dB in both cases.

By taking all overhead due to channel estimation into account and considering the same transmission time for both the OFDM and OFDM-OQAM systems, the theoretical achievable bit rates for the MIMO systems in the 1000 m horizontal and vertical UACs are provided in Table 4.1. The parameters as used for the simulations are considered. In the horizontal channel, the 2×12 and 4×12 OFDM-OQAM systems achieve about 40% and 67% higher bit rate than the 2×12 and 4×12 OFDM systems, respectively. As for the vertical channel, the increase in bit rate with OFDM-OQAM compared to OFDM is around 17% and 25% for the 2×12 and 4×12 systems, respectively.

Table 4.1: Theoretical Bit Rates (kbps) Achieved with MIMO OFDM-OQAM Systems.

Channel Configuration	OFDM		OFDM-OQAM	
	n_t=2	n_t=4	n_t=2	n_t=4
Horizontal	52.6	63.2	73.7	105.3
Vertical	63.2	84.2	73.7	105.3

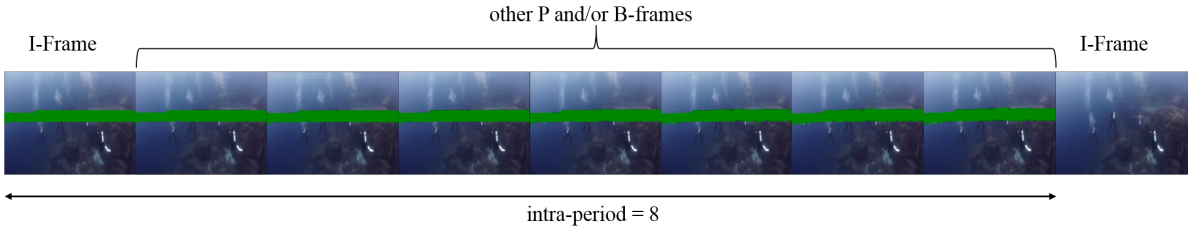


Figure 4.9: Error propagation within group of pictures (GOP).

4.4 Video Transmission using MIMO OFDM-OQAM Systems

In this section, the transmission of a SD video is evaluated over the 1000 m vertical time-varying UAC using the MIMO OFDM-OQAM systems. The bit rate values for the MIMO OFDM-OQAM systems in Table 4.1 are considered for video transmission. An uncompressed SD video with a resolution of 640×480, frame rate of 30 fps and duration 12 s is considered. The video stream (365 frames) is compressed to target bit rates of 70 kbps and 100 kbps for the $2\times n_r$ and $4\times n_r$ OFDM-OQAM systems, respectively. Frequent coding of the I-frames avoids propagation of errors between frames, thereby improving the error resilience of the video streams. Hence, a short intra-period of 8 is used. If for example there are bit errors causing packet loss in the first I-frame, these errors will only propagate in the trailing frames within the intra-period as shown in Figure 4.9. The maximum packet size is set at 800 bytes and the number of reference frames for video coding is set to 5.

4.4.0.1 Frame Rate Computation

The average number of bits per frame for the video streams compressed at 70 kbps and 100 kbps are 2355 and 3364, respectively. It is to be noted that the I-frames consist of more bits than the P and/or B video frames. Assuming the maximum supported capacity of the systems, the calculated frame rate for the $2\times n_r$ and $4\times n_r$ OFDM-OQAM systems is 31.3 fps in both cases. This value is greater than the original frame rate of 30 fps, implying real-time transmission is theoretically feasible.

4.4.0.2 Packet Loss Rate

The H.264/AVC bitstreams are encoded in 502 and 543 video packets for the $2\times n_r$ and $4\times n_r$ OFDM-OQAM systems, respectively. Figure 4.10 shows the PLR for the two systems. The PLR is zero at E_b/N_0

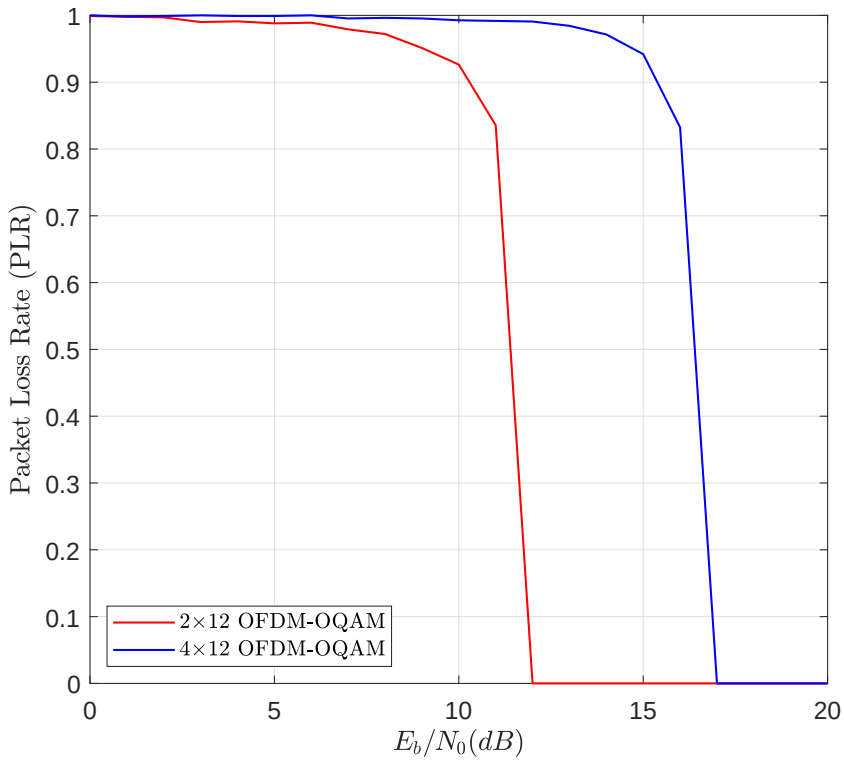


Figure 4.10: Packet Loss Rate for the 2×12 and 4×12 OFDM-OQAM systems.

values above 12 dB and 17 dB for $2\times n_r$ and $4\times n_r$ systems, respectively.

4.4.0.3 Peak-Signal-to-Noise-Ratio (PSNR)

The graph in Figure 4.11 shows the received video quality for the systems under investigation. The received video quality is assessed using the PSNR metric. The frame copy error concealment technique is used in order to cope with the partial or complete loss of frames due to packet losses. For low bit rate video compression, it is very likely that a whole video frame is encoded within a single packet. In this case, a number of bit errors distributed across the video stream may cause many packet losses (and therefore frame losses), making error concealment techniques ineffective. The range of E_b/N_0 values shown in Figure 4.11 denotes a maximum PLR of 25% for each system. Beyond this value, the video quality is so severely degraded that no useful information can be extracted from the received videos. The higher the supported bit rate of the system, the better is the video quality, as can be observed for the $4\times n_r$ system when packet loss is negligible.

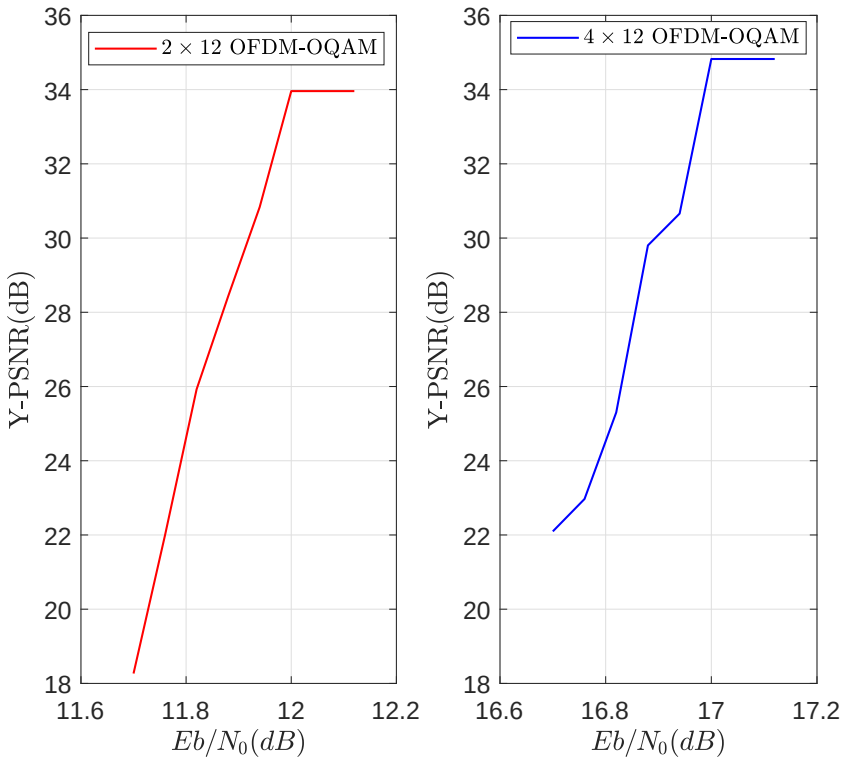


Figure 4.11: Received video quality in terms of PSNR (dB) for the 2×12 and 4×12 OFDM-OQAM systems.

4.5 Conclusion

In this chapter, it was shown that MIMO OFDM-OQAM provides a better error performance than MIMO-OFDM in both horizontally and vertically configured time-varying UACs. This shows that OFDM-OQAM can provide robust performance in doubly-dispersive channels by using prototype filters with good TFL properties. The channel configuration also determines the optimum parameters that can be used. For instance, it was shown that in a horizontally-configured channel the error rate is higher than in a vertical channel due to the more severe time dispersion. The absence of a CP in OFDM-OQAM implies that more useful bits can be transmitted per second. For instance, the theoretical bit rates of 73.7 kbps and 105.3 kbps obtained using the 2×12 and 4×12 OFDM-OQAM systems, respectively, were found to be sufficiently high to transmit a SD video with acceptable quality in real-time over the 1 km vertical acoustic link.

CHAPTER 5

SINGLE-USER AND MULTI-USER UNDERWATER ACOUSTIC COMMUNICATION USING FBMC-BASED MASSIVE MIMO SYSTEMS

5.1 Introduction

Massive MIMO technology is currently being considered for the Fifth Generation (5G) terrestrial wireless communication systems in order to provide additional data capacity and enhance the system's robustness to channel impairments. This technology is also being investigated for UWA communication for enhanced bandwidth efficiency. For instance in [98], an FBMC-based MU massive MIMO system with 32 subcarriers is proposed with 4 transmitting stations and 40 receiving hydrophones at the base-station. The system's bandwidth is 5 kHz and the carrier frequency is 8 kHz. Each of the four transmitting stations is equipped with a single loudspeaker. CMT is selected as the FBMC system. The short symbol duration allows tracking of channel variations using a blind equalisation method based on Constant Modulus Algorithm (CMA). Hence pilot symbols are not used, thereby achieving a high bandwidth efficiency (8 b/s/Hz). A matched filter estimator was used at the receiver. The self-equalisation property of the massive MIMO system allowed 32 wide subcarriers to be used for a system's bandwidth of 5 kHz and a carrier frequency of 8 kHz. Simulation results showed that the blind tracking algorithm allowed data to be transmitted without the need for pilot symbols. Furthermore, the SINR for the four users increased from approximately 12 dB to 18 dB with increasing number of base-station hydrophones (from 20 to 100). The authors of [135] derived a closed-form expression for the uplink achievable rate in a massive MIMO-OFDM system with carrier aggregation. ZF equalisation was assumed. A simulated UAC was considered, with a water depth of 5000 m and a fixed receiver array depth of 5 m. The distance between the transmitter (4 transducers) and receiver (100 hydrophones) was varied from 1 km to 25 km in the simulations. A system bandwidth of 20.48 kHz was considered with 2048 subcarriers and a CP duration of 20 ms was assumed. A frequency range from 10 to 500 kHz

was allocated for the uplink transmission. Simulated capacity curves showed that the carrier aggregated massive MIMO system can offer high data rates for both short/medium and long distances compared to systems without carrier aggregation.

In this chapter massive MIMO reception is proposed to improve the communication reliability over a 1 km vertically-configured time-varying UAC and at the same time boost the achievable bit rate for both SU and MU scenarios. For the SU scenario, the BER performance of Turbo-coded massive MIMO systems based on FBMC modulation and OFDM are evaluated over the UAC. The transmission setup consists of a large array of 100 receiving hydrophones at a surface vessel (BS) while the underwater ROV is equipped with at most four transducers. To overcome the imaginary interference in the conventional OFDM-OQAM system and making its application to MIMO straightforward, a modified OFDM-OQAM system as proposed in [136] is considered whereby complex data symbols are transmitted instead of real symbols. This is made possible by spreading the data symbols in time while still ensuring maximum bandwidth efficiency and low complexity. Scattered pilot-based channel estimation is applied to both the massive MIMO OFDM and OFDM-OQAM systems. It is shown that the coded FBMC systems outperform the coded OFDM systems both in terms of error performance and bit rate for the same massive MIMO configuration. The transmission of a 480p video which is compressed using the H.264/AVC standard is also evaluated in the UAC with the systems that achieve the highest bit rates. With bit rates as high as 133 kbps and 242 kbps for the 2×100 and 4×100 FBMC systems, respectively, it is shown that real-time transmission is possible with acceptable video quality over the long range acoustic link.

The MU scenario consists of four underwater ROVs, each transmitting with four transducers to the surface BS which is equipped with 100 receiving hydrophones. The performance of the 4-user Turbo-coded FBMC-based massive MIMO system is assessed over the simulated 1 km time-varying UAC in terms of BER, Packet Error Rate (PER) and maximum achievable throughput. It is demonstrated that the implementation of massive MIMO at the receiver allows all the users (ROVs) to use the same bandwidth to transmit their information reliably to the surface BS. Furthermore, it is shown that each ROV can achieve a bit rate of approximately 221 kbps, which is high enough to support real-time video transmission with adequate quality.

The main contributions of this chapter are published in [43, 46]. The rest of this chapter is organised as follows. Section 5.2 provides the basic concepts of a massive MIMO system. The discrete-time baseband model for the modified FBMC-OQAM system is presented in Section 5.3. Section 5.4 presents the performance evaluation of the SU OFDM and FBMC based systems in the vertically-configured UAC in terms of BER, theoretical achievable bit rate and received video quality. In Section 5.5, the performance of the 4-user FBMC-based massive MIMO system is evaluated over the 1 km vertical UAC in terms of BER, PER and maximum achievable bit rate. Finally, concluding remarks are drawn in Section 5.6.

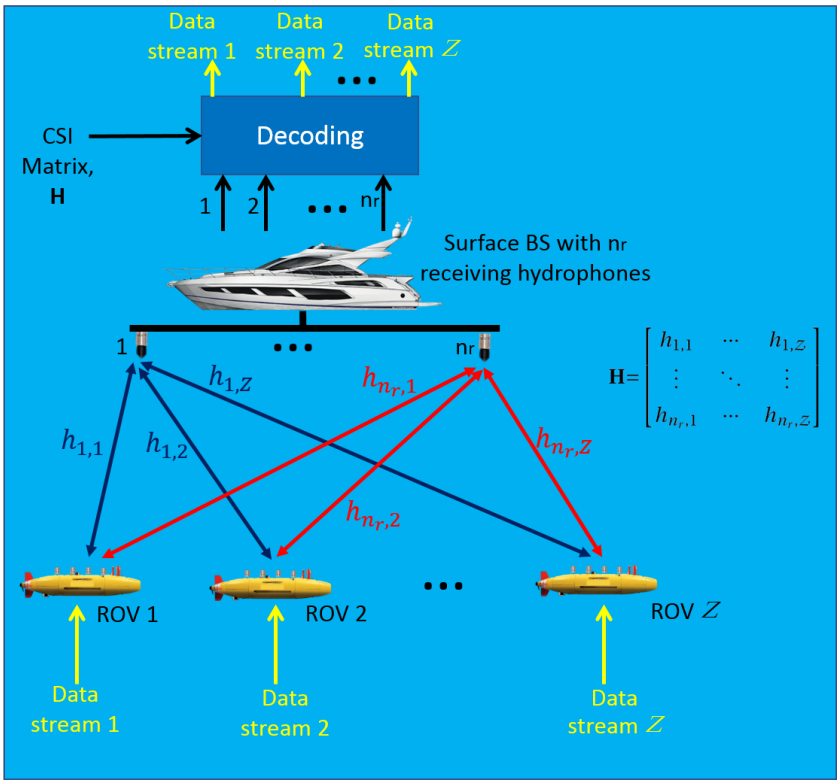


Figure 5.1: Illustration of uplink massive MIMO operation.

5.2 Overview of Massive MIMO

Massive MIMO technology is expected to significantly improve the throughput and reliability not only in future terrestrial communication systems but also for UWA communication. In contrast to conventional MU-MIMO, Massive MIMO uses a multitude of antennas at the BS which is orders of magnitude more than the number of users being served by the BS [98]. It is a scalable technology that improves both the spectral and power efficiencies and simplifies the signal processing as the number of antennas increases at the BS [137]. It uses spatial division multiplexing whereby the multiple information streams are transmitted in the same time and frequency resources.

In the downlink operation, the BS must ensure that each user terminal only receives the information that is meant for it [137]. For the BS to perform the multiplexing process, it must have knowledge of the channel response between each user and each of its antennas. This response is commonly known as the Channel State Information (CSI). In the downlink operation, the CSI is used in the pre-coding process to map the information streams that will be transmitted through the BS antennas. The higher the number of antennas, the narrower the signal beams become and these can be focused more sharply or selectively to the intended user [137, 138].

The uplink operation involves the transmission of information streams from the users to the BS over the same frequency and time resources. An example of the uplink operation in a massive MIMO system

Table 5.1: Uplink and Downlink Signal Processing in a Massive MIMO System.

	Decoding (uplink)	Pre-coding (downlink)
MRC	$\mathbf{H}^{\mathsf{H}}$	$\mathbf{H}^{*}$
ZF	$(\mathbf{H}^{\mathsf{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathsf{H}}$	$\mathbf{H}^{*}(\mathbf{H}^{\mathsf{T}}\mathbf{H}^{*})^{-1}$
MMSE	$(\mathbf{H}^{\mathsf{H}}\mathbf{H}+\frac{\mathcal{Z}}{\rho_{\mathrm{u}}}\mathbf{I}_{\mathcal{Z}})^{-1}\mathbf{H}^{\mathsf{H}}$	$\mathbf{H}^{*}(\mathbf{H}^{\mathsf{T}}\mathbf{H}^{*}+\frac{\mathcal{Z}}{\rho_{\mathrm{d}}}\mathbf{I}_{\mathcal{Z}})^{-1}$
	where ρ_{u} is the uplink average SNR	where ρ_{d} is the downlink average SNR

assuming single-antenna users (ROVs) is shown in Figure 5.1. The large antenna (hydrophone) array at the BS captures the signals transmitted from the different users and the BS uses the CSI to decode (de-multiplex) the composite received signal to finally obtain the individual user signals.

Pilots are used to estimate the channels between the BS antennas and user terminals to obtain the CSI which is used for the pre-coding and decoding processes. The CSI can be regarded as a group of spatial transfer functions between each BS antenna and user terminal [139] and which can be arranged in a matrix ($\mathbf{H}$) as shown in Figure 5.1. For a Time-Division Duplex (TDD) massive MIMO system, the channels between the BS antennas and users are assumed to be reciprocal, which implies that a given channel is the same in both the uplink and downlink [138, 139]. This channel reciprocity also implies that the channels need to be estimated in one direction only. To this end, uplink channel estimation is preferred since fewer pilot signals need to be transmitted from the users to the BS. Furthermore, the channel estimation complexity is dependent only on the number of users and not the number of BS antennas [139]. This fact is of high importance in an environment where the users are in constant motion and therefore the channels need to be estimated more frequently. Another major advantage of uplink channel estimation is that this computationally intensive task, as well as other signal processing, are only performed at the BS [139].

The pre-coding and decoding operations in the downlink and uplink, respectively, can be performed using linear techniques such as ZF, MMSE and Maximum-Ratio Combining (MRC). The latter is a simple approach that tries to maximise the received SNR at the expense of multi-user interference. In contrast, ZF nullifies the interference between users but neglects the effect of noise. MMSE aims to achieve a balance between noise enhancement and inter-user interference suppression [140] and therefore, as compared to ZF and MRC, it maximises the received SINR [141]. Considering a massive MIMO system with n_r BS antennas and $\mathcal{Z}$ users with single transmitting antenna, the pre-coding and decoding matrices for MRC, ZF and MMSE are given in Table 5.1. With increasing number of BS antennas, linear pre-coding and decoding techniques are nearly optimal [137, 138].

Massive MIMO is considered to be a scalable technology because, firstly, only the BS has to learn the matrix, $\mathbf{H}$; secondly, the number of BS antennas is much higher than the number of user terminals and thirdly, simple linear processing techniques can be used for both the uplink and downlink operations [138]. Massive MIMO also ensures simple signal processing due to an effect known as *channel hardening*, where the effects caused by small-scale fading are minimised as the number of BS antenna increases [138]. Small-scale fading can be broadly defined as the rapid fluctuations in

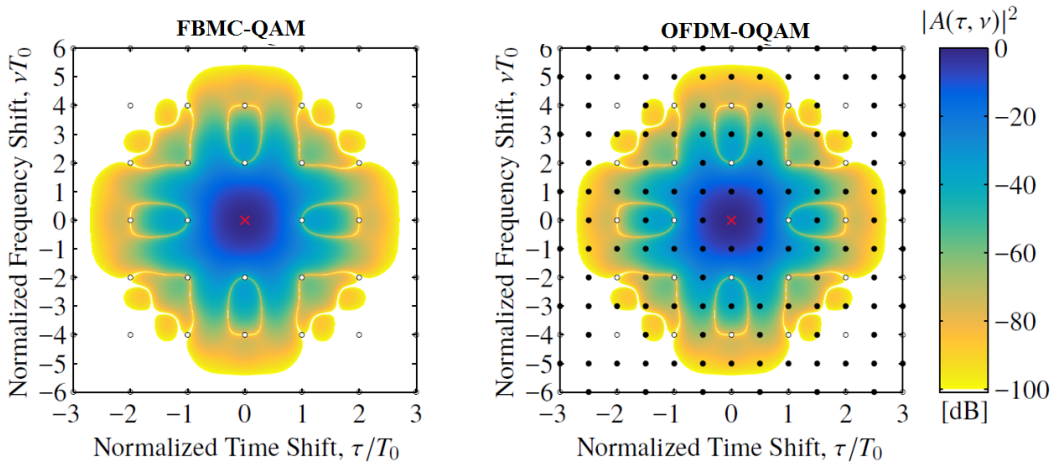


Figure 5.2: Ambiguity function (see equation (3.31)) of Hermite filter in the case of FBMC-QAM and OFDM-QAM systems [119].

the channel gain due to very small changes in the propagation channel [142]. These fluctuations can sometimes result in the channel gain to be very small, causing the transmitted signal to be received in error. Small-scale fading can be mitigated by exploiting the diversity provided by multiple transmitting and receiving elements. Owing to the large number of antennas in a massive MIMO system, the channel hardens and hence the user terminals do not need to perform channel estimation as the channel seen by each one of them behaves as if it is a non-fading one [138, 142].

5.3 Restoring Complex Orthogonality in OFDM-OQAM

Recall that in a conventional OFDM-OQAM system real symbols are transmitted, allowing prototype filters with a good TFL property to be used. However, its orthogonality condition only holds in the real field. It should be noted that OFDM-OQAM is related to FBMC-QAM but has the same symbol density as OFDM without the CP [119]. There is no unique definition for FBMC-QAM as some authors prefer to sacrifice frequency localisation, yielding a worse performance than OFDM in terms of out-of-band radiation while others sacrifice orthogonality to achieve both time and frequency localisation and obtain a time-frequency spacing of $TF{\approx}1$ [119]. The working principle of an OFDM-OQAM system can be described as follows [119]:

1. Design a prototype filter with $g(t){=}g(-t)$, which guarantees orthogonality for frequency and time spacings of $F{=}2/T_0$ and $T{=}T_0$, respectively, such as the Hermite filter in equation (4.17).

2. Reduce both the time and frequency spacings by a factor of 2 such that $TF{=}0.5$ in terms of real-valued symbols, which is equivalent to $TF{=}1$ for complex symbols.

3. This compression in time and frequency causes purely imaginary interference due to the phase shift of $\frac{\pi}{2}(m{+}\tilde{l})$.

This process is illustrated in Figure 5.2 where the left part of the figure shows an FBMC-QAM system designed using a Hermite filter for a time-frequency spacing of TF=2. This large time-frequency spacing ensures both time and frequency localisation and complex orthogonality at the expense of a reduced spectral efficiency. Improved robustness can be achieved in doubly-selective channels in this case [119]. For the case of OFDM-OQAM in the right part of Figure 5.2, the frequency and time spacings are reduced to $F=1/T_0$ and $T=T_0/2$, respectively. The black dots represent the induced interference due to the squeezing in time and frequency and it is purely imaginary due to the phase shift of $\frac{\pi}{2}(m+\tilde{l})$. The authors of [136] proposed to spread the symbols either in time or frequency to completely eliminate the imaginary interference. This technique not only allows complex symbols to be transmitted but also ensures a low complexity since it is based on Hadamard matrices. For an OFDM-OQAM system with M subcarriers and N symbols, the continuous-time transmitted signal can be expressed as

$$s(t)=\sum_{\tilde{l}=1}^{N}\sum_{m=1}^{M} g_{m,\tilde{l}}(t)b_{m,\tilde{l}}, \tag{5.1}$$

where $b_{m,\tilde{l}}$ are the transmitted symbols on the mth subcarrier at the $\tilde{l}$th timing instant. $g_{m,\tilde{l}}(t)$ is the time and frequency shifted version of the prototype filter $g(t)$ which is given by

$$g_{m,\tilde{l}}(t)=g(t-\tilde{l}T)e^{j2\pi mF(t-\tilde{l}T)}e^{j\varrho_{m,\tilde{l}}}, \tag{5.2}$$

where $\varrho_{m,\tilde{l}}=\frac{\pi}{2}(m+\tilde{l})$ is the phase shift. The received symbols $y_{m,\tilde{l}}$ are obtained by projecting the received signal $r(t)$ on the received basis pulses, $g^{\text{an}}_{m,\tilde{l}}(t)$ as follows [136]

$$y_{m,\tilde{l}}=\langle r(t),g^{\text{an}}_{m,\tilde{l}}(t)\rangle=\int_{-\infty}^{\infty} r(t)\left(g^{\text{an}}_{m,\tilde{l}}\right)^{*}(t)\,\mathrm{d}t. \tag{5.3}$$

5.3.1 Matrix-based System Model of Conventional OFDM-OQAM System

Equation (5.1) can be formulated in matrix notation as

$$\mathbf{s}=\mathbf{G}\mathbf{b}, \tag{5.4}$$

where each column of $\mathbf{G}$ is a vector of the sampled basis pulses $g_{m,\tilde{l}}(t)$:

$$\mathbf{G}=\left[\mathbf{g}_{1,1}\ \cdots\ \mathbf{g}_{M,1}\ \mathbf{g}_{1,2}\ \cdots\ \mathbf{g}_{M,N}\right], \tag{5.5}$$

and $\mathbf{b}$ is the transmitted symbol vector which is represented as

$$\mathbf{b}=\left[b_{1,1}\ \cdots\ b_{M,1}\ b_{1,2}\ \cdots\ b_{M,N}\right]^{\mathsf{T}}. \tag{5.6}$$

The sampled basis pulses $g^{\text{an}}_{m,\tilde{l}}(t)$ at the receiver can also be stacked as column vectors to obtain the matrix

$$\mathbf{G}^{\text{an}}=\left[\mathbf{g}^{\text{an}}_{1,1}\ \cdots\ \mathbf{g}^{\text{an}}_{M,1}\ \mathbf{g}^{\text{an}}_{1,2}\ \cdots\ \mathbf{g}^{\text{an}}_{M,N}\right]. \tag{5.7}$$

The two different transmit and receive pulse matrices in equations (5.5) and (5.7), respectively, provide a more general description of multicarrier systems such as CP-OFDM and filtered/windowed OFDM.

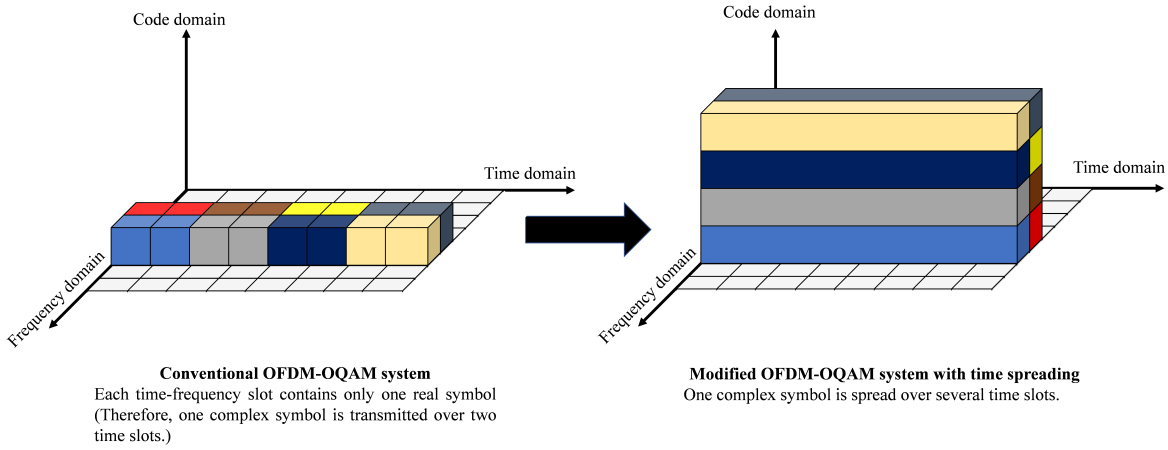


Figure 5.3: Block time-spreading of OFDM-OQAM.

However, it should be noted that FBMC usually makes use of a matched filter, i.e., $\mathbf{G}=\mathbf{G}^{\text{an}}$ [120]. The received symbols in (5.3) can be reformulated as

$$\mathbf{y}=(\mathbf{G}^{\text{an}})^{\mathsf{H}}\mathbf{r}=(\mathbf{G}^{\text{an}})^{\mathsf{H}}\mathbf{H}\mathbf{G}\mathbf{b}+\boldsymbol{\eta}, \tag{5.8}$$

where $\mathbf{H}$ is the channel convolution matrix, $\mathbf{r}$ is the sampled received signal vector and $\boldsymbol{\eta}$ is the random noise vector. Due to the orthogonality condition in OFDM, $(\mathbf{G}^{\text{an}})^{\mathsf{H}}\mathbf{G}=\mathbf{I}_{MN}$. However, this is not the case in OFDM-OQAM where the orthogonality holds only in the real field. In order to obtain an identity matrix, only the real part of the transmission matrix needs to be considered, i.e., $\mathscr{R}\{(\mathbf{G}^{\text{an}})^{\mathsf{H}}\mathbf{G}\}=\mathscr{R}\{\mathbf{G}^{\mathsf{H}}\mathbf{G}\}=\mathbf{I}_{MN}$. Using eigendecomposition, it can be found that $\mathbf{G}^{\mathsf{H}}\mathbf{G}$ contains exactly $\frac{MN}{2}$ non-zero eigenvalues, each having a value of two as $\{M,N\}\rightarrow\infty$ [136]. This means that only $\frac{MN}{2}$ complex symbols can be transmitted (or equivalently MN real symbols). For OFDM, the channel equalisation can be easily performed using a ZF equaliser, $\left[(\mathbf{G}^{\text{an}})^{\mathsf{H}}\mathbf{H}\mathbf{G}\right]^{-1}$ or an MMSE equaliser. In conventional OFDM-OQAM, this direct matrix inversion is not practical due to the imaginary interference and the fact that $(\mathbf{G}^{\text{an}})^{\mathsf{H}}\mathbf{H}\mathbf{G}$ is not a full rank matrix [120].

5.3.2 OFDM-OQAM based on Block Spreading

To overcome the shortcomings of conventional OFDM-OQAM, the complex orthogonality has to be restored and this can be achieved when the symbols are spread either in frequency [143] or time [136]. Figure 5.3 illustrates the concept of block spreading for OFDM-OQAM in the time domain. The left part of Figure 5.3 represents the conventional OFDM-OQAM system where each time-frequency position consists of one real symbol (therefore one complex symbol is transmitted in two time slots). The time-spreading process, as shown in the right part of Figure 5.3, consists of spreading one complex symbol over several time slots, thereby cancelling the imaginary interference and restoring the complex orthogonality. In fact, an additional dimension (coding) is involved in this process besides the time and frequency dimensions. OFDM-OQAM with block spreading is closely related to the Code Division

Multiple Access (CDMA) technique as used in the Third Generation (3G) mobile communication system [120]. However, a Root Raised Cosine (RRC) filter or RAKE receiver are not required in this case. Instead, simple one-tap equalisation can be performed assuming that the channel remains approximately flat in time or frequency depending on the type of block spreading used [120]. Usually, the spreading increases the sensitivity to both frequency and time selectiveness which occur due to multipath propagation and Doppler effect, respectively. However, if the delay and Doppler spreads are not significant, the interference caused by the channel can be neglected and in this case each data symbol can be allowed to occupy more than only a small time-frequency resource of TF=1 [120]. The spreading process can be performed using the Walsh-Hadamard method where the complex orthogonality can be restored with lower computational complexity as compared to the FFT spreading method [136]. One of the main reason to perform block spreading is to allow the integration of MIMO with OFDM-OQAM with approximately the same complexity as MIMO-OFDM. Furthermore, all receiver signal processing methods commonly used in MIMO-OFDM can be straightforwardly applied to its OFDM-OQAM counterpart using block spreading [136].

5.3.2.1 Block Spreading Procedure

The complex-valued data symbols, denoted by $\mathbf{a}\in\mathbb{C}^{\frac{MN}{2}\times 1}$, are precoded using a spreading matrix $\mathbf{C}\in\mathbb{C}^{MN\times\frac{MN}{2}}$ such that the transmitted symbols $\mathbf{b}\in\mathbb{C}^{MN\times 1}$ are given by [136]

$$\mathbf{b}=\mathbf{C}\mathbf{a}. \tag{5.9}$$

In order to eliminate the imaginary interference and restore orthogonality, the spreading matrix $\mathbf{C}$ should be selected to satisfy the following condition [136]

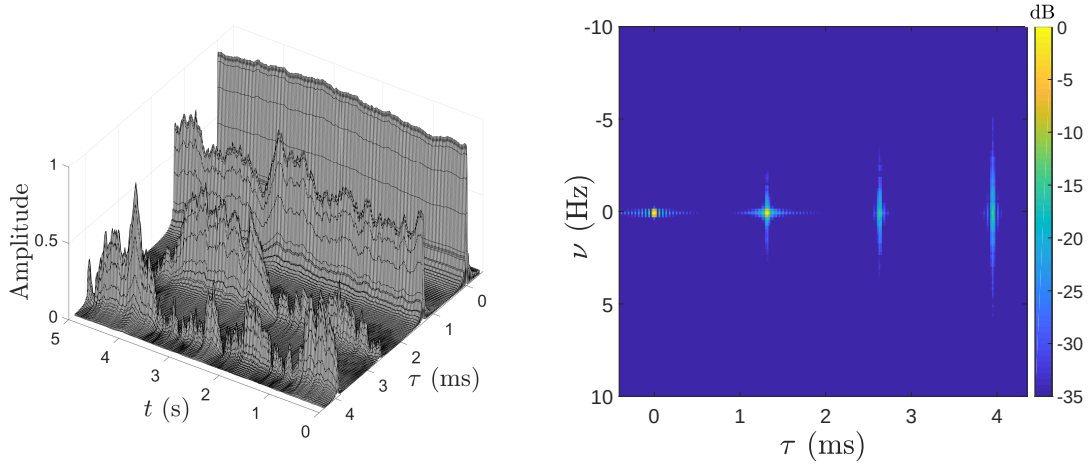
$$\mathbf{C}^{\mathsf{H}}(\mathbf{G}^{\text{an}})^{\mathsf{H}}\mathbf{G}\mathbf{C}=\mathbf{I}_{\frac{MN}{2}}. \tag{5.10}$$

At the receiver, decoding is performed such that the received complex data symbols $\bar{\mathbf{y}}\in\mathbb{C}^{\frac{MN}{2}\times 1}$ are obtained as follows

$$\bar{\mathbf{y}}=\mathbf{C}^{\mathsf{H}}\mathbf{y}. \tag{5.11}$$

Let the transmission matrix be denoted by $\mathbf{T}_{\text{mat}}=(\mathbf{G}^{\text{an}})^{\mathsf{H}}\mathbf{G}$. The condition in (5.10) can be satisfied by performing the eigenvalue decomposition of the transmission matrix $\mathbf{T}_{\text{mat}}$, i.e., $\mathbf{T}_{\text{mat}}=\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$ where $\mathbf{\Lambda}$ is a diagonal matrix of the eigenvalues and $\mathbf{U}$ represents a unitary matrix, and then selecting the first $\frac{MN}{2}$ column vectors of $\mathbf{U}$ as the spreading matrix $\mathbf{C}$ [136]. This method is however not practical since it will result in a high computational complexity [136]. A more efficient method is to select $\frac{N}{2}$ appropriate column vectors from a Hadamard matrix of order N and then spread the complex symbols over several time slots. This process is done for all subcarriers until a spreading matrix of size $MN\times\frac{MN}{2}$ that satisfies the condition in (5.10) is obtained [136].

One major advantage of using Hadamard spreading is its low complexity processing as multiplications are not involved and only additions/subtractions are required at the transmitter and receiver [136].



(a) (b)

Figure 5.4: Example of vertical channel responses (a) Channel Impulse Response (b) Channel Scattering Function.

The only problem of using Hadamard matrices for spreading is that the spreading length needs to be a power of two [144]. Hence, integrating OFDM-OQAM with block spreading into an existing terrestrial communication system may represent a challenge [136]. On the other hand, this will not be an issue for an underwater communication system since it has to be designed from scratch for a given transmission scenario and thus the problem of compatibility with existing systems does not arise.

In theory, the number of FBMC symbols can approach infinity but due to the time spreading over N time slots, this is not practical because of latency constraints. Moreover, for a time-varying channel the subcarrier orthogonality can be lost. Hence, block-wise transmission should be performed. However, inter-block interference may occur due to small overlapping between blocks. In conventional OFDM-OQAM, this overlapping would not be a major concern due to the real orthogonality condition. Nevertheless, this is no longer valid if time spreading is applied within a transmission block [136]. One guard slot per block may not completely eliminate the interference between neighbouring blocks. On the other hand using more guard slots will decrease the spectral efficiency. By using a higher spreading length in a time-invariant channel, the interference can be spread over a higher number of symbols, thereby improving the SIR [136]. For a time-varying channel, the robustness of the systems can be improved by using a smaller spreading length.

5.4 Performance Evaluation of a SU Massive MIMO System in an UAC

5.4.1 BER Performance

A near-vertical channel with a water depth of 1000 m is considered. The transmitter and receiver are immersed at a depth of 998 m and 1 m, respectively. The horizontal separation between them is assumed

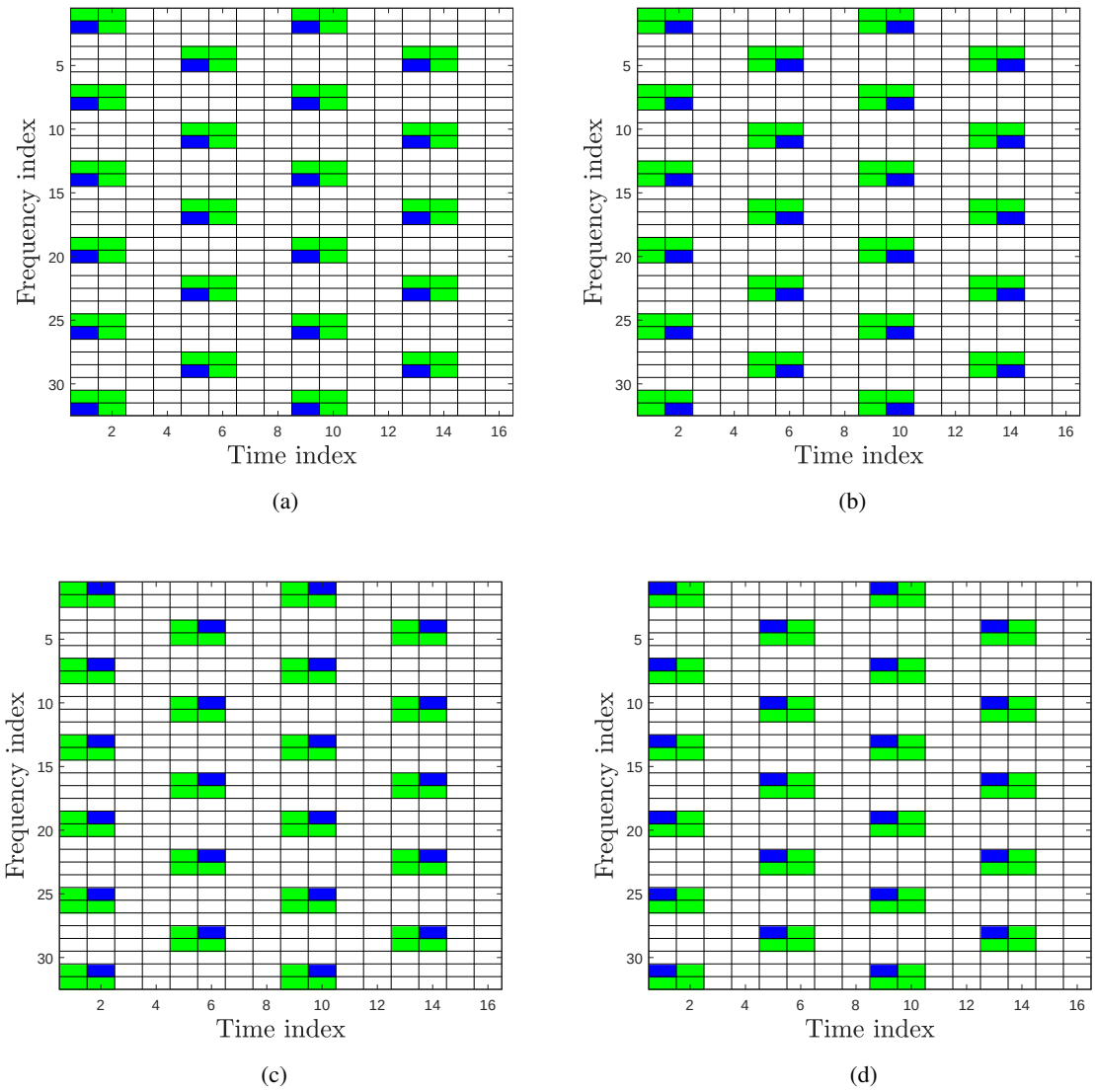


Figure 5.5: Example of pilot pattern for a 4-transmitting element setup (a) TX 1 (b) TX 2 (c) TX 3 (d) TX 4 (Blue - Pilot Symbol, Green - Zero Symbol, White - Data Symbol).

to be 5 m. The system's bandwidth is set at 25 kHz and the carrier frequency f_c is 32.5 kHz. The maximum RMS Doppler spread in the vertical UAC is 2 Hz. Typical channel responses as observed between a given transmitting element and receiving hydrophone in this channel configuration are shown in Figure 5.4, where the maximum delay spread in the UAC is 3.9 ms. The number of subcarriers is set at 256 for both the FBMC and OFDM systems. Coloured noise is also used in the simulations instead of AWGN to better represent the real-world noise in an UAC. The other system parameters include 64-QAM modulation and rate-$\frac{1}{2}$ punctured Turbo codes for the coded MIMO systems (with n_t transmit elements and n_r receive hydrophones). Due to the time spreading, block-wise transmission is considered with 64 FBMC symbols (real) and a zero time-slot per block (to minimise the inter-block interference).

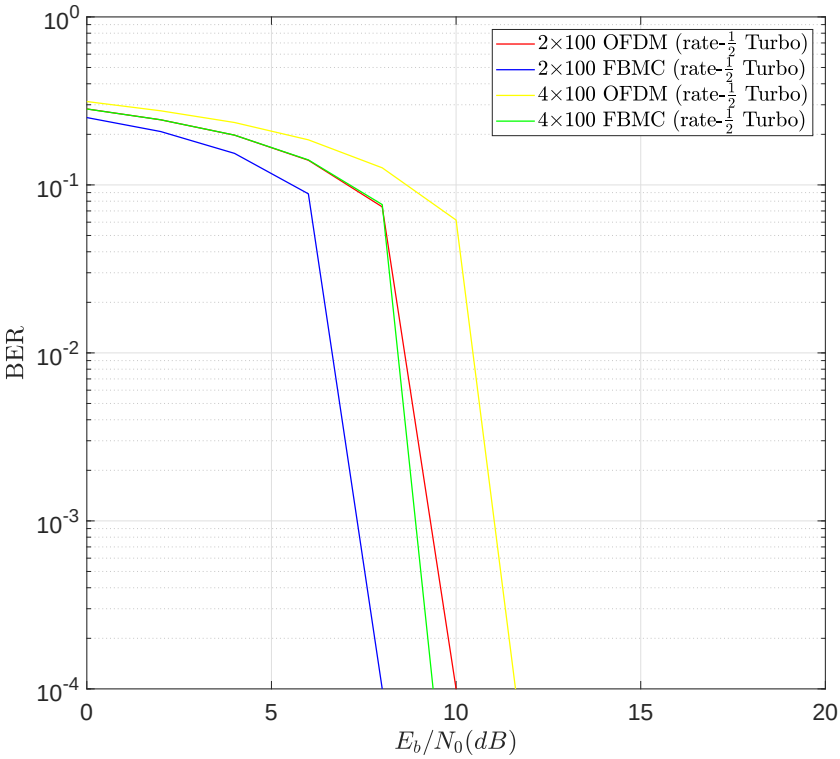


Figure 5.6: BER performance of rate-$\frac{1}{2}$ Turbo-coded OFDM and OFDM-OQAM systems using 64-QAM and massive MIMO reception (100 elements) in a 1000 m vertical UAC.

With time spreading, 32 complex symbols are available per subcarrier per block for the FBMC system. Considering a CP duration of 5.12 ms and assuming the same transmission time as in the OFDM-OQAM system, the OFDM system requires 22 complex symbols per subcarrier per block. No guard slot is required between blocks for the OFDM case. Pilot-based channel estimation is considered for both the OFDM and OFDM-OQAM systems with the pilots arranged as per the Long Term Evolution (LTE) standard and reproduced in Figure 5.5. For the same transmission time, the OFDM-OQAM and OFDM systems transmit 340 pilots and 255 pilots, respectively, per transducer. To avoid interference, at the pilot positions of one transmitting element, the other elements should transmit zero symbols. A large receiver array of 100 hydrophones is assumed. This is perfectly feasible for the scenario of interest since the receiver will be attached to a surface vessel and can accommodate a large number of receiving hydrophones. In order to ensure decorrelated received signals, the inter-element spacing must be greater than at least half a wavelength [23], which is equal to 2.3 cm (assuming a sound velocity of 1500 m/s).

The BER performance for massive MIMO OFDM and FBMC systems in the vertical UAC is shown in Figure 5.6. The receiver processing consists of least-square (LS) channel estimation with linear interpolation. Assuming that the receiver has prior knowledge of the transmitted pilots, then the LS

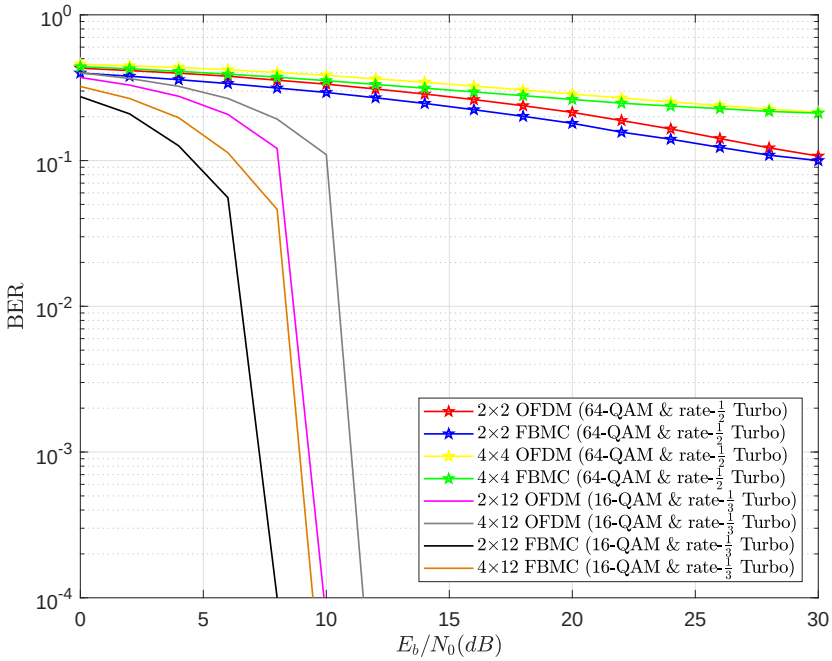


Figure 5.7: BER performance of Turbo-coded MIMO-OFDM and MIMO OFDM-OQAM systems in a 1000 m vertical UAC.

Table 5.2: Theoretical Achievable Bit Rate (kbps) with Single-User (SU) MIMO and Massive MIMO Systems.

System Configuration ($n_t \times n_r$)	OFDM	OFDM-OQAM
2×12	40	59
4×12	73	108
2×100	91	133
4×100	164	242

channel estimates at the pilot positions $(m,\bar{l})\in\mathscr{P}$ can be simply obtained as

$$\hat{h}_{\mathscr{P}_i}=\frac{\bar{y}_{\mathscr{P}_i}}{\bar{b}_{\mathscr{P}_i}}, \tag{5.12}$$

where m and $\bar{l}$ are the subcarrier and complex-symbol indices, respectively, $\bar{b}_{\mathscr{P}_i}$ and $\bar{y}_{\mathscr{P}_i}$ are the transmitted and received symbols at the ith pilot position, $\mathscr{P}_i$, respectively. Using a linear interpolation method such as spline, linear MMSE or nearest-neighbour, the LS estimates at the data positions are given by

$$\hat{h}_{m,\bar{l}}=\mathbf{w}^{\mathsf{H}}_{m,\bar{l}}\hat{\mathbf{h}}_{\mathscr{P}}, \tag{5.13}$$

where $\hat{\mathbf{h}}_{\mathscr{P}}\in\mathbb{C}^{|\mathscr{P}|\times 1}$ is a column vector of the LS estimates obtained in (5.12) and $\mathbf{w}_{m,\bar{l}}\in\mathbb{C}^{|\mathscr{P}|\times 1}$ is a vector function which describes the interpolation method used. For instance, the linear MMSE channel

interpolation can be expressed as [120]

$$\mathbf{w}_{m,\bar{l}}^{\text{LMMSE}} = \arg \min_{\mathbf{w}_{m,\bar{l}}} \mathbb{E}\left\{|h_{m,\bar{l}} - \mathbf{w}_{m,\bar{l}}^{\mathsf{H}} \hat{\mathbf{h}}_{\mathscr{P}}|^2\right\} \tag{5.14}$$

$$= \mathbf{R}_{\hat{\mathbf{h}}_{\mathscr{P}}}^{-1} \mathbf{r}_{\hat{\mathbf{h}}_{\mathscr{P}}, h_{m,\bar{l}}}^{\text{corr}}, \tag{5.15}$$

where $\mathbf{R}_{\hat{\mathbf{h}}_{\mathscr{P}}} = \mathbb{E}\left\{\hat{\mathbf{h}}_{\mathscr{P}} \left(\hat{\mathbf{h}}_{\mathscr{P}}\right)^{\mathsf{H}}\right\}$ is the correlation matrix and $\mathbf{r}_{\hat{\mathbf{h}}_{\mathscr{P}}, h_{m,\bar{l}}}^{\text{corr}} = \mathbb{E}\left\{\hat{\mathbf{h}}_{\mathscr{P}} h_{m,\bar{l}}^*\right\}$ is the correlation vector.

A Hermite filter with an overlapping factor of 4 is considered for the OFDM-OQAM system. Since the spreading is performed in time, the use of Hermite filter instead of the PHYDYAS prototype filter is justified as it provides a better localisation in the time domain. If frequency spreading was considered, then the PHYDYAS filter would have been a better option due to its better localisation in frequency.

Considering the Turbo coded systems in Figure 5.6, the 2×100 and 4×100 OFDM-OQAM systems outperform the OFDM systems with the same configuration by 2 dB and 2.2 dB, respectively, at a BER of 10^{-4}. For comparison purposes, the BER performance for standard symmetric MIMO configurations (2×2 and 4×4) in the same channel is shown in Figure 5.7. The parameters are similar to the massive MIMO systems. Even with the use of Turbo codes, the 2×2 and 4×4 OFDM-OQAM and OFDM systems achieve a high error rate. The application of massive MIMO reception to the OFDM and OFDM-OQAM systems greatly improve the BER performance under the same system parameters. Usually for UWA communication, asymmetric MIMO configuration is used where $n_r > n_t$ (see, e.g., [23]). Hence, the performance of coded 2×12 and 4×12 MIMO systems are included in Figure 5.7. In order to achieve comparable performance to the coded massive MIMO systems in Figure 5.6, a lower modulation order (16-QAM) and a lower FEC code rate (1/3) with more decoding iterations were used. As such, the achievable bit rates are much lower than in the massive MIMO systems.

5.4.2 Theoretical Achievable Bit Rate

The same system parameters as used for the simulations described in the previous subsection are considered for the bit rate computation. The zero guard slots which are inserted between blocks in the OFDM-OQAM system cause a bandwidth efficiency loss of $\frac{1}{N+1}$ [136]. Hence, for a spreading length of 64, the loss in bandwidth efficiency is 1.54%. Taking into account this loss for the OFDM-OQAM system and all the overheads due to channel estimation, the theoretical bit rates that can be achieved with the standard and massive MIMO systems in the 1000 m vertically configured UAC are given in Table 5.4.1. The 2×100 and 4×100 OFDM-OQAM systems achieve around 46% and 48% higher bit rate, respectively, than the OFDM systems with the same configurations. Compared to the n_t×12 MIMO systems, more than 50% increase in bit rate is achieved with the n_t×100 massive MIMO systems in the same channel.

5.4.3 Video Evaluation

For the video transmission part, the 2×100 and 4×100 Turbo-coded OFDM-OQAM systems are considered since they achieve a higher bit rate and better error performance than the OFDM systems.

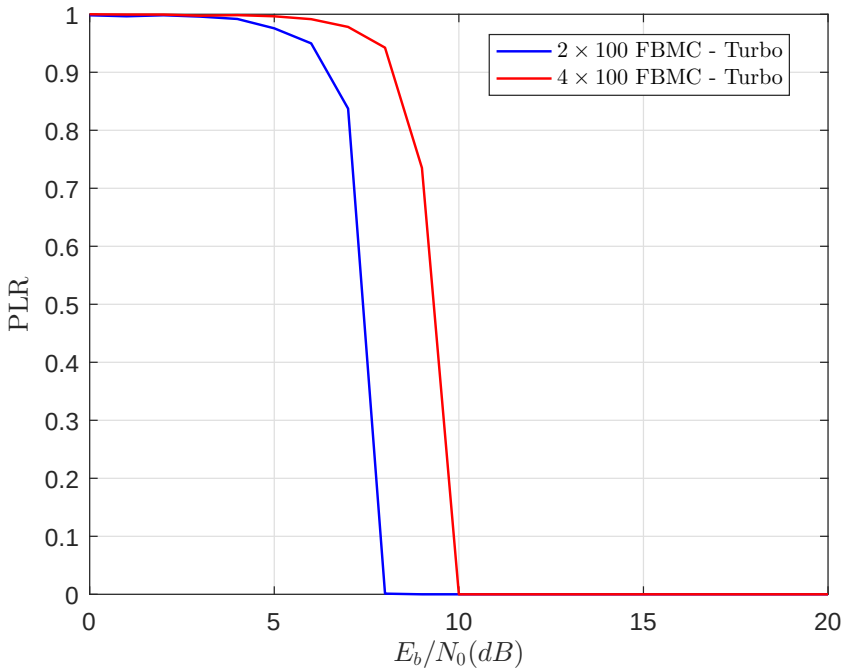


Figure 5.8: PLR for transmission of 480p compressed video using 2×100 and 4×100 OFDM-OQAM systems.

As previously stated, H264/AVC is considered as the video coding standard. An uncompressed 480p (854×480) video with a frame rate of 30 fps and duration 12 s is compressed to 130 kbps and 240 kbps for the 2×100 and 4×100 OFDM-OQAM systems, respectively. The total number of frames in the video is 365. An intra-period of 8 was selected to provide increased error resilience. It is noted that if the period is too small the target bit rate becomes difficult to achieve. The maximum packet length was set to 800 bytes and the number of reference frames was set to 6 to ensure a good compression efficiency.

5.4.3.1 Calculated Frame Rate

The number of bits per frame varies in the video sequence with the I-frames consisting of more bits than others. Hence an average number of bits per frame is used to compute the frame rate of the encoded video streams. The H.264/AVC videos compressed at 130 kbps and 240 kbps have an average number of bits per frame of 4324 and 8024, respectively. Considering the maximum achievable bit rates in Table 5.4.1, the calculated frame rates for the $2\times n_r$ and $4\times n_r$ OFDM-OQAM systems are 30.76 fps and 30.16 fps, respectively. Hence, real-time video transmission is theoretically possible with both systems.

5.4.3.2 Packet Loss Rate (PLR)

Figure 5.8 shows the PLR for the systems under investigation. The H.264/AVC bitstreams are encoded in 867 and 1008 video packets for the 2×100 and 4×100 systems, respectively. The E_b/N_0 values that

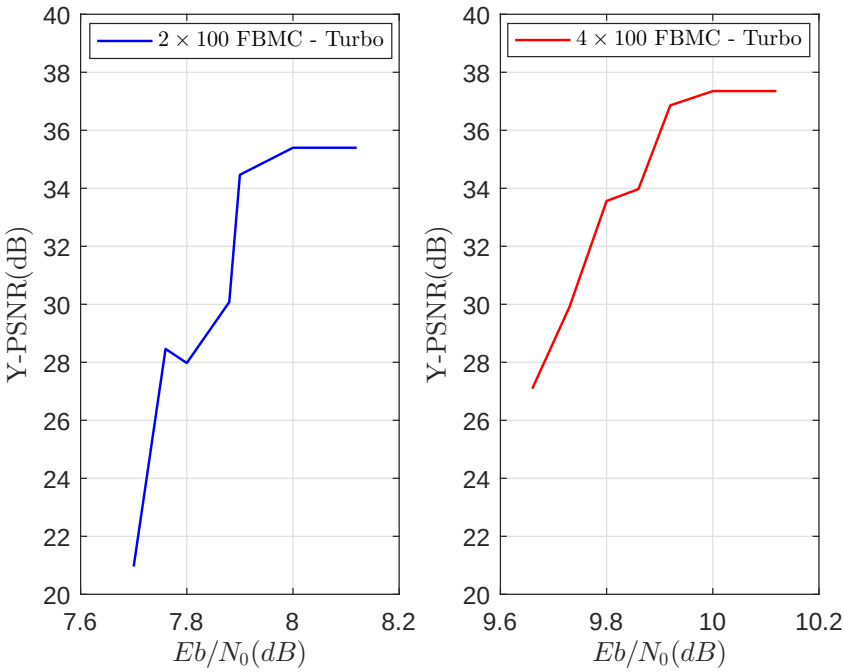


Figure 5.9: PSNR versus E_b/N_0 for received video streams for 2×100 and 4×100 OFDM-OQAM systems.

are required to achieve negligible bit errors and hence packet loss are 8 dB and 10 dB for the 2×100 and 4×100 systems, respectively.

5.4.3.3 Video Quality

Figure 5.9 shows the PSNR plot of the received H.264/AVC video streams. For each system, the packet loss rate varies between 0 and 25%. Higher values of packet loss were not considered since the video streams became practically impossible to decode. With increasing E_b/N_0 values, it is observed that the video quality for each system has a general tendency to increase as the number of packet loss decreases. Assuming no packet losses, the 4×100 OFDM-OQAM system achieves a higher PSNR value than the 2×100 system.

5.5 Performance Evaluation of a MU Massive MIMO System in an UAC

5.5.1 MU Massive MIMO System Model

The present work focuses on the uplink where data is transmitted from the users (ROVs) to the BS. In the uplink, the BS requires CSI to decode the received signals from multiple users. Hence, the users

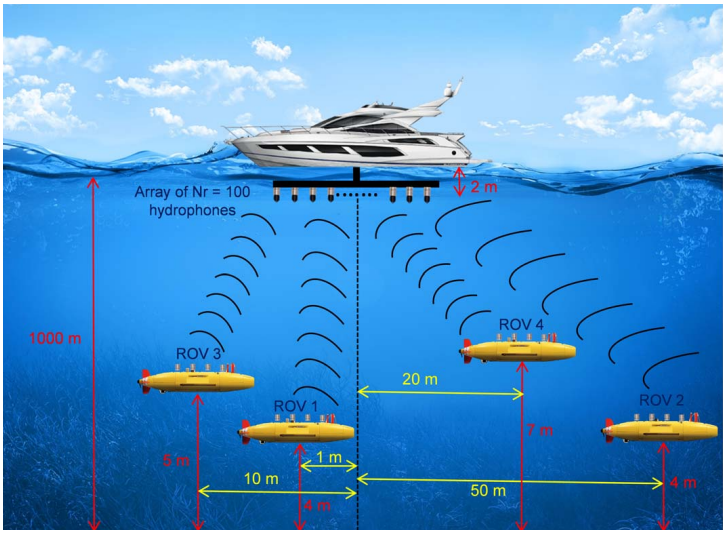


Figure 5.10: MU-massive MIMO UWA transmission scenario.

Table 5.3: MU FBMC-based Massive MIMO System Simulation Parameters.

Parameters	ROV 1	ROV 2	ROV 3	ROV 4
Bandwidth	25 kHz			
Carrier frequency	32.5 kHz			
Water depth	1000 m			
TX height from sea-floor	4 m	4 m	5 m	7 m
RX height from sea-floor	998 m			
Number of transmitting transducers per ROV, n_t	4			
Number of receiving hydrophones at BS, n_r	100			
Number of subcarriers	256			
Delay spread	5.3 ms	5.3 ms	6.6 ms	9.2 ms
Modulation	64-QAM			
Turbo code rate	1/2			
Max. RMS Doppler spread	2 Hz			
FBMC filter	Hermite prototype filter			
Filter overlapping factor	4			
FBMC spreading factor	16			

transmit orthogonal pilots to the BS which in turn performs channel estimation using the received pilot signals together with linear decoding techniques.

Consider an uplink MU-massive MIMO system with a BS having n_r hydrophones and serving Z users (ROVs) each equipped with n_t=4 transmitting elements. Considering also the FBMC-OQAM physical layer waveform with M subcarriers and N symbols, the received signal vector at the BS,

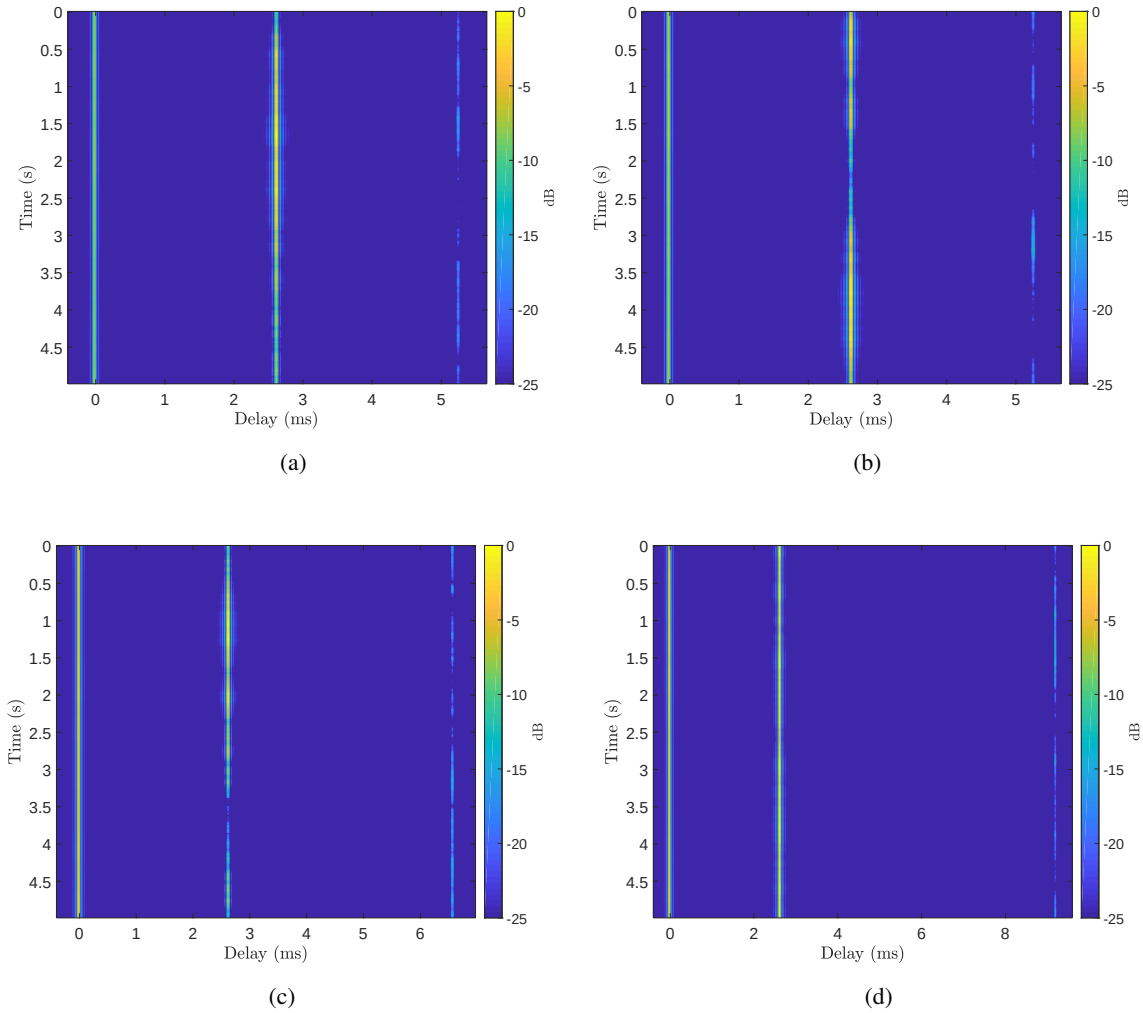


Figure 5.11: Typical CIRs as observed between the BS and four ROVs (a) ROV 1 (b) ROV 2 (c) ROV 3 (d) ROV 4.

$\tilde{\mathbf{y}}_{\tilde{l}} \in \mathbb{C}^{Mn_r \times 1}$ for the $\tilde{l}$th symbol can be expressed as

$$\tilde{\mathbf{y}}_{\tilde{l}} = \sum_{z=0}^{Z-1} \tilde{\mathbf{H}}_{\tilde{l}}^{z} \ddot{\mathbf{s}}_{\tilde{l}}^{z} + \tilde{\boldsymbol{\eta}}_{\tilde{l}}, \tag{5.16}$$

where $\tilde{\boldsymbol{\eta}}_{\tilde{l}} \in \mathbb{C}^{Mn_r \times 1}$ is the noise vector, $\tilde{\mathbf{H}}_{\tilde{l}}^{z} \in \mathbb{C}^{Mn_r \times Mn_t}$ is the channel matrix for user z and $\ddot{\mathbf{s}}_{\tilde{l}}^{z} \in \mathbb{C}^{Mn_t \times 1}$ is the time-domain FBMC signal vector for user z which is obtained after pre-coding the complex-valued transmitted symbols with the spreading matrix $\mathbf{C}$ as in (5.9) and then performing FBMC modulation. At the receiver, FBMC demodulation is performed followed by the de-spreading process as in (5.11). Let the resulting signal vector for the $\bar{l}$th complex symbol timing instant be denoted by $\bar{\mathbf{y}}_{\bar{l}}$. Assuming that each user transmits pilots so that the BS can perform channel estimation, the frequency domain channel

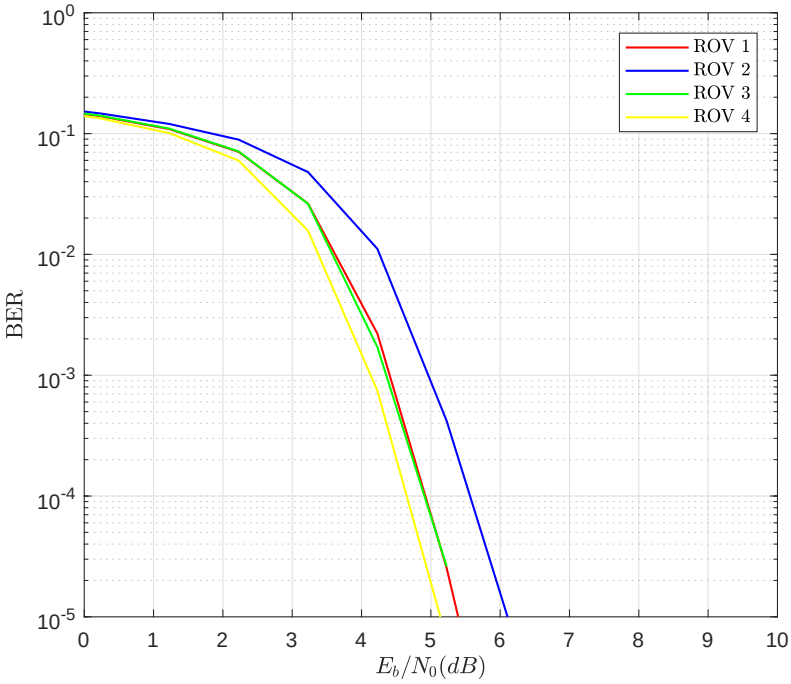


Figure 5.12: BER performance of Turbo-coded massive MIMO-FBMC/OQAM systems in the UAC.

matrix for user z, denoted by $\mathbf{H}_{\bar{l}}^{z} \in \mathbb{C}^{Mn_r \times Mn_t}$, can be written as

$$\mathbf{H}_{\bar{l}}^{z} = (\mathbf{I}_{n_r} \otimes \mathbf{F}_M)\hat{\mathbf{H}}_{\bar{l}}^{z}(\mathbf{I}_{n_t} \otimes \mathbf{F}_M^{\mathsf{H}}), \tag{5.17}$$

where $\mathbf{F}_M$ is the normalised M-point DFT matrix and $\hat{\mathbf{H}}_{\bar{l}}^{z}$ is the estimated time-domain matrix for user z. Considering the mth subcarrier and $\bar{l}$th symbol in (5.17), the frequency domain channel matrix for user z can be denoted as $\mathbf{H}_{m,\bar{l}}^{z} \in \mathbb{C}^{n_r \times n_t}$. Using the elements of the latter matrix for each user, an overall matrix denoted as $\ddot{\mathbf{H}}_{m,\bar{l}} \in \mathbb{C}^{n_r \times (n_t \times Z)}$ can be constructed. As an example, consider a 2-user transmission scenario where each user is equipped with two transmitting elements. In this case, the matrix $\ddot{\mathbf{H}}_{m,\bar{l}}$ is defined as

$$\ddot{\mathbf{H}}_{m,\bar{l}} = \begin{bmatrix} {}^{0,0}H_{m,\bar{l}}^{0} & {}^{0,0}H_{m,\bar{l}}^{1} & {}^{0,1}H_{m,\bar{l}}^{0} & {}^{0,1}H_{m,\bar{l}}^{1} \\ {}^{1,0}H_{m,\bar{l}}^{0} & {}^{1,0}H_{m,\bar{l}}^{1} & {}^{1,1}H_{m,\bar{l}}^{0} & {}^{1,1}H_{m,\bar{l}}^{1} \\ \vdots & \vdots & \vdots & \vdots \\ {}^{n_r-1,0}H_{m,\bar{l}}^{0} & {}^{n_r-1,0}H_{m,\bar{l}}^{1} & {}^{n_r-1,1}H_{m,\bar{l}}^{0} & {}^{n_r-1,1}H_{m,\bar{l}}^{1} \end{bmatrix}, \tag{5.18}$$

where ${}^{r,t}H_{m,\bar{l}}^{z}$ is the frequency domain channel coefficient for the $\bar{l}$th symbol and mth subcarrier between the t-th transmit element and r-th receive hydrophone for user z. It is known that simple linear processing techniques such as MMSE or ZF provide near-optimal performance in massive MIMO systems [145]. Hence, assuming ZF technique, the equalised signal $\hat{\mathbf{y}}_{m,\bar{l}} \in \mathbb{C}^{(n_t \times Z) \times 1}$ can be expressed as

$$\hat{\mathbf{y}}_{m,\bar{l}} = (\ddot{\mathbf{H}}_{m,\bar{l}})^{+}\bar{\mathbf{y}}_{m,\bar{l}}, \tag{5.19}$$

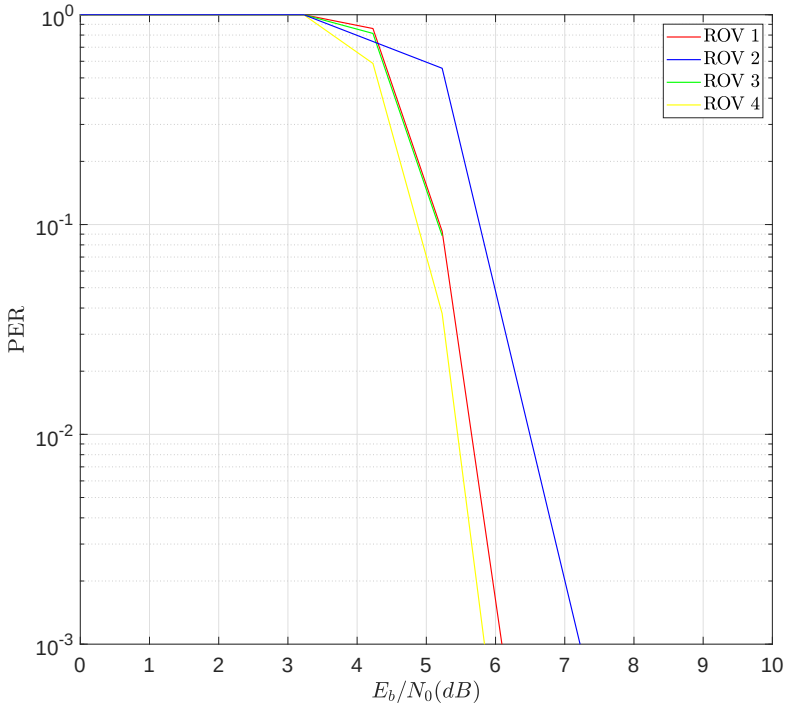


Figure 5.13: PER performance of Turbo-coded massive MIMO-FBMC/OQAM systems in the UAC.

where $\bar{\mathbf{y}}_{m,\bar{l}} \in \mathbb{C}^{n_r \times 1}$ is the received signal vector (after FBMC demodulation and de-spreading) for the $\bar{l}$th symbol and mth subcarrier and $(\ddot{\mathbf{H}}_{m,\bar{l}})^+$ is the Moore-Penrose Pseudoinverse of matrix $\ddot{\mathbf{H}}_{m,\bar{l}}$. Considering the 2-user scenario each with two transmitting elements, $\hat{\mathbf{y}}_{m,\bar{l}}$ is given by

$$\hat{\mathbf{y}}_{m,\bar{l}} = \begin{bmatrix} {}^0\hat{y}^0_{m,\bar{l}} \\ {}^0\hat{y}^1_{m,\bar{l}} \\ {}^1\hat{y}^0_{m,\bar{l}} \\ {}^1\hat{y}^1_{m,\bar{l}} \end{bmatrix}, \tag{5.20}$$

where ${}^t\hat{y}^z_{m,\bar{l}}$ is the equalised signal for the t-th transmit element of user z for a symbol $\bar{l}$ and subcarrier m. By separating each user's signal from $\hat{\mathbf{y}}_{m,\bar{l}}$ and then stacking the results for all subcarriers and symbols, each user's signal vector can be denoted as $\hat{\mathbf{y}}^z \in \mathbb{C}^{\frac{MN}{2} n_t \times 1}$.

5.5.2 Simulation Results

The physical setup of the MU-massive MIMO UWA communication scenario is shown in Figure 5.10 while the relevant simulation parameters are summarised in Table 5.3. The BS receiver array consists of 100 receiving hydrophones which can be attached to a surface vessel such as a ship or oil platform. Four ROVs are deployed at various depths and separation distances and each one is equipped with four transmitting elements. The ROVs transmit their data simultaneously to the surface vessel in the uplink.

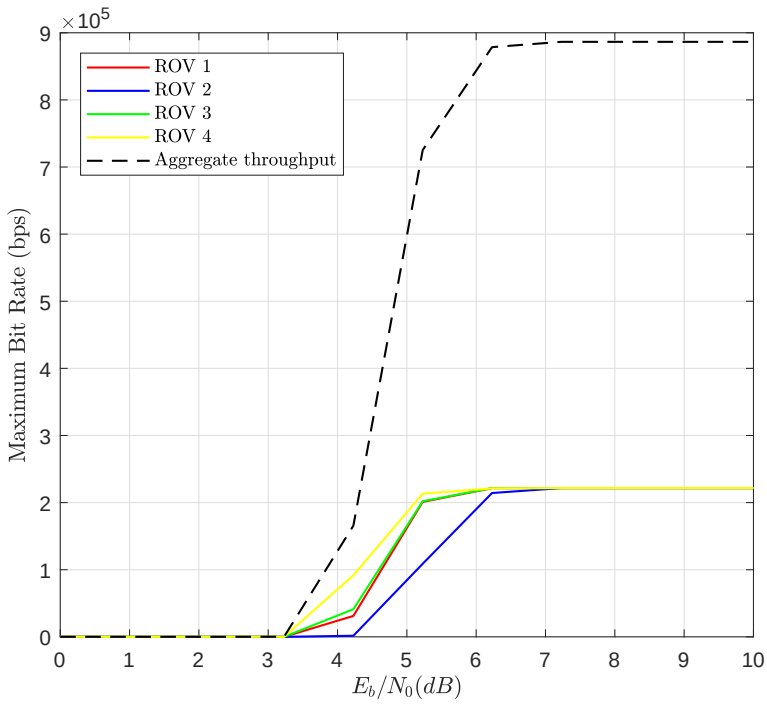


Figure 5.14: Maximum achievable bit rates for the four ROVs in the UAC.

This scenario can be of interest in the offshore oil and gas industry to carry out time-critical maintenance or intervention work on sub-sea infrastructure. This underwater communication scenario can also allow a larger search area to be covered on the ocean floor when looking for sunken objects. The typical CIRs as observed by the four users (ROVs) are illustrated in Figure 5.11. To approximate a real-world UAC, coloured noise is assumed in the simulations. The transmission of data is organised in packets where a zero guard interval is inserted between them to avoid interference in the OFDM-OQAM systems. It is assumed that the four users send pilot signals and hence imperfect CSI is available at the BS to decode the users' signals. A ZF linear detector is used in the system to recover the transmitted symbols from each user at the BS.

Figures 5.12, 5.13 and 5.14 show the performance of the four users in the UAC in terms of BER, PER and throughput, respectively. The results show that the 4×100 OFDM-OQAM systems can achieve robust performance in the time-varying UAC. Furthermore, the massive MIMO system allows all the ROVs (each equipped with four transmitting elements) to transmit their data simultaneously to the surface BS over the same frequency and time resources and still achieve a very good error performance as can be observed in Figure 5.12 and Figure 5.13. In fact, in a massive MIMO FBMC system, it not essential for the subcarriers to experience approximately flat channel gain [146]. This means that a smaller number of subcarriers can be used, which in turn reduces the latency or delay caused by the synthesis and analysis filter-banks. A greater subcarrier spacing also implies that the system will be less susceptible to carrier frequency offsets. Furthermore, larger modulation orders can be used to

further boost the bandwidth efficiency as the linear combination of the signals received on the multiple hydrophones at the BS smooths the channel distortions across the subcarriers, thereby equalising the channel gain across each one of them [146, 147]. Figure 5.14 shows the E_b/N_0 value that is required to achieve a given level of throughput for each user. It can be inferred that each ROV can reach their maximum throughput at a low E_b/N_0 value. Moreover, the fact that OFDM-OQAM is used instead of OFDM implies that each user achieves a higher bit rate since no bandwidth resources are wasted in the transmission of a CP. The overall aggregate throughput in this four-user massive MIMO system is approximately 886 kbps.

5.6 Conclusion

OFDM-OQAM achieves a higher bit rate than OFDM due to the absence of a CP, making it attractive for underwater wireless video transmission. However, conventional OFDM-OQAM systems suffer from inherent imaginary interference. By spreading the symbols in time (or frequency), this interference can be eliminated and complex orthogonality restored without any increase in complexity. This enables the application of MIMO to OFDM-OQAM and the use of pilot-based channel estimation as in OFDM. The use of massive MIMO reception increases the reliability of the communication link, allowing higher order modulation schemes to be used with higher FEC code rates. This allows a higher bit rate to be achieved, promoting real-time video transmission. In this chapter, it was shown that the bit rates of 133 kbps and 242 kbps obtained with the 2×100 and 4×100 FBMC systems, respectively, in the SU scenario are sufficiently high to achieve an acceptable video quality which conveys useful information in real-time. Furthermore, the application of massive MIMO in a MU scenario allows multiple ROVs to be deployed simultaneously and the fact that they can all share the same time and frequency resources makes this technology very attractive for the UWA environment where the bandwidth is extremely limited, especially for long range links. In this chapter, it was further shown that all the ROVs achieve very good error performance and high bit rates (221.5 kbps each) over the 1 km UAC using the OFDM-OQAM waveform shaping. Hence, OFDM-OQAM represents a better candidate than OFDM to be integrated in a massive MIMO system to satisfy the high data traffic and reliability demands of future underwater applications.

CHAPTER 6

INVESTIGATING VIDEO TRANSMISSION USING OFDM-BASED OTFS SYSTEMS FOR SINGLE-USER AND MULTI-USER UWA COMMUNICATION

6.1 Introduction

OFDM can achieve a high bandwidth efficiency in a linear time-invariant channel [148], but this is however not the case in a channel characterised by both frequency and time dispersions, also known as a doubly-dispersive channel. In this case, the performance of OFDM is severely degraded due to the significant interference caused by the channel Doppler spread [149]. In this regard, OTFS modulation is proposed for UWA communication. OTFS is a newly proposed multiplexing scheme to address channels characterised by extreme Doppler effects, for example in high speed vehicular environments [150–152]. OTFS typically transforms a time-varying fading channel into a 2-Dimensional (2D) non-fading and time-independent channel in the DD domain, allowing all the modulation symbols, e.g., QAM, which are spread across time and frequency to experience the same complex channel gain [152]. The diversity of the channel can be fully exploited in this scheme, enabling a linear scaling in spectral efficiency for a MIMO system with increasing number of antennas [151]. Although full diversity can be achieved in the Time-Frequency (TF) domain using a properly designed equaliser [152], OTFS however provides a more robust and practical approach in exploiting the diversity as the DD functions are sparse in nature, where most of the energy due to channel reflectors with associated Doppler is concentrated in a few bins (due to the lower channel variability in this domain). Another benefit of OTFS is that pilot signals can be flexibly and densely packed in the DD domain. This is desirable in both MIMO and massive MIMO systems where many pilot signals need to be multiplexed to simultaneously estimate the channels between the BS and the users [151].

In OTFS, QAM symbols are indexed by points on a grid in the DD domain. They are then transformed

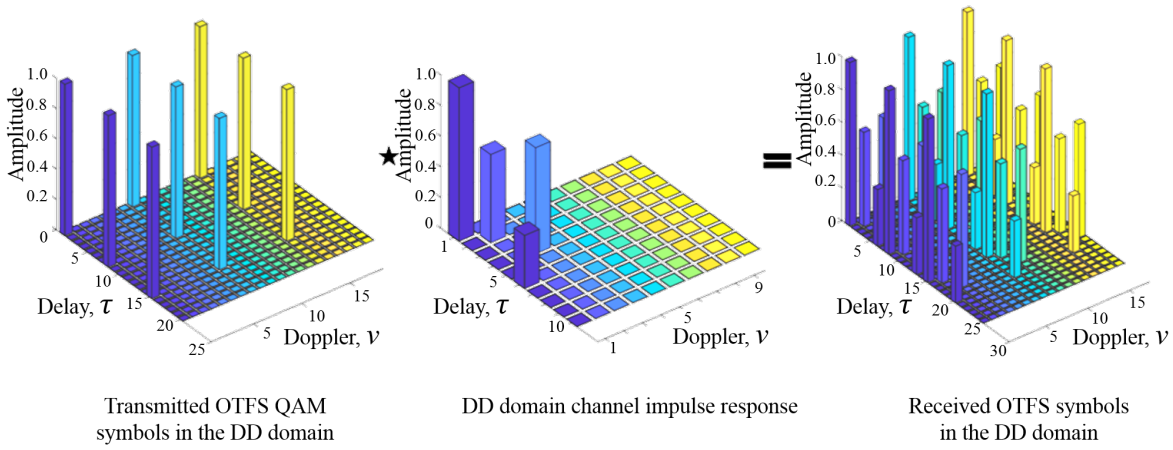


Figure 6.1: OTFS channel-symbol coupling [153].

into a TF domain signal using the Inverse Symplectic Finite Fourier Transform (ISFFT) and windowing [152]. This is familiar to the TF domain in which OFDM symbols reside. However, in OTFS, each symbol is spread across the whole TF plane. Then the Heisenberg transform is applied to the TF modulated signal to obtain a time domain signal which is transmitted over the channel [152]. The receiver performs the reverse operations. The OTFS channel-symbol interaction can be expressed as a 2D convolution between the DD domain CIR and the QAM symbols as shown in Figure 6.1.

Most of the studies in UWA communication have considered SU scenarios. However, for some time-critical applications such as repairing sub-sea oil pipes or searching for debris due to plane crashes or ship wrecks, multiple ROVs need to be deployed. This chapter investigates video transmission using OFDM-based OTFS systems for UWA communication. The BER performances of coded and un-coded spatially multiplexed MIMO and massive MIMO systems for SU and MU scenarios, respectively, are evaluated over a 1 km vertically-configured time-varying UAC. Furthermore, the performance of the systems is assessed in terms of their maximum achievable bit rates. The transmitted videos are compressed using the H.264/AVC standard and the received video quality is analysed in terms of PSNR. It is shown that the OFDM-based OTFS systems outperform the conventional OFDM systems in a dynamic UAC with a short channel coherence time using a high modulation order such as 64-QAM and frequency-domain pilot-based channel estimation. It is also demonstrated that the implementation of massive MIMO allows multiple underwater vehicles to be deployed simultaneously and to use the same frequency and time resources to transmit their information reliably to a single surface station. Bit rates as high as 109.7 kbps (with 2 transmitting elements) and 198.7 kbps (with 4 transmitting elements) are achieved in the SU-MIMO and MU-massive MIMO systems using frequency-domain pilot multiplexing. These maximum bit rate values are found to be sufficiently high to transmit a video with acceptable quality in real-time to a surface station over the 1 km vertical acoustic link. It is further shown that by

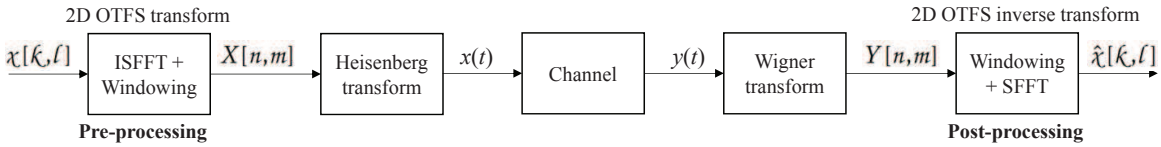


Figure 6.2: Illustration of the OTFS modulation.

multiplexing much fewer pilots in the DD domain than in the frequency domain to estimate the channel, a further increase in bit rate is obtained.

The work presented in this chapter has been submitted in [49, 50]. The rest of this chapter is organised as follows. The mathematical formulations for both SU MIMO and MU massive MIMO OFDM-based OTFS systems where the ROVs are equipped with multiple transmitting elements are provided in Section 8.2. A thorough investigation is carried out in Section 8.3 by comparing the performances of the different systems in terms of BER, maximum achievable bit rate and received video quality over the simulated 1 km vertically-configured time-varying UAC. Section 8.4 introduces the concept of DD domain pilot multiplexing and detection in OTFS systems and provides the corresponding mathematical formulations. Furthermore, the performance of the OFDM-based OTFS systems using DD domain pilot multiplexing and detection is evaluated. Finally, concluding remarks are drawn in Section 8.5.

6.2 System Model

The block diagram of the OTFS modulation scheme is shown in Figure 6.2. The pre- and post-processing blocks refer to two 2D transforms that are performed at the transmitter and receiver, respectively, to implement the OTFS modulation technique. At the transmitter, the data symbols, $x[k,l]$, residing in the DD domain are converted to a TF domain signal, $X[n,m]$, using the 2D ISFFT and windowing and this is known as the OTFS transform [152]. Note that windowing is used to further improve the sparsity of the DD domain channel [154]. The Heisenberg transform then converts the TF signal, $X[n,m]$, to a time domain signal, $x(t)$, which is subsequently transmitted through the channel. The inverse operations take place at the receiver side, wherein the received signal, $y(t)$, is converted to a TF domain signal, $Y[n,m]$, using the Wigner transform and then to the DD domain using windowing and Symplectic Finite Fourier Transform (SFFT) [152].

Consider a block of $\mathcal{N}\times\mathcal{M}$ QAM symbols, $x[k,l]$, residing in a DD grid where $k=0,\cdots,\mathcal{N}-1$ and $l=0,\cdots,\mathcal{M}-1$. The SFFT of a TF domain signal, $X[n,m]$, is expressed as

$$x[k,l]=\sum_{n=0}^{\mathcal{N}-1}\sum_{m=0}^{\mathcal{M}-1}X[n,m]e^{-j2\pi(\frac{nk}{\mathcal{N}}-\frac{ml}{\mathcal{M}})}, \tag{6.1}$$

and the inverse SFFT (SFFT^{-1}) of the information symbols, $x[k,l]$, is given by

$$X[n,m]=\frac{1}{\sqrt{\mathcal{M}\mathcal{N}}}\sum_{k=0}^{\mathcal{N}-1}\sum_{l=0}^{\mathcal{M}-1}x[k,l]e^{j2\pi(\frac{nk}{\mathcal{N}}-\frac{ml}{\mathcal{M}})}. \tag{6.2}$$

6.2.1 Formulation for a SISO OFDM-based OTFS System

Considering an OFDM-based OTFS system, the TF samples $X[n,m]$ obtained after the ISFFT operation are fed to an OFDM transmitter to obtain the OTFS transmit signal [149],

$$\bar{\mathbf{S}}=\mathbf{A}_{\mathrm{CP}}\mathbf{F}_{\mathcal{M}}^{\mathsf{H}}\mathbf{X}^{\mathsf{T}}, \tag{6.3}$$

where $\mathbf{X}\in\mathbb{C}^{\mathcal{N}\times\mathcal{M}}$ contains the TF samples $X[n,m]$ and is given by

$$\mathbf{X}=[\mathbf{F}_{\mathcal{M}}\check{\mathbf{X}}^{\mathsf{T}}\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}]^{\mathsf{T}}, \tag{6.4}$$

where $\check{\mathbf{X}}\in\mathbb{C}^{\mathcal{N}\times\mathcal{M}}$ contains the DD domain data symbols $\chi[k,l]$. From (6.4) it can be observed that the ISFFT operation is implemented using $\mathcal{M}$-point Discrete Fourier Transform (DFT) and $\mathcal{N}$-point IDFT on the columns and rows of $\check{\mathbf{X}}^{\mathsf{T}}$, respectively [155]. $\mathbf{F}_{\mathcal{M}}$ is the normalised $\mathcal{M}$-point DFT matrix. $\mathbf{A}_{\mathrm{CP}}=[\mathbf{G}_{\mathrm{CP}}^{\mathsf{T}},\mathbf{I}_{\mathcal{M}}^{\mathsf{T}}]^{\mathsf{T}}$ is the matrix for appending a CP to the signal and $\mathbf{G}_{\mathrm{CP}}\in\mathbb{C}^{N_{\mathrm{CP}}\times\mathcal{M}}$ is constructed using the last N_{CP} (CP length) rows of an identity matrix, $\mathbf{I}_{\mathcal{M}}$ [154]. Finally, the columns of $\bar{\mathbf{S}}\in\mathbb{C}^{(\mathcal{M}+N_{\mathrm{CP}})\times\mathcal{N}}$ contain the time-domain OFDM signals [154]. $\mathbf{X}$ in (6.4) can be represented in vectorised form, $\mathbf{x}$, as

$$\mathbf{x}=(\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}\otimes\mathbf{F}_{\mathcal{M}})\check{\mathbf{x}}, \tag{6.5}$$

where $\mathbf{x}=\mathrm{vec}(\mathbf{X}^{\mathsf{T}})$ and $\check{\mathbf{x}}=\mathrm{vec}(\check{\mathbf{X}}^{\mathsf{T}})$. Stacking the transmitted OTFS signal matrix, $\bar{\mathbf{S}}$, as a vector, $\bar{\mathbf{s}}\in\mathbb{C}^{\mathcal{N}(\mathcal{M}+N_{\mathrm{CP}})\times 1}$, the received signal can be expressed as

$$\mathbf{y}=\mathbf{H}\bar{\mathbf{s}}+\boldsymbol{\eta}, \tag{6.6}$$

where $\mathbf{H}\in\mathbb{C}^{\mathcal{N}(\mathcal{M}+N_{\mathrm{CP}})\times\mathcal{N}(\mathcal{M}+N_{\mathrm{CP}})}$ is the CIR matrix and $\boldsymbol{\eta}$ is the noise vector. The received OFDM symbols are free from ISI provided that $N_{\mathrm{CP}}\geq N_{\mathrm{mp}}$. The receiver then partitions $\mathbf{y}$ into $\mathcal{N}$ blocks and performs CP removal for each block. After this process, the output vectors are stacked into one vector $\tilde{\mathbf{y}}\in\mathbb{C}^{\mathcal{N}\mathcal{M}\times 1}$ given by [155]

$$\tilde{\mathbf{y}}=(\mathbf{I}_{\mathcal{N}}\otimes\mathbf{R}_{\mathrm{CP}})\mathbf{H}(\mathbf{I}_{\mathcal{N}}\otimes\mathbf{A}_{\mathrm{CP}})\tilde{\mathbf{s}}+\boldsymbol{\eta} \tag{6.7}$$

$$=\tilde{\mathbf{H}}\tilde{\mathbf{s}}+\boldsymbol{\eta}, \tag{6.8}$$

where $\mathbf{R}_{\mathrm{CP}}=[\mathbf{0}_{\mathcal{M}\times N_{\mathrm{CP}}},\mathbf{I}_{\mathcal{M}}]$ is the CP removal matrix and $\tilde{\mathbf{s}}$ is given by

$$\tilde{\mathbf{s}}=(\mathbf{I}_{\mathcal{N}}\otimes\mathbf{F}_{\mathcal{M}}^{\mathsf{H}})\mathbf{x}. \tag{6.9}$$

From (6.8), $\tilde{\mathbf{H}}=(\mathbf{I}_{\mathcal{N}}\otimes\mathbf{R}_{\mathrm{CP}})\mathbf{H}(\mathbf{I}_{\mathcal{N}}\otimes\mathbf{A}_{\mathrm{CP}})$ represents a block diagonal matrix of size $\mathcal{M}\mathcal{N}\times\mathcal{M}\mathcal{N}$ which is given by [155]

$$\tilde{\mathbf{H}}=\begin{bmatrix}\tilde{\mathbf{H}}_0 & \mathbf{0}_{\mathcal{M}} & \cdots & \mathbf{0}_{\mathcal{M}}\\ \mathbf{0}_{\mathcal{M}} & \tilde{\mathbf{H}}_1 & \cdots & \mathbf{0}_{\mathcal{M}}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0}_{\mathcal{M}} & \mathbf{0}_{\mathcal{M}} & \cdots & \tilde{\mathbf{H}}_{\mathcal{N}-1}\end{bmatrix}, \tag{6.10}$$

where $\tilde{\mathbf{H}}_n \in \mathbb{C}^{\mathcal{M}\times\mathcal{M}}$ is the CIR matrix of the nth OFDM symbol. After OFDM demodulation, the TF signal vector of length $\mathcal{MN}$ is expressed as

$$\check{\mathbf{y}} = (\mathbf{I}_{\mathcal{N}} \otimes \mathbf{F}_{\mathcal{M}})\tilde{\mathbf{y}}. \tag{6.11}$$

If equalisation is performed in the frequency domain, the overall block diagonal frequency domain channel matrix, $\tilde{\mathbf{H}}_{\mathbf{f}}$ is given by

$$\tilde{\mathbf{H}}_{\mathbf{f}} = (\mathbf{I}_{\mathcal{N}} \otimes \mathbf{F}_{\mathcal{M}})\tilde{\mathbf{H}}(\mathbf{I}_{\mathcal{N}} \otimes \mathbf{F}_{\mathcal{M}}^{\mathsf{H}}). \tag{6.12}$$

If ZF equalisation is used, the equalised TF signal, $\hat{\mathbf{y}}_{\mathbf{eq}}$ is given by

$$\hat{\mathbf{y}}_{\mathbf{eq}} = (\tilde{\mathbf{H}}_{\mathbf{f}})^{+}\check{\mathbf{y}}, \tag{6.13}$$

where $(\tilde{\mathbf{H}}_{\mathbf{f}})^{+}$ is the Moore-Penrose Pseudoinverse of matrix $\tilde{\mathbf{H}}_{\mathbf{f}}$. Finally, the equalised TF signal is converted to the DD domain and the transmitted symbol estimates, i.e, $\hat{x}[k,l]$, are obtained in vectorised form as [155]

$$\hat{\bar{\mathbf{x}}} = (\mathbf{F}_{\mathcal{N}} \otimes \mathbf{F}_{\mathcal{M}}^{\mathsf{H}})\hat{\mathbf{y}}_{\mathbf{eq}}. \tag{6.14}$$

6.2.2 Formulation for a MIMO OFDM-based OTFS System

Consider an $n_t \times n_r$ MIMO system with n_t transmitting elements and n_r receiving hydrophones. The DD data symbols at the transmitter are stacked for all elements to form an $\mathcal{M}n_t \times \mathcal{N}$ matrix as [155]

$$\bar{\check{\mathbf{X}}} = \begin{bmatrix} \grave{\mathbf{X}}^0 \\ \grave{\mathbf{X}}^1 \\ \vdots \\ \grave{\mathbf{X}}^{n_t-1} \end{bmatrix}, \tag{6.15}$$

where $\grave{\mathbf{X}}^t$ is the $\mathcal{M}\times\mathcal{N}$ matrix of DD domain data symbols of the t-th transmitting element. Performing ISFFT on each sub-matrix $\grave{\mathbf{X}}^t$ yields

$$\bar{\mathbf{X}} = (\mathbf{I}_{n_t} \otimes \mathbf{F}_{\mathcal{M}})\bar{\check{\mathbf{X}}}\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}, \tag{6.16}$$

where $(\mathbf{I}_{n_t} \otimes \mathbf{F}_{\mathcal{M}})$ is given by

$$(\mathbf{I}_{n_t} \otimes \mathbf{F}_{\mathcal{M}}) = \begin{bmatrix} \mathbf{F}_{\mathcal{M}}^0 & \mathbf{0}_{\mathcal{M}} & \cdots & \mathbf{0}_{\mathcal{M}} \\ \mathbf{0}_{\mathcal{M}} & \mathbf{F}_{\mathcal{M}}^1 & \cdots & \mathbf{0}_{\mathcal{M}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{\mathcal{M}} & \mathbf{0}_{\mathcal{M}} & \cdots & \mathbf{F}_{\mathcal{M}}^{n_t-1} \end{bmatrix}, \tag{6.17}$$

i.e., $\mathbf{F}_{\mathcal{M}}$ is repeated to cover the data matrices of all the transmitting elements. Equation (6.16) can be represented in vectorised form as

$$\bar{\mathbf{x}} = (\mathbf{F}_{\mathcal{N}}^{\mathsf{H}} \otimes \mathbf{I}_{n_t} \otimes \mathbf{F}_{\mathcal{M}})\bar{\check{\mathbf{x}}}, \tag{6.18}$$

where $\bar{\check{\mathbf{x}}}\in\mathbb{C}^{\mathcal{MN}n_t\times 1}$ is the vectorised form of $\bar{\check{\mathbf{X}}}\in\mathbb{C}^{\mathcal{M}n_t\times\mathcal{N}}$ which is expressed as [155]

$$\bar{\check{\mathbf{x}}}=\begin{bmatrix}\check{\mathbf{x}}_0\\ \check{\mathbf{x}}_1\\ \vdots\\ \check{\mathbf{x}}_{\mathcal{N}-1}\end{bmatrix},\quad \check{\mathbf{x}}_n=\begin{bmatrix}\check{\mathbf{x}}_n^0\\ \check{\mathbf{x}}_n^1\\ \vdots\\ \check{\mathbf{x}}_n^{n_t-1}\end{bmatrix},\quad \check{\mathbf{x}}_n^t=\begin{bmatrix}\check{x}_{0,n}^t\\ \check{x}_{1,n}^t\\ \vdots\\ \check{x}_{\mathcal{M}-1,n}^t\end{bmatrix}, \tag{6.19}$$

and $\check{x}_{m,n}^t$, where $t=0,1,\cdots,n_t-1$, is the symbol transmitted on the t-th transmitting element. Following the ISFFT operation, the TF signal goes through an OFDM modulator where it is partitioned into $\mathcal{N}$ blocks, $\bar{\mathbf{x}}_n\in\mathbb{C}^{\mathcal{M}n_t\times 1}$ which is then converted to a time domain signal $\bar{\bar{\mathbf{s}}}_n\in\mathbb{C}^{\mathcal{M}n_t\times 1}$ as

$$\bar{\bar{\mathbf{s}}}_n=(\mathbf{I}_{n_t}\otimes\mathbf{F}_{\mathcal{M}}^{\mathsf{H}})\bar{\mathbf{x}}_n, \tag{6.20}$$

and stacking in a vector $\bar{\bar{\mathbf{s}}}\in\mathbb{C}^{\mathcal{MN}n_t\times 1}$ yields

$$\bar{\bar{\mathbf{s}}}=(\mathbf{I}_{\mathcal{N}n_t}\otimes\mathbf{F}_{\mathcal{M}}^{\mathsf{H}})\bar{\mathbf{x}}. \tag{6.21}$$

The OFDM modulator appends a CP to the time domain signal which is then transmitted over the channel. At the receiver, following the CP removal, the received signal for the nth symbol, $\bar{\bar{\mathbf{y}}}_n\in\mathbb{C}^{\mathcal{M}n_r\times 1}$ can be written as [155]

$$\bar{\bar{\mathbf{y}}}_n=(\mathbf{I}_{n_r}\otimes\mathbf{R}_{\mathrm{CP}})\bar{\mathbf{H}}_n(\mathbf{I}_{n_t}\otimes\mathbf{A}_{\mathrm{CP}})\bar{\bar{\mathbf{s}}}_n+\bar{\boldsymbol{\eta}}_n \tag{6.22}$$

$$=\bar{\bar{\mathbf{H}}}_n\bar{\bar{\mathbf{s}}}_n+\bar{\boldsymbol{\eta}}_n, \tag{6.23}$$

where $\bar{\bar{\mathbf{H}}}_n=(\mathbf{I}_{n_r}\otimes\mathbf{R}_{\mathrm{CP}})\bar{\mathbf{H}}_n(\mathbf{I}_{n_t}\otimes\mathbf{A}_{\mathrm{CP}})\in\mathbb{C}^{\mathcal{M}n_r\times\mathcal{M}n_t}$ can be expressed as

$$\bar{\bar{\mathbf{H}}}_n=\begin{bmatrix}\bar{\bar{\mathbf{H}}}_n^{0,0} & \bar{\bar{\mathbf{H}}}_n^{0,1} & \cdots & \bar{\bar{\mathbf{H}}}_n^{0,n_t-1}\\ \bar{\bar{\mathbf{H}}}_n^{1,0} & \bar{\bar{\mathbf{H}}}_n^{1,1} & \cdots & \bar{\bar{\mathbf{H}}}_n^{1,n_t-1}\\ \vdots & \vdots & \ddots & \vdots\\ \bar{\bar{\mathbf{H}}}_n^{n_r-1,0} & \bar{\bar{\mathbf{H}}}_n^{n_r-1,1} & \cdots & \bar{\bar{\mathbf{H}}}_n^{n_r-1,n_t-1}\end{bmatrix}, \tag{6.24}$$

where $\bar{\bar{\mathbf{H}}}_n^{r,t}\in\mathbb{C}^{\mathcal{M}\times\mathcal{M}}$ denotes the CIR matrix between the t-th transmit element and r-th receive hydrophone for the nth symbol. The output vectors $\bar{\bar{\mathbf{y}}}_n$ are stacked to form $\bar{\bar{\mathbf{y}}}\in\mathbb{C}^{\mathcal{MN}n_r\times 1}$

$$\bar{\bar{\mathbf{y}}}=\bar{\bar{\mathbf{H}}}\bar{\bar{\mathbf{s}}}+\bar{\boldsymbol{\eta}}, \tag{6.25}$$

where

$$\bar{\bar{\mathbf{H}}}=\begin{bmatrix}\bar{\bar{\mathbf{H}}}_0 & \mathbf{0}_{\mathcal{M}n_r\times\mathcal{M}n_t} & \cdots & \mathbf{0}_{\mathcal{M}n_r\times\mathcal{M}n_t}\\ \mathbf{0}_{\mathcal{M}n_r\times\mathcal{M}n_t} & \bar{\bar{\mathbf{H}}}_1 & \cdots & \mathbf{0}_{\mathcal{M}n_r\times\mathcal{M}n_t}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0}_{\mathcal{M}n_r\times\mathcal{M}n_t} & \mathbf{0}_{\mathcal{M}n_r\times\mathcal{M}n_t} & \cdots & \bar{\bar{\mathbf{H}}}_{\mathcal{N}-1}\end{bmatrix}, \tag{6.26}$$

and $\bar{\bar{\mathbf{H}}}$ is an $\mathcal{MN}n_r\times\mathcal{MN}n_t$ diagonal matrix. The frequency domain channel matrix for the nth symbol is given by

$$\bar{\bar{\mathbf{H}}}_{\mathbf{f}n}=(\mathbf{I}_{n_r}\otimes\mathbf{F}_{\mathcal{M}})\bar{\bar{\mathbf{H}}}_n(\mathbf{I}_{n_t}\otimes\mathbf{F}_{\mathcal{M}}^{\mathsf{H}}). \tag{6.27}$$

At the receiver, the nth symbol vector after the OFDM demodulation process is expressed as

$$\hat{\bar{\mathbf{y}}}_n = (\mathbf{I}_{n_r} \otimes \mathbf{F}_{\mathcal{M}}) \bar{\tilde{\mathbf{y}}}_n. \tag{6.28}$$

Assuming ZF equalisation, the equalised TF signal vector for the nth symbol is given by

$$\hat{\bar{\mathbf{y}}}_{\mathbf{eq}_n} = (\bar{\tilde{\mathbf{H}}}_{\mathbf{f}n})^{+} \hat{\bar{\mathbf{y}}}_n. \tag{6.29}$$

The equalised vectors $\hat{\bar{\mathbf{y}}}_{\mathbf{eq}_n}$ are stacked for all symbols to form $\hat{\bar{\mathbf{y}}}_{\mathbf{eq}} \in \mathbb{C}^{\mathcal{MN}n_t \times 1}$. Finally, the TF vector $\hat{\bar{\mathbf{y}}}_{\mathbf{eq}}$ is converted to the DD domain and the transmitted symbol estimates are acquired as follows

$$\hat{\bar{\tilde{\mathbf{x}}}} = (\mathbf{F}_{\mathcal{N}} \otimes \mathbf{I}_{n_t} \otimes \mathbf{F}_{\mathcal{M}}^{\mathsf{H}}) \hat{\bar{\mathbf{y}}}_{\mathbf{eq}}. \tag{6.30}$$

6.2.3 Formulation for a Multi-User (MU) Massive MIMO OFDM-based OTFS System

Consider an uplink MU-massive MIMO system with a BS equipped with n_r receiving hydrophones and serving $\mathcal{Z}$ users, each transmitting with n_t multiple transducers and $n_r \gg n_t$. The baseband received signal vector $\ddot{\mathbf{y}}_n \in \mathbb{C}^{\mathcal{M}n_r \times 1}$ at the BS for the nth symbol can be expressed as

$$\ddot{\mathbf{y}}_n = \sum_{z=0}^{\mathcal{Z}-1} \bar{\tilde{\mathbf{H}}}_n^z \bar{\tilde{\mathbf{s}}}_n^z + \bar{\boldsymbol{\eta}}_n, \tag{6.31}$$

where $\bar{\tilde{\mathbf{H}}}_n^z$ and $\bar{\tilde{\mathbf{s}}}_n^z$ are the nth symbol convolution matrix (as in equation (6.24)) and time-domain signal vector (as in equation (6.20)), respectively, for user z. $\bar{\boldsymbol{\eta}}_n \in \mathbb{C}^{\mathcal{M}n_r \times 1}$ is the noise vector. After OFDM demodulation, the TF signal vector at the BS for the nth symbol, denoted as $\bar{\tilde{\mathbf{y}}}_n \in \mathbb{C}^{\mathcal{M}n_r \times 1}$, can be expressed as

$$\bar{\tilde{\mathbf{y}}}_n = (\mathbf{I}_{n_r} \otimes \mathbf{F}_{\mathcal{M}}) \ddot{\mathbf{y}}_n. \tag{6.32}$$

The frequency-domain channel matrix for user z, denoted as $\bar{\tilde{\mathbf{H}}}_{\mathbf{f}n}^{z} \in \mathbb{C}^{\mathcal{M}n_r \times \mathcal{M}n_t}$, can be written as

$$\bar{\tilde{\mathbf{H}}}_{\mathbf{f}n}^{z} = (\mathbf{I}_{n_r} \otimes \mathbf{F}_{\mathcal{M}}) \bar{\tilde{\mathbf{H}}}_n^z (\mathbf{I}_{n_t} \otimes \mathbf{F}_{\mathcal{M}}^{\mathsf{H}}). \tag{6.33}$$

Let the frequency-domain channel matrix for user z for the nth symbol and mth subcarrier be denoted by $\bar{\tilde{\mathbf{H}}}_{\mathbf{f}m,n}^{z} \in \mathbb{C}^{n_r \times n_t}$. The latter can be constructed by stacking every $(\mathcal{M}-1)$th element row-wise and column-wise from matrix $\bar{\tilde{\mathbf{H}}}_{\mathbf{f}n}^{z}$ for $m=0,\cdots,\mathcal{M}-1$.

Next, we define a matrix $\ddot{\mathbf{H}}_{m,n} \in \mathbb{C}^{n_r \times (n_t \times \mathcal{Z})}$, constructed from the elements of $\bar{\tilde{\mathbf{H}}}_{\mathbf{f}m,n}^{z}$ for each user and which will be used for the equalisation process. As an example, consider a 2-user scenario each equipped with two transmitting elements. In this case, the matrix $\ddot{\mathbf{H}}_{m,n}$ can be defined as

$$\ddot{\mathbf{H}}_{m,n} = \begin{bmatrix} {}^{0,0}\bar{\tilde{h}}_{m,n}^{0} & {}^{0,0}\bar{\tilde{h}}_{m,n}^{1} & {}^{0,1}\bar{\tilde{h}}_{m,n}^{0} & {}^{0,1}\bar{\tilde{h}}_{m,n}^{1} \\ {}^{1,0}\bar{\tilde{h}}_{m,n}^{0} & {}^{1,0}\bar{\tilde{h}}_{m,n}^{1} & {}^{1,1}\bar{\tilde{h}}_{m,n}^{0} & {}^{1,1}\bar{\tilde{h}}_{m,n}^{1} \\ \vdots & \vdots & \vdots & \vdots \\ {}^{n_r-1,0}\bar{\tilde{h}}_{m,n}^{0} & {}^{n_r-1,0}\bar{\tilde{h}}_{m,n}^{1} & {}^{n_r-1,1}\bar{\tilde{h}}_{m,n}^{0} & {}^{n_r-1,1}\bar{\tilde{h}}_{m,n}^{1} \end{bmatrix}, \tag{6.34}$$

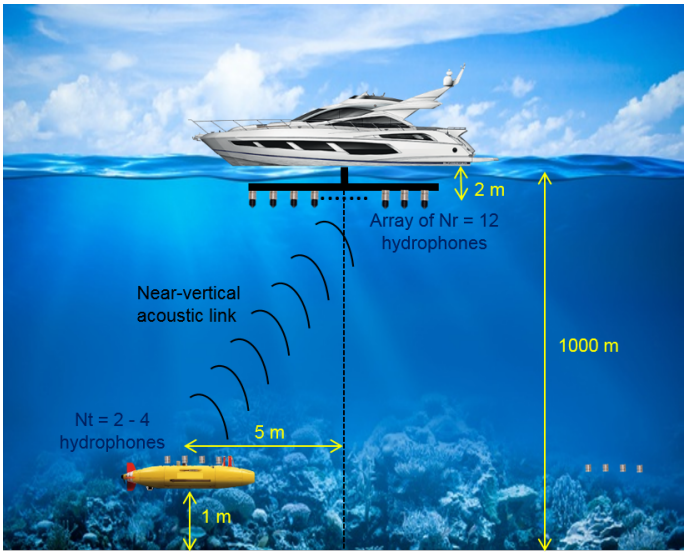


Figure 6.3: Target scenario for single-user UWA communication.

where ${}^{r,t}\bar{\bar{h}}^{z}_{m,n}$ is the frequency-domain channel coefficient for the nth symbol and mth subcarrier between the t-th transmit element and r-th receive hydrophone for user z. It is known that simple linear processing techniques such as MMSE or ZF provide near-optimal performance in massive MIMO systems [145]. Hence, assuming ZF technique, the equalised signal $\hat{\ddot{\mathbf{y}}}_{\mathbf{eq}_{m,n}} \in \mathbb{C}^{(n_t \times \mathcal{Z}) \times 1}$ can be expressed as

$$\hat{\ddot{\mathbf{y}}}_{\mathbf{eq}_{m,n}} = (\ddot{\mathbf{H}}_{m,n})^{+} \bar{\bar{\mathbf{y}}}_{m,n}, \tag{6.35}$$

where $\bar{\bar{\mathbf{y}}}_{m,n} \in \mathbb{C}^{n_r \times 1}$ is the received TF signal vector for the nth symbol and mth subcarrier which is constructed from $\bar{\mathbf{y}}_n$ by stacking every $(\mathcal{M}-1)$th element for $m=0,\cdots,\mathcal{M}-1$. Considering the same example as previously, i.e., the 2-user scenario each with two transmitting elements, in this case $\hat{\ddot{\mathbf{y}}}_{\mathbf{eq}_{m,n}}$ is given by

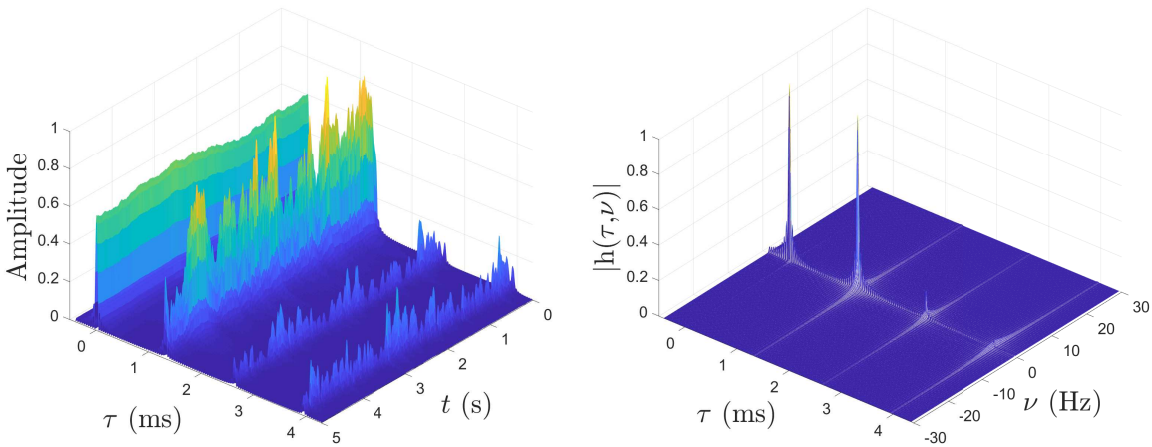
$$\hat{\ddot{\mathbf{y}}}_{\mathbf{eq}_{m,n}} = \begin{bmatrix} {}^{0}\hat{\ddot{y}}^{0}_{m,n} \\ {}^{0}\hat{\ddot{y}}^{1}_{m,n} \\ {}^{1}\hat{\ddot{y}}^{0}_{m,n} \\ {}^{1}\hat{\ddot{y}}^{1}_{m,n} \end{bmatrix}, \tag{6.36}$$

where ${}^{t}\hat{\ddot{y}}^{z}_{m,n}$ is the equalised signal of the t-th transmit element of user z for a given symbol n and subcarrier m. By separating each user's signal from $\hat{\ddot{\mathbf{y}}}_{\mathbf{eq}_{m,n}}$ and then stacking the results for all subcarriers and symbols, each user's TF signal vector can be denoted as $\hat{\ddot{\mathbf{y}}}^{z}_{\mathbf{eq}} \in \mathbb{C}^{\mathcal{M}\mathcal{N}n_t \times 1}$. Finally, the DD domain estimate of the transmitted vector for each user z is obtained as

$$\hat{\ddot{\mathbf{x}}}^{z} = (\mathbf{F}_{\mathcal{N}} \otimes \mathbf{I}_{n_t} \otimes \mathbf{F}^{\mathsf{H}}_{\mathcal{M}})\, \hat{\ddot{\mathbf{y}}}^{z}_{\mathbf{eq}}. \tag{6.37}$$

Table 6.1: MIMO OFDM-based OTFS System Simulation Parameters.

Bandwidth	25 kHz
Carrier frequency	32.5 kHz
Water depth	1000 m
TX height from sea-floor	1 m
RX height from sea-floor	998 m
Number of transmitting transducers, n_t	2 & 4
Number of receiving hydrophones, n_r	12
Subcarriers	512
Cyclic prefix duration	5.12 ms
Max. delay spread	3.9 ms
Modulation	64-QAM
Turbo code rate	1/2



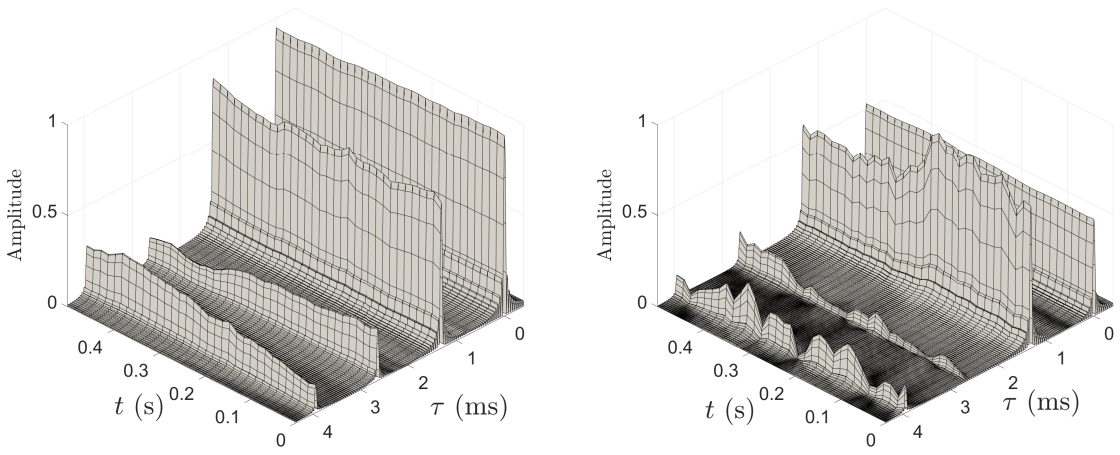
(a) (b)

Figure 6.4: Channel representations (a) Time-delay (b) Doppler-delay (max. RMS Doppler spread = 7 Hz).

6.3 Simulation Results

6.3.1 Performance Evaluation of SU-MIMO Systems

An illustration of the transmission scenario for the SU-MIMO systems is shown in Figure 6.3 and the simulation parameters are given in Table 6.1. The RMS Doppler spread value increases linearly from 0.5 Hz to a maximum value with the tap delay. The systems' performances are investigated using different maximum RMS Doppler spread values. An example of the observed channel response between a given transmitting element and receiving hydrophone for a maximum RMS Doppler spread of 7 Hz is shown in Figure 6.4 (in terms of two representations). The maximum delay spread is 3.9 ms in this



(a) (b)

Figure 6.5: Typical time-variant channel response at a maximum RMS Doppler spread value of (a) 1 Hz (b) 7 Hz.

case. Each transmit transducer is assigned a set of pilots arranged in a comb-type fashion. This pilot arrangement is chosen to track the rapid channel variations on all the symbols. For both the $2\times n_r$ and $4\times n_r$ systems, each transducer transmits 22 equally spaced pilot subcarriers for each timing instant. Furthermore, considering the $2\times n_r$ MIMO system, at the pilot positions of the first transmitting element, zero symbols are sent by the second transmitting element to prevent interference and vice versa. Thus, for a 2-transmitting element system, the overhead due to the pilot and zero symbols is around 8.6% for a transmission block of one second. As for a 4-transmitting element system, the percentage overhead is approximately 17.2%. There are multiple channels between the transmitter and receiver elements and it is assumed that the channels are independent of each other. The Least Square (LS) channel estimator and spline interpolation method are used to estimate the channel at the pilot and data subcarriers, respectively. The time-variant channel responses as observed between a given transmitting element and receiving hydrophone under two different RMS Doppler spread values are depicted in Figure 6.5 for a duration of 500 ms. As the RMS Doppler spread value increases, the channel is observed to vary more rapidly within this lapse of time. It is stated that the UAC coherence time can be considered to be in the order of hundreds of milliseconds for practical designs [7]. For example, shallow water sea trials were performed in [156] where the horizontal transmission distance varied between 250 m and 1500 m. A 5-element receiver array (1.75 m separation between elements) was deployed at sea occupying a water column between 9 m to 16 m. Two source depths of 10 m and 12 m were considered. A carrier frequency of 85 kHz was used and the maximum bandwidth was 21.25 kHz. For the two transducers, the channel coherence time became higher as the transmission range was increased from 250 m to 1500 m. The recorded coherence times for the transducer at a depth of 10 m were 107.1 ms and 123.5 ms for 250 m and 1500 m respectively. As for the transducer which was at a depth of 12 m, the coherence times

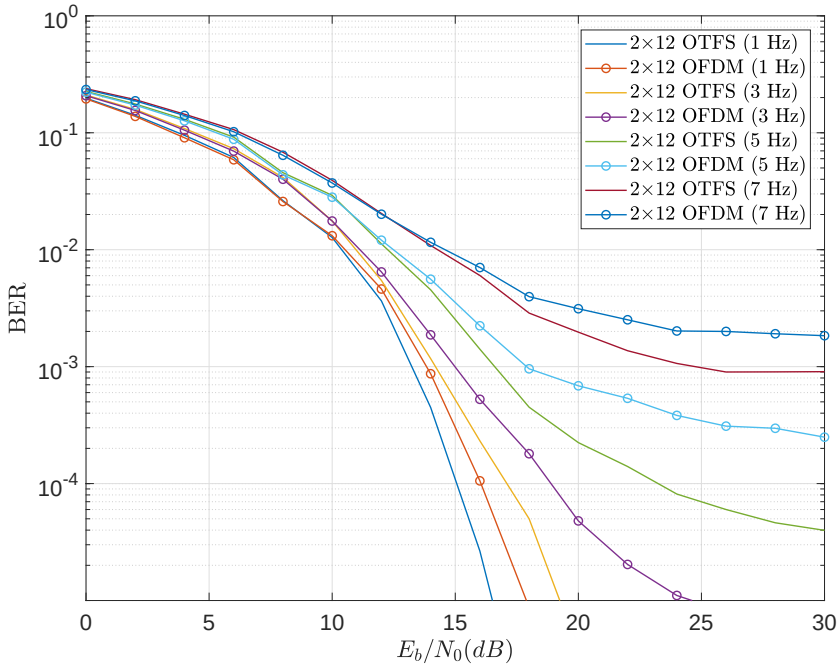


Figure 6.6: BER performance of un-coded 2×12 OFDM & OTFS systems in the UAC with AWGN for different maximum RMS Doppler spread values.

were 327.9 ms and 383.1 ms for 250 m and 1500 m, respectively. Furthermore, the average received SNRs over a distance of 1000 m were 26.1 dB and 32.6 dB with the 12 m and 10 m deep transducers, respectively. The maximum achievable data rate was 34 kbps at 1500 m using QPSK modulation. It was also observed that as the transmission range increased, the impulse responses had smaller delay span. It has also been stated in [156] that the channel characteristics, in terms of arrival paths and coherence time, observed with a system having a carrier frequency as high as 85 Hz showed much similarity with those achieved in the commonly used 10 to 50 kHz band for the same distance. Hence, for a practical vertical transmission scenario of similar range, the coherence time is expected to be higher because the transmission is least affected by reflections due to the surface and sea floor (hence very small grazing angles and attenuation). In this case, if the channel response is considered for an RMS Doppler spread of 1 Hz, the channel coherence time can be considered to be around 100 ms, which is analogous to a very dynamic sea environment. For larger RMS Doppler spread values, the coherence times are much smaller. Higher values of RMS Doppler spread are included to investigate the behaviour of the different systems in extremely harsh channel conditions and to determine whether the receiver processing can cope with such rapid variations.

The BER performances for the 2×12 OFDM and OTFS un-coded systems in the vertical UAC with AWGN and coloured noise are shown in Figure 6.6 and Figure 6.7, respectively. This type of MIMO configuration is often used for SU UWA communication (see, e.g., [19, 22, 157]). Assuming the speed

of sound is 1500 m/s, the transmitting and receiving elements should be spaced apart by at least half a wavelength to ensure proper de-correlation between the MIMO channels [23]. Thus, for the MIMO scenario in Figure 6.3, the separation should be at least 2.3 cm considering a carrier frequency of 32.5 kHz. So the ROV can be easily equipped with 2 to 4 transmit transducers. By comparing Figure 6.6 and Figure 6.7, it is observed that the OFDM and OTFS systems in AWGN perform better than those in the coloured noise. Considering a maximum RMS Doppler spread value of 1 Hz, the OTFS and OFDM systems in AWGN outperform those in coloured noise by 7.5 dB and 11.7 dB, respectively, at a BER of 10^{-4}. Hence, coloured noise impacts on the system's performance more than AWGN due to the frequency-dependent nature of coloured noise. However, assuming AWGN for an UAC is not a good representation of noise in the real-world underwater environment.

Considering the 2×12 systems in coloured noise in Figure 6.7, the superiority of OTFS over OFDM can be observed. For example, at a maximum RMS Doppler spread value of 1 Hz, the OTFS system outperform OFDM by 4 dB at a BER of 10^{-3}. The performance improvement with OTFS compared to OFDM at a Doppler spread value of 3 Hz is 3.3 dB for the same BER. If the blue dotted curve with perfect CSI is considered, it can be observed that the difference in E_b/N_0 value at a BER of 10^{-3} is 2.4 dB compared to the OTFS system at an RMS Doppler spread value of 1 Hz (imperfect CSI). The difference in the E_b/N_0 value becomes larger with increasing RMS Doppler spread values. Moreover, as the RMS Doppler spread increases, it is observed that the OTFS and OFDM system performance approaches each other.

For comparison purposes, OTFS and OFDM systems with QPSK and 16-QAM modulation are included in Figure 6.7. The RMS Doppler spread value considered for these two modulation schemes is 7 Hz. As expected, the performance of OFDM and OTFS with QPSK is better than those with 16-QAM and 64-QAM. Moreover, with QPSK modulation, a distinguishable improvement of OTFS over OFDM is noted (4.7 dB at a BER of 10^{-3}) although an RMS Doppler spread value as high as 7 Hz is considered. As for the systems with 16-QAM modulation, the performance improvement of OTFS as compared to OFDM at a BER of 10^{-3} is 2 dB. This implies that the comb-type pilots can more effectively track the channel variations at a lower modulation order. ICI may cause higher degradation with increasing modulation order. High modulation orders are also known to be more affected by phase errors due to the points in the constellation diagram being more densely packed and as the modulation order increases, the effect of even small phase errors is magnified [158]. Since 64-QAM symbols have less phase difference between them as compared to QPSK, it can be deduced that for the same phase error, 64-QAM shows a greater performance degradation. However, as evaluated in [152], the performance of OTFS can be improved for high modulation orders by using a non-linear equaliser such as an iterative DFE.

Next, the BER performances of un-coded 4×12 OFDM and OTFS systems are evaluated in coloured noise as illustrated in Figure 6.8. The same simulation parameters as the 2×12 systems are considered. Comparing the 2×12 systems in Figure 6.7 with the 4×12 systems in Figure 6.8 for the coloured noise case, it can be observed that as the number of transmitting elements increases, so does the error rate. For instance, at an RMS Doppler spread value of 1 Hz, the 2×12 OTFS system outperform the 4×12

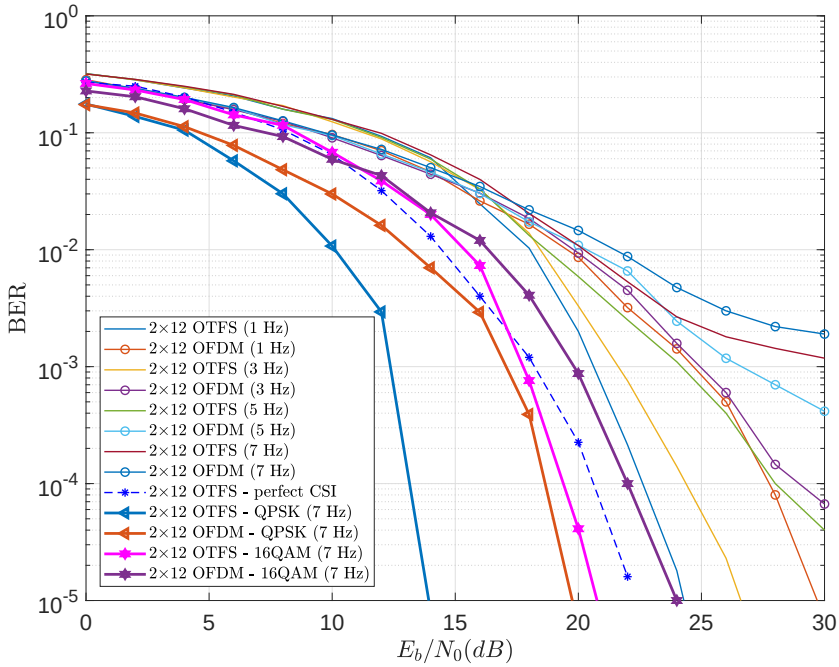


Figure 6.7: BER performance of un-coded 2×12 OFDM & OTFS systems in the UAC with coloured noise for different maximum RMS Doppler spread values.

OTFS system by 5.7 dB at a BER of 10^{-3} while for the OFDM case, the difference is 5.1 dB. This is an expected behavior since the focus is on the spatial multiplexing gain of the MIMO systems instead of the reliability improvement. In Figure 6.8, the performance of the 4×12 OTFS and OFDM systems approaches each other as from 5 Hz. It is however realised that in practical sea experiments longer coherence times are expected. The higher RMS Doppler spread values can be regarded as an upper limit at which no further improvement is achieved with OTFS as compared to OFDM when 64-QAM modulation is used.

The probability of errors in a time-varying UAC tends to be very high and hence the use of FEC coding becomes a necessity. However, this lowers the useful bit rate due to the added redundancy. In this work, Turbo code is integrated in the MIMO OFDM-based OTFS system and the performance of the coded system is shown in Figure 6.9. As can be observed in the latter Figure, the Turbo code greatly reduces the BER over all the RMS Doppler spread values considered for both the 2×12 anf 4×12 systems in the UAC with coloured noise. For instance at an RMS Doppler spread of 1 Hz, the 2×12 Turbo-coded OTFS system outperforms its un-coded counterpart by 7.75 dB at a BER of 10^{-3}. As for the 4×12 Turbo-coded OTFS system, it outperforms its un-coded counterpart by 9.6 dB at the same BER value. When Turbo codes are used and the SNR is sufficiently high, the Turbo decoder can progress beyond the interference limitation and convergence to the noise limitation fixed point becomes possible [159]. This phenomenon usually occurs at a sharp SNR threshold and this explains the waterfall BER

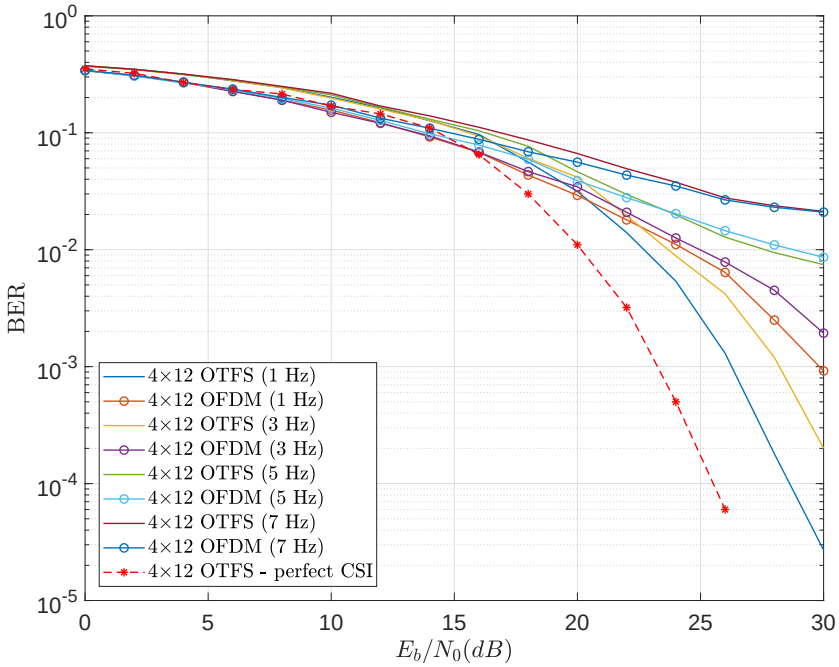


Figure 6.8: BER performance of un-coded 4×12 OFDM & OTFS systems in the UAC with coloured noise for different maximum RMS Doppler spread values.

performance after an arbitrary number of decoding iterations.

For the bit rate computation, the parameters as given in Table 6.1 are considered and the overhead due to pilot-based channel estimation is taken into account. The theoretical maximum achievable bit rate (computed using equation (4.28)) for the 2×12 and 4×12 OTFS and OFDM systems considering an RMS Doppler spread value of 7 Hz is shown in Figure 6.10. The maximum bit rates that can be achieved with the 2×12 and 4×12 systems are 109.7 kbps and 198.7 kbps, respectively. Hence, increasing the number of transmitting elements from 2 to 4 yields a bit rate increase of 81%. It is also observed that the OTFS systems tend to reach their maximum bit rate at a lower E_b/N_0 value than the OFDM systems. Subsequently, one can expect that for the lower RMS Doppler spread values (longer channel coherence times) as expected in practical systems, the difference in E_b/N_0 at which both systems attain their maximum bit rate will be larger.

6.3.2 Performance Evaluation of MU-Massive MIMO Systems

The MU-massive MIMO setup considered in this work is depicted in Figure 6.11. The BS receiver array consists of 100 receiving hydrophones which can be attached to a surface vessel such as a ship or oil platform. Four ROVs are deployed at various depths and separation distances and each one is equipped with 2 to 4 transmitting elements. The ROVs transmit their data simultaneously to the surface vessel in the uplink. The relevant simulation parameters for the massive MIMO systems are summarised in Table

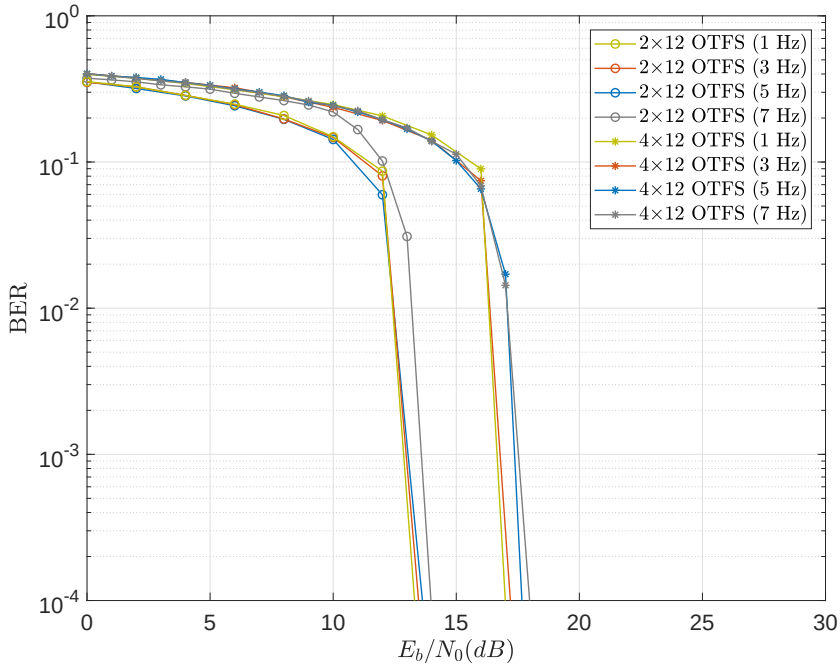


Figure 6.9: BER performance of Turbo-coded MIMO-OTFS systems in the UAC with coloured noise for different maximum RMS Doppler spread values.

6.2. The typical CIRs and CSFs as observed between the BS and the four users (ROVs) are shown in Figure 6.12. Note that the latter figure illustrates an example of the channel response between only one transmitting element and one receiving hydrophone. It should be noted that for UWA communication, the problem of pilot contamination is not as critical as in a terrestrial cellular communication system since the number of users is much less and there is only one surface station which communicate with the ROVs.

The BER performances of un-coded 2×100 OFDM and OTFS systems in an UAC with coloured noise are given in Figure 6.13(a) and Figure 6.13(b), respectively, for two RMS Doppler spread values. For the RMS Doppler spread value of 1 Hz, it can be observed that the OTFS system performs better than its OFDM counterpart. For example, considering only UE1 (ROV 1) the OTFS system outperforms the OFDM system with the same MIMO configuration by 4.2 dB at a BER of 10^{-4}. Moreover, for the same RMS Doppler spread of 1 Hz, the overall BER performance is better than for a 2×12 system. In a massive MIMO system, larger modulation orders can be used to further boost the bandwidth efficiency. This is because, as mentioned previously, the linear combination of the signals received on the multiple hydrophones at the BS results in the channel distortions to be averaged out and the channel gain across each subcarrier is equalised [146, 147]. As for the RMS Doppler spread value of 7 Hz, both systems achieve comparable performance.

Next, the BER performances of un-coded 4×100 OFDM and OTFS systems are evaluated in an UAC

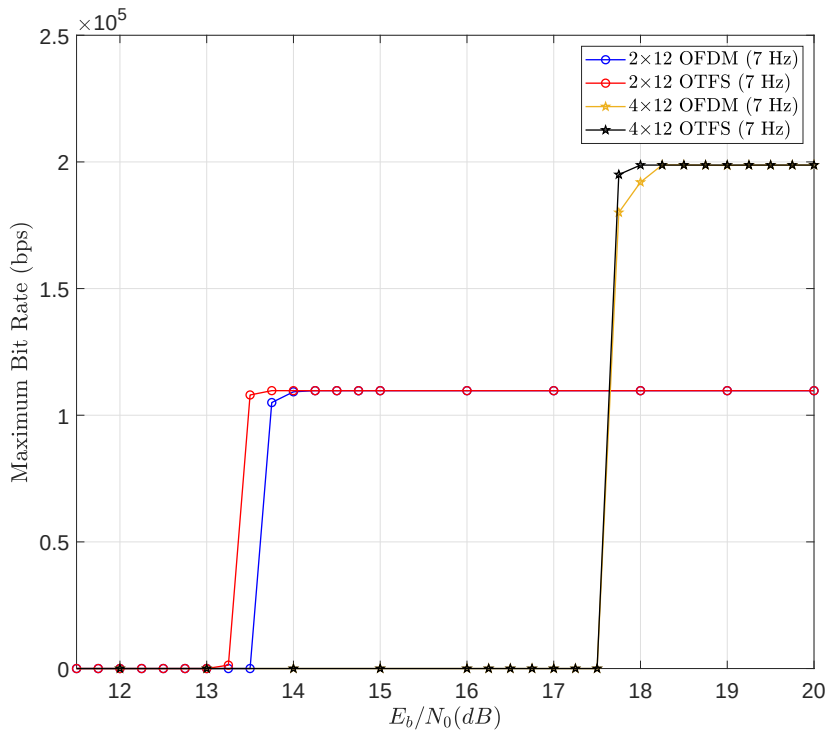


Figure 6.10: Maximum bit rate performance of Turbo-coded MIMO systems in the UAC with coloured noise for a RMS Doppler spread value of 7 Hz.

with coloured noise as depicted in Figure 6.14(a) and Figure 6.14(b), respectively. Considering UE1 (ROV 1) and an RMS Doppler spread value of 1 Hz, the 4×100 OTFS system outperforms its OFDM counterpart by 4.6 dB at a BER of 10^{-4}. Comparing the 4×100 OTFS system in Figure 6.14(b) with the 2×100 OTFS system in Figure 6.13(b) for the same RMS Doppler spread value and user (ROV), a better performance by 3.1 dB is observed with the latter system at a BER of 10^{-4}. With Turbo code, a significant improvement is noted as compared to the un-coded systems as shown in Figure 6.15 for an RMS Doppler spread value of 7 Hz.

The maximum bit rates that can be achieved for each user with the 2×100 and 4×100 OTFS systems are 109.7 kbps and 198.7 kbps, respectively, as shown in Figure 6.16. In fact, if these values are compared with those in Figure 6.10, the same bit rates are achieved as the SU-MIMO case for the same number of transmitting elements. However, the massive MIMO system allows all the ROVs to transmit their data simultaneously to the surface BS over the same frequency and time resources and still achieve a very good error performance.

6.3.3 Video Evaluation over the UAC using OFDM-based MIMO-OTFS Systems

The video transmission is evaluated over the 1 km time-varying UAC using the Turbo-coded MIMO-OTFS systems. The massive MIMO-OTFS systems are not considered in this section since the ROVs

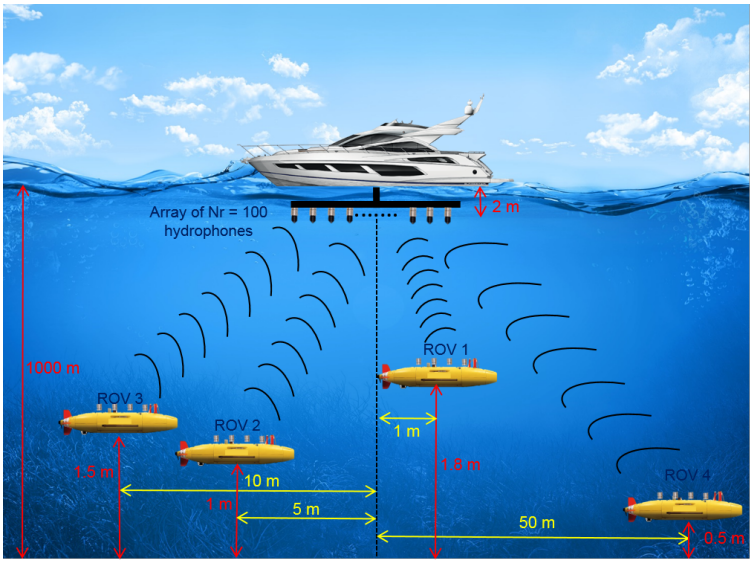
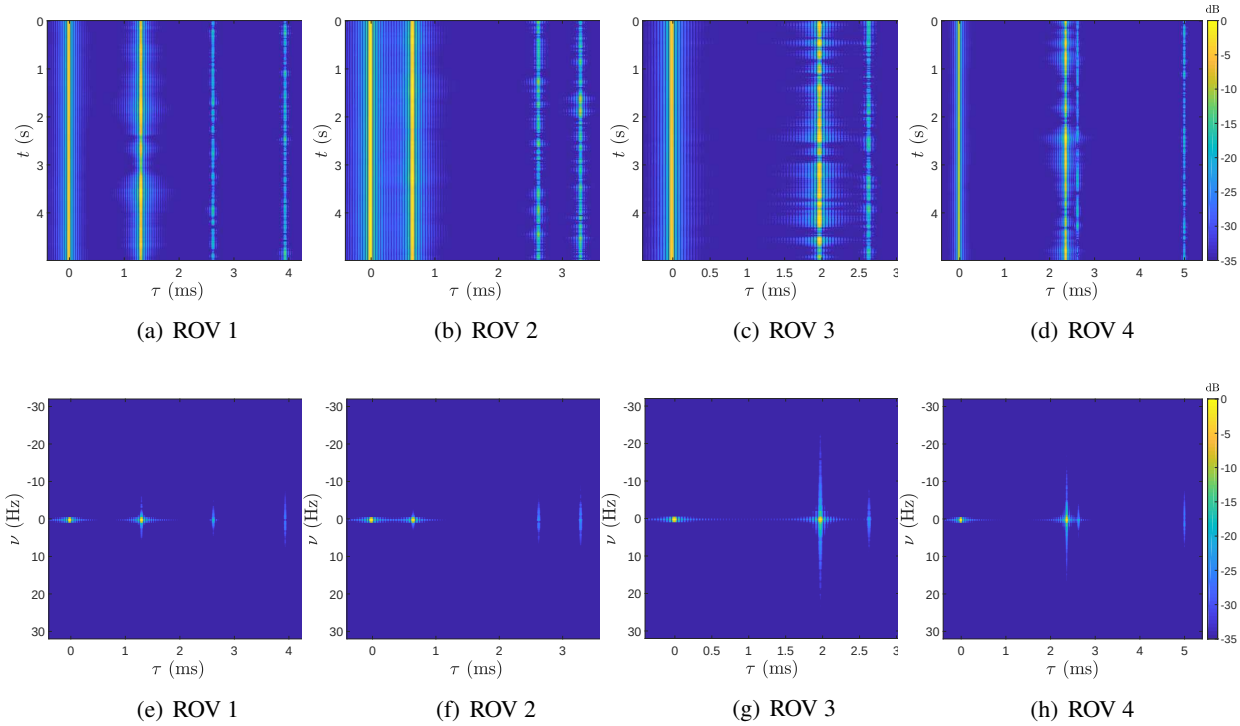


Figure 6.11: Transmission scenario for multi-user (MU) UWA communication.

Table 6.2: Massive MIMO OFDM-based OTFS System Simulation Parameters.

Parameters	ROV 1	ROV 2	ROV 3	ROV 4
Bandwidth	25 kHz			
Carrier frequency	32.5 kHz			
Water depth	1000 m			
TX height from sea-floor	1.8 m	1 m	1.5 m	0.5 m
RX height from sea-floor	998 m			
Number of transmitting transducers per ROV, n_t	2 & 4			
Number of receiving hydrophones at BS, n_r	100			
Subcarriers	512			
Cyclic prefix duration	5.12 ms			
Delay spread	3.9 ms	3.3 ms	2.6 ms	5.0 ms
Modulation	64-QAM			
Turbo code rate	1/2			

achieve the same maximum bit rates as in the SU-MIMO case. The video transmission parameters are summarised in Table 6.3. The compressed video is organised into NAL units which can be regarded as video packets that are transmitted through the UAC. With the aim of error resiliency, the frames are also encoded as slices with a fixed number of bytes. Thus, if a corrupted slice is received for a given frame, the visual degradation is limited to that frame slice and possibly to the subsequent frames' slices within the intra-period. The slice size determines the size of the video packet and consequently has an impact on the video loss rate. The values in Table 6.3 for intra-period and slice size were chosen with the aim to implement the video encoder in real-time for a low-bit rate UAC but may not necessarily represent



(a) ROV 1 (b) ROV 2 (c) ROV 3 (d) ROV 4

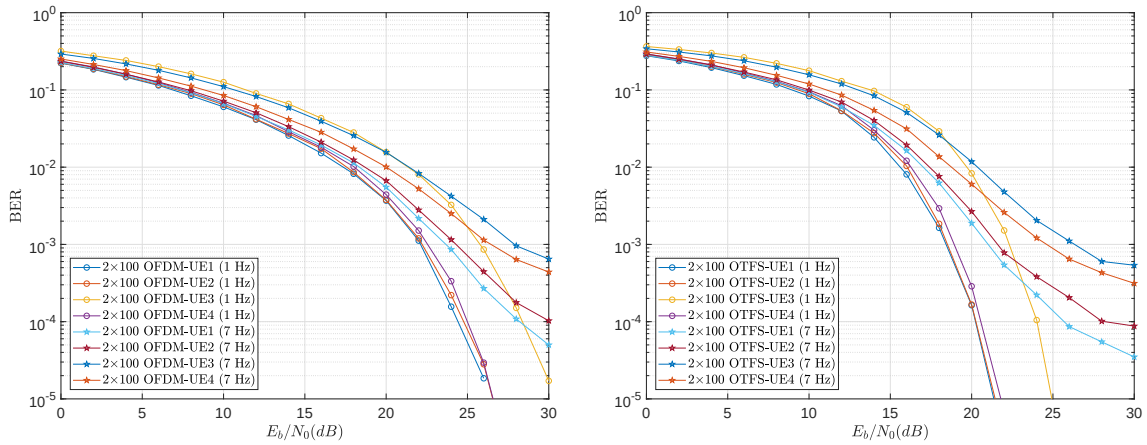
(e) ROV 1 (f) ROV 2 (g) ROV 3 (h) ROV 4

Figure 6.12: Typical channel impulse responses (first row) and channel scattering functions (second row) between BS and ROVs for a maximum RMS Doppler spread value of 7 Hz.

the optimum values. An example of a video frame (not to scale) compressed at 100 kbps and 190 kbps is shown in Figure 6.17. As expected, the video quality at a higher bit rate is better as more details can be extracted for the video. However, as can be observed from Figure 6.17, a video compressed at 100 kbps will also provide enough useful visual information to the end-user to actually differentiate between objects in the video scene. The fact that in an underwater scene, there is little or negligible motion between consecutive frames, the video encoding can be performed more efficiently, enabling a high compression ratio without serious perceptible quality loss.

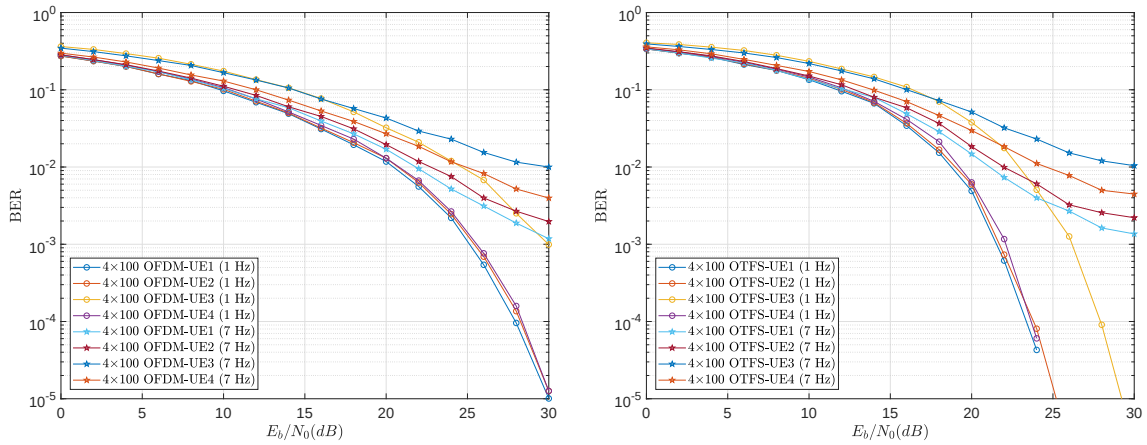
6.3.3.1 Frame Rate Computation

For a video compressed at 100 kbps, the average number of bits per frame is 3320. So considering the maximum bit rate of the 2×12 system, the calculated frame rate is 33.04 fps. As for the video compressed at 190 kbps, the average number of bits per frame is 6359. The calculate frame rate for the 4×12 system is thus 31.25 fps. These values are greater than the original recorded frame rate of 30 fps, implying that real-time transmission is possible. It can also be assumed that the ROVs will have higher computational power than hand-held devices such as mobiles phones to process the video in real-time.



(a) OFDM (b) OTFS

Figure 6.13: BER performance of un-coded 2×100 OFDM and OTFS systems in the UAC with coloured noise for different maximum RMS Doppler spread values.



(a) OFDM (b) OTFS

Figure 6.14: BER performance of un-coded 4×100 OFDM and OTFS systems in the UAC with coloured noise for different maximum RMS Doppler spread values.

6.3.3.2 Packet Loss Rate

Figure 6.18 shows the PLR for the 2×12 and 4×12 OTFS systems. The number of video packets refers to the number of NAL units in each encoded video stream. For both systems, the PLR drops significantly within 0.2 dB owing to the effectiveness of the Turbo codes used.

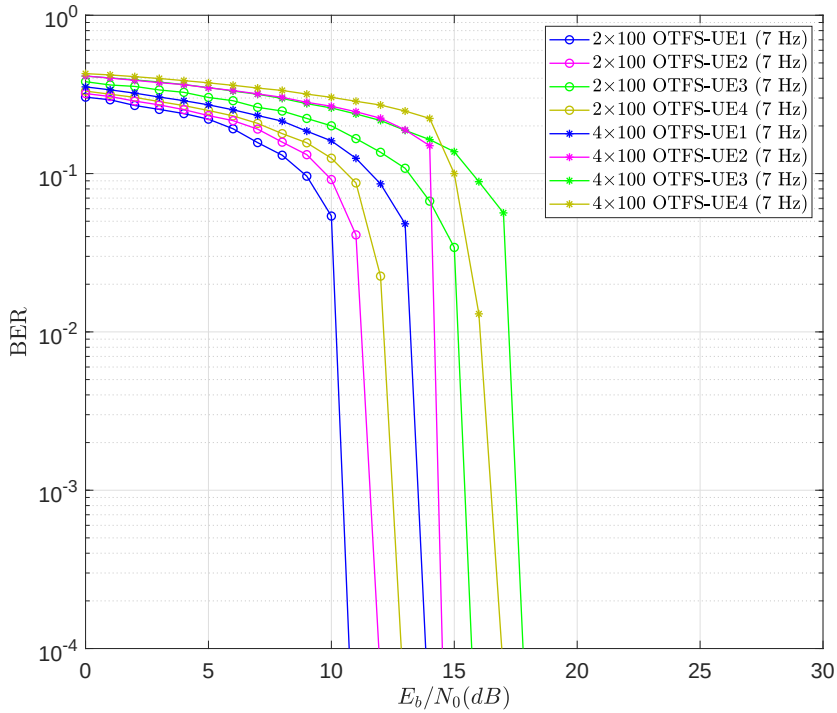


Figure 6.15: BER performance of Turbo-coded massive-MIMO OTFS systems in the UAC with coloured noise for a RMS Doppler spread of 7 Hz.

6.3.3.3 Video Quality Assessment

The video quality is evaluated using the PSNR metric. Figure 6.19 shows the PSNR plot computed from the raw uncompressed 480p video and the received H.264/AVC video streams after being transmitted through the UAC. PLR values between 0 and 60% are considered for the systems. However, it should be noted that videos affected by a PLR beyond 20-25% does not provide any useful information to the end-user since most of the frames are severely degraded. Furthermore, the video may suffer from frame losses for high PLR values because for low bit rate video coding, it is likely a whole frame is encoded within a single packet. For a zero PLR, a higher PSNR value is obtained with the 4×12 OTFS system due to the higher supported bit rate of the system. Moreover, the 4×12 OTFS system will suffer lesser visual degradation than the 2×12 OTFS system for the same PLR as the video is made up of a larger number of video packets. Hence, the probability of frame losses with the 4×12 OTFS system is smaller than the 2×12 system.

For further investigation and considering only the 4×12 OTFS system, a video is encoded with different intra-periods and slice sizes and the video quality of the received streams are analysed as shown in Table 6.4. Eight sequences are tested at a PLR of 6%. An intra-period of 0 means that only the first frame of the video stream is an I-frame while the other frames are P- and B-frames. In this case, the video

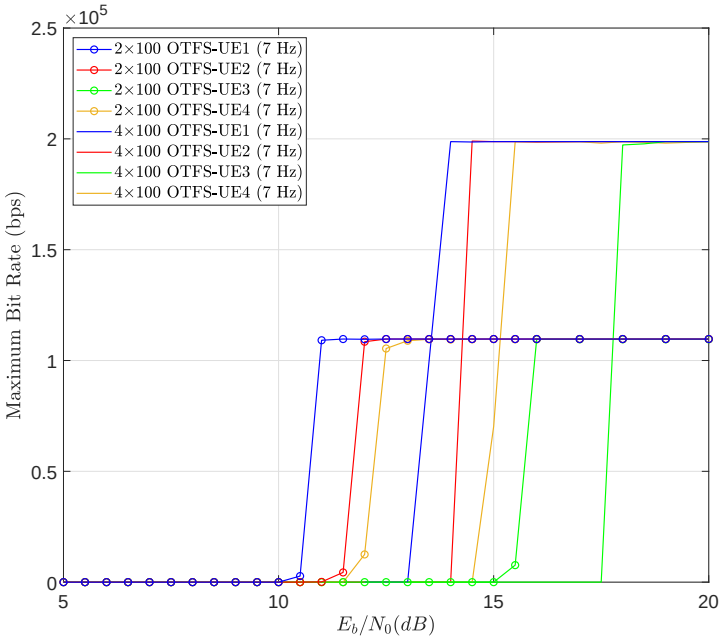


Figure 6.16: Maximum bit rate performance of Turbo-coded massive-MIMO OTFS systems in the UAC with colored noise for a RMS Doppler spread of 7 Hz.

can be encoded more efficiently but the downside is that errors occurring in one frame will propagate to subsequent frames in the video stream. This effect can be observed in test sequences 1 and 2 in Table 6.4 where the PSNR values are lower than the other sequences. As for frequent coding of I-frames, an error in one frame is propagated only to the other frames within the GOP. From Table 6.4, it can also be observed that as the slice size decreases, the number of NAL units increases. Furthermore, there is only a marginal improvement in PSNR with smaller slice sizes for the same intra-period. Therefore in this case it is better to use the bigger slice sizes since they introduce a lower packet header overhead (especially when considering higher layer protocols) and increase the efficiency of the available bit rate. In addition, considering test sequences 4, 6 and 8 for different intra-periods but same slice size, it can be observed than there is no significant difference in the PSNR values. Hence, considering the fact that the use of I-frames results in a higher bandwidth requirement in limited-resource environments and also makes the target bit rate difficult to achieve when using the rate control algorithm in the video encoder, it would be preferable to use longer intra-periods and in so-doing increase the video coding efficiency.

6.4 Improving Bit Rate using Delay-Doppler Domain Pilot Multiplexing

A linear time-varying channel may be represented in various forms such as TF, time-delay or DD representations [160]. The CIR is localised in the DD domain unlike the TF or time-delay representations. In the latter two representations, the rate of the channel variation depends on the operating frequency and

Table 6.3: Video Encoding Parameters for 2×12 and 4×12 OTFS Systems.

Video Resolution	640×480
Frame Count	360
Video duration	12 s
Original frame rate	30 fps
Compression	Uncompressed - 162000 KB Compressed 280 KB (190 kbps) Compressed 146 KB (100 kbps) Ratio - 579:1 (190 kbps) & 1110:1 (100 kbps)
NAL count	614 (100 kbps) 793 (190 kbps)
Time taken to encode	2.03 s (100 kbps) 2.15 s (190 kbps)
Intra-period	8
Slice size	680 bytes

Table 6.4: Received Video PSNR with Different Intra-periods and Slice Sizes for a 4×12 OTFS System with a PLR of 6%.

Test sequence	Intra-period	Slice size (bytes)	NAL count	PSNR (dB)
1	0	170	1937	29.68
2	0	340	1023	29.53
3	4	340	1290	32.99
4	4	680	691	32.75
5	8	340	1186	33.20
6	8	680	793	32.77
7	16	170	1979	32.69
8	16	680	732	32.61

mobility [160]. When either of these two factors is high, the channel varies rapidly and hence channel estimation becomes more challenging. On the other hand, the channel is said to be sparse in the DD domain, where the peaks associated with a delay and Doppler value represent the physical geometry of the channel [160]. This sparsity can be exploited for efficient channel estimation, especially in channels characterised by high Doppler effect.

In an OTFS system, channel estimation can be performed by multiplexing the data symbols with pilot symbols in the TF domain and then performing ZF or MMSE equalisation, as investigated in Section 6.2. However, the implementation of an OTFS system in this fashion does not provide additional diversity [161]. A key aspect of OTFS modulation is its ability to convert a time-varying channel into a compact and sparse DD domain representation where the taps are essentially time-invariant for a large observation time [151]. The 2D circular convolution of the data symbols with such a channel as in Figure 6.1 ensures that the symbols in a given transmission block encounter the same channel gain

[160]. The authors of [150] stated that the OTFS transform allows full diversity to be obtained in the DD domain, where full diversity signifies the number of multipath components, or alternatively, the number of reflectors that can be separated in either the Doppler or delay domain [152].

Several equalisers have been recently proposed for DD domain equalisation for OTFS systems [160–163]. For instance, the authors of [162] proposed a low-complexity message passing algorithm which takes advantage of the channel sparsity to recover the data symbols in an OTFS system. The algorithm was later extended to a MIMO-OTFS system in [163] and channel estimation was also considered using interleaved pilots in the DD domain. The iterative message passing algorithm proposed in [163] is described in section 6.4.2. In [160], an OTFS signal detector based on the Markov chain Monte Carlo sampling method is proposed along with a DD domain channel estimation technique using pseudo-random noise pilots. The authors of [161] proposed an ISI cancellation technique which uses soft-symbol feedback to lower the complexity of the maximum-likelihood detection algorithm. More recently, [164] presented a DD domain embedded pilot-based channel estimation scheme for SISO, MIMO and MU-MIMO systems where guard symbols are used to minimise interference between the pilot and data symbols.

6.4.1 Effective Channel Matrix

Rewriting equation (6.3), the time domain transmitted signal is given by [165]

$$\bar{\mathbf{S}}=\mathbf{G}_{\text{tx}}\mathbf{F}_{\mathcal{M}}^{\mathsf{H}}(\mathbf{F}_{\mathcal{M}}\check{\mathbf{X}}\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}) \tag{6.38}$$

$$=\mathbf{G}_{\text{tx}}\check{\mathbf{X}}\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}, \tag{6.39}$$

where $\mathbf{G}_{\text{tx}}=\text{diag}\big[g_{\text{tx}}(0),g_{\text{tx}}(T/\mathcal{M}),\cdots,g_{\text{tx}}((\mathcal{M}-1)T/\mathcal{M})\big]\in\mathbb{C}^{\mathcal{M}\times\mathcal{M}}$ is a diagonal matrix consisting of pulses $g_{\text{tx}}(t)$ of duration $[0,\ T]$. It is assumed that the transmitted OTFS frame is of duration $\mathcal{N}T$ and the sampling period is $T/\mathcal{M}$ [165]. If rectangular waveforms are considered, then $\mathbf{G}_{\text{tx}}$ is simplified to an identity matrix $\mathbf{I}_{\mathcal{M}}$. Equation (6.39) can be represented as an $\mathcal{MN}\times 1$ vector as

$$\bar{\mathbf{s}}=\text{vec}(\bar{\mathbf{S}})=(\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}\otimes\mathbf{G}_{\text{tx}})\check{\mathbf{x}}, \tag{6.40}$$

where $\check{\mathbf{x}}=\text{vec}(\check{\mathbf{X}}^{\mathsf{T}})\in\mathbb{C}^{\mathcal{MN}\times 1}$. It is assumed that a CP of length N_{CP} is appended to the signal $\bar{\mathbf{s}}$ before it is transmitted through the channel. For a signal, $\bar{s}(t)$, transmitted over the DD domain channel, $h(\tau,\nu)$, the received signal, $r(t)$, can be expressed as [152]

$$r(t)=\int_{\nu}\int_{\tau}h(\tau,\nu)\bar{s}(t-\tau)e^{j2\pi\nu(t-\tau)}d\tau d\nu, \tag{6.41}$$

where the sparse DD domain channel with N_{mp} propagation paths is given by [162]

$$h(\tau,\nu)=\sum_{i=1}^{N_{\text{mp}}}h_i\delta(\tau-\tau_i)\delta(\nu-\nu_i), \tag{6.42}$$

with ν_i and τ_i denoting the Doppler shift and delay taps for the ith path, respectively, and h_i representing the channel coefficient. From equations (6.41) and (6.42), the received signal following the removal of

the CP can be represented in discrete form as [165]

$$r[l]=\sum_{i=1}^{N_{\text{mp}}} h_i e^{j2\pi\frac{k_{\nu_i}(l-l_{\tau_i})}{\mathcal{MN}}}\bar{s}\Big(\big[l-l_{\tau_i}\big]_{\mathcal{MN}}\Big)+\omega(l)\ \text{ for }\ 0\leq l\leq \mathcal{MN}-1, \tag{6.43}$$

where $\omega(l)$ is the noise term and it is assumed that the Doppler shift and delay taps for the ith path are integer multiples of $\frac{1}{\mathcal{N}T}$ and $\frac{1}{\mathcal{M}\Delta f}$, respectively, i.e., $\nu_i=\frac{k_{\nu_i}}{\mathcal{N}T}$ and $\tau_i=\frac{l_{\tau_i}}{\mathcal{M}\Delta f}$ [165]. In vector notation, (6.43) can be written as [165]

$$\mathbf{r}=\dot{\mathbf{H}}\bar{\mathbf{s}}+\boldsymbol{\omega}, \tag{6.44}$$

where $\dot{\mathbf{H}}$ is an $\mathcal{MN}\times\mathcal{MN}$ matrix which is given by

$$\dot{\mathbf{H}}=\sum_{i=1}^{N_{\text{mp}}} h_i \beth^{(l_{\tau_i})}\daleth^{(k_{\nu_i})}, \tag{6.45}$$

with $\beth$ denoting a forward cyclic shift matrix given by

$$\beth=\begin{bmatrix} 0 & \cdots & 0 & 1 \\ 1 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}_{\mathcal{MN}\times\mathcal{MN}}, \tag{6.46}$$

and $\daleth$ is a diagonal matrix of the form

$$\daleth=\begin{bmatrix} e^{j2\pi\frac{k_{\nu_i}(0)}{\mathcal{MN}}} & 0 & \cdots & 0 \\ 0 & e^{j2\pi\frac{k_{\nu_i}(1)}{\mathcal{MN}}} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{j2\pi\frac{k_{\nu_i}(\mathcal{MN}-1)}{\mathcal{MN}}} \end{bmatrix}_{\mathcal{MN}\times\mathcal{MN}}. \tag{6.47}$$

$\beth$ and $\daleth$ model the delay shift (as in OFDM) and the Doppler shift, respectively. The received signal in DD domain is then expressed as

$$\check{\mathbf{y}}=(\mathbf{F}_{\mathcal{N}}\otimes\mathbf{G}_{\text{rx}})\mathbf{r} \tag{6.48}$$

$$=(\mathbf{F}_{\mathcal{N}}\otimes\mathbf{G}_{\text{rx}})\dot{\mathbf{H}}(\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}\otimes\mathbf{G}_{\text{tx}})\check{\mathbf{x}}+(\mathbf{F}_{\mathcal{N}}\otimes\mathbf{G}_{\text{rx}})\boldsymbol{\omega} \tag{6.49}$$

$$=\mathbf{H}_{\text{eff}}\check{\mathbf{x}}+\bar{\boldsymbol{\omega}}, \tag{6.50}$$

where $\mathbf{H}_{\text{eff}}=(\mathbf{F}_{\mathcal{N}}\otimes\mathbf{G}_{\text{rx}})\dot{\mathbf{H}}(\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}\otimes\mathbf{G}_{\text{tx}})$ is the effective channel matrix and $\bar{\boldsymbol{\omega}}=(\mathbf{F}_{\mathcal{N}}\otimes\mathbf{G}_{\text{rx}})\boldsymbol{\omega}$ is the noise vector. $\mathbf{G}_{\text{rx}}$ is the diagonal matrix which contains the pulses $g_{\text{rx}}(t)$ for the filtering operation at the receiver and is similarly defined as $\mathbf{G}_{\text{tx}}$. For arbitrary pulses, the effective channel matrix is given by [153]

$$\mathbf{H}_{\text{eff}}=(\mathbf{I}_{\mathcal{N}}\otimes\mathbf{G}_{\text{rx}})(\mathbf{F}_{\mathcal{N}}\otimes\mathbf{I}_{\mathcal{M}})\dot{\mathbf{H}}(\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}\otimes\mathbf{I}_{\mathcal{M}})(\mathbf{I}_{\mathcal{N}}\otimes\mathbf{G}_{\text{tx}}). \tag{6.51}$$

As for rectangular pulses, $\mathbf{G}_{\text{tx}}=\mathbf{G}_{\text{rx}}=\mathbf{I}_{\mathcal{M}}$. Therefore, in this case the effective channel matrix is given by

$$\mathbf{H}_{\text{eff}}^{\text{rect}}=\sum_{i=1}^{N_{\text{mp}}} h_i \underbrace{\Big[(\mathbf{F}_{\mathcal{N}}\otimes\mathbf{I}_{\mathcal{M}})\beth^{(l_{\tau_i})}(\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}\otimes\mathbf{I}_{\mathcal{M}})\Big]}_{\text{(delay)}}\underbrace{\Big[(\mathbf{F}_{\mathcal{N}}\otimes\mathbf{I}_{\mathcal{M}})\daleth^{(k_{\nu_i})}(\mathbf{F}_{\mathcal{N}}^{\mathsf{H}}\otimes\mathbf{I}_{\mathcal{M}})\Big]}_{\text{(Doppler)}}. \tag{6.52}$$

In order to recover the information symbols, $\chi[k,l]$ in (6.50), a deconvolution operation has to be performed which typically consists of solving a linear system of $\mathcal{NM}$ equations [153].

For a MIMO-OTFS system with n_t transmitting elements and n_r receiving hydrophones, the channel between the tth transmit element and rth receive hydrophone is given by [163]

$$h^{r,t}(\tau,\nu)=\sum_{i=1}^{N_{\text{mp}}} h_i^{r,t}(\tau,\nu)\delta(\tau-\tau_i)\delta(\nu-\nu_i), \tag{6.53}$$

for $t=1,2,\cdots,n_t$ and $r=1,2,\cdots,n_r$. The MIMO-OTFS system can be formulated in matrix notation as [163]

$$\check{\mathbf{y}}^{\text{MIMO}}=\mathbf{H}_{\text{eff}}^{\text{MIMO}}\check{\mathbf{x}}^{\text{MIMO}}+\bar{\boldsymbol{\omega}}^{\text{MIMO}}, \tag{6.54}$$

where

$$\check{\mathbf{x}}^{\text{MIMO}}=\begin{bmatrix}\check{\mathbf{x}}_1\\ \check{\mathbf{x}}_2\\ \vdots\\ \check{\mathbf{x}}_{n_t}\end{bmatrix},\quad \check{\mathbf{y}}^{\text{MIMO}}=\begin{bmatrix}\check{\mathbf{y}}_1\\ \check{\mathbf{y}}_2\\ \vdots\\ \check{\mathbf{y}}_{n_r}\end{bmatrix},\quad \bar{\boldsymbol{\omega}}^{\text{MIMO}}=\begin{bmatrix}\bar{\boldsymbol{\omega}}_1\\ \bar{\boldsymbol{\omega}}_2\\ \vdots\\ \bar{\boldsymbol{\omega}}_{n_r}\end{bmatrix}. \tag{6.55}$$

$\check{\mathbf{x}}^{\text{MIMO}}\in\mathbb{C}^{n_t\mathcal{NM}\times 1}$ is the DD domain transmitted symbol vector over n_t transmitting elements, $\check{\mathbf{y}}^{\text{MIMO}}\in\mathbb{C}^{n_r\mathcal{NM}\times 1}$ is the received symbol vector over n_r receiving hydrophones and $\bar{\boldsymbol{\omega}}^{\text{MIMO}}\in\mathbb{C}^{n_r\mathcal{NM}\times 1}$ is the noise vector over n_r receiving hydrophones. $\mathbf{H}_{\text{eff}}^{\text{MIMO}}\in\mathbb{C}^{n_r\mathcal{NM}\times n_t\mathcal{NM}}$ is expressed as

$$\mathbf{H}_{\text{eff}}^{\text{MIMO}}=\begin{bmatrix}\mathbf{H}_{\text{eff}}^{1,1} & \mathbf{H}_{\text{eff}}^{1,2} & \cdots & \mathbf{H}_{\text{eff}}^{1,n_t}\\ \mathbf{H}_{\text{eff}}^{2,1} & \mathbf{H}_{\text{eff}}^{2,2} & \cdots & \mathbf{H}_{\text{eff}}^{2,n_t}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{H}_{\text{eff}}^{n_r,1} & \mathbf{H}_{\text{eff}}^{n_r,2} & \cdots & \mathbf{H}_{\text{eff}}^{n_r,n_t}\end{bmatrix}, \tag{6.56}$$

where $\mathbf{H}_{\text{eff}}^{r,t}$ is the effective channel matrix between the r-th receive hydrophone and t-th transmit element. The matrix $\mathbf{H}_{\text{eff}}^{\text{MIMO}}$ contains only $n_t\times N_{\text{mp}}$ non-zero elements in each row and $n_r\times N_{\text{mp}}$ in each column.

6.4.2 Message Passing Detection for MIMO-OTFS

Since $\mathbf{H}_{\text{eff}}^{\text{MIMO}}$ is sparse in nature, equation (6.54) can be implemented using a sparsely connected factor graph having $n_t\mathcal{NM}$ variable and $n_r\mathcal{NM}$ observation nodes which correspond to the number of elements in $\check{\mathbf{x}}^{\text{MIMO}}$ and $\check{\mathbf{y}}^{\text{MIMO}}$, respectively [163]. Let ζ_a and ζ_b denote the groups of non-zero positions in the ath column and bth row of the matrix $\mathbf{H}_{\text{eff}}^{\text{MIMO}}$, respectively. Each observation node $\check{y}_b$ is linked to the set of variable nodes $\{\check{x}_c, c\in\zeta_b\}$ and each variable node $\check{x}_a$ is linked to the set of observation nodes $\{\check{y}_c, c\in\zeta_a\}$ [163]. Applying the maximum a posteriori (MAP) rule to equation (6.54) yields the following [163]

$$\hat{\check{\mathbf{x}}}^{\text{MIMO}}=\underset{\check{\mathbf{x}}^{\text{MIMO}}\in\mathbb{S}^{n_t\mathcal{NM}}}{\text{argmax}}\ \mathbf{Pr}\left(\check{\mathbf{x}}^{\text{MIMO}}|\check{\mathbf{y}}^{\text{MIMO}},\mathbf{H}_{\text{eff}}^{\text{MIMO}}\right), \tag{6.57}$$

where $\mathbb{S}$ denotes the modulation alphabet. An optimal performance can be achieved using (6.57). However, it results in exponential complexity [163]. An alternative method is to use the symbol-by-symbol MAP detection for $0\le a\le n_t\mathcal{NM}-1$ as follows

$$\hat{\check{x}}_a=\underset{a_i\in\mathbb{S}}{\text{argmax}}\ \mathbf{Pr}\left(\check{x}_a=a_i|\check{\mathbf{y}}^{\text{MIMO}},\mathbf{H}_{\text{eff}}^{\text{MIMO}}\right) \tag{6.58}$$

$$=\underset{a_i\in\mathbb{S}}{\text{argmax}}\frac{1}{|\mathbb{S}|}\mathbf{Pr}\left(\check{\mathbf{y}}^{\text{MIMO}}|\check{x}_a=a_i,\mathbf{H}_{\text{eff}}^{\text{MIMO}}\right) \tag{6.59}$$

$$\approx\underset{a_i\in\mathbb{S}}{\text{argmax}}\prod_{c\in\zeta_a}\mathbf{Pr}\left(\check{y}_c|\check{x}_a=a_i,\mathbf{H}_{\text{eff}}^{\text{MIMO}}\right). \tag{6.60}$$

The message which is passed from $\check{x}_a$ for $a=\{0,1,\cdots,n_t\mathcal{N}\mathcal{M}-1\}$ to $\check{y}_b$ for $b\in\zeta_a$ is the probability mass function (pmf), defined as $\mathbf{p}_{ab}=\{p_{ab}(a_i)|a_i\in\mathbb{S}\}$, of the symbols in $\mathbb{S}$. The steps in the message passing algorithm are as follows [163]:

1. **Inputs:** $\check{\mathbf{y}}^{\text{MIMO}}$, $\mathbf{H}_{\text{eff}}^{\text{MIMO}}$, maximum number of iterations, N_{iter}.

2. **Initialisation:** Iteration index $iter=0$, $\mathbf{p}_{ab}^{(0)}=1/|\mathbb{S}|\ \forall\ a\in\{0,1,\cdots,n_t\mathcal{N}\mathcal{M}-1\}$ and $b\in\zeta_a$.

3. **Messages from $\check{y}_b$ to $\check{x}_a$:** The mean and variance of the interference component I_{ba} are passed from $\check{y}_b$ to $\check{x}_a$. The term I_{ba} is expressed as

$$I_{ba}=\sum_{c\in\zeta_b,c\neq a}\check{x}_c H_{\text{eff}}^{b,c}+\bar{\omega}_b, \tag{6.61}$$

where $H_{\text{eff}}^{b,c}$ corresponds to the element in the b-th row and c-th column of $\mathbf{H}_{\text{eff}}^{\text{MIMO}}$. The mean of I_{ba} is given by

$$\mu_{ba}^{(iter)}=\mathbb{E}\{I_{ba}\}=\sum_{c\in\zeta_b,c\neq a}\sum_{i=1}^{|\mathbb{S}|}p_{cb}^{(iter)}(a_i)\,a_i\,H_{\text{eff}}^{b,c}, \tag{6.62}$$

while the variance of I_{ba} is computed as

$$(\sigma_{ba}^{(iter)})^2=\sum_{c\in\zeta_b,c\neq a}\left(\sum_{i=1}^{|\mathbb{S}|}p_{cb}^{(iter)}(a_i)|a_i|^2\left|H_{\text{eff}}^{b,c}\right|^2-\left|\sum_{i=1}^{|\mathbb{S}|}p_{cb}^{(iter)}(a_i)\,a_i\,H_{\text{eff}}^{b,c}\right)\right|^2+\sigma^2. \tag{6.63}$$

4. **Messages from $\check{x}_a$ to $\check{y}_b$:** The messages that are passed from $\check{x}_a$ to $\check{y}_b$ is the pmf vector $\mathbf{p}_{ab}^{(iter+1)}$ whose elements are obtained as

$$p_{ab}^{(iter+1)}=\nabla p_{ab}^{(iter)}(a_i)+(1-\nabla)\,p_{ab}^{(iter-1)}(a_i), \tag{6.64}$$

where $\nabla\in(0,1]$ is known as the damping factor which is used to improve the convergence rate and

$$p_{ab}^{(iter)}\propto\prod_{c\in\zeta_a,c\neq b}\mathbf{Pr}\left(\check{y}_c|\check{x}_a=a_i,\mathbf{H}_{\text{eff}}^{\text{MIMO}}\right), \tag{6.65}$$

where

$$\mathbf{Pr}\left(\check{y}_c|\check{x}_a=a_i,\mathbf{H}_{\text{eff}}^{\text{MIMO}}\right)\propto\exp\left(\frac{-\left|\check{y}_c-\mu_{ca}^{(iter)}-H_{\text{eff}}^{c,a}a_i\right|^2}{\sigma_{ca}^{2\ (iter)}}\right). \tag{6.66}$$

5. **Stopping criteria:** Steps 3 and 4 are repeated until either of the following conditions is satisfied:

$$\max_{a,b,a_i}\left|p_{ab}^{(iter+1)}(a_i)-p_{ab}^{(iter)}(a_i)\right|<\varepsilon\ , \tag{6.67}$$

where ε represents a small value or the maximum iteration count, N_{iter}, is reached.

6. **Output:** Finally, the symbol estimate is given by

$$\hat{\check{x}}_a = \underset{a_i \in \mathbb{S}}{\operatorname{argmax}} \;\; p_a(a_i), \;\; a \in \{0,1,\cdots,n_t \mathcal{N}\mathcal{M}-1\}, \tag{6.68}$$

where

$$p_a(a_i) = \prod_{c \in \zeta_a} \mathbf{Pr}\Big(\check{y}_c | \check{x}_a = a_i, \mathbf{H}_{\text{eff}}^{\text{MIMO}}\Big). \tag{6.69}$$

6.4.3 Channel Estimation in Delay-Doppler Domain

Consider an OTFS system with its input-output relation given by [164]

$$y[k,l] = \sum_{k'=-k_v}^{k_v} \sum_{l'=0}^{l_\tau} b[k',l'] \hat{h}[k',l'] x\big[[k-k']_{\mathcal{N}}, [l-l']_{\mathcal{M}}\big] + \bar{\omega}[k,l], \tag{6.70}$$

where $\hat{h}[k',l'] = h[k',l'] e^{-j2\pi \frac{k'}{\mathcal{N}T} \frac{l'}{\mathcal{M}\Delta f}}$ and $b[k',l'] \in \{0,1\}$ represents a path indicator such that if $b[k',l']=1$, this implies that there exists a path with delay, l', and Doppler, k', with a path gain, $\hat{h}[k',l']$ [164]. On the other hand, if $b[k',l']=0$, no such path exists and hence $\hat{h}[k',l']=0$. The total number of paths in the channel can be expressed as

$$\sum_{k'=-k_v}^{k_v} \sum_{l'=0}^{l_\tau} b[k',l'] = N_{\text{mp}}. \tag{6.71}$$

Figure 6.20 illustrates the transmitted pilots for a $4 \times n_r$ system. Note that the high peaks imply that a large pilot power is used as compared to the data symbols. According to [164], the peak transmit power is not affected by the high pilot power as the ISFFT operation in OTFS spreads each DD domain symbol in the entire TF grid. The pilot for each transducer is surrounded by a number of guard symbols as shown in Figure 6.21. The transmitted symbols can be expressed as [164]

$$x[k,l] = \begin{cases} x_{\mathscr{P}}, & k = k_{\mathscr{P}}, l = l_{\mathscr{P}}, \\ 0, & k_{\mathscr{P}} - 4k_v \le k \le k_{\mathscr{P}} + 4k_v, l_{\mathscr{P}} - l_\tau \le l \le l_{\mathscr{P}} + l_\tau \\ x_d[k,l] & \text{otherwise} \end{cases} \tag{6.72}$$

where k_v and l_τ are the taps which correspond to the maximum Doppler and delay, respectively, $x_{\mathscr{P}}$ denotes the pilot symbol, $x_d[k,l]$ represents the data symbols located at $[k,l]$ in the DD OTFS grid and 0 represents a guard symbol. The latter are included to prevent the received symbols, which are used for data detection and channel estimation, from interfering with one another [164]. Consequently, more accurate channel estimates are obtained for a given OTFS frame. The received symbols that are used for channel estimation can be expressed as [164]

$$y[k,l] = b[k-k_{\mathscr{P}}, l-l_{\mathscr{P}}] \hat{h}[k-k_{\mathscr{P}}, l-l_{\mathscr{P}}] x_{\mathscr{P}} + \bar{\omega}[k,l], \tag{6.73}$$

for $k \in [k_{\mathscr{P}} - 4k_v, k_{\mathscr{P}} + 4k_v]$ and $l \in [l_{\mathscr{P}}, l_{\mathscr{P}} + l_\tau]$. By using a positive detection threshold $\mathcal{T}$, if $y[k,l] \ge \mathcal{T}$, this implies that a path exists with a delay, $l - l_{\mathscr{P}}$, and Doppler, $k - k_{\mathscr{P}}$, i.e., $b[k-k_{\mathscr{P}}, l-l_{\mathscr{P}}]=1$, and

Table 6.5: Achievable Bit Rates with TF and DD Domain Equalisations for MIMO-OTFS Systems.

	TF domain equalisation	DD domain equalisation
2×12 OTFS system	109.7 kbps	119.1 kbps
4×12 OTFS system	198.7 kbps	238.1 kbps

therefore $y[k,l]=\hat{h}[k-k_{\mathcal{P}},l-l_{\mathcal{P}}]x_{\mathcal{P}}+\bar{\omega}[k,l]$. Otherwise, $b[k-k_{\mathcal{P}},l-l_{\mathcal{P}}]=\hat{h}[k-k_{\mathcal{P}},l-l_{\mathcal{P}}]=0$. Therefore, the existence of a path implies that the received symbol is the scaled pilot plus noise, otherwise it is solely noise [164]. The threshold $\mathcal{T}$ can be varied to prevent incorrect path detections. The received symbols for data detection can thus be expressed as

$$y[k,l]=\sum_{k'=-k_v}^{k_v}\sum_{l'=0}^{l_\tau} b[k',l']\hat{h}[k',l']x_d\big[[k-k']_{\mathcal{N}},[l-l']_{\mathcal{M}}\big]+\bar{\omega}[k,l], \tag{6.74}$$

for $k\notin[k_{\mathcal{P}}-4k_v,k_{\mathcal{P}}+4k_v]$ and $l\notin[l_{\mathcal{P}},l_{\mathcal{P}}+l_\tau]$.

6.4.4 Performance Evaluation of OFDM-based MIMO-OTFS Systems using Delay-Doppler Domain Equalisation

The BER performance of OFDM-based MIMO-OTFS systems using DD domain equalisation is shown in Figure 6.22. The maximum RMS Doppler spread is set at 2 Hz. The other simulation parameters are the same as those used in Figures 6.7 and 6.8. As shown in Figure 6.20(e), each transmitting and receiving element pair experiences a different channel, which only extends to the amount of the channel delay and Doppler spreads [163]. If the transmitted pilot symbols are sufficiently spaced in the DD domain, then they should arrive at the receiver without overlapping with each other [163], as shown in Figure 6.20(e). As can be observed from Figure 6.22, the uncoded 2×12 and 4×12 OTFS systems with imperfect channel estimation deviate by approximately 3.6 dB and 7.5 dB, respectively, from the systems with perfect channel estimation at a BER of 10^{-3}. The performance with imperfect channel estimation is worse possibly due to the interference between symbols used for channel estimation and data detection. As evaluated in [160], systems with a larger $\mathcal{N}$ provide a better performance as the channel estimates become more accurate. However, a larger value for $\mathcal{N}$ is not suitable for applications requiring a small latency, such as real-time underwater video transmission. The Turbo-coded 2×12 and 4×12 OTFS systems provide approximately 8.4 dB and 11.7 dB better performance, respectively, than their uncoded counterparts at a BER of 10^{-3}. For channels with larger delay and Doppler spreads, more guard symbols should be used along the delay and Doppler axis to accurately estimate the channel [164]. However, this will cause a loss in bandwidth efficiency. By multiplexing pilots in the DD domain instead of the frequency domain, approximately 9% and 20% increase in bit rate is obtained with the 2×12 and 4×12 OTFS systems, respectively, as shown in Table 6.5.

6.5 Conclusion

In this chapter it was demonstrated through various simulations that OTFS provides a better alternative to the more commonly used OFDM to cope with dynamic UACs characterised by large Doppler spreads. It was also shown that using a high modulation order and increasing the number of transmitting elements, high bit rates get can be achieved with the Turbo-coded OFDM-based OTFS systems with good error performance over a 1 km vertically-configured time-varying UAC. The application of massive MIMO reception allows multiple ROVs to be deployed simultaneously and the fact that they can all share the same time and frequency resources makes this technology very attractive for the UWA environment where the bandwidth is extremely limited, especially for long range acoustic links. Using frequency domain pilot multiplexing, bit rates as high as 109.7 kbps and 198.7 kbps were obtained with the 2×12 and 4×12 OTFS systems, respectively, over the 1 km vertical acoustic link. Furthermore, multiple ROVs deployed in a massive MIMO-OTFS system can achieve the same bit rates with a very good error performance (but better bandwidth efficiency). These bit rates are sufficiently high to transmit a 480p video in real-time. The bit rates also enable an acceptable video quality to be achieved using the H.264/AVC compression standard. Moreover, it was shown that further to FEC coding the video encoding parameters can be adapted to provide an improved error resiliency in an UAC. To further increase the achievable bit rates, fewer pilots can be transmitted in the DD domain to estimate the sparse channel in this domain. Approximately 9% and 20% increase in bit rate is obtained with the 2×12 and 4×12 OTFS systems, respectively, over the 1 km vertically-configured time-varying UAC as compared to the case where pilots are multiplexed in the frequency domain. These bit rates will further improve the overall video quality when no bit errors occur for a given SNR.

(a)

(b)

(c)

Figure 6.17: Example of a video frame (not to scale) (a) Original (average PSNR = 35.5 dB) (b) Compressed at 100 kbps (average PSNR = 32.7 dB) (c) Compressed at 190 kbps (average PSNR = 34.4 dB).

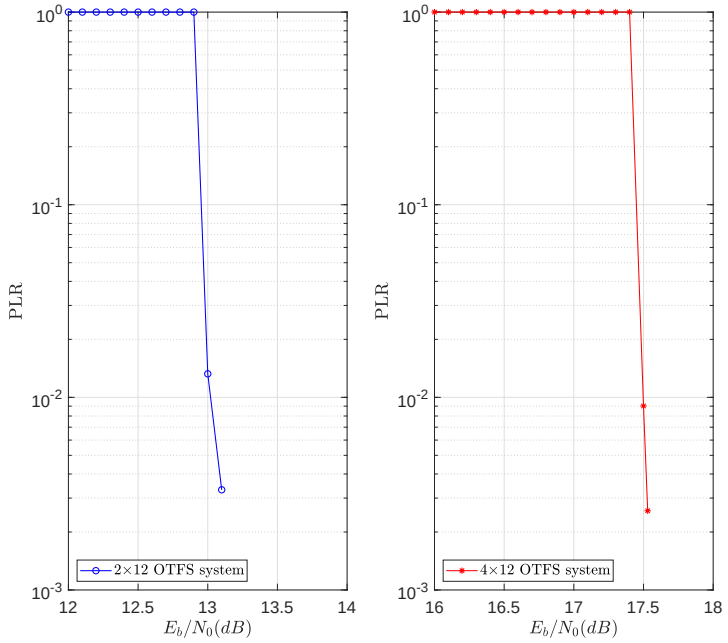


Figure 6.18: Video PLR with Turbo-coded MIMO-OTFS systems (RMS Doppler spread 7 Hz).

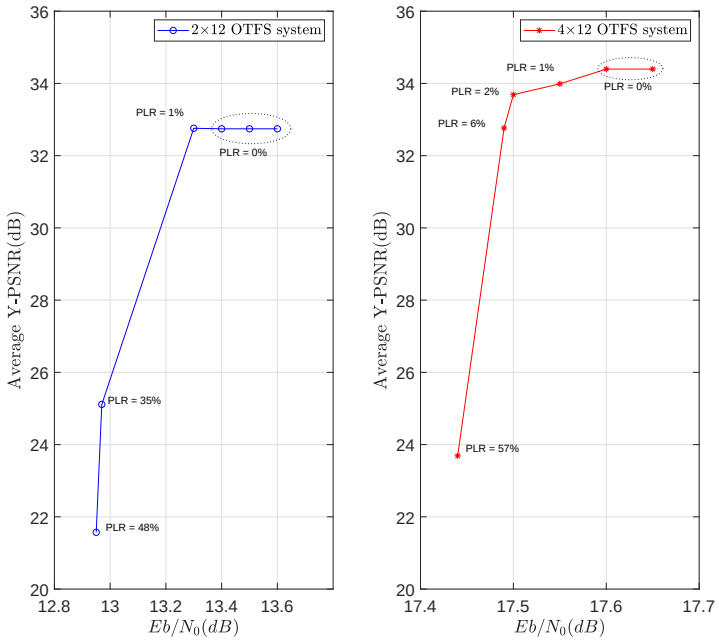
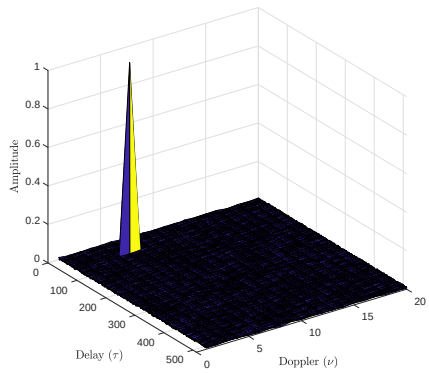
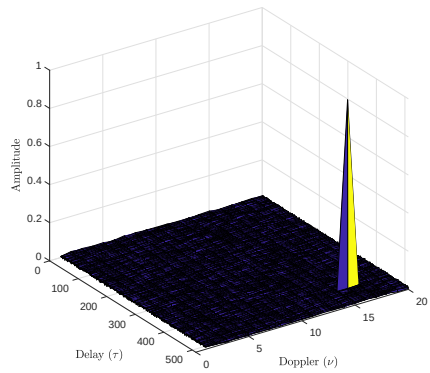


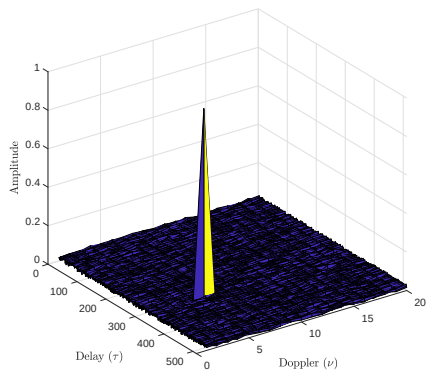
Figure 6.19: Received video PSNR with Turbo-coded MIMO-OTFS systems (RMS Doppler spread 7 Hz).



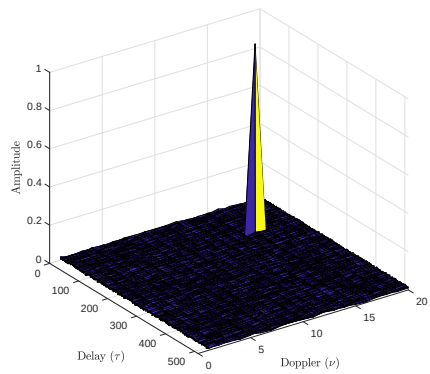
(a) Transmit element 1



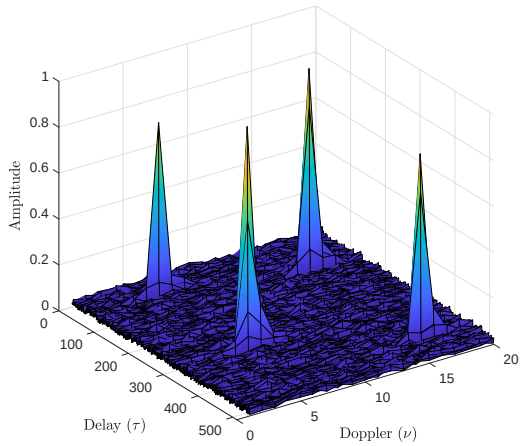
(b) Transmit element 2



(c) Transmit element 3



(d) Transmit element 4



(e) Receive hydrophone 12

Figure 6.20: Transmitted and received pilot symbols in the DD domain for a 4×12 OTFS system.

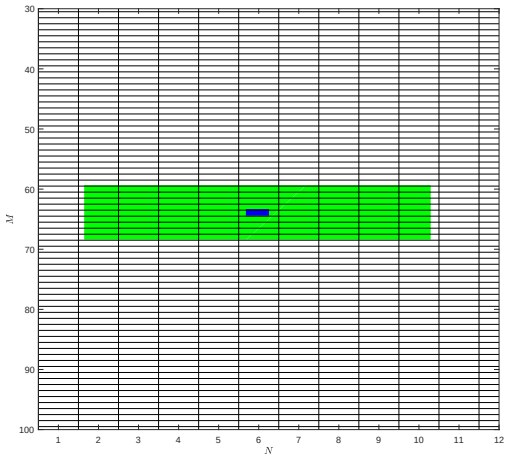


Figure 6.21: Example of symbols arrangement in DD domain for transmitting element 1 (Blue - Pilot symbol, Green - Guard symbols, White - data symbols).

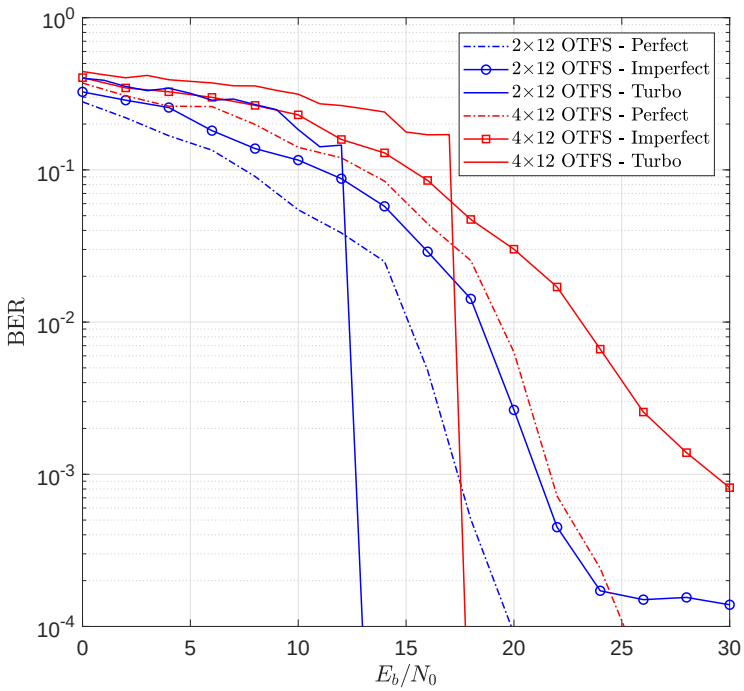


Figure 6.22: BER performance of MIMO-OTFS systems with DD domain equalisation (RMS Doppler spread = 2 Hz).

CHAPTER 7

CONCLUSION AND FUTURE WORK

7.1 Thesis Summary and Conclusions

Since the bandwidth is severely limited in an UAC for long distances (≥ 1 km), real-time video transmission represents a serious challenge. Therefore, the focus of this thesis was to improve the achievable bit rates over long range acoustic links and at the same time address the feasibility of real-time video transmission with acceptable quality. In order to capitalise on the restricted acoustic bandwidth, this thesis investigated and compared bandwidth-efficient physical layer waveforms such as FBMC and the more popular OFDM modulation. FBMC is a potential candidate for 5G terrestrial communication systems since it provides a better bandwidth efficiency than OFDM as no redundant CP is used. Similar to terrestrial communication links, MIMO technology with spatial multiplexing gain is proposed to be used in conjunction with the physical layer waveforms to boost the bit rates in SU UWA communication systems.

In some applications several underwater vehicles need to be deployed. For instance, in some time-critical applications such as carrying out maintenance work or repairing leaks on sub-sea infrastructure in the offshore oil and gas industry, deploying multiple ROVs can give an edge over SU systems. Another advantage of using several ROVs would be in pollution control or covering a larger search area when investigating for sunken objects such as ship wrecks, plane crash, etc. However, while deploying several ROVs simultaneously, all of which are to communicate with a single surface central node, we do not want to reduce the throughput per user (ROV) in an already-bandwidth limited UAC. In this regard, massive MIMO and NOMA techniques are proposed and investigated for MU UWA communication. The advantage of using technologies such as massive MIMO or NOMA for any communication system is that they support many users and all of them can communicate with a central node (or base-station) over the same time and frequency resource. This results in significant improvement in spectrum efficiency

and is therefore very attractive for UWA communication.

One propagation phenomenon that can hinder the performance of any communication system is the Doppler effect and in an UAC this factor can be quite severe. In this respect, this research work also investigated the new modulation technique known as OTFS. This waveform is specifically designed to address channels with high Doppler Effect. Using the OTFS technique, a time-varying fading channel is converted into a time-independent one where all the symbols experience an approximately constant channel gain and, moreover, the diversity of the channel can be fully exploited in both the frequency and time domains using an appropriate equaliser.

In Chapter 3, the FBMC modulation technique was introduced and the advantages it could bring to UWA communication as compared to the already well-established OFDM system were investigated. The BER performance of a SISO system based on FMT and OFDM-OQAM was first evaluated in AWGN and Rayleigh fading channels and compared to OFDM. Assuming perfect channel estimation, it was found that the FBMC systems achieve the same performance as CP-OFDM in these channels. Since the FBMC system based on OFDM-OQAM modulation achieves 100% bandwidth efficiency as compared to FMT where the subcarriers are non-overlapping, it is more suited for UWA video transmission. The BER performance comparison between the SISO OFDM-OQAM and CP-OFDM systems was then carried out in two time-invariant UAC scenarios of link distance 200 m. Assuming that the noise in the UACs was AWGN, the uncoded OFDM-OQAM and CP-OFDM systems once again achieved the same performance. Due to the greater time dispersion in the horizontally-configured UAC, the probability of bit errors was higher than in the vertically-configured UAC. However, with the use of FEC codes such as Reed Solomon and Turbo codes, the error performance was greatly improved in both channel configurations. For the same code rate (1/3), the Turbo-coded systems proved to be more robust since they outperformed the Reed Solomon coded systems in both UAC scenarios. Considering a bandwidth of 115 kHz for the acoustic link distance of 200 m, 16-QAM modulation and 8192 subcarriers, a theoretical bit rate of 138 kbps was achieved with the Turbo-coded OFDM-OQAM system as compared to 121 kbps for the Turbo-coded CP-OFDM system. Given the severe limitation on the acoustic bandwidth, the increase in bit rate achieved with OFDM-OQAM modulation makes this scheme better suited for UWA video transmission where high data rates are desirable for delivering good quality and resolution real-time video.

Chapter 4 proposed to apply MIMO technology to FBMC modulation in order to boost the bit rate in UWA communication. In this chapter, the BER performances of coded and un-coded spatially-multiplexed MIMO OFDM-OQAM and MIMO-OFDM were evaluated over 1 km time-varying horizontal and vertical UACs. Coloured noise was considered in the simulations instead of AWGN to better reflect the real-world underwater noise. Channel estimation was carried out using the IAM. Based on the acoustic link distance of 1 km, the system's bandwidth was set at 25 kHz. The other simulation parameters consisted of 16-QAM modulation and based on the maximum delay spread in the UACs, the number of subcarriers was set to 1024 and 512 for the horizontal and vertical channels, respectively. It was demonstrated that the MIMO OFDM-OQAM systems achieve better error performance than

the MIMO-OFDM systems in both the horizontal and vertical UACs for the same transmission time. As demonstrated previously in Chapter 3, the systems' performance in the vertical UAC was better than in the horizontal UAC due to the lower time dispersion. The 2×12 and 4×12 rate-$\frac{1}{2}$ Turbo-coded OFDM-OQAM systems achieved maximum bit rates of 73.7 kbps and 105.3 kbps, respectively, in both channel configurations. The improvement in bit rate with the 2×12 and 4×12 OFDM-OQAM systems as compared to OFDM with the same MIMO configuration was 40% and 67%, respectively, in the horizontal channel and 17% and 25%, respectively, in the vertically-configured UAC. In this chapter, the transmission of a SD video was also evaluated over the 1 km vertically-configured UAC. The raw video of resolution 640×480 and frame rate 30 fps was compressed to target bit rates of 70 kbps and 100 kbps for the 2×12 and 4×12 Turbo-coded OFDM-OQAM systems, respectively. The H.264/AVC standard was used ahead of HEVC because as investigated in Chapter 2, the received H.264/AVC stream is more resilient to video packet loss than the HEVC encoded stream, especially for low bit rate channels. Furthermore, HEVC involves more computational overheads than H.264/AVC since it uses a complex prediction process to remove redundant information during video encoding and this may cause a delay in video transmission. The calculated frame rates for the 2×12 and 4×12 Turbo-coded OFDM-OQAM systems were found to be greater than the original video frame rate of 30 fps, implying real-time transmission is theoretically feasible with these bit rates. The 2×12 and 4×12 OFDM-OQAM systems also achieved a zero video PLR at E_b/N_0 values of 12 dB and 17 dB, respectively, with corresponding PSNR values of 34 dB and 35 dB. This means that the bit rates obtained with the MIMO OFDM-OQAM systems are sufficiently high for the ROV to transmit a SD video with acceptable quality and in real-time to a surface vessel over a 1 km vertical acoustic link. This chapter also showed that OFDM-OQAM can provide robust performance in doubly-dispersive UAC without having to use a CP as in OFDM. In fact, since the complex orthogonality condition is relaxed to the real field in the conventional OFDM-OQAM system, pulse shape filters with good TFL properties can be used such as the Hermite filter.

In Chapter 5, massive MIMO reception was proposed to improve the communication reliability over a 1 km vertically-configured time-varying UAC and at the same time boost the achievable bit rate for both SU and MU systems. In the SU scenario, the BER performance of Turbo-coded massive MIMO systems based on FBMC modulation and OFDM were evaluated over the UAC. To overcome the imaginary interference which is inherent in the conventional OFDM-OQAM system and make its application to MIMO straightforward, a modified FBMC system as proposed in [121] was considered whereby complex data symbols are transmitted instead of real-valued ones. This method involves spreading the data symbols in time without sacrificing the bandwidth efficiency and ensuring a low system's complexity. The transmission set-up consisted of a large array of 100 hydrophones at the surface BS while the underwater ROV was equipped with a maximum of 4 transmitting elements. Scattered pilot-based channel estimation was used for the massive MIMO OFDM and OFDM-OQAM systems. It was shown that the Turbo-coded FBMC systems outperform the Turbo-coded OFDM systems both in terms of error performance and bit rate for the same massive MIMO configuration. For instance, using a system's bandwidth of 25 kHz, 64-QAM modulation, 256 subcarriers and rate-$\frac{1}{2}$ Turbo code, bit

rates as high as 133 kbps and 242 kbps were theoretically possible with the 2×100 and 4×100 FBMC systems, respectively, as compared to 91 kbps and 164 kbps for OFDM systems with the same massive MIMO configuration. Furthermore, the transmission of a 480p video which was compressed using the H.264/AVC standard was evaluated in the UAC using the massive MIMO OFDM-OQAM systems. The raw video of resolution 854×480 and frame rate 30 fps was compressed to target bit rates of 130 kbps and 240 kbps for the 2×100 and 4×100 OFDM-OQAM systems, respectively, using the rate control algorithm in the H.264/AVC encoder. The E_b/N_0 values that were required to achieve no bit errors and hence a zero video PLR were 8 dB and 10 dB for the 2×100 and 4×100 systems, respectively, with corresponding received video PSNR values of 35.5 dB and 37.5 dB. This chapter also investigated the application of massive MIMO to a MU UWA communication system. The transmission scenario consisted of four underwater ROVs, each one transmitting with four transducers. It was shown that the implementation of massive MIMO at the receiver with an array of 100 hydrophones allowed all the ROVs to use the same frequency bandwidth to transmit their information reliably to the receiver (very good error performance achieved for all ROVs). Furthermore, considering the simulation parameters as in the SU scenario but a lower FBMC spreading length, each ROV could transmit at a bit rate of approximately 221.5 kbps over the simulated 1 km time-varying UAC. This chapter showed that the use of massive MIMO reception increased the reliability of the long range vertical UWA communication link, and also enabled a higher modulation order to be used together with a higher FEC code rate. It is to be noted that the FBCM-OQAM modulation technique with time-spreading allows for a lower pilot overhead for channel estimation as compared to preamble-based channel estimation in conventional OFDM-OQAM systems. As a result, high bit rates were achieved over the 1 km vertically-configured time-varying UAC, which consequently made real-time video transmission with acceptable quality feasible. Furthermore, the advantage of an FBMC-based massive MIMO system is that it allows a smaller number of subcarriers to be used. This in turn reduces the delay caused by the synthesis and analysis filter banks and at the same time makes the system more robust to carrier frequency offsets. Moreover, higher modulation orders can be utilised to increase the bandwidth efficiency as the linear combination of the signals received on the multiple hydrophones at the BS smooths the channel distortions across the subcarriers and therefore equalises the channel gain across each one of them.

Finally, Chapter 6 investigated video transmission using OFDM-based OTFS systems for UWA communication. The performance of coded and un-coded spatially multiplexed MIMO and massive MIMO systems for SU and MU scenarios, respectively, was evaluated over a 1 km vertically-configured time-varying UAC. Using frequency domain pilot multiplexing, bit rates of 109.7 kbps and 198.7 kbps were obtained using the Turbo-coded 2×12 and 4×12 MIMO-OTFS systems, respectively, over the 1 km vertical acoustic link. Furthermore, four ROVs deployed in a Turbo-coded massive MIMO-OTFS system achieved the same bit rates with a very good error performance. These bit rates were found to be sufficiently high to transmit a 480p video, compressed using the H.264/AVC standard, in real-time and with acceptable quality. It was also shown that further to FEC coding the video encoding parameters could be tweaked to provide an improved error resiliency in an UAC. The DD domain of the

OTFS modulation scheme was also exploited to multiplex much fewer pilots than would be required in the frequency domain. This allowed bit rates as high as 119.1 kbps and 238.1 kbps to be achieved using the Turbo-coded 2×12 and 4×12 MIMO-OTFS systems, respectively, over the same UAC. This chapter demonstrated that OTFS represents a better alternative than OFDM to cope with dynamic UACs characterised by large Doppler spreads and also to improve the overall bit rates of the systems over a long-range UWA link.

7.2 Future Work

In this thesis, the FBMC physical layer waveform implemented with the time spreading method has been found to be very attractive for UWA video transmission since it provides a higher bit rate than OFDM and performs well in doubly-dispersive UACs. Furthermore, the application of massive MIMO to next-generation UWA communication systems would be highly advantageous since this technology allows multiple underwater vehicles to be deployed simultaneously and share the same time and frequency resources to transmit their data to a surface station with excellent communication performance (in terms of achievable bit rates and link reliability). It was also shown that the OTFS modulation scheme can cope well with UACs with high Doppler spreads.

The first recommendation for future work is to assess the combination of FBMC and OTFS using the DD pilot multiplexing and detection. Based on the desired underwater application, MIMO and massive MIMO can then be applied to the FBMC-based OTFS systems for SU and MU scenarios, respectively. Furthermore, the sparsity of the UAC using OTFS modulation can be investigated using different types of transmit and receive pulses/windowing. In Chapter 6, $\mathcal{N}$ CPs were used in one OTFS frame. In the scope to improve the overall spectral efficiency of an OFDM-based OTFS system, it may be worth appending only one CP to the whole OTFS frame as in [165] and evaluating the system's performance over the UAC. Moreover, if frequency-domain equalisation is considered for the OTFS systems, non-linear equalisers such as iterative DFE can be used to further improve the performance of OTFS for high modulation orders [152].

Underwater images usually exhibit low contrast, especially with increasing depths. This means that their information content is more acute at the lower frequencies. Thus, by categorising the information for each video frame into low and high frequencies and then encoding the information at the lower frequencies with a higher accuracy, a better compression efficiency can be obtained, for example using the DWT which may be much faster that the software implementations of H.264/AVC and HEVC standards. Alternatively, following video decoding at the receiver side, the DWT technique may be used to enhance the quality of each frame in the low contrast decoded video sequence.

For promoting real-time underwater video transmission, the Scalable Video Coding (SVC) version of H.264/AVC or HEVC can be investigated for an adaptive UWA communication system where the bandwidth, and therefore bit rates, vary for different transmission distances. Hence, in this case the received video quality will depend on how good the channel is for a specific transmission range.

A lot of research in this field usually depends on field experiments which are however expensive and time-consuming. Hence, UAC simulators are used to design the UWA communication systems and test their link and system level performances. However, these simulators do not usually take into account the beam patterns of the transducers and the combined frequency response of the amplifiers and transducers. Therefore, in a future contribution these characteristics may be considered in order to develop a more realistic simulation platform. Furthermore, the author aims to test and validate the techniques proposed in this thesis on a real-life set-up. Firstly, practical experiments can be performed, for example in a swimming pool, using a scaled-down version of the proposed UWA communication system. Then, the experiments can be later extended to sea-trials by designing and implementing a modem using the Software Defined Radio (SDR) technology but which is however adapted to support UWA frequencies and hardware components.

APPENDIX A

MULTI-USER UNDERWATER ACOUSTIC COMMUNICATION USING AN FBMC-BASED NOMA SYSTEM

A.1 Introduction

In order to meet the demand for future underwater applications where urgent intervention is required, multiple unmanned underwater vehicles may be deployed simultaneously. Multiple-access techniques based on spread spectrum have been investigated for UWA communication to realise MU communication in an UAC (see, e.g., [166, 167]). CDMA has been used since it provides good interference suppression capability between users and is robust against multipath effects. However, the challenge in a MU scenario is to obtain a proper trade-off between user fairness and system throughput [144]. Orthogonal Multiple Access (OMA) techniques based on CDMA, FDMA, TDMA and OFDMA have been and are still being used in terrestrial mobile radio communication systems. As compared to OMA schemes such as Orthogonal Frequency Division Multiple Access (OFDMA), in NOMA multiple users use the same time and frequency resources, hence achieving a superior overall spectral efficiency. This makes the NOMA technology very attractive for MU UWA communication where the bandwidth is extremely limited. It has also been shown in many literatures for terrestrial mobile communication that the users in a NOMA system achieve a higher throughput than those using the OMA technique [168].

The basic idea behind NOMA is that the users' signals are superposed in the power domain by exploiting the difference in their channel gain and the power allocation is performed such that the users who experience favourable and poor channel conditions are allocated less power and more power, respectively [169]. At the receiver, Successive Interference Cancellation (SIC) is then used for MU detection and decoding. OFDM or FBMC signalling combined with non-orthogonal user multiplexing can be used to generate the transmit signals for the users.

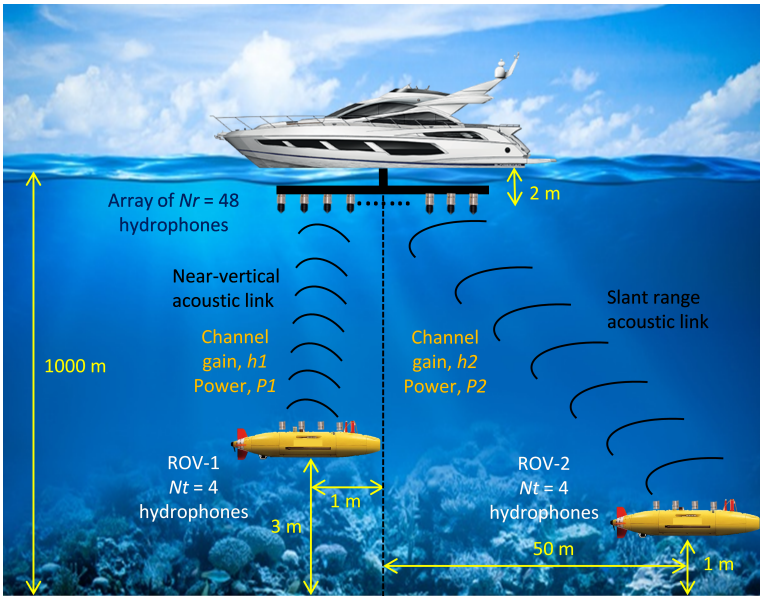


Figure A.1: Target scenario for multi-user UWA communication.

In this contribution, the power domain NOMA technique is investigated for MU UWA communication. The scenario of interest is depicted in Figure A.1 where two underwater ROVs communicate simultaneously with a surface vessel over a range of approximately 1 km. Each ROV is equipped with 4 transmitting elements while the surface vessel has an array of 48 receiving hydrophones. The downlink is used for sending control commands to the ROVs while the uplink is used solely for transmitting real-time video to the surface vessel. The NOMA scheme can be implemented using either OFDM or FBMC modulation for waveform shaping. In order to boost the throughput over the 1 km time-varying UAC, spatially multiplexed MIMO systems are considered. The systems are also Turbo-coded to improve the reliability of the communication link. The BER, PER and maximum bit rate performances of the Turbo-coded NOMA-OFDM and NOMA-FBMC systems are evaluated for the 2-user scenario where both users utilise the same frequency bandwidth. An FBMC system based on OFDM-OQAM modulation is considered since the subcarriers maximally overlap and hence 100% bandwidth efficiency is achievable. Moreover, to cope with the imaginary interference in conventional OFDM-OQAM and make its application to MIMO straightforward, a modified OFDM-OQAM system as proposed in [136] is considered (See Section 5.3). It is shown that while both the NOMA-OFDM and NOMA-FBMC systems show comparable performance in terms of BER and PER, the MIMO NOMA-FBMC system however achieves a higher bit rate than the OFDM-based MIMO-NOMA system (125.5 kbps with FBMC compared to 88.72 kbps with OFDM).

The main contributions of this work are reported in [44]. The rest of this appendix is organised as follows. Section A.2 provides the mathematical formulations of the NOMA systems for the 2-user scenario. The performance evaluation of the systems in terms of BER, PER and maximum achievable bit rate is presented in Section A.3. Finally, conclusions are drawn in Section A.4.

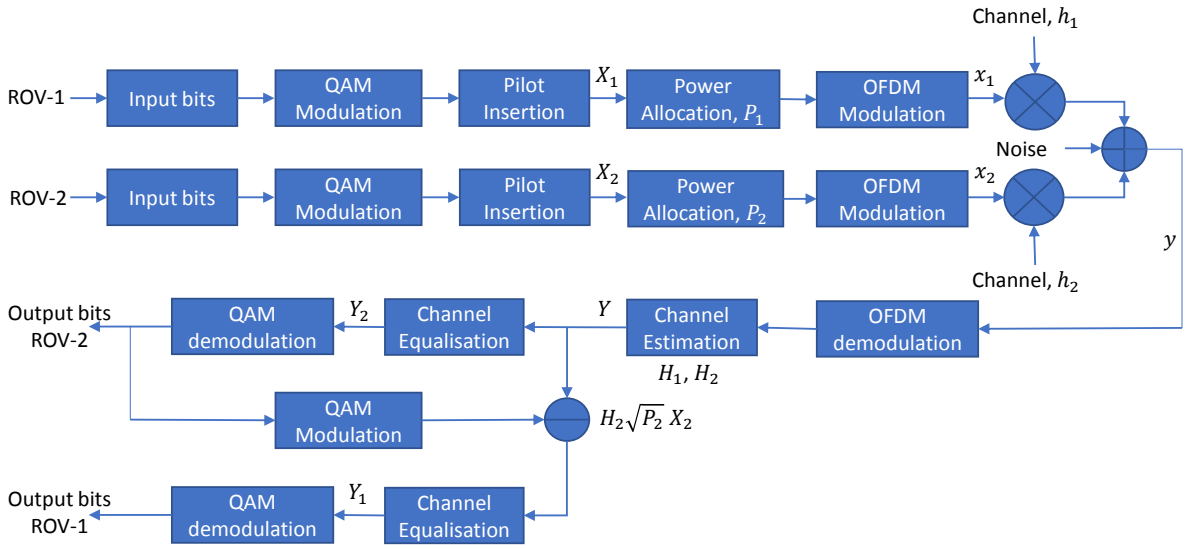


Figure A.2: Block diagram of a 2-user uplink OFDM-based NOMA system.

A.2 System Model

A.2.1 SISO-NOMA

NOMA is a potential candidate for the 5G mobile communication systems since it can provide a better bandwidth efficiency than OMA techniques such as OFDMA [168]. In NOMA, all the subcarriers in the system are assigned to each user. Hence the bandwidth resources which are assigned to the user in poor channel conditions are also available to the one in strong channel conditions, significantly improving the spectral efficiency [170]. Compared to conventional power allocation strategies where a user experiencing strong channel conditions is allocated more power, NOMA on the other hand allocates more power to the user in less favourable channel conditions [169]. SIC is then used at the receiver for detection and decoding, whereby the strongest signal is decoded first (treating other signals as interference) which is then subtracted from the overall received signal to get the next strongest signal [171]. Similarly, the latter is decoded and subtracted from the combined signal and this process is repeated until all information signals are decoded [171]. In our scenario, each ROV utilises the same spectrum resources to transmit simultaneously in the uplink to a BS which is located at the surface. Depending on the channel gain differences, each ROV transmits at a maximum or a controlled power level. The superposed signal received at the BS from the different ROVs interfere with each other. Hence, the BS applies SIC to decode the message signal from each ROV. The block diagram of a 2-user SISO OFDM-based NOMA system is shown in Figure A.2. At the BS, the received superposed signal is given by

$$y = h_1\sqrt{P_1}\star x_1 + h_2\sqrt{P_2}\star x_2 + \omega, \tag{A.1}$$

Table A.1: FBMC-based NOMA System Simulation Parameters.

Parameters	ROV 1	ROV 2
Bandwidth	25 kHz	
Carrier frequency	32.5 kHz	
Water depth	1000 m	
TX height from sea-floor	3 m	1 m
RX height from sea-floor	998 m	
Number of transmitting transducers per ROV, n_t	4	
Number of receiving hydrophones at BS, n_r	48	
Subcarriers	256	
Max. delay spread	2.6 ms	3.9 ms
OFDM CP duration	5.12 ms	
Modulation	16-QAM	
Turbo code rate	1/3	
Maximum RMS Doppler spread	2 Hz	
Average channel gain	≈-27 dB	≈-32 dB
NOMA power allocation factor	0.4	0.6

where the signal transmitted by ROV-i (i=1, 2) is denoted by x_i ($\mathbb{E}[|x_i|^2]$=1), P_i is the allocated power for ROV-i, h_i is the channel coefficient between the BS and ROV-i, and ω represents the channel noise with a PSD of N_0. The frequency-domain representation of (A.1) is given by

$$Y = H_1\sqrt{P_1}X_1 + H_2\sqrt{P_2}X_2 + \Omega, \tag{A.2}$$

where H_i is the frequency-domain channel coefficient of ROV-i which can be estimated using the received pilot symbols from the two ROVs at the BS. In Figure A.1, ROV-1 experiences a higher channel gain than ROV-2 and hence $P_1 < P_2$. At the BS, the equalised signal for ROV-2 is obtained as

$$Y_2 = \frac{Y}{H_2\sqrt{P_2}} = X_2 + \frac{H_1\sqrt{P_1}X_1}{H_2\sqrt{P_2}} + \frac{\Omega}{H_2\sqrt{P_2}}. \tag{A.3}$$

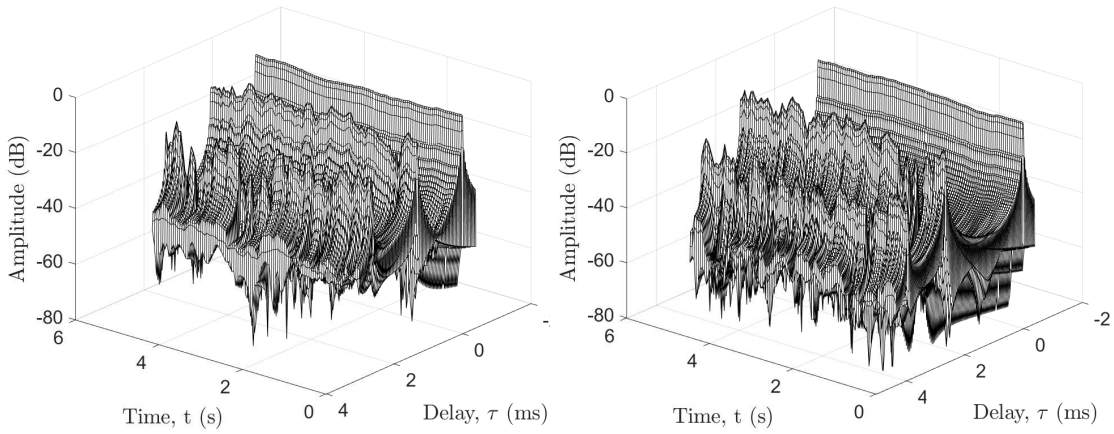
The equalised signal for ROV-1 is then computed as

$$Y_1 = \frac{Y - H_2\sqrt{P_2}X_2}{H_1\sqrt{P_1}} = X_1 + \frac{\Omega}{H_1\sqrt{P_1}}. \tag{A.4}$$

The estimated transmitted bits of ROV-1 and ROV-2 are then obtained by de-mapping Y_1 and Y_2, respectively.

A.2.2 MIMO-NOMA

MIMO technology can be applied to NOMA to exploit both the spatial and power domains, thus further improving the performance of the system. In this case, the users are separated in the power domain while the spatial domain is used for multiplexing the information signal of each user on multiple transmitting



(a) (b)

Figure A.3: A typical channel response between BS and (a) ROV 1 (b) ROV 2.

elements. Consider an uplink MIMO system with n_r receive hydrophones at the BS and n_t transmitting elements at each ROV. Thus, the received signal vectors before equalisation for the two ROVs at the BS can be expressed as

$$\mathbf{y}_2 = \mathbf{H}_1\sqrt{P_1}\mathbf{x}_1 + \mathbf{H}_2\sqrt{P_2}\mathbf{x}_2 + \boldsymbol{\eta}, \tag{A.5}$$

$$\mathbf{y}_1 = \mathbf{y}_2 - (\mathbf{H}_2\sqrt{P_2}\hat{\mathbf{x}}_2) + \boldsymbol{\eta}, \tag{A.6}$$

where $\mathbf{y}_i$ is the $n_r \times 1$ received signal vector from ROV-i, $\mathbf{x}_i$ is the $n_t \times 1$ transmitted symbol vector for ROV-i, $\mathbf{H}_i$ is the $n_r \times n_t$ channel matrix for ROV-i, $\boldsymbol{\eta}$ is the $n_r \times 1$ noise vector and P_i is the power allocation factor for ROV-i. At the receiver side, a ZF or MMSE detection matrix can be used to decode the signals, $\hat{\mathbf{x}}_i$. The detection matrix can be generated by the BS using the CSI from the two ROVs. Assuming MMSE, the detection matrix of the channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ is given by

$$\mathbf{W}_i = \left(\mathbf{H}_i\mathbf{H}_i^{\mathsf{H}} + \rho_{\mathrm{rx}}^{-1}\mathbf{I}_{n_t}\right)^{-1}, \tag{A.7}$$

where $\mathbf{I}_{n_t}$ is an $n_t \times n_t$ identity matrix and ρ_{rx} is the receive SNR at the BS. The detected symbol vectors for ROV-i is thus given by

$$\hat{\mathbf{x}}_i = \mathbf{W}_i\mathbf{y}_i. \tag{A.8}$$

A.3 Simulation Results

The simulation parameters are summarised in Table A.1 while the physical setup of the transmission scenario is illustrated in Figure A.1. Typical channel responses as observed by the two ROVs are illustrated in Figure A.3. The BER and PER performances for the two ROVs using Turbo-coded 4×48 NOMA-OFDM and NOMA OFDM-OQAM signalling are shown in Figure A.4 and Figure A.5,

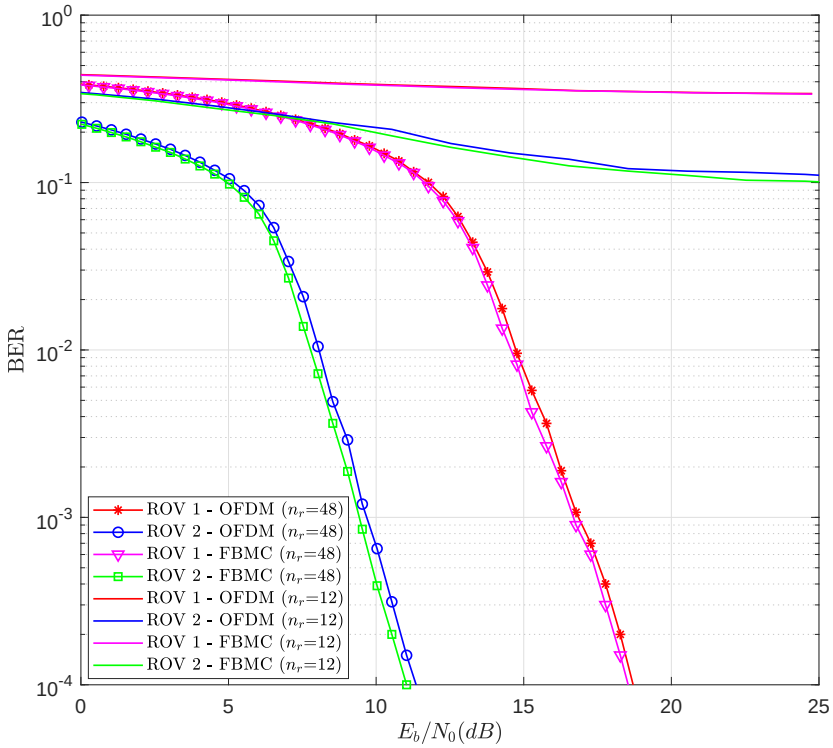


Figure A.4: BER performance of Turbo-coded NOMA OFDM-OQAM and NOMA-OFDM systems in the UAC.

respectively. For comparison purposes only, the BER performances of 4×12 NOMA systems in the same channel and using the same parameters as the 4×48 systems are shown in Figure A.4. This type of MIMO configuration is often used for SU UWA communication (e.g., [22]). In order to better represent the real-world UAC, coloured noise is used in the simulations instead of AWGN. Packet-based transmission is considered for both the OFDM-OQAM and OFDM-based NOMA systems. For the modified OFDM-OQAM system (See Section 5.3), a Hermite-based prototype filter with an overlapping factor of 4 is considered. Also, the time spreading can cause interference between packets and hence a guard (zero) time-slot is included between each packet. A spreading factor of 16 FBMC (real) symbols is considered.

It can be observed from Figure A.4 that for both ROVs, the Turbo-coded 4×12 NOMA systems achieve a high error rate. This justifies the use of a larger receiving array (n_r=48) for the scenario of interest. Furthermore, the SINR for each user (ROV) can be improved when the number of receiving hydrophones is increased [98]. In Figure A.4 and Figure A.5, the FBMC system achieves a marginal improvement over the OFDM system. Considering the 4×48 NOMA-FBMC system in Figure A.4, ROV-2 achieves around 7.5 dB better performance than ROV-1 at a BER of 10^{-4}. It can be inferred that although ROV-2 suffers from weaker channel conditions, the BS can successfully decode its message signal using SIC as the interference from ROV-1 can be compensated for using a greater power allocation

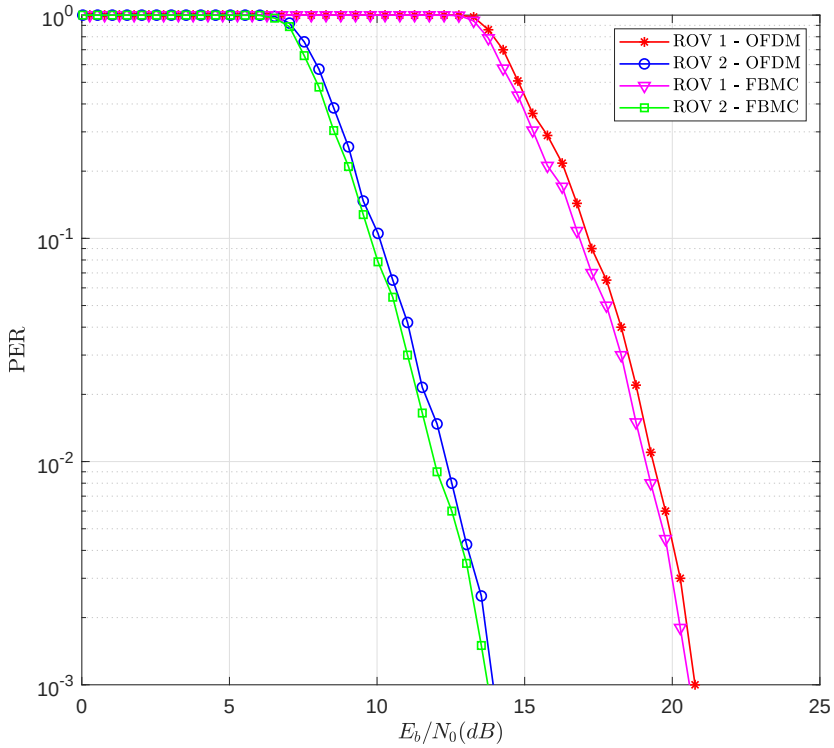


Figure A.5: PER performance of Turbo-coded NOMA OFDM-OQAM and NOMA-OFDM systems in the UAC.

for ROV-2. Using the PER, the maximum achievable bit rate for the two ROVs can be calculated as [168]

$$\text{bit rate} = \frac{(1-\text{PER}) \times N_{\text{bits/packet}}}{T_{\text{packet}}(s)}, \tag{A.9}$$

where $N_{\text{bits/packet}}$ represents the number of bits per packet and T_{packet} is the packet duration. Considering the simulation parameters in Table A.1 and the PER performance in Figure A.5, the maximum throughput that can be achieved for the two ROVs using OFDM and OFDM-OQAM signalling is shown in Figure A.6. The guard time-slots which are inserted in the modified OFDM-OQAM system cause a bandwidth efficiency loss of $\frac{1}{N+1}$. Despite this loss, it can be observed from Figure A.6 that for both ROVs, the improvement in bit rate with NOMA OFDM-OQAM (125.5 kbps) compared to NOMA-OFDM (88.7 kbps) is approximately 41.5%.

A.4 Conclusion

This work examined the feasibility of the NOMA technique for MU UWA communication whereby all ROVs utilised the same available spectrum resources to transmit data simultaneously in the uplink to a surface BS. The BER, PER and maximum bit rate performances of Turbo-coded NOMA-OFDM

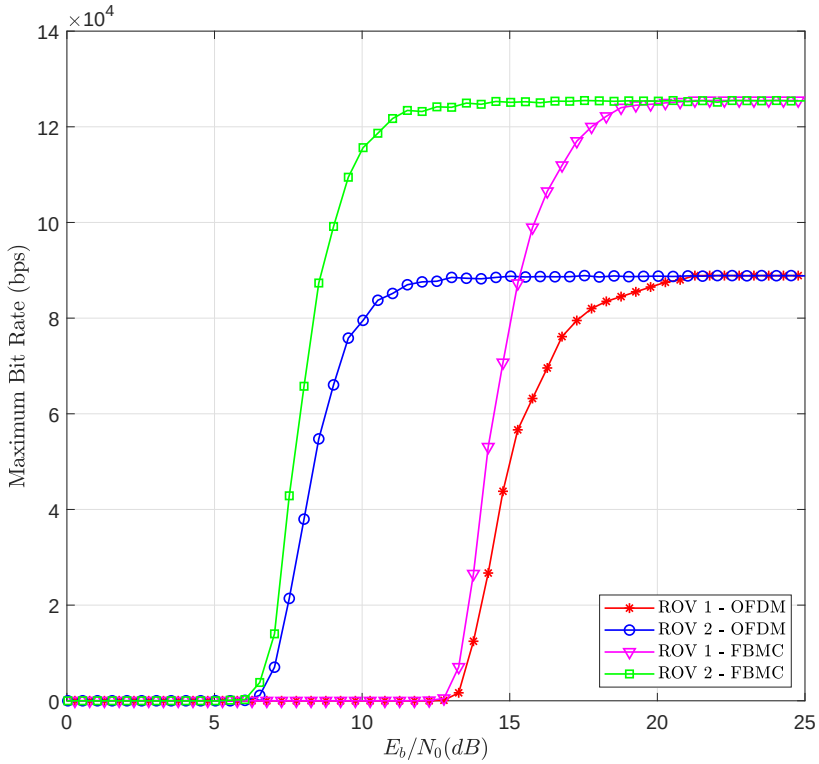


Figure A.6: Maximum achievable bit rates for the two ROVs in the UAC.

and NOMA-FBMC systems were evaluated for a 2-user scenario over a 1 km vertically-configured time-varying UAC. The FBMC system was implemented using the time spreading method as in Chapter 5 to make the application of MIMO straightforward and keep the system's complexity low. It was shown that while both the NOMA-OFDM and NOMA-FBMC systems showed comparable performance in terms of BER and PER, the MIMO NOMA-FBMC system however achieved higher bit rates than the OFDM-based NOMA system. More specifically, using a signal bandwidth of 25 kHz, 256 subcarriers, 16-QAM modulation and rate-$\frac{1}{3}$ Turbo code, the 4×48 NOMA OFDM-OQAM system achieved a bit rate of 125.5 kbps as compared to 88.72 kbps for the 4×48 NOMA OFDM system over the 1 km vertical UAC. The performance of 4×12 MIMO-NOMA systems were also evaluated for the 2-user scenario. However, they achieved a high error rate even with the use of Turbo-codes, justifying the use of a larger receiving array for this particular scenario. If the surface BS does not have any specific requirements on the number of hydrophones it can accommodate, then having more receiving hydrophones is desirable since the SINR for each transmitting ROV can be consequently improved. It could be inferred that although one ROV suffered from less favourable channel conditions, the BS could successfully decode its message signal using the SIC algorithm as it was allocated a greater power allocation, thus minimising the interference from the ROV in better channel conditions. By using the power domain NOMA technique, a good trade-off between user fairness and system throughput was achieved for the ROVs which were

located in different channel conditions. It is to be noted that for the NOMA technique to work properly, there should be sufficient channel gain difference between the ROVs. Else, error which occurs during the SIC decoding of one ROV may cause the decoding of other ROVs' signals to become erroneous. This downside usually limits the number of users in a given system. However, for UWA communication this may not be a problem as the number users (ROVs) will be much less than in a RF terrestrial system and they are expected to be deployed at different locations for a given application.

BIBLIOGRAPHY

[1] NOAA, "How deep is the ocean?" https://oceanservice.noaa.gov/facts/oceandepth.html, Jun. 2018, (Accessed on 10/16/2018).

[2] J. Liang, "Ahmed Gabr Breaks Scuba Diving World Record ," https://www.deeperblue.com/ahmed-gabr-breaks-scuba-diving-world-record/, Sep. 2014, (Accessed on 10/16/2018).

[3] H. Esmaiel and D. Jiang, "Review article: Multicarrier communication for underwater acoustic channel," *Int'l J. of Communications, Network and System Sciences*, vol. 6, no. 08, p. 361, 2013.

[4] T. Melodia, H. Kulhandjian, L.-C. Kuo, and E. Demirors, "Advances in underwater acoustic networking," *Mobile ad hoc networking: Cutting edge directions*, vol. 852, 2013.

[5] M. Doniec, A. Xu, and D. Rus, "Robust real-time underwater digital video streaming using optical communication," in *Proc. IEEE Int. Conf. Robotics and Automation*, May 2013, pp. 5117–5124.

[6] H. Kaushal and G. Kaddoum, "Underwater optical wireless communication," *IEEE Access*, vol. 4, pp. 1518–1547, 2016.

[7] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 84–89, Jan. 2009.

[8] C. Pelekanakis, M. Stojanovic, and L. Freitag, "High rate acoustic link for underwater video transmission," in *OCEANS 2003. Proceedings*, vol. 2, Sep. 2003, pp. 1091–1097 Vol.2.

[9] Y. Zhang, S. Negahdaripour, and Q. Li, "Error-resilient coding for underwater video transmission," in *OCEANS 2016 MTS/IEEE Monterey*. IEEE, 2016, pp. 1–7.

[10] A. M. Demirtas *et al.*, "Performance of H.264 with isolated bit error: Packet decode or discard?" in *2011 18th IEEE International Conference on Image Processing*, Sept 2011, pp. 949–952.

[11] M. Suzuki, K. Nemoto, T. Tsuchiya, and T. Nakanishi, "Digital acoustic telemetry of color video information," in *OCEANS '89. Proceedings*, vol. 3, Sept 1989, pp. 893–896.

[12] G. Ayela and J. Coudeville, "TIVA: A long range, high baud rate image/data acoustic transmission system for underwater applications," *Underwater Defense Technology Conference*, Paris, France, 1991.

[13] J. Gomes, V. Barroso, G. Ayela, and P. Coince, "An overview of the ASIMOV acoustic communication system," in *OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings (Cat. No.00CH37158)*, vol. 3, 2000, pp. 1633–1637 vol.3.

[14] J. Kojima, T. Ura, H. Ando, and K. Asakawa, "High-speed acoustic data link transmitting moving pictures for autonomous underwater vehicles," in *Proceedings of the 2002 Interntional Symposium on Underwater Technology (Cat. No.02EX556)*, 2002, pp. 278–283.

[15] M. Stojanovic, "Low complexity OFDM detector for underwater acoustic channels," in *Proc. OCEANS 2006*, Sep. 2006, pp. 1–6.

[16] J. Ribas *et al.*, "Underwater wireless video transmission for supervisory control and inspection using acoustic OFDM," in *OCEANS 2010*, Sep. 2010, pp. 1–9.

[17] L. D. Vall *et al.*, "Towards underwater video transmission," in *Proceedings of the Sixth ACM International Workshop on Underwater Networks*. ACM, 2011, p. 4.

[18] K. Pelekanakis and A. B. Baggeroer, "Exploiting space -time -frequency diversity with MIMO-OFDM for underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 36, no. 4, pp. 502–513, Oct. 2011.

[19] E. V. Zorita and M. Stojanovic, "Space-frequency block coding for underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 2, pp. 303–314, Apr. 2015.

[20] S. Roy, T. M. Duman, V. McDonald, and J. G. Proakis, "High-rate communication for underwater acoustic channels using multiple transmitters and space time coding: Receiver structures and experimental results," *IEEE Journal of Oceanic Engineering*, vol. 32, no. 3, pp. 663–688, Jul. 2007.

[21] B. Li and M. Stojanovic, "A simple design for joint channel estimation and data detection in an Alamouti OFDM system," in *Proc. OCEANS 2010 MTS/IEEE SEATTLE*, Sep. 2010, pp. 1–5.

[22] B. Li *et al.*, "MIMO-OFDM for high-rate underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 34, no. 4, pp. 634–644, Oct. 2009.

[23] M. Stojanovic, "MIMO OFDM over underwater acoustic channels," in *Proc. Systems and Computers 2009 Conf. Record of the Forty-Third Asilomar Conf. Signals*, Nov. 2009, pp. 605–609.

[24] J. Ling, X. Tan, T. Yardibi, J. Li, M. L. Nordenvaad, H. He, and K. Zhao, "On bayesian channel estimation and FFT-based symbol detection in MIMO underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 39, no. 1, pp. 59–73, Jan. 2014.

[25] H. Ochi, Y. Watanabe, T. Shimura, and T. Hattori, "The acoustic communication experiment at 1,600 m depth using QPSK and 8PSK," in *OCEANS 2010 MTS/IEEE SEATTLE*, Sept 2010, pp. 1–5.

[26] A. Song, M. Badiey, V. K. McDonald, and T. C. Yang, "Time reversal receivers for high data rate acoustic multiple-input -multiple-output communication," *IEEE Journal of Oceanic Engineering*, vol. 36, no. 4, pp. 525–538, Oct. 2011.

[27] A. D. Bowen, D. R. Yoerger, C. Taylor, R. McCabe, J. Howland, D. Gomez-Ibanez, J. C. Kinsey, M. Heintz, G. McDonald, D. B. Peters, B. Fletcher, C. Young, J. Buescher, L. L. Whitcomb, S. C. Martin, S. E. Webster, and M. V. Jakuba, "The Nereus hybrid underwater robotic vehicle for global ocean science operations to 11,000m depth," in *OCEANS 2008*, Sept 2008, pp. 1–10.

[28] D. Levin. (2013, Jan.) Wireless communication system for real-time rov control. [Online]. Available: http://www.oedigital.com/production/item/1539-wireless-communication-system-for-real-time-rov-control

[29] A. Ben-Artzi, "Startup develops underwater acoustic modem," Jan. 2008. [Online]. Available: https://www.eetimes.com/document.asp?doc_id=1307616

[30] Underwater wireless acoustic video communications channel. [Online]. Available: http://www.baltrobotics.com/index.php/products-services-mnu/item/269-underwater-wireless-acoustic-video-communications-channel

[31] Sonardyne International Ltd, https://www.sonardyne.com/, (Accessed on 09/15/2019).

[32] Aquatec Group Ltd, http://www.aquatecgroup.com/, (Accessed on 09/15/2019).

[33] Tritech - Outstanding Performance in Underwater Technology, https://www.tritech.co.uk/, (Accessed on 09/15/2019).

[34] University of York, "Underwater networks," https://www.york.ac.uk/electronic-engineering/research/communication-technologies/underwater-networks/, (Accessed on 09/15/2019).

[35] Newcastle University, "Underwater Communications," https://www.ncl.ac.uk/engineering/research/eee/isc/comms/underwater-communications/, (Accessed on 09/15/2019).

[36] Newcastle University, "Improving Underwater Communications," https://www.ncl.ac.uk/research/impact/casestudies/underwater/#discovermore, (Accessed on 09/15/2019).

[37] E. Demirors, G. Sklivanitis, G. E. Santagati, T. Melodia, and S. N. Batalama, "A high-rate software-defined underwater acoustic modem with real-time adaptation capabilities," *IEEE Access*, vol. 6, pp. 18 602–18 615, 2018.

[38] P. J. Beaujean, E. A. Carlson, J. Spruance, and D. Kriel, "Hermes - a high-speed acoustic modem for real-time transmission of uncompressed image and status transmission in port environment and very shallow water," in *OCEANS 2008*, Sept 2008, pp. 1–9.

[39] "Evologics s2c datasheet," https://www.evologics.de/files/DataSheets/EvoLogics_S2CM-HS-Product_Information.pdf, (Accessed on 10/10/2018).

[40] "LinkQuest," http://www.link-quest.com/html/uwm2200.htm, (Accessed on 10/10/2018).

[41] "S2CR 48/78," https://www.evologics.de/files/DataSheets/EvoLogics_S2CR_4878_Product_Information.pdf, (Accessed on 10/10/2018).

[42] "MATS 3G specifications," http://www.sercel.com/products/Lists/ProductSpecification/Mats3G_specifications_Sercel_EN.pdf, (Accessed on 10/10/2018).

[43] M. J. Bocus, A. Doufexi, and D. Agrafiotis, "MU-massive MIMO for UWA communication," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, August 2018, pp. 1–5.

[44] M. J. Bocus, D. Agrafiotis, and A. Doufexi, "Non-orthogonal multiple access (NOMA) for underwater acoustic communication," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, August 2018, pp. 1–5.

[45] M. J. Bocus, D. Agrafiotis, and A. Doufexi, "Underwater acoustic video transmission using MIMO-FBMC," in *2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*, May 2018, pp. 1–6.

[46] M. J. Bocus, D. Agrafiotis, and A. Doufexi, "Real-time video transmission using massive MIMO in an underwater acoustic channel," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2018, pp. 1–6.

[47] M. J. Bocus, A. Doufexi, and D. Agrafiotis, "Performance evaluation of MIMO-OFDM/OQAM in time-varying underwater acoustic channels," in *OCEANS 2017 - Anchorage*, Sept 2017, pp. 1–6.

[48] M. J. Bocus, A. Doufexi, and D. Agrafiotis, "Performance evaluation of filterbank multicarrier systems in an underwater acoustic channel," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2016, pp. 1–6.

[49] M. J. Bocus, D. Agrafiotis, and A. Doufexi, "Investigating video transmission using OFDM-based MIMO OTFS systems for UWA communication," in *IEEE Journal of Oceanic Engineering*, 2019.

[50] M. J. Bocus, A. Doufexi, and D. Agrafiotis, "On the performance of OFDM-based massive MIMO OTFS systems for UWA communication," in *IET Communications*, 2019.

[51] M. E. Khedr, "Wireless communication fading channel overview," 2008. [Online]. Available: http://webmail.aast.edu/~khedr/Courses/Graduate/Wireless%20Communications_F08/Lecture%20four%20channel%20II.pdf

[52] A. Grami, "Chapter 12 - wireless communications," in *Introduction to Digital Communications*, A. Grami, Ed. Boston: Academic Press, 2016, pp. 493 – 527. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780124076822000120

[53] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: research challenges," *Ad hoc networks*, vol. 3, no. 3, pp. 257–279, 2005.

[54] S. Zhou and Z. Wang, *OFDM for Underwater Acoustic Communications*, 1st ed. Wiley Publishing, 2014.

[55] W. H. Thorp, "Analytic description of the low-frequency attenuation coefficient," *The Journal of the Acoustical Society of America*, vol. 42, no. 1, pp. 270–270, 1967.

[56] F. Fisher and V. Simmons, "Sound absorption in sea water," *The Journal of the Acoustical Society of America*, vol. 62, no. 3, pp. 558–564, 1977.

[57] M. C. Domingo, "Overview of channel models for underwater wireless communication networks," *Physical Communication*, vol. 1, no. 3, pp. 163–182, 2008.

[58] R. E. Francois and G. R. Garrison, "Sound absorption based on ocean measurements: Part i: Pure water and magnesium sulfate contributions," *The Journal of the Acoustical Society of America*, vol. 72, no. 3, pp. 896–907, 1982. [Online]. Available: https://doi.org/10.1121/1.388170

[59] R. E. Francois and G. R. Garrison, "Sound absorption based on ocean measurements. part ii: Boric acid contribution and equation for total absorption," *The Journal of the Acoustical Society of America*, vol. 72, no. 6, pp. 1879–1890, 1982. [Online]. Available: https://doi.org/10.1121/1.388673

[60] K. V. Mackenzie, "Nine-term equation for sound speed in the oceans," *The Journal of the Acoustical Society of America*, vol. 70, no. 3, pp. 807–812, 1981. [Online]. Available: https://doi.org/10.1121/1.386920

[61] A. Zielinski, R. Coates, L. Wang, and A. Saleh, "High rate shallow water acoustic communication," in *Proceedings of OCEANS '93*, Oct 1993, pp. III432–III437 vol.3.

[62] A. Zielinski, Y.-H. Yoon, L. Wu *et al.*, "Performance analysis of digital acoustic communication in a shallow water channel," *IEEE journal of Oceanic Engineering*, vol. 20, no. 4, pp. 293–299, 1995.

[63] R. Coates, "An empirical formula for computing the Beckmann-Spizzichino surface reflection loss coefficient," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 35, no. 4, pp. 522–523, July 1988.

[64] P. Qarabaqi and M. Stojanovic, "Statistical modeling of a shallow water acoustic communication channel," in *Proc. Underwater Acoustic Measurements Conference, Nafplion, Greece*. Citeseer, 2009, pp. 1341–1350.

[65] M. Stojanovic, "Underwater acoustic communications: Design considerations on the physical layer," in *Wireless on Demand Network Systems and Services, 2008. WONS 2008. Fifth Annual Conference on*. IEEE, 2008, pp. 1–10.

[66] M. Stovanovic and L. Freitag, "Integrated doppler tracking and efficient resampling for phase coherent acoustic communication," *IEEE of Oceanic Engineering*, 1994.

[67] M. Caley and A. Duncan, "Investigation of underwater acoustic multi-path doppler and delay spreading in a shallow marine environment," *Acoustics Australia*, vol. 41, no. 1, pp. 20–28, 2013.

[68] M. L. Seto, *Marine robot autonomy.* Springer Science & Business Media, 2012.

[69] L. Wan, Z. Wang, S. Zhou, T. Yang, and Z. Shi, "Performance comparison of Doppler scale estimation methods for underwater acoustic OFDM," *Journal of Electrical and Computer Engineering*, vol. 2012, p. 1, 2012.

[70] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts," *IEEE Journal of Oceanic Engineering*, vol. 33, no. 2, pp. 198–209, Apr. 2008.

[71] H. Li, Z. D. Deng, and T. J. Carlson, "Piezoelectric materials used in underwater acoustic transducers," *Sensor Letters*, vol. 10, no. 3-4, pp. 679–697, 2012.

[72] B. Benson, Y. Li, B. Faunce, K. Domond, D. Kimball, C. Schurgers, and R. Kastner, "Design of a low-cost underwater acoustic modem," *IEEE Embedded Systems Letters*, vol. 2, no. 3, pp. 58–61, 2010.

[73] International Transducer Corporation, "Deep Water Omnidirectional Transducer - Model ITC-1032," https://docs.wixstatic.com/ugd/8aa48d_4a88b62a500647d5be95a99d53c43520.pdf, (Accessed on 09/17/2019).

[74] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 4, pp. 701–717, Oct. 2013.

[75] "Ray Models," http://oalib.hlsresearch.com/Rays/, (Accessed on 10/10/2018).

[76] F. X. Socheleau *et al.*, "A maximum entropy framework for statistical modeling of underwater acoustic communication channels," in *Proc. OCEANS 2010 IEEE - Sydney*, May 2010, pp. 1–7.

[77] J. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standardsóincluding high efficiency video coding (HEVC)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1669–1684, Dec 2012.

[78] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, Dec 2012.

[79] M. Wien, *High Efficiency Video Coding - Coding Tools and Specification*. Springer Berlin, Heidelberg, Oct. 2014.

[80] L. D. Vall, "Towards underwater video transmission," Master's thesis, Massachusetts Institute of Technology, 2011.

[81] W. Gao and S. Ma, *Video Coding Basic Principle*. Springer, Jan. 2014, ch. 2, pp. 17–34.

[82] B. Juurlink, M. Alvarez-Mesa, C. C. Chi, A. Azevedo, C. Meenderinck, and A. Ramirez, "Understanding the application: An overview of the H. 264 standard," in *Scalable Parallel Programming Applied to H. 264/AVC Decoding*. Springer, 2012, pp. 5–15.

[83] D. Nagamalai, E. Renault, and M. Dhanuskodi, *Advances in Digital Image Processing and Information Technology: First International Conference on Digital Image Processing and Pattern Recognition*, 1st ed. Springer Publishing Company, Incorporated, Sep. 2011.

[84] M. A. Shaikh and S. S. Badnerkar, "Video compression algorithm using motion compensation technique," in *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, vol. 3, no. 6, Jun. 2014.

[85] Y. Ye and M. Karczewicz, "Improved h.264 intra coding based on bi-directional intra prediction, directional transform, and adaptive coefficient scanning," in *2008 15th IEEE International Conference on Image Processing*, Oct 2008, pp. 2116–2119.

[86] *Cisco Video and TelePresence Architecture Design Guide*, Cisco Systems, Inc., Mar. 2012. [Online]. Available: https://www.cisco.com/c/en/us/td/docs/voice_ip_comm/uc_system/design/guides/videodg/vidguide.pdf

[87] M. Shahid, "Methods for objective and subjective video quality assessment and for speech enhancement," Ph.D. dissertation, Department of Applied Signal Processing, Blekinge Institute of Technology, Sweden, Dec. 2014.

[88] B. Bing, *Next-Generation Video Coding and Streaming*. John Wiley & Sons, Inc, Sep. 2015.

[89] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electronics Letters*, vol. 44, no. 13, pp. 800–801, June 2008.

[90] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[91] *Advanced video coding for generic audiovisual services*, Rec. ITU-T H.264 and ISO/IEC 14496-10 Std., Rev. 12.0, Apr. 2017. [Online]. Available: https://www.itu.int/rec/T-REC-H.264-201704-I/en

[92] D. Marpe, T. Wiegand, and G. J. Sullivan, "The H.264/MPEG4 advanced video coding standard and its applications," *IEEE Communications Magazine*, vol. 44, no. 8, pp. 134–143, Aug 2006.

[93] P. Lambert, W. D. Neve, Y. Dhondt, and R. V. de Walle, "Flexible macroblock ordering in H.264/AVC," *Journal of Visual Communication and Image Representation*, vol. 17, no. 2, pp. 358 – 375, 2006, introduction: Special Issue on emerging H.264/AVC video coding standard. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1047320305000623

[94] B. Hosking, D. Agrafiotis, D. R. Bull, and N. Easton, "Spatial resampling of IDR frames for low bitrate video coding with HEVC," in *Visual Information Processing and Communication*, San Francisco, California, USA, Feb. 2015. [Online]. Available: https://doi.org/10.1117/12.2083582

[95] A. Atayero, O. I. Sheluhin, and Y. A. Ivanov, *Modeling, Simulation and Analysis of Video Streaming Errors in Wireless Wideband Access Networks*. Springer, 09 2013, vol. 170, ch. 2, pp. 15–28.

[96] A. Aldahdooh *et al.*, "Spatio-temporal error concealment technique for high order multiple description coding schemes including subjective assessment," in *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, June 2016, pp. 1–6.

[97] Q. He and A. Schmeink, "Comparison and evaluation between FBMC and OFDM systems," in *WSA 2015; 19th International ITG Workshop on Smart Antennas; Proceedings of*, Mar. 2015, pp. 1–7.

[98] A. Aminjavaheri and B. Farhang-Boroujeny, "UWA massive MIMO communications," in *OCEANS 2015 - MTS/IEEE Washington*, Oct 2015, pp. 1–6.

[99] M. Alard, "Construction of a multicarrier signal," Aug. 21 2001, US Patent 6,278,686.

[100] R. Haas and J.-C. Belfiore, "A time-frequency well-localized pulse for multiple carrier transmission," *Wireless personal communications*, vol. 5, no. 1, pp. 1–18, 1997.

[101] R. Chang, "High-speed multichannel data transmission with bandlimited orthogonal signals," *Bell Sys. Tech. J*, vol. 45, no. 10, pp. 1775–1796, 1966.

[102] B. Saltzberg, "Performance of an efficient parallel data transmission system," *IEEE Transactions on Communication Technology*, vol. 15, no. 6, pp. 805–811, Dec. 1967.

[103] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 92–112, May 2011.

[104] M. Bellanger and J. Daguet, "TDM-FDM transmultiplexer: Digital polyphase and fft," *IEEE Transactions on Communications*, vol. 22, no. 9, pp. 1199–1205, September 1974.

[105] B. Hirosaki, "An orthogonally multiplexed QAM system using the discrete Fourier transform," *IEEE Transactions on Communications*, vol. 29, no. 7, pp. 982–989, July 1981.

[106] J. Gomes and M. Stojanovic, "Performance analysis of filtered multitone modulation systems for underwater communication," in *OCEANS 2009, MTS/IEEE Biloxi-Marine Technology for Our Future: Global and Local Challenges*. IEEE, 2009, pp. 1–9.

[107] P. Amini *et al.*, "Filterbank multicarrier communications for underwater acoustic channels," *Oceanic Engineering, IEEE Journal of*, vol. 40, no. 1, pp. 115–130, 2015.

[108] A. Aminjavaheri *et al.*, "Frequency spreading Doppler scaling compensation in underwater acoustic multicarrier communications," in *Communications (ICC), 2015 IEEE International Conference on*, Jun. 2015, pp. 2774–2779.

[109] J. G. Andrews, "Chapter 4: Orthogonal frequency division multiplexing (OFDM)," Mar. 2006. [Online]. Available: https://pdfs.semanticscholar.org/5517/85bcbf6e109e798324aba30686fec0d5853a.pdf

[110] M. Bellanger, "FS-FBMC: A flexible robust scheme for efficient multicarrier broadband wireless access," in *2012 IEEE Globecom Workshops*, Dec 2012, pp. 192–196.

[111] D. Mattera, M. Tanda, and M. Bellanger, "Frequency-spreading implementation of OFDM/OQAM systems," in *2012 International Symposium on Wireless Communication Systems (ISWCS)*, Aug 2012, pp. 176–180.

[112] M. Aldababseh and A. Jamoos, "Estimation of FBMC/OQAM fading channels using dual Kalman filters," *The Scientific World Journal*, vol. 2014, 2014.

[113] M. Bellanger, D. Le Ruyet, D. Roviras, M. Terré, J. Nossek, L. Baltar, Q. Bai, D. Waldhauser, M. Renfors, T. Ihalainen *et al.*, "FBMC physical layer: a primer," *PHYDYAS, January*, 2010.

[114] A. Viholainen, M. Bellanger, and M. Huchard, "Prototype filter and structure optimization," *website: www. ict-phydyas. org: Document D*, vol. 5, 2009.

[115] J. Louveaux *et al.*, "Equalization and demodulation in the receiver (single antenna)," *ICT-211887 PHYDYAS deliverable D*, vol. 3, 2008.

[116] I. Berenguer and I. J. Wassell, "FMT modulation: receiver filter bank definition for the derivation of an efficient implementation," in *Proc. 7th International OFDM Workshop*, 2002, pp. 158–162.

[117] J. Du and S. Signell, "Classic OFDM systems and pulse shaping OFDM/OQAM systems," 2007.

[118] H. G. Feichtinger and T. Strohmer, Eds., *Gabor Analysis and Algorithms: Theory and Applications*, 1st ed. Secaucus, NJ, USA: Birkhauser Boston, Inc., 1997.

[119] R. Nissel, "Filter bank multicarrier modulation for future wireless systems," Ph.D. dissertation, Dissertation, TU Wien, 2017.

[120] R. Nissel, S. Schwarz, and M. Rupp, "Filter bank multicarrier modulation schemes for future mobile communications," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 8, pp. 1768–1782, Aug 2017.

[121] R. Nissel and M. Rupp, "Enabling low-complexity MIMO in FBMC-OQAM," in *2016 IEEE Globecom Workshops (GC Wkshps)*, Dec 2016, pp. 1–6.

[122] "LTE; evolved universal terrestrial radio access (E-UTRA); multiplexing and channel coding," 3GPP TS 36.212 version 14.2.0 Release 14, Tech. Rep., Apr. 2017.

[123] M. Synthia and M. S. Ali, "Performance study of Turbo code with interleaver design," *International Journal of Scientific & Engineering Research*, vol. 2, no. 7, 2011.

[124] T. Duman and A. Ghrayeb, *Coding for MIMO Communication Systems*. John Wiley & Sons, Ltd., Oct. 2007.

[125] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct 1998.

[126] T. M. A.-Q. Shawbow Abdulkarim Nizamaldeen, "A comparison between STBC-OFDM and SFBC-OFDM," *International Journal of Arts & Sciences*, vol. 5, no. 4, 2012.

[127] C. Lélé *et al.*, "Channel estimation methods for preamble-based OFDM/OQAM modulations," *European Transactions on Telecommunications*, vol. 19, pp. 741–750, 2008.

[128] C. Lele, P. Siohan, and R. Legouable, "2 dB better than CP-OFDM with OFDM/OQAM for preamble-based channel estimation," in *Proc. IEEE Int. Conf. Communications*, May 2008, pp. 1302–1306.

[129] E. Kofidis and D. Katselis, "Preamble-based channel estimation in MIMO-OFDM/OQAM systems," in *Proc. IEEE Int Signal and Image Processing Applications (ICSIPA) Conf*, Nov. 2011, pp. 579–584.

[130] R. Nissel and M. Rupp, "On pilot-symbol aided channel estimation in FBMC-OQAM," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 3681–3685.

[131] E. Kofidis, D. Katselis, A. Rontogiannis, and S. Theodoridis, "Preamble-based channel estimation in OFDM/OQAM systems: A review," *Signal Processing*, vol. 93, no. 7, pp. 2038–2054, 2013.

[132] J. Du and S. Signell, "Novel preamble-based channel estimation for OFDM/OQAM systems," in *Proc. IEEE Int. Conf. Communications*, Jun. 2009, pp. 1–6.

[133] R. P. Hebbar and P. G. Poddar, "Generalized frequency division multiplexing for acoustic communication in underwater systems," in *Proc. and Communications (CCUBE) 2017 Int. Conf. Circuits, Controls*, Dec. 2017, pp. 86–90.

[134] G. Palou and M. Stojanovic, "Underwater acoustic MIMO OFDM: An experimental analysis," in *OCEANS 2009*, Oct 2009, pp. 1–8.

[135] X. Zhao and D. Pompili, "AMMCA: Acoustic massive MIMO with carrier aggregation to boost the underwater communication data rate," in *Proceedings of the 10th International Conference on Underwater Networks & Systems*. ACM, 2015, p. 5.

[136] R. Nissel and M. Rupp, "Enabling low-complexity MIMO in FBMC-OQAM," in *2016 IEEE Globecom Workshops (GC Wkshps)*, Dec 2016, pp. 1–6.

[137] T. L. Marzetta, "Massive MIMO: An introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015.

[138] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, Nov. 2016, ch. 1, pp. 5–18.

[139] C. Masterson, "Massive MIMO and beamforming: The signal processing behind the 5G buzzwords," *Analog Dialogue*, vol. 51, no. 06, p. 10, Jun. 2017.

[140] F. A. de Figueiredo, J. P. Miranda, F. L. Figueiredo, and F. A. Cardoso, "Uplink performance evaluation of massive MU-MIMO systems," *arXiv preprint arXiv:1503.02192*, 2015.

[141] H. Q. Ngo, "Massive MIMO: Fundamentals and system designs," Ph.D. dissertation, Division of Communication Systems, Department of Electrical Engineering, Linköping University, 2015.

[142] E. Björnson, "Channel hardening makes fading channels behave as deterministic," Jan. 2017. [Online]. Available: https://ma-mimo.ellintech.se/2017/01/25/channel-hardening-makes-fading-channels-behave-as-deterministic/

[143] R. Nissel, J. Blumenstein, and M. Rupp, "Block frequency spreading: A method for low-complexity MIMO in FBMC-OQAM," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Sapporo, Japan*, 2017.

[144] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple access techniques for 5G wireless networks and beyond*. Springer, 2019.

[145] X. Li, E. Björnson, S. Zhou, and J. Wang, "Massive MIMO with multi-antenna users: When are additional user antennas beneficial?" in *Proc. 23rd Int. Conf. Telecommunications (ICT)*, May 2016, pp. 1–6.

[146] A. Farhang, N. Marchetti, L. E. Doyle, and B. Farhang-Boroujeny, "Filter bank multicarrier for massive MIMO," in *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*. IEEE, 2014, pp. 1–7.

[147] A. Farhang, N. Marchetti, F. Figueiredo, and J. P. Miranda, "Massive MIMO and waveform design for 5th generation wireless communication systems," in *1st International Conference on 5G for Ubiquitous Connectivity*, Nov 2014, pp. 70–75.

[148] P. Raviteja, K. T. Phan, Y. Hong, and E. Viterbo, "Interference cancellation and iterative detection for orthogonal time frequency space modulation," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6501–6515, Oct 2018.

[149] A. Farhang, A. RezazadehReyhani, L. E. Doyle, and B. Farhang-Boroujeny, "Low complexity modem structure for OFDM-based orthogonal time frequency space modulation," *CoRR*, vol. abs/1710.00655, 2017. [Online]. Available: http://arxiv.org/abs/1710.00655

[150] R. Hadani and A. Monk, "OTFS: A new generation of modulation addressing the challenges of 5G," *CoRR*, vol. abs/1802.02623, 2018. [Online]. Available: http://arxiv.org/abs/1802.02623

[151] A. Monk, R. Hadani, M. Tsatsanis, and S. Rakib, "OTFS - orthogonal time frequency space," *CoRR*, vol. abs/1608.02993, 2016. [Online]. Available: http://arxiv.org/abs/1608.02993

[152] R. Hadani, S. Rakib, M. Tsatsanis, A. Monk, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, "Orthogonal time frequency space modulation," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, Mar. 2017, pp. 1–6.

[153] A. Chockalingam, Y. Hong, and E. Viterbo, "Orthogonal time frequency space (OTFS) modulation," http://www.ece.iisc.ernet.in/~achockal/pdf_files/vtc2018fall_otfs_tutorial_presentation_1.pdf, August 2018, (Accessed on 11/27/2018).

[154] A. Farhang, A. RezazadehReyhani, L. E. Doyle, and B. Farhang-Boroujeny, "Low complexity modem structure for OFDM-based orthogonal time frequency space modulation," *IEEE Wireless Communications Letters*, vol. 7, no. 3, pp. 344–347, Jun. 2018.

[155] A. RezazadehReyhani, A. Farhang, M. Ji, R. Chen, and B. Farhang-Boroujeny, "Analysis of discrete-time MIMO OFDM-based orthogonal time frequency space modulation," *CoRR*, vol. abs/1710.07900, 2017. [Online]. Available: http://arxiv.org/abs/1710.07900

[156] Y. Zhou, A. Song, and F. Tong, "Underwater acoustic channel characteristics and communication performance at 85 khz," *The Journal of the Acoustical Society of America*, vol. 142, no. 4, pp. EL350–EL355, 2017.

[157] J. Huang, S. Zhou, J. Huang, J. Preisig, L. Freitag, and P. Willett, "Progressive MIMO-OFDM reception over time-varying underwater acoustic channels," in *Proc. Systems and Computers 2010 Conf. Record of the Forty Fourth Asilomar Conf. Signals*, Nov. 2010, pp. 1324–1329.

[158] J. A. Smith, "Doppler effects on single-carrier signals operating in fading channels," Ph.D. dissertation, Monterey, California: Naval Postgraduate School, 2009.

[159] K. Sripimanwat, *Turbo Code Applications: A Journey from a Paper to Realization*, 1st ed. Springer Publishing Company Incorporated, 2010.

[160] K. R. Murali and A. Chockalingam, "On OTFS modulation for high-Doppler fading channels," in *Proc. Information Theory and Applications Workshop (ITA)*, Feb. 2018, pp. 1–10.

[161] T. Zemen, M. Hofer, and D. Loeschenbrand, "Low-complexity equalization for orthogonal time and frequency signaling (OTFS)," *CoRR*, vol. abs/1710.09916, 2017. [Online]. Available: http://arxiv.org/abs/1710.09916

[162] P. Raviteja, K. T. Phan, Q. Jin, Y. Hong, and E. Viterbo, "Low-complexity iterative detection for orthogonal time frequency space modulation," *CoRR*, vol. abs/1709.09402, 2017. [Online]. Available: http://arxiv.org/abs/1709.09402

[163] M. K. Ramachandran and A. Chockalingam, "MIMO-OTFS in high-doppler fading channels: Signal detection and channel estimation," *CoRR*, vol. abs/1805.02209, 2018. [Online]. Available: http://arxiv.org/abs/1805.02209

[164] P. Raviteja, K. T. Phan, and Y. Hong, "Embedded pilot-aided channel estimation for OTFS in delay-Doppler channels," 2018.

[165] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Practical pulse-shaping waveforms for reduced-cyclic-prefix OTFS," *IEEE Transactions on Vehicular Technology*, 2018.

[166] S. A. Aliesawi, C. C. Tsimenidis, B. S. Sharif, and M. Johnston, "Iterative multiuser detection for underwater acoustic channels," *IEEE Journal of Oceanic Engineering*, vol. 36, no. 4, pp. 728–744, Oct 2011.

[167] C. He, J. Huang, Z. Yan, and Q. Zhang, "M-ary CDMA multiuser underwater acoustic communication and its experimental results," *Science China Information Sciences*, vol. 54, no. 8, pp. 1747–1755, 2011.

[168] X. Su, H. Yu, W. Kim, C. Choi, and D. Choi, "Interference cancellation for non-orthogonal multiple access used in future wireless mobile networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, p. 231, 2016.

[169] S. Ali, E. Hossain, and D. I. Kim, "Non-orthogonal multiple access (NOMA) for downlink multiuser MIMO systems: User clustering, beamforming, and power allocation," *IEEE Access*, vol. 5, pp. 565–577, 2017.

[170] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, I. Chih-Lin, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185–191, 2017.

[171] F. Al Rabee, K. Davaslioglu, and R. Gitlin, "The optimum received power levels of uplink non-orthogonal multiple access (NOMA) signals," in *Wireless and Microwave Technology Conference (WAMICON), 2017 IEEE 18th*. IEEE, 2017, pp. 1–4.